# Original Research

## Title: COVID-19 lung disease shares driver AT2 cytopathic features with Idiopathic pulmonary fibrosis

**Authors:** Saptarshi Sinha[1], Vanessa Castillo[1], Celia R. Espinoza[1], Courtney Tindle[1], Ayden G. Fonseca[1], Jennifer M. Dan[4, 5], Gajanan D. Katkar[1], Soumita Das[6] Debashis Sahoo[2, 3†] and Pradipta Ghosh[1, 7†]

**Affiliations:**

[1]*Department of Cellular and Molecular Medicine, University of California San Diego, La Jolla, CA, 92093*
[2]*Department of Pediatrics, University of California San Diego, La Jolla, CA, 92093.*
[3]*Department of Computer Science and Engineering, Jacobs School of Engineering, University of California San Diego, La Jolla, CA, 92093.*
[4]*Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA, USA.*
[5]*Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA, USA.*
[6]*Department of Pathology, University of California San Diego, La Jolla, CA, 92093.*
[7]*Department of Medicine, University of California San Diego, La Jolla, CA, 92093.*

† Co-Corresponding

**Corresponding authors**:

**Debashis Sahoo, Ph.D.;** Associate Professor, Department of Pediatrics, University of California San Diego; 9500 Gilman Drive, MC 0703, Leichtag Building 132; La Jolla, CA 92093-0703. **Phone:** 858-246-1803: **Fax:** 858-246-0019: **Email:** dsahoo@ucsd.edu

**Pradipta Ghosh, M.D.;** Professor, Departments of Medicine and Cell and Molecular Medicine, University of California San Diego; 9500 Gilman Drive (MC 0651), George E. Palade Bldg, Rm 232; La Jolla, CA 92093. **Phone**: 858-822-7633:
**Email:** prghosh@ucsd.edu

## ABSTRACT

**Background**: In the aftermath of Covid-19, some patients develop a fibrotic lung disease, i.e., *p*ost-*C*OVID-19 *l*ung *d*isease (PCLD), for which we currently lack insights into pathogenesis, disease models, or treatment options.

**Method**: Using an AI-guided approach, we analyzed > 1000 human lung transcriptomic datasets associated with various lung conditions using two viral pandemic signatures (ViP and sViP) and one covid lung-derived signature. Upon identifying similarities between COVID-19 and idiopathic pulmonary fibrosis (IPF), we subsequently dissected the basis for such similarity from molecular, cytopathic, and immunologic perspectives using a panel of IPF-specific gene signatures, alongside signatures of alveolar type II (AT2) cytopathies and of prognostic monocyte-driven processes that are known drivers of IPF. Transcriptome-derived findings were used to construct protein-protein interaction (PPI) network to identify the major triggers of AT2 dysfunction. Key findings were validated in hamster and human adult lung organoid (ALO) pre-clinical models of COVID-19 using immunohistochemistry and qPCR.

**Findings**: COVID-19 resembles IPF at a fundamental level; it recapitulates the gene expression patterns (ViP and IPF signatures), cytokine storm (IL15-centric), and the AT2 cytopathic changes, e.g., injury, DNA damage, arrest in a transient, damage-induced progenitor state, and senescence-associated secretory phenotype (SASP). These immunocytopathic features were induced in pre-clinical COVID models (ALO and hamster) and reversed with effective anti-CoV-2 therapeutics in hamsters. PPI-network analyses pinpointed ER stress as one of the shared early triggers of both diseases, and IHC studies validated the same in the lungs of deceased subjects with COVID-19 and SARS-CoV-2-challenged hamster lungs. Lungs from *tg*-mice, in which ER stress is induced specifically in the AT2 cells, faithfully recapitulate the host immune response and alveolar cytopathic changes that are induced by SARS-CoV-2.

**Interpretation:** Like IPF, COVID-19 may be driven by injury-induced ER stress that culminates into progenitor state arrest and SASP in AT2 cells. The ViP signatures in monocytes may be key determinants of prognosis. The insights, signatures, disease models identified here are likely to spur the development of therapies for patients with IPF and other fibrotic interstitial lung diseases.



**Funding:** This work was supported by the National Institutes for Health grants R01- GM138385 and AI155696 and funding from the Tobacco-Related disease Research Program (R01RG3780).

**One Sentence Summary:** Severe COVID-19 triggers cellular processes seen in fibrosing Interstitial Lung Disease

**Key Words:**

- Artificial Intelligence/Machine Learning
- Immune response
- Alveolar Type II Pneumocytes
- Damage-associated transient progenitor (DATP)
- Senescence
- ER-stress
- Interleukin 15 (IL15)

## RESEARCH IN CONTEXT

### Evidence before this study

In its aftermath, the COVID-19 pandemic has left many survivors, almost a third of those who recovered, with a mysterious long-haul form of the disease which culminates in a fibrotic form of interstitial lung disease (post-COVID-19 ILD). Post-COVID-19 ILD remains a largely unknown entity. Currently, we lack insights into the core cytopathic features that drive this condition.

### Added value of this study

Using an AI-guided approach, which involves the use of sets of gene signatures, protein-protein network analysis, and a hamster model of COVID-19, we have revealed here that COVID-19 -lung fibrosis resembles IPF, the most common form of ILD, at a fundamental level—showing similar gene expression patterns in the lungs and blood, and dysfunctional AT2 processes (ER stress, telomere instability, progenitor cell arrest, and senescence). These findings are insightful because AT2 cells are known to contain an elegant quality control network to respond to intrinsic or extrinsic stress; a failure of such quality control results in diverse cellular phenotypes, of which ER stress appears to be a point of convergence, which appears to be sufficient to drive downstream fibrotic remodeling in the lung.

### Implications of all the available evidence

Because unbiased computational methods identified the shared fundamental aspects of gene expression and cellular processes between COVID-19 and IPF, the impact of our findings is likely to go beyond COVID-19 or any viral pandemic. The insights, tools (disease models, gene signatures, and biomarkers), and mechanisms identified here are likely to spur the development of therapies for patients with IPF and, other fibrotic interstitial lung diseases, all of whom have limited or no treatment options. To dissect the validated prognostic biomarkers to assess and track the risk of pulmonary fibrosis and develop therapeutics to halt fibrogenic progression.

## INTRODUCTION

As the acute phase of the COVID-19 pandemic winds down, the chronic diseases in its aftermath have begun to emerge. For example, many survivors are suffering from a mysterious long-haul form of the disease which culminates in a fibrotic form of interstitial lung disease (post-COVID-19 ILD)(1-9). The actual prevalence of post-COVID-19 ILD (henceforth, *PCLD*) is still emerging; early analysis indicates that more than a third of the survivors develop fibrotic abnormalities. One of the important determinants for PCLD[10] is the duration of disease; ~4% of patients with a disease duration of less than 1 week, ~24% of patients with a disease duration between 1-3 weeks, and ~61% of patients with a disease duration > 3 weeks developed fibrosis. Disability among survivors of severe COVID-19 has been primarily attributed to reduced lung capacities (11). While COVID-19–affected lungs and post-Influenza lungs share the histologic pattern of diffuse alveolar damage (DAD)(12) with perivascular T-cell infiltration, key differences have been observed (13). The lungs from patients with Covid-19 also showed (13): (i) widespread thrombosis with microangiopathy; (ii) severe endothelial injury associated with the presence of intracellular virus and disrupted cell membranes; and (iii) neovascularization predominantly through a mechanism of intussusceptive angiogenesis. These findings highlight that despite being a viral pandemic, COVID-19, but not influenza impacts the lungs in ways that go beyond acute DAD, a common phenomenon in progressive ILDs (12). Although a recent study (14) implicated pathogenic subsets of respiratory $CD8^+$ T cells in contributing to persistent tissue microenvironment in PCLD, what remains largely unknown are some of the earliest cellular and molecular mechanisms that fuel the progression of fibrosis in COVID-19 lung (but not influenza).

