Tracheal aspirate RNA sequencing identifies distinct immunological features of COVID-19 ARDS

Aartik Sarma1, Stephanie A. Christenson1,47, Ashley Byrne2,47, Eran Mick1,2,3,47, Angela Oliveira Pisco1,47, Catherine DeVoe3, Thomas Deiss2, Rajani Ghale1,3, Beth Shoshana Zha1, Alexandra Tsitsiklis1,3, Alejandra Jauregui1, Farzad Moazed1, Angela M. Detweiler3, Natasha Spottiswoode4, Pratik Sinha5, Norma Neff2, Michelle Tan2, Paula Hayakawa Serpa3, Andrew Willmore1, K. Mark Ansel6,7, Jennifer G. Wilson8, Aleksandra Leligdowicz1,9,10, Emily R. Siegel11, Marina Sirota12, Joseph L. DeRisi2,13, Michael A. Matthay1,10,14, COMET Consortium*, Carolyn M. Hendrickson1, Kirsten N. Kangelaris4, Matthew F. Krummel15, Prescott G. Woodruff1,7, David J. Erle1,4,16,17, Carolyn S. Calfee1,10,14,48 & Charles R. Langelier2,3,48

The immunological features that distinguish COVID-19-associated acute respiratory distress syndrome (ARDS) from other causes of ARDS are incompletely understood. Here, we report the results of comparative lower respiratory tract transcriptional profiling of tracheal aspirate from 52 critically ill patients with ARDS from COVID-19 or from other etiologies, as well as controls without ARDS. In contrast to a “cytokine storm,” we observe reduced proinflammatory gene expression in COVID-19 ARDS when compared to ARDS due to other causes. COVID-19 ARDS is characterized by a dysregulated host response with increased PTEN signaling and elevated expression of genes with non-canonical roles in inflammation and immunity. In silico analysis of gene expression identifies several candidate drugs that may modulate gene expression in COVID-19 ARDS, including dexamethasone and granulocyte colony stimulating factor. Compared to ARDS due to other types of viral pneumonia, COVID-19 is characterized by impaired interferon-stimulated gene (ISG) expression. The relationship between SARS-CoV-2 viral load and expression of ISGs is decoupled in patients with COVID-19 ARDS when compared to patients with mild COVID-19. In summary, assessment of host gene expression in the lower airways of patients reveals distinct immunological features of COVID-19 ARDS.
In its most severe form, coronavirus disease 2019 (COVID-19) can precipitate the acute respiratory distress syndrome (ARDS), which is characterized by low arterial oxygen concentrations, alveolar injury, and a dysregulated inflammatory response in the lungs. Early reports hypothesized that COVID-19 ARDS was driven by a "cytokine storm" based on the detection of higher circulating inflammatory cytokine levels in critically ill COVID-19 patients compared to those with mild disease or healthy controls. Recent studies, however, have found that patients with COVID-19 ARDS have lower plasma cytokine levels compared to those with ARDS due to other causes, highlighting a need to understand the underlying mechanisms of COVID-19 ARDS.

Clinical trials have demonstrated a significant mortality benefit for dexamethasone in COVID-19 patients with ARDS, implicating a role for dysregulated inflammation in COVID-19 pathophysiology given the immunomodulatory effects of corticosteroids. In contrast, clinical trials of corticosteroids for ARDS prior to the SARS-CoV-2 pandemic have had mixed results, ranging from benefit to possible harm. These differences suggest distinct, corticosteroid-responsive biology in COVID-19 ARDS, with important implications for pathogenesis and treatment.

While several studies have assessed host lower respiratory tract gene expression in patients with SARS-CoV-2, none has compared COVID-19 ARDS to other causes of ARDS. Here, we perform this comparison in a prospective cohort of critically ill adults with ARDS from COVID-19 or from other etiologies, as well as controls without ARDS. From RNA sequencing (RNA-seq) of tracheal aspirate (TA), we identify distinct immunologic features of COVID-19 ARDS.

**Results**

We conducted a prospective case-control study of 52 adults requiring mechanical ventilation (Table 1, Supplementary Data 1, Supplementary Fig. 1) for ARDS from COVID-19 (COVID-ARDS, n = 15), ARDS from other etiologies (Other-ARDS, n = 32), or for airway protection in the absence of pulmonary disease (No-ARDS, n = 5). Other ARDS etiologies included pneumonia, aspiration, sepsis, and transfusion reaction. Patients were enrolled at two tertiary care hospitals in San Francisco, California under research protocols approved by the University of California San Francisco Institutional Review Board. We excluded immunosuppressed patients to avoid confounding the measurement of host inflammatory responses. TA was collected within 5 days of intubation and underwent RNA-seq.

We began by comparing TA gene expression between COVID-ARDS and Other-ARDS patients (Fig. 1a, Supplementary Fig. 2, and Supplementary Data 2) and identified 793 differentially expressed genes at a Benjamini–Hochberg false discovery rate (FDR) < 0.1, as well as differentially activated pathways using Ingenuity Pathway Analysis (IPA) (Fig. 1b). The IPA upstream regulator analysis predicted activation of PTEN and inhibition of PI3K in COVID-ARDS versus Other-ARDS patients (Fig. 1c and Supplementary Data 4). Applying this approach to upstream cytokines additionally predicted activation of IFNγ and CNTF, and inhibition of IL-10 in COVID-ARDS versus Other-ARDS patients (Fig. 1c).