As for how the fibrotic sequel of COVID-19 is managed at present, based on the shared pathological features of overt fibrosis in both end-stage COVID-19 and ILDs, lung transplantation remains the mainstay option for these patients (15-17). Although corticosteroids appear to improve the risk of PCLD(2), beyond that, there is no existing therapeutic option.

Here we seek to unravel the fundamental molecular mechanisms underlying PCLD, identifying key disease drivers (cellular processes, immune pathways, and the signaling cascades that support those pathways). We use artificial intelligence (AI) and machine learning derived gene signatures, the *Vi*ral *P*andemic, ViP and, severe(s) ViP signatures that are induced in all respiratory viral pandemics(18). Besides the ViP signatures, we also use gene signatures induced in the lungs of patients with severe COVID-19. These approaches, in conjunction with experimental validation (in pre-clinical disease models and in human lung tissue), not only helped identify which lung pathology shares fundamental molecular features with COVID-19 lung but also revealed key mechanistic insights into the pathogenesis of PCLD. The findings also provide clues into how to navigate, prognosticate and recapitulate in pre-clinical models this emergent mysterious condition. The objectivity and precision of the AI-guided unbiased approaches enhance the translational potential of our findings.

137 # METHODS

138 ## KEY RESOURCE TABLE:

| MATERIALS & REAGENTS | | | |
|---|---|---|---|
| **ANTIBODIES USED FOR IMMUNOCYTOCHEMISTRY** | | | |
| **Name** | **Manufacturer** | **Catalog number** | **Dilution factor** |
| **p14 ARF (E3X6D)** | Cell Signaling | 74560S | 1:250 |
| **GRP78/BIP Rabbit polyclonal antibody** | Proteintech | 11587-1-AP | 1:500 |
| **P53 Rabbit polyclonal antibody** | Proteintech | 21891-1-AP | 1:200 |
| **Cytokeratin 8 Rabbit polyclonal antibody** | Proteintech | 10384-1-AP | 1:2000 |
| **Claudin 4-specific Rabbit polyclonal antibody** | Proteintech | 16195-1-AP | 1:200 |
| **p21 Waf1 / Cip1 (12D1)** | Cell Signaling | 2947S | 1:50 |
| **CD14** | Invitrogen | MA5-32248 | 1:100 |
| **SARS-CoV2 Nucleoprotein** | Sino Biological | 40143-MM05 | 1:250 |
| **anti-Rabbit IgG Secondary Antibody; AlexaFluor 594 nm** | Invitrogen | A-11012 | 1:500 |
| **anti-Mouse IgG Secondary Antibody, AlexaFluor 488 nm** | Invitrogen | A-31560 | 1:500 |
| | | | |
| **INSTRUMENTS** | | | |
| **Leica DMI4000B (Automated Inverted Microscope)** | Leica Microsystems | DMI4000B | |
| **Power Pressure Cooker XL** | Tristar Products | | |
| **MESO QuickPlex SQ 120** | PBL Assay Science | | |
| | | | |
| **SOFTWARE & PACKAGES** | | | |
| **ImageJ** | https://imagej.nih.gov/ij/index.html | | |
| **GraphPad Prism** | https://www.graphpad.com/scientific-software/prism/ | | |
| **MSD® DISCOVERY WORKBENCH 4.0** | https://www.mesoscale.com/en/products_and_services/software/ | | |
| **NetworkX** | https://networkx.org | | |
| **Cytoscape** | https://cytoscape.org | | |
| | | | |
| **KITS, ENZYMES, CHEMICALS, AND REAGENTS** | | | |
| **DAPI** | Invitrogen | D1306 | 1:500 |
| **Zinc Formalin** | Fisher Scientific | 23-313096 | |
| **Xylene** | VWR | XX0060-4 | |
| **Hematoxylin** | BioGenex | HK100-9K | |
| **Two-Component DAB** | BioGenex | HK542-XAKE | 2 drops per 1mL |
| **Ethanol** | Koptec | UN1170 | |
| **Sodium Citrate** | Sigma-Aldrich | W302600 | |
| **Tris Base** | Fisher Scientific | BP154-1 | |
| **Ethylenediaminetetraacetic acid (EDTA)** | Fischer Scientific | BP118-500 | |
| **3% Hydrogen Peroxide** | Target | 245-07-3628 | |

| | | | |
|---|---|---|---|
| **ImmPRESS HRP Horse Anti-Rabbit IgG Polymer Detection Kit, Peroxidase** | Vector Labs | MP-7401 | |
| **Goat serum 2.5%** | Vector Laboratories | S-10120 | |
| **Normal Antibody Diluent** | MP Biomedicals, LLC | 980641 | |
| | | | |

139

## REAGENT VALIDATION:

There are no cell lines used in this work. All the antibodies used for immunocytochemistry in this work have been previously validated by commercial entities (in many instances, by confirming loss of signal by target depletion). These details are included in **Supplemental Information 8** (Reagent Validation file).

## DETAILED METHODS:

### ***Computational Methods***

#### ***StepMiner* analysis**

*StepMiner* is an algorithm that identifies step-wise transitions using step function in a time-series data(19). *StepMiner* undergoes an adaptive regression scheme to verify the best possible up and down steps based on sum-of-square errors. The steps are placed between time points at the sharpest change between expression levels, which gives us the information about timing of the gene expression-switching event. To fit a step function, the algorithm evaluates all possible steps for each position and computes the average of the values on both sides of a step for the constant segments. An adaptive regression scheme is used that chooses the step positions that minimize the square error with the fitted data. Finally, a regression test statistic is computed as follows:

$$F\ stat = \frac{\sum_{i=1}^{n}(\widehat{X}_i - \bar{X})^2/(m-1)}{\sum_{i=1}^{n}(X_i - \widehat{X}_i)^2/(n-m)}$$

Where $X_i$ for $i = 1$ to $n$ are the values, $\widehat{X}_i$ for $i = 1$ to $n$ are fitted values. m is the degrees of freedom used for the adaptive regression analysis. $\bar{X}$ is the average of all the values: $\bar{X} = \frac{1}{n} * \sum_{j=1}^{n} X_j$. For a step position at k, the fitted values $\widehat{X}_i$ are computed by using $\frac{1}{k} * \sum_{j=1}^{n} X_j$ for $i = 1$ to $k$ and $\frac{1}{(n-k)} * \sum_{j=k+1}^{n} X_j$ for $i = k+1$ to $n$.

#### **Composite gene signature analysis using Boolean Network Explorer (BoNE)**

Boolean network explorer (BoNE) provides an integrated platform for the construction, visualization and querying of a gene expression signature underlying a disease or a biological process in three steps (**Fig 2, 4, 6D, 8D, S1**): First, the expression levels of all genes in these datasets were converted to binary values (high or low) using the StepMiner algorithm. Second, Gene expression values were normalized according to a modified Z-score approach centered around *StepMiner* threshold (formula = (expr - SThr)/3*stddev). Third, the normalized expression values for every genes were added together to create the final score for the gene signature. The samples were ordered based on the final signature score. Classification of sample categories using this ordering is measured by ROC-AUC (Receiver Operating Characteristics Area Under The Curve) values. Welch's Two

170 Sample t-test (unpaired, unequal variance (equal_var=False), and unequal sample size) parameters were used to
171 compare the differential signature score in different sample categories. Violin, Swarm and Bubble plots are
172 created using python seaborn package version 0.10.1. Pathway analysis of gene lists were carried out via the
173 reactome database and algorithm(20).

174

175 **ViP and severe (s)ViP signatures**

176 ViP (Viral Pandemic) signature(18) is derived from a list of 166 genes using Boolean Analysis of large viral
177 infection datasets (training datasets: GSE47963, n = 438; GSE113211, n = 118). This 166-gene signature was
178 conserved in all viral pandemics, including COVID-19, inspiring the nomenclatures ViP signature. A subset of
179 20-genes classified disease severity called severe-ViP signature using an additional cohort (GSE101702, n =
180 159)(18). No interstitial lung disease (ILD) datasets, or for that matter any other lung dataset was used during the
181 process of discovery and validation of the ViP/sViP signatures. To compute the ViP signature, first the genes
182 present in this list were normalized according to a modified Z-score approach centered around *StepMiner*
183 threshold (formula = (expr -SThr)/3*stddev). The normalized expression values for every probeset for 166 genes
184 were added together to create the final ViP signature. The severe ViP signature is computed similarly using 20
185 genes. The samples were ordered finally based on both the ViP and severe-ViP signature. A color-coded bar plot
186 is combined with a violin plot to visualize the gene signature-based classification.

187

188 **Protein-protein interaction network (PPIN) construction**

189 PPIN has been constructed using the gene signatures as a set of seed nodes. The nodes between the seed nodes
190 were fetched using the connecting shortest paths and their components from the human protein interaction dataset
191 of the STRING database(21). A high cutoff of STRING interaction score has been chosen based on the proteins
192 present in the signature list to neglect the false positive interactions.

193

194 **Transcriptomic datasets and data analysis**

195 Several publicly available microarrays and RNASeq databases were downloaded from the National Center for
196 Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) server (22-24). Gene expression
197 summarization was performed by normalizing Affymetrix platforms by RMA (Robust Multichip Average) and
198 RNASeq platforms by computing TPM (Transcripts Per Millions) (25, 26) values whenever normalized data were
199 not available in GEO. We used log2(TPM +1) as the final gene expression value for analyses. GEO accession
200 numbers are reported in figures and text. A catalog of all datasets analyzed in this work can be found in
201 **Supplemental Information 1**. Patient demographics, where available, are curated in **Supplemental Information**

202 **6** and show that age, gender, and ethnicity are reflective of a wide population of patients impacted by the COVID-
203 19 pandemic hence, the findings and conclusions of the current study may maintain relevance in a similar
204 population. *StepMiner* analysis and methodologies for composite gene signatures, outcome analyses, and violin plots
205 are detailed in Supplementary Online Methods.
206
207 **Survival outcome analysis**
208
209 Kaplan-Meier analysis was done for three different gene signatures: sViP (18), 52 gene IPF signature (27) and
210 COVID-lung signature (13) over two different survival datasets of COVID patients (GSE157103) and IPF
211 patients (GSE28221). Hospital-free days analysis (45 days followup) of COVID-19 patients (GSE157103) were
212 limited to less than 70 years old patients. The high and low group are separated based on *StepMiner* threshold on
213 the composite score of the gene expression values. The IPF dataset (GSE28221) included 114 patients with
214 outcome was divided into male (n = 87) and female (n = 27) population for checking the gender bias. The survival
215 data for the IPF dataset was also censored at 3 years for comparative analysis. The statistical significance of KM
216 plots was assessed by log rank test.
217
218 **Single Cell RNA Seq analysis**
219 Single Cell RNASeq data from GSE145926 was downloaded from Gene Expression Omnibus (GEO) in the HDF5
220 Feature Barcode Matrix Format. The filtered barcode data matrix was processed using Seurat v3 R package. B
221 cells (CD19, MS4A1, CD79A), T cells (CD3D, CD3E, CD3G), CD4 T cells (CCR7, CD4, IL7R, FOXP3,
222 IL2RA), CD8 T cells (CD8A, CD8B), Natural killer cells (KLRF1), Macrophages, Monocytes and DCs
223 (TYROBP, FCER1G), Epithelial (SFTPA1, SFTPB, AGER, AQP4, SFTPC, SCGB3A2, KRT5, CYP2F1,
224 CCDC153, TPPP3) cells were identified using relevant gene markers using SCINA algorithm. Pseudo bulk
225 datasets were prepared by adding counts from the different cell subtypes and normalized using log2(CPM+1). All
226 other single-cell RNA Seq datasets GSE159354, GSE132914, GSE146981, and GSE149878 were processed
227 using scanpy python package (v1.5.1) and converted to pseudo-bulk samples by adding counts from all cells and
228 normalized using log2(CPM+1). The downstream BoNE composite score analyses were performed similarly to
229 regular bulk RNASeq dataset.
230
231 **Statistical analyses**
232 Gene signature is used to classify sample categories and the performance of the multi- class classification is
233 measured by ROC-AUC (Receiver Operating Characteristics Area Under the Curve) values. A color-coded bar
234 plot is combined with a density plot to visualize the gene signature-based classification.All statistical tests were

performed using R version 3.2.3 (2015-12-10). Standard t-tests were performed using python scipy.stats.ttest_ind package (version 0.19.0) with Welch's Two Sample t-test (unpaired, unequalvariance (equal_var=False), and unequal sample size) parameters. Multiple hypothesis correction was performed by adjusting *p* values with statsmodels.stats.multitest.multipletests (fdr_bh: Benjamini/Hochberg principles). Sample number of each analysis is provided with associated plots beside each GSE ID no. or sample name. Pathway analysis of gene lists were carried out via the Reactome database and CluGo algorithm. Reactome identifiessignaling and metabolic molecules and organizes their relations into biological pathways and processes. Kaplan- Meier analysis is performed using lifelines python package version 0.14.6. The statistical significance of KM plots was assessed by log-rank test. Violin, Swarm and Bubble plots are created using python seaborn package version 0.10.1.

**Multivariate analysis**

Multivariate regression has been performed on the single cell (GSE132914) and bulk (GSE150910) sequence IPF datasets of various covid and IPF signatures (in Figure S1). Multivariate analysis models the healthy vs IPF samples as a linear combination of composite scores of sViP, ViP, CoV-lung and six independent IPF signatures. Here, the statsmodels module from python has been used to perform Ordinary least-squares (OLS) regression analysis of each of the variables. The choice of these datasets was driven by the criteria that they are high quality datasets with maximal unique patient samples. For single cell dataset (GSE132914), cell types has been considered as an additional variable. The p-value for each term tests the null hypothesis that the coefficient is equal to zero (no effect).

***Experimental Methods***

**Rapid autopsy procedure for tissue collection**

The lung specimens from the COVID-19 positive human subjects were collected as described in detail previously (18, 28) using autopsy (study was IRB Exempt). All donations to this trial were obtained after telephone consent followed by written email confirmation with next of kin/power of attorney per California state law (no in-person visitation could be allowed into our COVID-19 ICU during the pandemic). The team member followed the CDC guidelines for COVID-19 and the autopsy procedures[8, 9]. Lung specimens were collected in 10% Zinc-formalin and stored for 72 h before processing for histology. Autopsy #2 was a standard autopsy performed by anatomical pathology in the BSL3 autopsy suite. The patient expired and his family consented for autopsy. After 48 hours, lungs were removed and immersion fixed whole in 10% formalin for 48 hours and then processed further. Lungs were only partially fixed at this time (about 50% fixed in thicker segments) and were sectioned further into small 2-4cm chunks and immersed in 10% formalin for further investigation. Autopsies #4 and #5 were collected from

267 rapid postmortem lung biopsies. The procedure was performed in the Jacobs Medical Center ICU (all the ICU
268 rooms have a pressure-negative environment, with air exhausted through HEPA filters [Biosafety Level 3 (BSL3)]
269 for isolation of SARS-CoV-2 virus). Biopsies were performed 2-4 hours after patient expiration. Ventilator was
270 shut off to reduce aerosolization of viral particles at least 1 hour after loss of pulse and before the sample
271 collection. Every team member had personal protective equipment in accordance with the University policies for
272 procedures on patients with COVID-19 (N95 mask + surgical mask, hairnet, full face shield, surgical gowns,
273 double surgical gloves, booties). Lung biopsies were obtained after L-thoracotomy in the 5th intercostal space by
274 our cardiothoracic surgery team. Samples were taken from the left upper lobe (LUL) and left lower lobe (LLL)
275 and then sectioned further.