Next, we asked whether existing pharmaceuticals could counter the dysregulated gene expression in COVID-19 ARDS by comparing genes that were differentially expressed in COVID-ARDS and No-ARDS groups against the IPA database of 12,981 drug treatment-induced transcriptional signatures derived from human studies and cell culture experiments (Fig. 1d and Supplementary Data 5). Dexamethasone was the compound predicted to most significantly counter-regulate the genes expressed in COVID-ARDS patients compared to No-ARDS patients. This finding was striking given that dexamethasone has a mortality benefit in patients with severe COVID-19 in clinical trials. Granulocyte colony-stimulating factor (G-CSF), which was also found to reduce COVID-19 mortality in a recent clinical trial, was also significant. Other corticosteroids (fluticasone, prednisolone), as well as omega-3 fatty acids (eicosapentaenoic and docosahexaenoic acids), were additionally predicted to shift the transcriptional profile of COVID-ARDS toward No-ARDS controls (Fig. 1d and additional candidates in Supplementary Data 5).

To identify genes that might underlie the established therapeutic benefit of dexamethasone, we examined genes differentially expressed in COVID-ARDS that were also predicted to be regulated by dexamethasone (Supplementary Data 6). Interestingly, both dexamethasone and G-CSF were predicted to modulate the expression of several genes differentially expressed between COVID-ARDS and controls. Many of these genes have established roles in immunity, inflammation, and interferon responses (Supplementary Fig. 3). For instance, both drugs were predicted to inhibit the expression of P2RY14, which regulates interferon-α secretion in plasmacytoid dendritic cells, and mediates chemotaxis of hematopoietic stem cells, EPSTI1, which promotes M1 macrophage polarization, and STAT1, which induces chemokine expression, regulates differentiation of hematopoietic cells and promotes reactive oxygen species production.

Since TA contains a heterogeneous mix of cells from the airways and alveoli, we conducted additional analyses to understand which cells were contributing to the measured gene expression. In silico prediction of cell-type composition demonstrated that monocytes/macrophages and neutrophils were the most abundant cell types in patients with COVID-ARDS as well as Other-ARDS (Fig. 2a, Supplementary Fig. 1, and Supplementary Data 7). While no differences in lymphocyte, macrophage, or neutrophil populations were observed, decreased proportions of type-2 alveolar epithelial cells and increased proportions of goblet cells were found in COVID-ARDS compared to Other-ARDS subjects. This may reflect alveolar epithelial injury, airway remodeling, and/or preferential SARS-CoV-2 infection of cells with the highest expression of ACE2 and TMPRSS2.
Table 1 Clinical and demographic characteristics of patients with ARDS due to COVID-19 (COVID-ARDS), control patients with ARDS due to other etiologies (Other-ARDS), and intubated control patients without ARDS (No-ARDS).

|                | COVID-ARDS | Other-ARDS | P     | No-ARDS | P     |
|----------------|------------|------------|-------|---------|-------|
| N              | 15         | 32         | 0.205 | 5       | 0.190 |
| Age (median [IQR]) | 54.8 [42.5, 67.5] | 61.4 [47.3, 71.5] | 0.005 | 66.2 [62.0, 82.0] | 0.546 |
| Male           | 9 (60.0)   | 20 (62.5)  | 1.000 | 2 (40.0) | 0.795 |
| 30-day mortality | 3 (20.0)   | 11 (34.4)  | 0.508 | 2 (40.0) | 0.511 |
| APACHE III     | 97 [88, 106] | 95 [78, 126] | 0.900 | 51 [50, 71] | 0.011 |
| Days since start of COVID-19 symptoms | 10 [7, 17] | - | - | - | - |
| Duration of hospitalization (days) | 24 [19, 40] | 14 [8, 25] | 0.006 | 7 [6, 7] | 0.002 |
| Race (%)       |            |            |       |         | 0.029 |
| Black          | 0 (0.0)    | 2 (6.2)    | 0.003 | 0 (0.0) | 0.141 |
| Asian          | 3 (20.0)   | 4 (12.5)   | 0.003 | 1 (20.0) | 0.003 |
| White          | 1 (6.7)    | 23 (71.9)  | 0.003 | 3 (60.0) | 0.003 |
| Other          | 11 (73.3)  | 3 (9.4)    | 0.003 | 1 (20.0) | 0.003 |
| Hispanic ethnicity | 8 (53.3) | 3 (9.4) | 0.003 | 0 (0.0) | 0.141 |
| PaO2/FiO2 (median [IQR]) | 74 [60, 115] | 96 [67, 148] | 0.141 | 296 [216, 367] | 0.003 |
| ARDS etiology (%) | 0.109 | 0.001 | <0.001 | <0.001 |<0.001 |
| Aspiration     | 0 (0.0)    | 5 (15.6)   | 0.003 | 0 (0.0) | 0.003 |
| LRTI           | 15 (100.0) | 20 (62.5)  | 0.003 | 0 (0.0) | 0.003 |
| Sepsis         | 0 (0.0)    | 4 (12.5)   | 0.003 | 0 (0.0) | 0.003 |
| Transfusion    | 0 (0.0)    | 2 (6.2)    | 0.003 | 0 (0.0) | 0.003 |
| Unknown        | 0 (0.0)    | 1 (3.1)    | 0.003 | 0 (0.0) | 0.003 |
| None           | 0 (0.0)    | 0 (0.0)    | 0.003 | 5 (100.0) | 0.003 |
| LRTI type (%)  | <0.001     | <0.001     | <0.001 | <0.001 |<0.001 |
| Bacterial      | 0 (0.0)    | 9 (28.1)   | 0.003 | 0 (0.0) | 0.003 |
| Viral          | 8 (60.0)   | 4 (12.5)   | 0.003 | 0 (0.0) | 0.003 |
| Viral + bacterial | 4 (20.0) | 3 (9.4) | 0.003 | 0 (0.0) | 0.003 |
| Viral + viral  | 3 (20.0)   | 0 (0.0)    | 0.003 | 0 (0.0) | 0.003 |
| Unknown        | 0 (0.0)    | 4 (12.5)   | 0.003 | 0 (0.0) | 0.003 |
| None           | 0 (0.0)    | 12 (37.5)  | 0.003 | 5 (100.0) | 0.003 |