276 The histologically confirmed normal lung samples (non-COVID-19 control) were obtained from
277 surgically removed lobes for indications such as neoplastic and non-neoplastic lung nodules, at sites distant from
278 the pathologic lesion.

279

280 **Immunofluorescence of normal and COVID-19 lung tissue samples**

281 FFP-embedded lung tissue sections underwent antigen retrieval immersed in Tris-EDTA buffer (pH 9.0) and
282 boiled at 100°C inside a pressure cooker for 10 min. Once sections returned to room temperature, samples were
283 washed in deionized (DI) water and then permeabilized and blocked for 2 h using a blocking buffer (2 mg/mL
284 BSA and 0.1% Triton X-100 in PBS); as described previously(29). Primary antibodies CD14 (1:100 dilution) and
285 SARS-COV2 Nucleoprotein (1:250 dilution) were diluted in blocking buffer and allowed to incubate overnight
286 at 4°C. Secondary antibodies and DAPI (1:500 dilution) were diluted in blocking buffer and allowed to incubate
287 for 2 h in the dark. ProLong Glass was used as a mounting medium. #1 Thick Coverslips were applied to slides
288 and sealed. Samples were stored at 4°C until imaged. As negative control (to determine the specificity of staining)
289 we eliminated the step of incubation with primary antibodies while maintaining all other steps.

290 Images were acquired at room temperature with Leica TCS SPE confocal and with DMI4000 B
291 microscope using the Leica LAS-AF Software. Images were taken with a 40× oil-immersion objective using 405-
292 ,488-, 561-nm laser lines for excitation. Z-stack images were acquired by successive Z-slices of 1μm in the desired
293 confocal channels. Fields of view that were of interest and demonstrated a positive immunofluorescence signal
294 were selected and imaged. Z-slices of a Z-stack were overlaid to create maximum intensity projection images; all
295 images were processed using FIJI (Image J) software. Power analyses for these studies have been presented in
296 **Supplemental Information 7**.

297

298 **Collection of blood from COVID-19 patients**

299 Blood from COVID-19 donors was either obtained at a UC San Diego Health clinic under the approved IRB
300 protocols of the University of California, San Diego (UCSD; 200236X) or recruited at the La Jolla Institute under
301 IRB approved (LJI; VD-214). COVID-19 donors were California residents, who were either referred to the study
302 by a health care provider or self-referred. Blood was collected in acid citrate dextrose (ACD) tubes (UCSD) or in
303 EDTA tubes (LJI) and stored at room temperature prior to processing for plasma collection. Seropositivity against
304 SARS-CoV-2 was confirmed by ELISA. At the time of enrollment, all COVID-19 donors provided written
305 informed consent to participate in the present and future studies. Patient characteristic is listed in **Supplemental**
306 **Information 3**.

307

308 **Human serum cytokines measurement**

309 Human serum cytokines measurement was performed using the V-PLEX Custom Human Biomarkers from MSD
310 platform [MESO QuickPlex SQ 120]. Human serum samples fractionated from peripheral blood of KD and MIS-
311 C patients (all samples collected prior to the initiation of treatments) were analyzed using customized standard
312 multiplex plates as per the manufacturer's instructions using recommended software [MSD® DISCOVERY
313 WORKBENCH 4.0].

314

315 **COVID-19 lung models in-a-dish**

316 Monolayers derived from adult lung organoids (ALOs), primary airway cells, or hiPSC-derived alveolar type II
317 (AT2) pneumocytes were infected with SARS-CoV-2 to create in vitro lung models of COVID-19 as described
318 previously(28). The raw data and processed data were deposited in Gene Expression Omnibus under accession
319 no. GSE157057. These datasets were analyzed here using various gene signatures.

320

321 **COVID-19 modeling in Syrian hamsters**

322 Lung samples from 8-week-old Syrian hamsters were generated from experiments conducted exactly as in a
323 previously published study (18, 30, 31). Briefly, the infected (IN) arm was challenged with an intranasal dose of
324 $10^6$ PFU SARS-CoV-2 on day 1. The indicated treatments (or corresponding vehicle controls) were administrated
325 either a day prior (monoclonal antibodies; 2 mg per animal, intraperitoneal route) or 4 h prior (EIDD-2801; 500
326 mg/kg, oral route) to the CoV-2 challenge as a single dose. Animals were scarified $5^{th}$ day after infection to collect
327 blood and lung tissues (18, 30, 31). Viral titers, weight and lung histology findings have been published elsewhere
328 (18, 30, 31). Animal studies were approved and performed in accordance with Scripps Research IACUC Protocol
329 #20-0003 (30, 31). Power analyses for these studies have been presented in **Supplemental Information 7**.

330 **Immunohistochemistry**

331 COVID-19 samples were inactivated by storing in 10% formalin for 2 days and then transferred to zinc-formalin
332 solution for another 3 days. The deactivated tissues were transferred to 70% ethanol and cassettes were prepared
333 for tissue sectioning. The slides containing hamster and human lung tissue sections were deparaffinized in xylene
334 (Sigma-Aldrich Inc., MO, USA; catalog# 534056) and rehydrated in graded alcohols to water. For GRP78/ BIP,
335 p53, Cytokeratin 8, and Claudin 4-specific polyclonal antigen retrieval, slides were immersed in Tris-EDTA
336 buffer (pH 9.0) and boiled for 10 minutes at 100°C. Slides were immersed in Sodium Citrate Buffer (pH 6.0) and
337 boiled for 10 minutes at 100°C, for p14 ARF (E3X6D) and p21 Waf1/Cip1(12D1) antigen retrieval. Endogenous
338 peroxidase activity was blocked by incubation with 3% H2O2 for 5-10 minutes. To block non-specific protein
339 binding either 2.5% goat or 2.5% horse serum (Vector Laboratories, Burlingame, USA; catalog# MP-7401 or S-
340 1012) was added. Tissues were then incubated with the following antibodies: rabbit GRP78/BIP polyclonal
341 antibody (1:500 dilution; proteintech®, Rosemont, IL, USA; catalog# 11587-1-AP), rabbit p53 polyclonal
342 antibody (1:200 dilution ; proteintech®, Rosemont, IL, USA; catalog# 21891-1-AP), rabbit Cytokeratin 8
343 polyclonal antibody (1:2000 dilution; proteintech®, Rosemont, IL, USA; catalog# 10384-1-AP), and rabbit
344 Claudin 4-specific polyclonal antibody (1:200 dilution, proteintech®, Rosemont, IL, USA; catalog# 16195-1-AP)
345 for 1.5 hours at room temperature in a humidified chamber then rinsed with TBS twice for 3 minutes each.
346 Antibodies were prepared in antibody diluent (MP Biomedicals, LLC., Catalog # 980641) for rabbit p14 ARF
347 (E3X6D) monoclonal antibody (1:250 dilution; Cell Signaling Technology, Danvers, MA, USA; catalog#
348 74560S) and rabbit p21 Waf1/Cip1 (12D1) monoclonal antibody (1:50 dilution; Cell Signaling Technology,
349 Danvers, MA, USA; catalog# 2947S) that were then incubated in a humidified chamber overnight at 4°C, then
350 rinsed with TBST twice for 5 minutes each. Sections were incubated with horse anti-rabbit (Vector Laboratories,
351 Burlingame, USA; catalog# MP-7401) secondary antibodies for 30 minutes at room temperature in a humidified
352 chamber and then washed with TBS or TBST 3x, 5 minutes each. Then incubated with DAB (BioGenex, Fremont,
353 CA, USA; catalog # HK542-XAKE) for 5-10 minutes and counterstained with hematoxylin (BioGenex, Fremont,
354 CA, USA; catalog# HK100-9K), dehydrated in graded alcohols, cleared in xylene, and cover slipped. All sides
355 were visualized by Leica DM1000 LED (Leica Microsystems, Germany).