IQR interquartile range.
P values represent comparisons versus COVID-ARDS. Reasons for intubation of No-ARDS patients included: hemorrhagic stroke, subdural hematoma, retroperitoneal hemorrhage, or other neurosurgical procedures. Statistical significance was determined using Fisher’s exact test (discrete variables) or by the Wilcoxon test (continuous variables).

To evaluate the immune cell landscape in COVID-19 ARDS more comprehensively, we performed single-cell RNA-seq (scRNAseq) on CD45+ enriched TA specimens from six COVID-ARDS patients (Fig. 2b). Monocytes, macrophages (in particular alveolar macrophages), and neutrophils were the most abundant cell types observed, consistent with bulk deconvolution results and in line with previous scRNAseq analyses of bronchial alveolar lavage (BAL) fluid from patients with severe COVID-19 pneumonia. We additionally observed significant populations of CD4+ and CD8+ T cells, which may interact with macrophages to drive dysregulated inflammatory responses in COVID-19.

ARDS is a heterogeneous syndrome caused by diverse infectious and noninfectious insults. To determine if COVID-ARDS had unique features compared to other types of infection-associated ARDS, we performed secondary analyses comparing the differential gene expression in COVID-ARDS without co-infections (n = 8) to that of ARDS caused exclusively by other viral (n = 4, Fig. 3a and Supplementary Data 8) or bacterial (n = 9, Fig. 3b and Supplementary Data 9) lower respiratory tract infections (LRTI). COVID-ARDS was characterized by lower expression of proinflammatory signaling pathways compared to bacterial LRTI-associated ARDS but higher levels of the same pathways compared to other viral LRTI-associated ARDS (Fig. 3c and Supplementary Data 9).

Although interferon-related gene expression was higher in COVID-ARDS compared to bacterial LRTI and No-ARDS controls, it was markedly attenuated in ARDS patients with COVID-19 versus those with other viral LRTI (Fig. 3d and Supplementary Data 10). Given prior findings of impaired interferon responses in patients with severe COVID-19, we evaluated ISGs more closely by comparing expression levels against SARS-CoV-2 viral load (Supplementary Fig. 6 and Supplementary Data 11). Prior studies found a strong correlation between viral load and ISG expression in the upper respiratory tract of patients with early, mild disease. In contrast, we observed decoupling of this relationship for several ISGs (Supplementary Fig. 6 and Supplementary Data 11).

Discussion

Our results challenge the “cytokine storm” model of COVID-19 ARDS. Instead, we observe a complex picture of host immune dysregulation that includes upregulation of genes with non-canonical roles in inflammation, immunity, and interferon signaling that are predicted to be attenuated by dexamethasone, G-CSF, and other potential therapeutics. Our transcriptomic data suggest that compared to other types of ARDS, COVID-19 ARDS is characterized by increased PTEN, interferon-γ, and CNTF-stimulated gene expression juxtaposed against inhibition of genes typically activated by IL-10. PTEN promotes inflammation in acute lung injury model. CNTF has been found to regulate B-cell differentiation and bind the IL-6 receptor, and IL-10 is a central anti-inflammatory cytokine, suggesting that a combination of inflammatory activation and dysregulated attenuation may drive COVID-19 respiratory pathobiology.

Trials of IL-6 receptor blockade in COVID-19 have had mixed results, with early trials showing no effect, but more recent studies demonstrating a mortality benefit in patients concomitantly receiving corticosteroids. Our analyses focused...
may be an important feature of disease responses in patients with severe COVID-19 pneumonia clinical guidelines, less invasive TA sampling is increasingly infections.

The interferon response may be a common feature of severe viral infections native specimen for transcriptional studies of the lower airways. Transfer RNA-seq datasets (Supplementary Data 12) and similarities in the predicted cellular composition of matched TA and mini-BAL specimens (Supplementary Data 13, Supplementary Fig. 7) also support this idea.

**Methods**

**Study design, clinical cohort, and ethics statement.** We conducted a case-control study of patients with ARDS due to COVID-19 (n = 15) versus two control groups. The first control group included patients with ARDS due to other causes (n = 32) and the second included patients intubated for airway protection without evidence of pulmonary pathology on imaging (n = 5). We studied patients who were enrolled in either of two prospective cohort studies of critically ill patients at only on subjects who did not receive treatment with immunomodulating therapies to avoid confounding the assessment of inflammatory gene expression. Future lower respiratory transcriptomic studies will thus be needed to directly assess the effects of dexamethasone at the transcriptional level and investigate the mechanisms of interaction between dexamethasone and IL-6 receptor antagonists.

Our findings build on recent reports that dysregulated interferon responses in patients with severe COVID-19 pneumonia may be an important feature of disease. This hypothesis is supported by recent findings of impaired interferon signaling in peripheral blood immune cells of patients with severe versus mild COVID-19, and a recent report suggesting that a dysregulated interferon response may be a common feature of severe viral infections.

Relatively few studies have evaluated lower airway specimens from COVID-19 patients using transcriptional profiling, and those to date have examined BAL fluid. Due in part to updates in clinical guidelines, less invasive TA sampling is increasingly employed for microbiologic diagnosis of pneumonia and offers the advantage of reducing unnecessary exposure to SARS-CoV-2 containing aerosols during bronchoscopy. The similarity in cellular populations in our TA scRNAseq data of critically ill COVID-19 patients compared to previously published scRNAseq data of BAL fluid suggests that TA may be a reasonable alternative specimen for transcriptional studies of the lower airways.

Significant overlap in comparative analyses of our bulk RNA-seq data against external BAL RNA-seq datasets (Supplementary Data 12) and similarities in the predicted cellular composition of matched TA and mini-BAL specimens (Supplementary Data 13, Supplementary Fig. 7) also support this idea.

This study has some limitations. First, TA contains a heterogeneous mix of cells from throughout the lower respiratory tract and thus does not intrinsically distinguish between airway and alveolar biological processes, and thus we cannot determine precisely where in the lung differences in observed gene expression are occurring. However, as discussed above, our comparative analyses suggest that TA has practical utility for assessing lower respiratory tract biology. Our sample size, particularly with respect to the Other Viral LRTI-ARDS and No-ARDS groups, may limit the generalizability of these findings, which require validation in a larger cohort. We were unable to directly measure protein expression in the lower airway, which limits the scope of our biological analysis. Pathway analyses in and silico drug discovery results require validation in experimental models. While our findings related to dexamethasone and G-CSF are supported by results from recent human clinical trials, additional studies will be required to verify that the candidate genes identified in our silico approach drive the observed clinical benefit.

In summary, comparative TA transcriptional profiling identified a lower respiratory gene expression signature of COVID-19 ARDS characterized by dysregulated inflammatory signaling different from other types of ARDS. Lower respiratory tract RNA-seq holds promise for advancing our understanding of other types of infectious and noninfectious ARDS, and for identifying potential new therapeutics.
**Fig. 2 Cellular landscape of COVID-19 ARDS.**

**a** In silico deconvolution of cell types from tracheal aspirate bulk RNA-sequencing data using lung single-cell signatures. The horizontal line inside the box denotes the median and the lower and upper hinges correspond to the first and third quartiles, respectively. Pairwise comparisons between patient groups were performed with a two-sided Mann–Whitney–Wilcoxon test followed by Bonferroni correction (n = 15 COVID-ARDS, n = 32 Other-ARDS, n = 5 No-ARDS). Data for other cell types examined are plotted in Supplementary Fig. 4 and tabulated in (Supplementary Data 7).

**b** UMAP demonstrating the immune cell landscape of COVID-19 ARDS from scRNAseq of TA specimens. Inset demonstrates cell-type proportions (n = 6, COVID-ARDS group). The bar plot denotes median and the error bars depict the interquartile range, respectively. Mac alveolar macrophage, mDC monocyte-derived dendritic cell, RBC red blood cell. Source data are provided in the Source Data file.
For both the COVID-19 and control cohorts, if a patient met inclusion criteria, a study coordinator or physician obtained written informed consent for enrollment from the patient or their surrogate. Patients or their surrogates were provided with detailed written and verbal information about the goals of the study, the data and specimens that would be collected, and the potential risks to the patient or the surrogate. Subjects were screened for enrollment from 7/2013 to 3/2020 in the first (pre-COVID-19) cohort and 3/2020-7/2020 in the second cohort. Tracheal aspirates were collected within five days of intubation. All consecutively enrolled patients were considered for inclusion in this study.