356

357 **IHC quantification**

358 IHC images were randomly sampled at different 300x300 pixel regions of interest (ROI). The ROIs were analyzed
359 using IHC Profiler(32). IHC Profiler uses a spectral deconvolution method of DAB/hematoxylin color spectra
360 by using optimized optical density vectors of the color deconvolution plugin for proper separation of the DAB
361 color spectra. The histogram of the DAB intensity was divided into 4 zones: high positive (0 to 60), positive (61
362 to 120), low positive (121 to 180) and negative (181 to 235). High positive, positive, and low positive percentages

were combined to compute the final percentage positive for each region of interest (ROI). The range of values for the percent positive is compared among different experimental groups. To calculate the statistical significance (determination of the p value), one way ANOVA has been performed between respective groups of samples.

**Quantification of collagen**

Due to acidophilic nature of collagen, it stains with eosin in H&E stanning. The relative extracellular collagen deposition was assessed using color deconvolution in ImageJ. The same color deconvolution technique was also used when collagen was stained using Masson's Trichrome (33, 34). To calculate the statistical significance (determination of the p value), one way ANOVA has been performed between respective groups of samples.

**Ethics statement**

Animal studies were performed in compliance with the ethical guidelines outlined in a Scripps Research Institutional Animal Care and Use Committee (IACUC)-approved protocol (Protocol #20–0003; PI: Tom Rogers) (31).

Blood from COVID-19 donors was either obtained at a UC San Diego Health clinic under the approved IRB protocols of the University of California, San Diego (UCSD; 200236X) or recruited at the La Jolla Institute under IRB approved (LJI; VD-214). An informed consent was obtained in all cases. COVID-19 donors were California residents who were either referred to the study by a health care provider or self-referred.

The lung specimens from the COVID 19 positive human subjects were collected using autopsy (study was IRB Exempt). All donations to this trial were obtained after telephone consent followed by written email confirmation with next of kin/power of attorney per California state law (no in-person visitation could be allowed into our COVID-19 ICU during the pandemic). The team member followed the CDC guidelines for COVID19 and the autopsy procedures.

**Role of funders**

Funders of the study had no role in study design, data collection, data analyses, interpretation, or writing of report.

389 **RESULTS**

390 **A study design that uses gene signatures as a computational framework to navigate COVID-19 lung disease**
391 Because PCLD is still an emergent illness that lacks insights into disease pathophysiology, we resorted to a study
392 design (**Figure 1**) that is geared to achieve three goals: (i) maximize rigor by using diverse transcriptomic datasets
393 cohorts (a total of 1825 unique samples, human: 1766; mouse: 41, hamster: 18; see **Supplemental Information**
394 **1**), but (ii) minimize bias by using well characterized sets of gene signatures from independent publications (see
395 **Supplemental Information 2**), and (iii) retain objectivity and precision by using artificial intelligence/machine
396 learning (AI/ML) tools that are supported by precise mathematical algorithms and statistical tools (see *Methods*).
397 First, to ensure that the findings maintain relevance to respiratory viral pandemics and COVID-19, we
398 used a recently discovered 166-gene signature that was found to be conserved in all viral pandemics (ViP),
399 including COVID-19, and a subset of 20-genes within that signature that classifies disease severity (18); the latter
400 was found to be consistently and significantly induced in the most severe cases of COVID-19 and prognostic of
401 poor outcome, e.g., death or invasive respiratory support (see **Supplemental Information 2**). These ViP
402 signatures were rigorously validated on ~45000 datasets representing the pandemics of the past (Training datasets:
403 Influenza and avian flu; GSE47963, n = 438; GSE113211, n = 118; Validated in numerous datasets as cited
404 before) and used without further training to prospectively analyze the samples from the current pandemic (i.e.,
405 COVID-19) (**Figure 1**, *Step 0*). The ViP signatures appeared to capture the 'invariant' host response, i.e., the
406 shared fundamental nature of the host immune response induced by all viral pandemics, including COVID-19.
407 Since they were discovered, these signatures have helped navigate the emergent syndrome of multisystem
408 inflammatory syndrome in children (MIS-C; a.k.a. PIMS-TS)(35). To avoid overreliance on one set of signatures
409 (i.e., ViP/sViP, which are tissue-agnostic signatures of host response), and increase the specificity for lung tissue,
410 an independent gene signature was included, one that is induced in COVID-19 lung tissues, but not in influenza
411 (13).
412 We began first by using the ViP signatures as a starting computational framework to navigate the
413 syndrome of COVID-19 lung disease, in conjunction with another independently reported (13) COVID-19-lung
414 derived signature to identify idiopathic pulmonary fibrosis (IPF)(36) as a computational match for COVID-19
415 lung disease (**Figure 1**, *Step 1*). The subsequent steps in the study use the ViP signatures alongside IPF-specific
416 PBMC-derived prognostic signatures (**Figure 1**, *Step 2*) and signatures of alveolar cytopathic changes (**Figure 1**,
417 *Step 3*) to systematically assess the degree of similarities between COVID-19 lung and IPF and the crossover of
418 pathophysiologic processes in the two conditions. In the final step (**Figure 1**, *Step 4*), a set of gene signature-
419 inspired protein-protein interaction (PPI) network analysis is used which pinpointed alveolar ER stress as a
420 potential early step in both IPF and PCLD, which is sufficient to recapitulate not just the ViP signatures, but also

421 the alveolar cytopathic features that are shared between the diseases. Key predictions are validated in human lung
422 tissues and plasma from COVID-19 subjects, hamster models of COVID-19, and human pre-clinical lung models
423 (**Figure 1**, *Steps 3-4*).
424

425 **COVID-19 lung disease and IPF induce a common set of gene expression signatures**

426 We first asked which pathologic lung condition comes closest to COVID-19 lung disease regarding the host
427 immune response. To this end, we used the ViP/sViP signatures (found to be induced in all CoV samples tested
428 so far (18)) and a second CoV-lung disease gene signature which was derived from a differential expression
429 analysis on lung samples from healthy controls vs. fatal COVID-19(13) (see **Supplemental Information 2**). As
430 for the lung diseases, we ensured that all four major pathologic conditions were represented: neoplastic (e.g.,
431 carcinoids and adenocarcinoma), granulomatous (e.g., tuberculosis/TB and sarcoidosis), allergic/infectious (e.g.,
432 severe asthma and community-acquired pneumonia/CAP and tuberculosis) and vasculopathic (e.g., pulmonary
433 hypertension/PHT and chronic thromboembolic pulmonary hypertension/CTEPH) (**Figure 2a**). As expected,
434 ViP/sViP signatures were induced in infectious diseases (CAP and tuberculosis), and to our surprise, these
435 signatures were upregulated also in sarcoidosis, a granulomatous condition that progresses to fibrosis in ~20% of
436 the patients (37, 38) (**Figure 2a**). Only two diseases (tuberculosis and IPF) showed significant induction of both
437 ViP and CoV-lung signatures, of which only one condition (IPF) manifests as a diffuse disease with patchy
438 parenchymal involvement resembling COVID-19 lung disease (**Figure 2a**).