For the COVID-19 and control cohorts, if a patient met inclusion criteria, then a study coordinator or physician obtained written informed consent for enrollment from the patient or their surrogate. Patients or their surrogates were provided with detailed written and verbal information about the goals of the study, the data and specimens that would be collected, and the potential risks to the patient or the surrogate. Subjects were screened for enrollment from 7/2013 to 3/2020 in the first (pre-COVID-19) cohort and 3/2020-7/2020 in the second cohort. Tracheal aspirates were collected within five days of intubation. All consecutively enrolled patients were considered for inclusion in this study.

For both the COVID-19 and control cohorts, if a patient met inclusion criteria, a study coordinator or physician obtained written informed consent for enrollment from the patient or their surrogate. Patients or their surrogates were provided with detailed written and verbal information about the goals of the study, the data and specimens that would be collected, and the potential risks to the patient or the surrogate. Subjects were screened for enrollment from 7/2013 to 3/2020 in the first (pre-COVID-19) cohort and 3/2020-7/2020 in the second cohort. Tracheal aspirates were collected within five days of intubation. All consecutively enrolled patients were considered for inclusion in this study.

For both the COVID-19 and control cohorts, if a patient met inclusion criteria, a study coordinator or physician obtained written informed consent for enrollment from the patient or their surrogate. Patients or their surrogates were provided with detailed written and verbal information about the goals of the study, the data and specimens that would be collected, and the potential risks to the patient or the surrogate. Subjects were screened for enrollment from 7/2013 to 3/2020 in the first (pre-COVID-19) cohort and 3/2020-7/2020 in the second cohort. Tracheal aspirates were collected within five days of intubation. All consecutively enrolled patients were considered for inclusion in this study.
In silico analysis of cell-type proportions

with a

In the IKB. Significant pathways and upstream regulators were defined as those with a Z-score absolute value greater than 2 or an overlap P value <0.05.

In silico analysis of cell-type proportions

Cell-type proportions were estimated from bulk host transcriptome data using the CIBERSORT X algorithm. We used the Human Lung Cell Atlas dataset to derive the single-cell signatures. The cell types estimated with this reference cover all expected cell types in the airway. The estimated proportions were compared between three patient groups: using a Mann–Whitney–Wilcoxon test (two-sided) with Bonferroni correction.

Quantification of SARS-CoV-2 viral load by mNGS

All samples were processed through a SARS-CoV-2-reference-based assembly pipeline that involved removing reads likely originating from the human genome or from other viral genomes annotated in RefSeq with Kraken v2.0.8.beta, and then aligning the remaining reads to the SARS-CoV-2 reference genome MN908947.3 using minimap2 v2.2.17. We calculated SARS-CoV-2 reads-per-million (rPM) by dividing the number of reads that aligned to the virus with mapq ≥ 20 by the total number of reads in the sample (excluding reads mapping to ERCC RNA standards).

Single-cell RNA sequencing and transcriptome analysis. After collection, fresh TA was transported to a BSL-3 laboratory at ambient temperature to improve neutrophil survival. In total, 3 mL of TA was dissociated in 40 mL of PBS with 50 µg/mL collagenase type 4 (Worthington) and 0.56 ku/mL of Dnase I (Worthington) for 10 min at room temperature, followed by passage through a 70-mm filter. Cells were pelleted at 350 × g for 4 min, resuspended in PBS with 2 mM EDTA and 0.5% BSA, and manually counted on a hemocytometer. Cells were stained with MojoSort Human CD45 and purified by the manufacturer’s protocol (Biolegend). After CD45-positive selection, cells were manually counted with trypan blue on a hemocytometer. Using a V(D)J v1.1 kit according to the manufacturer’s protocol, samples were loaded on a 10x Genomics Chip A without multiplexing, aiming to capture 10,000 cells (10x Genomics). Libraries underwent paired-end 150 base pair sequencing on an Illumina NovaSeq 6000.

Read count matrices were generated through the 10x Genomics Cell Ranger pipeline v3.0. Cell barcodes were then determined based upon UMI count distribution. Data were processed and analyzed using Scrapy v1.0.49. Cells that had less than 200 genes or had greater than 30,000 counts were filtered. Mitochondrial genes were removed and multi-sample integration was performed using Harmony v0.14.49.

Comparison against external datasets. No publicly available lower respiratory RNA-seq data were available to compare COVID-19 related ARDS to other types of ARDS. Thus, we alternatively compared differential gene expression between COVID-19 and No-ARDS against three previously published studies with comparisons of COVID-19 patients against controls (10,12,15). The first used Nanostring to assess transcript levels of angiogenesis and inflammation-associated genes in autopsy lung specimens that were differentially expressed between patients with severe COVID-19 and unaffected controls (15). The second study, a gene expression in BAL in a rhesus macaque model of SARS-CoV-2 infection. Gene expression data were downloaded from the Gene Expression Omnibus (GSE156701), and we used DESeq2 and apeglm to identify genes that were differentially expressed between baseline and day 2 (the day of peak inflammatory response in the macaque model). Third, we compared data against a study that performed RNA-seq of macrophages from BAL to study intubated patients with COVID-19 or controls. For this study, data were downloaded from the Gene Expression Omnibus (GSE155249), and we used DESeq2 and apeglm to identify genes that were differentially expressed. The TopGene suite was used to carry out functional enrichment analysis on overlapping genes differentially expressed at an FDR <0.1 in both our dataset and each external dataset.

Regression of ISG counts against viral load in TA and NP samples. We assembled a set of 100 interferon-stimulated genes based on the “Hallmark interferon-alpha response” gene set in MSigDB. We then performed robust regression of the quantile normalized gene counts (log2 scale), generated using the R package limma, against log10(rPM) of SARS-CoV-2. This was done in two separate datasets of COVID-19 patients: (i) the tracheal aspirate (TA) samples from patients with COVID-19 ARDS reported in this study (n = 5); and (ii) the nasopharyngeal swab (NP) samples from patients with mostly early and mild disease that we previously reported (n = 93). The analysis was performed using the R package robustbase v0.93.6, which implements MM-type estimators for linear regression. Model predictions were generated using the R package geepack v0.143.4 and used for display in the individual gene plots. Plots were generated using ggplot2 v3.3.3. Error bars represent normal distribution 95% confidence intervals around each prediction. Reported P values for significance of the difference of the regression coefficient from 0 are based on a t-statistic and calculated using the stats::t.test function. Reported R2 values represent the adjusted robust coefficient of determination.

Statistics and reproducibility. Statistical tests utilized for each analysis are described in the figure legends and in further detail in each respective methods section. The number of patient samples analyzed for each comparison is indicated in the figure legends. Data were generated from single sequencing runs without technical replicates.

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw sequencing data are protected due to data privacy restrictions from the IRB protocols governing patient enrollment in this study, which protect the release of raw genomic sequencing data from those patients enrolled under a waiver of consent.
Researchers who wish to obtain raw fastq files for the purposes of independently generating gene counts can contact the corresponding author (chaoy.labeler@ucsc.edu) to be added to the IRB protocols and sign a materials transfer agreement from UCSF ensuring that the data will be securely stored and only utilized for transcriptomic analyses. The processed gene count data are available from the National Center for Biotechnology Information Gene Expression Omnibus database under accession code GSE164426. The published human lung single-cell datasets used for cell-type proportion analyses can be obtained through Synapse under accessions syn21560510 and syn21560511. Source data are provided with this paper.

Code availability
Code for the differential expression, cell-type proportions, and scRNAseq analyses is available at https://doi.org/10.5281/zenodo.4990584.

Received: 5 January 2021; Accepted: 13 July 2021; Published online: 26 August 2021

References
1. Matthay, M. A. et al. Acute respiratory distress syndrome. Nat. Rev. Dis. Primers 5, 1–22 (2019).
2. Mehta, P. et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet 395, 1033–1034 (2020).
3. Blanco-Melo, D. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell https://doi.org/10.1016/j.cell.2020.04.026 (2020).
4. Giamarellos-Bourboulis, E. J. et al. Complex immune dysregulation in COVID-19 patients with severe respiratory failure. Cell Host Microbe 27, 992–1000.e3 (2020).
5. Sinha, P., Matthay, M. A. & Calfee, C. S. Is a “Cytokine Storm” relevant to COVID-19? JAMA Intern. Med. 180, 1152 (2020).
6. Wilson, J. G. et al. Cytokine profile in plasma of severe COVID-19 does not differ from ARDS and sepsis. https://insight.jcnn.org/articles/view/140289/sd/1 (2020).
7. Leisman, D. E. et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. Lancet Respir. Med. 8, 1233–1244 (2020).
8. The RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with covid-19—preliminary report. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2021436 (2020).
9. Bost, P. et al. Host-viral infection maps reveal signatures of severe COVID-19 patients. Cell 181, 1475–1488.e12 (2020).
10. The NU SCRIPT Study Investigators. et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. Nature 590, 635–641 (2021).
11. Liu, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat. Med. 26, 842–844 (2020).
12. Ackermann, M. et al. Pulmonary vascular endotheliitis, thrombosis, and angiogenesis in COVID-19. N. Engl. J. Med. 383, 120–128 (2020).
13. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches to be added to the IRB protocols and sign a materials transfer agreement from UCSF ensuring that the data will be securely stored and only utilized for transcriptomic analyses. The processed gene count data are available from the National Center for Biotechnology Information Gene Expression Omnibus database under accession code GSE164426. The published human lung single-cell datasets used for cell-type proportion analyses can be obtained through Synapse under accessions syn21560510 and syn21560511. Source data are provided with this paper.

Code availability
Code for the differential expression, cell-type proportions, and scRNAseq analyses is available at https://doi.org/10.5281/zenodo.4990584.