439 We next analyzed a set of six signatures, each independently derived from diverse IPF-centric studies (39-
440 42), on diverse CoV samples; we also analyzed the ViP/sViP(18) and CoV-lung (13) signatures on 4 independent
441 single-cell RNA sequence (scRNA seq) datasets from IPF explants. We found that CoV samples induced the IPF
442 signatures and IPF samples induced ViP/CoV-lung signatures (**Figure 2b**). It is noteworthy that such induction
443 of ViP/sViP signatures was undetectable in the bulk RNA seq studies on IPF explants (**Supplementary Figure**
444 **1a**); this is in keeping with numerous reports that scRNA-seq obtains cell-level resolution in IPF that bulk-RNA
445 seq misses(43-45).

446 Next, we created a multivariate model to decompose the covariance between ViP, sViP, CoV lung and all
447 other IPF signatures to estimate the amount of overlaps between their induction in the IPF samples. These studies
448 showed that the ViP signature (alongside a some IPF signatures) is an independent variable that is induced in IPF
449 samples (**Supplementary Figure 1b**). However, such induction was observed exclusively in scSeq datasets
450 (**Supplementary Figure 1b**), but not in bulk Seq dataset (**Supplementary Figure 1c**), which is also consistent
451 with **Supplementary Figure 1a**.

452 Finally, in a pooled dataset (GSE149878)(46) that included both CoV and IPF samples the ViP (**Figure**
453 **2c**), sViP (**Figure 2d**), IPF-lung (**Figure 2e**), and CoV-lung (**Figure 2f**) were indistinguishably induced in both

454 samples compared to healthy controls. These results indicate that CoV and IPF signatures have crossover utility
455 for sample classification and indicate that IPF is the closest computational match to COVID-19 lung disease
456 (**Figure 2g**). Consistent with the fact that PCLD occurs in survivors of severe COVID-19 that required
457 hospitalization(47), we found that sViP signature was induced significantly higher in hospitalized compared to
458 non-hospitalized subjects (**Figure 2h**). Finally, we asked what, if any relationship exists between another poorly
459 understood feature of long-haul COVID-19, i.e., severe fatigue(48, 49). To improve rigor, we analyzed a set of
460 other fatigue syndromes (e.g., gulf war illness, chronic fatigue syndrome, and idiopathic chronic fatigue) and for
461 consistency, restricted the analysis to only whole blood/PBMC samples (**Figure 2i**). ViP/sViP and CoV-lung
462 immune signatures were induced only in CoV samples, but not in the remaining fatigue-associated syndromes
463 (**Figure 2i**), indicating that gene expression signatures in long-haul COVID-19 are distinct from other fatigue-
464 associated syndromes. Instead, fatigue in long-haul COVID-19 may be of similar reasons/pathophysiology that
465 has been observed in IPF (50). It is noteworthy that profound fatigue is a shared phenomenon in both IPF and
466 sarcoidosis (50, 51), both carry the risk of progressive fibrosis, and both conditions induce the ViP/sViP response
467 (**Figure 2a**). These results show that COVID-19 lung is closest to IPF: both are characterized by lung fibrosis,
468 and share gene expression patterns, suggestive of shared molecular and cellular mechanisms of disease
469 pathogenesis.
470

471 **PBMC-derived prognostic signatures in COVID-19 and IPF show crossover utility in both diseases**

472 Previously we showed using sc-RNA seq as well as bulk-RNA seq on sorted cells, that ViP/sViP signatures are
473 primarily induced in COVID-19 by lung epithelial cells and myeloid cells (both alveolar macrophages and
474 PBMCs)(18). When assessed in whole blood and/or PBMCs, the sViP signature was prognostic in COVID-19—
475 a high expression of this 20-gene signature was associated with prolonged hospital stay and mechanical
476 ventilation (18). We noted that a PBMC-based 52-gene signature has previously been validated in numerous IPF
477 datasets (27). In the setting of IPF, the role of PBMC-derived lung macrophages in the activation of alveolar
478 fibroblasts has been cemented by several studies (52-57) (**Figure 3a**); the activated macrophages and
479 myofibroblasts are thought to cross-stimulate each other, resulting in a vicious cycle that assures propagation of
480 fibrosis throughout the lung in end-stage IPF (52, 58). Emerging studies on CoV lung have also reported the
481 presence of pro-fibrogenic alveolar macrophages in severe COVID-19 (15, 59, 60). Because
482 monocytes/macrophages induce ViP signatures, whereas fibroblasts do not (18, 30, 31), we focused on monocytes
483 and investigated if the PBMC-derived prognostic signatures for IPF also prognosticate outcome in CoV, and *vice*
484 *versa*.

485 We found that alveolar macrophages (marked by CD-14 (61)) in the lungs of deceased subjects contained
486 the viral nucleocapsid protein (**Figure 3b-c**); this is consistent with observations reported by others (25, 62). The

487 20-gene sViP signature was prognostic in the blood datasets from both COVID-19 (**Figure 3d**; *left*) and IPF
488 subjects (**Figure 3d**; *middle, right*); among IPF subjects, the signature was prognostic exclusively in male subjects
489 (**Figure 3d**; *middle*). The 52-gene 'IPF-specific' prognostic signature (27) retained its ability to prognosticate
490 outcomes in both male and female patients with IPF (as expected; **Figure 3e**; *middle, right*), but also did so in the
491 setting of COVID-19 (**Figure 3e**; *left*). By contrast, the CoV-lung tissue-derived signature that is known to
492 distinguish CoV from both healthy and influenza-affected lungs(63), did not prognosticate outcome in either
493 COVID-19 (**Figure 3f**; *left*) or IPF subjects (**Figure 3f**; *middle vs. right*). It indicates that the PBMC-derived
494 signatures in CoV and IPF retain their prognostic relevance in both conditions, with at least one key notable
495 difference; sViP, but not the IPF signature revealed the well-established gender specific differences in the
496 progression of IPF. These results indicate that PBMCs in both IPF and COVID-19 have similar gene induction
497 patterns and that the cross-over prognostic value of IPF- and sViP signatures induced in the PBMCs may reflect
498 the profibrogenic immune response that is shared in both conditions.
499

500 **Alveolar type II (AT2) telomere dysfunction and senescence are shared features in both COVID-19 and**
501 **IPF**.

502 Besides the contributions from PBMCs, a contemporary paradigm in IPF is that chronic injury to distal lung tissue
503 leads to either loss or altered function of epithelial stem cells (i.e., AT2 cells), which promote dysregulated repair
504 and pathogenic activation of fibroblasts (64-66). We asked if the AT2 dysfunction states previously reported in
505 IPF also occur in COVID-19. To this end, we analyzed a single cell dataset of bronchoalveolar lavage from
506 COVID-19 patients (GSE145926) for the induction of ViP/sViP, CoV-lung, and IPF signatures and other
507 signatures that are indicative of three important AT2 cytopathies: (i) Damage associated transient progenitor
508 (DATP(67)), a distinct *KRT8+/CLDN4+* AT2 lineage that has been independently shown, nearly simultaneously
509 by 3 independent groups(67-69) to be associated with activated myofibroblasts during lung fibrosis in the setting
510 of IPF(66). (ii) AT2-senescence signature, which was derived from another IPF-focused study (70) that used
511 human and mouse models to pinpoint the Doxycycline-induced program of p53-dependent cellular senescence,
512 AT2 cell depletion, and spontaneous, progressive pulmonary fibrosis; and (iii) Aging-associated telomerase
513 dysfunction (TERC_UP) signature, which was derived from aging telomerase knockout (*Terc*$^{-/-}$) mice. We found
514 that the entire panel of signatures analyzed were upregulated, most consistently, only in the epithelial cells (**Figure**
515 **4a**), whereas only some were induced and not others in the non-epithelial compartments.

516 We next confirmed that a composite score of *IL15/IL15RA*, AT2-senescence, and telomerase dysfunction
517 (TERC_UP) signatures were induced indistinguishably in both CoV and IPF samples compared to healthy
518 controls (**Figure 4b**). A correlation matrix (**Figure 4c**) on the entire cohort showed that while both IL15 and
519 IL15RA correlate with the ViP signature (as expected), IL15RA, but not IL15 correlate with TERC_UP and AT2-

520 senescence signatures. Both ViP and sViP signatures correlate with AT2 senescence. Furthermore, when we
521 tested AT2-senescence and telomere dysfunction in other CoV datasets, we found that these features were
522 conserved in most CoV samples across diverse cohorts (**Figure 4d**). An independent sc-RNA seq IPF dataset
523 further confirmed that, much like CoV lung, the induction of ViP/sViP in the alveolar epithelium (EpCAM+;
524 **Figure 4e**) is associated with IL15/IL15RA induction, AT2-senescence, and telomere dysfunction. These findings
525 support the existence of telomere dysfunction and DNA damage in both acute COVID-19 and IPF.
526

527 **Hamsters, but not mice recapitulate the host immune response and AT2 cytopathic features of COVID-19**
528 Evidence suggests that there is a wide gap between COVID-19 in humans and animal models(71); interspecies-
529 related differences, such as host specificity and divergent immune responses can lead to misinterpretation. To
530 objectively assess which model improves the translational potential of any discovery from this study, we curated
531 all publicly available transcriptomic datasets from pre-clinical rodent models of COVID-19 and analyzed them for
532 the induction of the ViP/sViP, IL15/IL15RA composite score, CoV-lung, IPF signatures and signatures of AT2
533 senescence and transition-state arrest (DATP) (**Supplementary Figure S2a**). We saw that the golden Syrian
534 hamsters, but not the humanized transgenic mouse models (**Supplementary Figure S2a**; *bottom*) induce most, if
535 not all the key signatures upon SARS-CoV-2 challenge (**Supplementary Figure S2a**; *top*). Most importantly, all
536 signatures (**Supplementary Figure S2a**; *top*), including that of AT2-damage associated transient progenitor
537 population (**Supplementary Figure S2b**; DATP) were effectively reversed with two therapeutic approaches. The
538 first approach was the use of N-hydroxycytidine, the parent of the prodrug MK-4482 (Molnupiravir, EIDD-2801)
539 which has not only proven as a potent and selective oral antiviral nucleoside analogue in mice, guinea pigs, ferrets
540 and human airway epithelium organoids (71-76) but also showed promise in Phase IIa trials in the treatment of
541 COVID-19 patients (NCT04405570(77)). The second approach was the use of SARS-CoV-2-neutralizing
542 antibodies whose design was inspired by monoclonal antibodies (mAbs) isolated from convalescent donors(31).
543 A specific isotype of this antibody, which binds to the receptor-binding domain (RBD-A) of SARS-CoV-2 spike
544 protein in a fashion that precludes binding to host ACE2, was demonstrated as effective in preventing infection
545 and weight loss symptoms, in cell-based and *in vivo* hamster models of infection, respectively.
546 These findings show that hamster lungs best recapitulate the host immune response, the nature of the
547 cytokine storm (IL15-centric), and the AT2 cytopathies that are shared between COVID-19 and IPF. It is also in
548 keeping with emerging data that the Syrian hamsters best emulate COVID-19(71); they also recapitulate
549 progressive fibrosis in IPF(78).
550

551 **Validation of the predicted AT2 cytopathic features of COVID-19 in hamster and human lungs**

552 We next sought to determine if the SARS-CoV-2 virus can induce the AT2 cytopathic changes and whether
553 effective therapeutics can abrogate those changes. We analyzed by RNA seq and trichrome stain the lungs from
554 SARS-CoV-2-challenged golden Syrian hamsters who were pre-treated either with EIDD-2801 or anti-Spike
555 (CoV-2) mAb or untreated controls (see *Methods* and study protocol in **Figure 5a**) (30, 31). Weight loss,
556 histologic confirmation for the presence/extent of inflammatory infiltrates in the lung and serum viral titers in
557 these experimental cohorts have been confirmed and reported previously(18, 30, 31). Key genes associated with
558 AT2 senescence (*Tp53* and *Cdkn2a*), ER stress (*Hspa5/Grp78*), DATP (*Krt8* and *Cldn4*) and fibrosis (*Col1a,*
559 *Col3a,* and *Serpine1*) were all induced in the infected lungs (compared to uninfected controls) but not in those
560 that were treated (**Figure 5b**). Trichrome staining confirmed that compared to uninfected controls, collagen
561 deposition was induced in the infected lungs, but not in those that were treated (**Figure 5c-d**). Immunohistology
562 studies confirmed that Cdkn2a, Krt8, Cldn4 and Tp53 proteins were expressed in the injured alveoli, but not in
563 the treated lungs (**Figure 5e-f**). Because the degree of viremia, the severity of the infection, and host immune
564 response (the IL15-centric ViP signatures) were significantly blunted by both treatments(18, 30, 31), the orally
565 administered anti-viral EIDD-2801 and the systemic administration of anti-spike mAb, we conclude that the AT2
566 cytopathic features are a consequence of severe CoV infection.

567 We next confirmed the extensive deposition of eosin-positive collagenous materials in lung tissues
568 obtained at rapid autopsies on deceased subjects with COVID-19 compared to normal lung tissues (**Figure 6a**;
569 *top*). KRT8 and TP53 were also induced in the COVID-19 lungs (**Figure 6a**; *middle* and *bottom*). Finally, in
570 plasma ELISA studies in a cohort of symptomatic COVID-19 patients who presented to the UC San Diego
571 Medical Center with varying disease severity, ranging from mild to fatal (see **Supplemental Information 3-4**),
572 we detected several cellular senescence-associated cytokines during the acute compared to the convalescent visit
573 (**Figure 6b; Supplementary Figure S3**). It is noteworthy that VEGF, TNFα, IL6, and IL1β, all cytokines that
574 have been implicated in the fibrotic process post-COVID-19 (reviewed in (79)) were significantly elevated only
575 in the most critical form of the disease. Findings are also consistent with the fact that VEGF-targeted therapeutics
576 (Bevacizumab) in COVD-19 patients improved lung oxygenation up to day 28 during follow-up(80).

577 We next asked if the key aspects of the host immune response and AT2 cytopathic changes are induced
578 in the various human pre-clinical *in vitro* models of COVID-19. We assessed our published models of COVID-
579 19-in-a-dish (28) (**Figure 6c, 6d-***top 4 rows*) for AT2-senescence and for their ability to mimic the unique milieu
580 in COVID-19 lungs(13). We also analyzed CoV2-challenged primary(p)AT2 and A549 cells (**Figure 6d-***bottom*
581 *3 rows*). We found that submerged monolayers of adult lung organoids (ALO mono), which were previously
582 found to outperform fetal lung organoids (81) and other models in recapitulating the gene expression signatures
583 observed in human COVID-19 lungs(28), also outperformed here. While most models induced some signatures,
584 the submerged ALO model induced them all (**Figure 6d**).