Received: 5 January 2021; Accepted: 13 July 2021; Published online: 26 August 2021

References
1. Matthay, M. A. et al. Acute respiratory distress syndrome. Nat. Rev. Dis. Primers 5, 1–22 (2019).
2. Mehta, P. et al. COVID-19: consider cytokine storm syndromes and immunosuppression. Lancet 395, 1033–1034 (2020).
3. Blanco-Melo, D. Imbalanced host response to SARS-CoV-2 drives development of COVID-19. Cell https://doi.org/10.1016/j.cell.2020.04.026 (2020).
4. Giamarellos-Bourboulis, E. J. et al. Complex immune dysregulation in COVID-19 patients with severe respiratory failure. Cell Host Microbe 27, 992–1000.e3 (2020).
5. Sinha, P., Matthay, M. A. & Calfee, C. S. Is a “Cytokine Storm” relevant to COVID-19? JAMA Intern. Med. 180, 1152 (2020).
6. Wilson, J. G. et al. Cytokine profile in plasma of severe COVID-19 does not differ from ARDS and sepsis. https://insight.jcnn.org/articles/view/140289/sd/1 (2020).
7. Leisman, D. E. et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. Lancet Respir. Med. 8, 1233–1244 (2020).
8. The RECOVERY Collaborative Group. Dexamethasone in hospitalized patients with covid-19—preliminary report. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2021436 (2020).
9. Bost, P. et al. Host-viral infection maps reveal signatures of severe COVID-19 patients. Cell 181, 1475–1488.e12 (2020).
10. The NU SCRIPT Study Investigators. et al. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. Nature 590, 635–641 (2021).
11. Liu, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat. Med. 26, 842–844 (2020).
12. Ackermann, M. et al. Pulmonary vascular endotheliitis, thrombosis, and angiogenesis in COVID-19. N. Engl. J. Med. 383, 120–128 (2020).
13. Krämer, A., Green, J., Pollard, J. & Tugendreich, S. Causal analysis approaches to be added to the IRB protocols and sign a materials transfer agreement from UCSF ensuring that the data will be securely stored and only utilized for transcriptomic analyses. The processed gene count data are available from the National Center for Biotechnology Information Gene Expression Omnibus database under accession code GSE164426. The published human lung single-cell datasets used for cell-type proportion analyses can be obtained through Synapse under accessions syn21560510 and syn21560511. Source data are provided with this paper.
Acknowledgements

This study was performed with support from the National Institute of Allergy and Infectious Diseases-sponsored Immune phenotyping Assessment in a COVID-19 Cohort (IMPACC) Network [N01AI707749] (DJL). This work was also supported by the National Heart, Lung and Blood Institute [R01 HL105026 (SGC), R35 HL105026 (CSC), K23HL138231 (AGI), R21 HL121784] (SM), and the Chan Zuckerberg Biohub (A.O.P., A.B., A.K., and J.L.D.). We thank Mark and Carrie Casey, Julie and Kevin Hartz, Carl Kawaja and Wendy Holcombe, Eric Kießman and Linda Nevin, Martin and Leesa Romo, Three Sisters Foundation, Diana Wagner and Jerry Yang and Akiko Yamazaki for their support and philanthropic contributions. We gratefully appreciate support and input from Amy Kistler, Jack Kam, Sahari Caldera, and Maira Phelps.

Author contributions

C.R.L., C.S.C., A.S., and S.C conceived and designed the study. T.D., R.G., P.H.S., N.M., M.T., and K.M.A. oversaw or performed sample processing, library preparation and sequencing. C.S.C., C.M.H., K.K.N., R.G., A.J., J.G.W., and E.R.S. coordinated or contributed to clinical operations including patient enrollment. C.D., A.S., C.R.L., F.M., T.D., and N.S. performed metadata collection or clinical chart review. A.S., E.M., A.O.P., A.T., C.R.L., and S.A.C. performed data analysis and interpretation. C.S.C., S.A.C., E.M., D.J.J., J.L.D., K.M.A., M.F.K., P.F.W., M.A.M., B.S.Z., J.G.W., A.L., B.F., P.S., and M.S. provided guidance, advice, and comments on the study design and manuscript. C.R.L., A.S., and C.S.C. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-25040-5.

Correspondence and requests for materials should be addressed to C.R.L.

Peer review information Nature Communications thanks Lievee Bos, Klaus Schughart and the other, anonymous, reviewer for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

COMET Consortium

Yumiko Abe-Jones18, K. Mark Ansley6,7, Saurabh Asthana15,19,20, Alexander Beagle4, Tanvi Bhakta21, Sharvari Bhide1, Cathy Ca122, Sahari Caldera3, Carolyn Calfee10,14, Maria Calvo1, Sidney Carrillo1, Adithya Cattamanchi1, Suzanna Chak1, Vincent Chan6, Nayvin Chew15,20,23,24, Stephanie Christenson1, Zachary Collins15,19,20, Alexis Combes15,20,23,24, Tristan Courau15,20,23,24, Spyros Darmanis25, Catherine DeVoe3, David Erle14,16,26, Armond Esmaili18, Gabriela K. Fragiadakis12,20,26, Rajani Ghale1, Jeremy Giberson1, Ana Gonzalez1, Paula Hayakawa Serpa2,3, Carolyn Hendrickson1, Kamir Hiam2,6,27,28,29, Kenneth Hu15, Billy Huang30, Alejandra Jauregui1, Chayne Jones1, Norman Jones31, Kirsten Kangelaris4, Matthew Krummel32,33,34, Nitasha Kumar31, Divya Kushnoor15,20,23,24, Charles R. Langelier9,3,34,35, Tasha Lea15, Deanna Lee1,10, David Lee12, Aleksandra Leligdowicz19,10, Kathleen D. Liu35,36, Yale Liu37, Salman Mahboob30, Michael Matthy110,14, Eran Mick1,2,3,47, Jeff Milush31, Priscila Muñoz-Sandoval6,7, Viet Nguyen110, Gabe Ortiz38, Randy Parada30, Maira Phelps2, Logan Pierce18, Priya Prasad18, Arjun Rao15,19,20, Sadeed Rashid28, Gabriella Reeder20,24,39, Nicklaus Rodriguez28, Bushra Samad15,19,20, Aartik Sarma1, Diane Scarlet40, Cole Shaw19,20,26, Alan Shen15,20,23,24, Austin Sigman1, Pratik Sinha5,
Matthew Spitzer\textsuperscript{2,6,27,28,29}, Yang Sun\textsuperscript{12}, Sara Sunshine\textsuperscript{13}, Kevin Tang\textsuperscript{28}, Luz Torres Altamirano\textsuperscript{28}, Alexandra Tsitsiklis\textsuperscript{3}, Jessica Tsui\textsuperscript{15,20,23,24}, Erden Tumurbaatar\textsuperscript{12}, Kathleen Turner\textsuperscript{41}, Alyssa Ward\textsuperscript{12}, Andrew Willmore\textsuperscript{1}, Michael Wilson\textsuperscript{42}, Juliane Winkler\textsuperscript{43}, Reese Withers\textsuperscript{41}, Kristine Wong\textsuperscript{22}, Prescott Woodruff\textsuperscript{1,7}, Jimmie Ye\textsuperscript{2,12,29,44,45,46}, Kimberly Yee\textsuperscript{1}, Michelle Yu\textsuperscript{1}, Shoshana Zha\textsuperscript{1}, Jenny Zhan\textsuperscript{22}, Mingyue Zhou\textsuperscript{22} & Wandi S. Zhu\textsuperscript{6,7}

\textsuperscript{18}Division of Hospital Medicine, University of California, San Francisco, CA, USA. \textsuperscript{19}Data Science CoLab, University of California, San Francisco, CA, USA. \textsuperscript{20}Bakar ImmunoX Initiative, University of California, San Francisco, CA, USA. \textsuperscript{21}Department of Nursing, Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco, CA, USA. \textsuperscript{22}Biospecimen Resource Program, University of California, San Francisco, CA, USA. \textsuperscript{23}Department of Anatomy, University of California, San Francisco, CA, USA. \textsuperscript{24}Disease 2 Biology CoLab, University of California, San Francisco, CA, USA. \textsuperscript{25}Microchemistry, Proteomics and Lipidomics Department, Genentech Inc, 1 DNA Way, South San Francisco, CA, USA. \textsuperscript{26}CoLabs, University of California, San Francisco, CA, USA. \textsuperscript{27}Department of Otolaryngology, University of California, San Francisco, CA, USA. \textsuperscript{28}Helen Diller Family Comprehensive Cancer Center, University of California, San Francisco, CA, USA. \textsuperscript{29}Parker Institute for Cancer Immunotherapy, San Francisco, CA, USA. \textsuperscript{30}Department of Orofacial Sciences, School of Dentistry, University of California, San Francisco, CA, USA. \textsuperscript{31}Core Immunology Laboratory, Division of Experimental Medicine, University of California, San Francisco, CA, USA. \textsuperscript{32}ImmuX Initiative, University of California, San Francisco, CA, USA. \textsuperscript{33}Departments of Medicine, University of California, San Francisco, CA, USA. \textsuperscript{34}Department of Laboratory Medicine, University of California, San Francisco, CA, USA. \textsuperscript{35}Division of Nephrology, University of California San Francisco, San Francisco, CA, USA. \textsuperscript{36}Division of Critical Care Medicine, Department of Anesthesia, University of California San Francisco, San Francisco, CA, USA. \textsuperscript{37}Department of Dermatology, University of California, San Francisco, CA, USA. \textsuperscript{38}Department of Medicine, Zuckerberg San Francisco General Hospital and Trauma Center, University of California, San Francisco, CA, USA. \textsuperscript{39}Biomedical Sciences Graduate Program, University of California, San Francisco, CA, USA. \textsuperscript{40}Zuckerberg San Francisco General Hospital and Trauma Center, San Francisco, CA, USA. \textsuperscript{41}Department of Nursing, University of California, San Francisco, CA, USA. \textsuperscript{42}Weill Institute for Neurosciences, Department of Neurology, University of California, San Francisco, CA, USA. \textsuperscript{43}School of Dentistry, University of California, San Francisco, CA, USA. \textsuperscript{44}Institute for Human Genetics, University of California, San Francisco, CA, USA. \textsuperscript{45}Department of Epidemiology and Biostatistics, University of California, San Francisco, CA, USA. \textsuperscript{46}Institute of Computational Health Sciences, University of California, San Francisco, CA, USA.