585

586 **Protein-protein network analysis predicts AT2 ER stress as an early initiating event in COVID-19**

587 To identify some of the earliest AT2 cell processes that might be triggered as a consequence of the host immune
588 response in acute COVID-19, which may precede the development of transient state arrest and senescence, we
589 resorted to a PPI network analysis. We used the genes in the sViP (18), AT2 senescence(70) and DATP(67)
590 signatures as 'seed' nodes to fetch additional connecting nodes from human protein-protein interactions in
591 STRING database (21) (**Figure 7a;** see **Supplemental Information 5)**. Once built, we assessed the network for
592 degrees of connectivity (**Figure 7b**; *left*) and representation of cellular processes in the most connected nodes by
593 using two approaches: CluGo analysis(82) (**Figure 7b**; *right*) and Reactome pathway analysis(20) (**Figure 7c**).
594 While CluGO analysis showed an overrepresentation of proteins associated with unfolded protein response, ER
595 stress, regulation of telomerase, cell aging and, cell cycle (G1/S phase transition) (**Figure 7b**; *right*), Reactome
596 analysis showed enrichment of pathways associated with senescence (**Figure 7c**; *red*), platelet activation and
597 degranulation (**Figure 7c**; *blue*), and, Tp53-dependent DNA damage response (**Figure 7b**; *green*). That platelet
598 activation (indicative of a prothrombotic state) is captured within the PPI network is in keeping with what appears
599 to be a hallmark of PCLD(13). Numerous independent observations in diverse cohorts of patients have now
600 confirmed that platelet dysfunction amplifies endotheliopathy in COVID-19 (83) and that a vicious cycle of
601 thromboinflammation, endothelial dysfunction, and immunothrombosis are key pathogenic mechanisms in
602 COVID-19(84, 85). It is noteworthy, that we found that levels of IL6 and platelets showed a significant negative
603 correlation (**Supplementary Figure 3**). Because IL6 has been specifically implicated in platelet hyper-reactivity
604 and thrombosis(86), as well as in thrombus resolution (87), the strong negative correlation may suggest the
605 presence of thromboinflammation; the latter is one of the first and most distinctive pathological features of CoV-
606 lung to be reported since the beginning of this pandemic(13).

607 As for AT2-centric early events/processes, we decided to focus on ER stress because the PPI network
608 revealed so, and because there is fundamental support for its role in cytokine elaboration, apoptosis, and disrupted
609 progenitor cell function as upstream drivers of lung fibrosis(88). AT2 ER stress has been shown to serve as a
610 causal role for profibrogenic epithelial states (SASP) that is shared between aging-associated pulmonary fibrosis
611 and IPF(89), raising the possibility that it may also be a common underlying factor in fibrotic COVID-19 lung
612 disease.

613

614 **ER stress in AT2 cells is observed in COVID-19 and is sufficient to recapitulate the host immune response**
615 **in COVID-19 and IPF**

616 We next sought to investigate if alveolar ER stress is a feature in CoV-lung. To this end, we analyzed infected
617 human (**Figure 8a**) and hamster (**Figure 8b**) lungs for the ER chaperone *Grp78*, a key regulator of ER

618 homeostasis expression by IHC. We found that compared to normal uninfected lungs, infected lungs showed a
619 significant increase in Grp78 staining and that the intensity of such staining was highest in the epithelial cells
620 lining the alveolar spaces in both species (**Figure 8a-b**).

621 To determine if ER stress in AT2 cells may contribute to the epithelial dysfunction and fibrosis that is
622 observed as a sequel of COVID-19, we leveraged a previously characterized *tg*-mouse model of AT2-specific
623 conditional [tamoxifen (Tmx)-inducible] deletion of Grp78 (*Sftpc*$^{+/creERT2}$;*Grp78*$^{f/f}$)(89). These mice are known to
624 lose weight and die (more pronounced in male and old mice) due to lung inflammation, and spatially
625 heterogeneous fibrosis characterized by fibroblastic foci and hyperplastic AT2 cells, all features of IPF(89). We
626 found that compared to their littermate controls, lungs from these mice (unlike other transgenic mouse models of
627 COVID-19; see **Figure S2a**) induced the ViP/sViP, CoV2-lung, and IPF signatures, as well as the AT2 cytopathic
628 changes such as DATP and AT2 senescence (**Figure 8c**; *bottom row*). In fact, the pattern of gene expression was
629 similar to SARS-CoV-2-challenged hamster lungs (**Figure 8c**; *top row*). These findings indicate that ER stress is
630 sufficient to faithfully recapitulate the host immune response and alveolar cytopathic changes that are observed
631 in COVID-19 lung disease and IPF.

632 **DISCUSSION**

633 The major discovery we report here is that lung disease in severe COVID-19 resembles IPF, the most common
634 form of ILD, at a fundamental level—showing similar gene expression patterns (ViP and IPF signatures), shared
635 prognostic signatures (in circulating PBMCs), dysfunctional cell states (AT2 and monocytes), and dysfunctional
636 AT2 processes (ER stress, telomere instability, and senescence)]. It is also noteworthy that this semblance
637 between COVID-19 and IPF was identified through a comprehensive and unbiased computational approach (see
638 **Figure 1**) which compared gene expression signatures in COVID-19 against transcriptomic datasets representing
639 a plethora of neoplastic, granulomatous, allergic/infectious, and vasculopathic pathologic conditions of the lung.
640 Although two other diseases (sarcoidosis and tuberculosis) that are characterized by focal fibrosis or lower
641 incidence of fibrosis were found to induce some of the COVID-19-associated lung signatures, IPF induced them
642 all and matched COVID-19 closely due to the diffuse nature of fibrosis. This finding suggests that the nature of
643 the host immune response to injury and the proximal AT2 cytopathies we report here are unlikely to be mere
644 sequelae of cellular and molecular events during any fibrotic remodeling of the lung and may instead be specific
645 to ILDs such as IPF and COVID-19.

646 Our findings support the following model (see **Figure 8d**): Upon injury (viral infection/inflammatory
647 cytokines), diffuse alveolar damage is associated with extensive loss of AT1 cells, which are duly replaced by
648 AT2 cells that serve as progenitors that are capable of self-renewal and differentiation into AT1 cells (90-92). But
649 such physiologic regeneration and healing process is impaired when AT2 progenitors are also injured and begin
650 to elicit ER stress responses, which induces telomerase activity (93, 94). Because severe COVID-19 has been
651 linked to short telomere lengths(94) [not as a cause, i.e, SARS-CoV-2 infection does not shorten telomeres, but
652 as a predisposition (95)], it is possible that patients who have short and/or dysfunctional telomeres fail to
653 adequately respond to ER stress, and instead, become senescent (through p53 activation), accumulate DNA
654 damage and enter a dysfunctional *KRT8*+ transitional stem cell state (96). The latter is a phenomenon that has
655 been independently reported by three groups, under different names: PATS, pre-alveolar type-1 transitional cell
656 state (69); ADI, alveolar differentiation intermediates (68); DATP, damage-associated transient progenitors (67).
657 All three groups reported their presence in fibrotic regions of IPF lungs; an increase of AT2 cells in the transitional
658 state was invariably accompanied by an increase in myofibroblasts. Our PPI network analyses suggested that AT2
659 senescence and progenitor arrested state in the setting of a ViP/sViP-immune response (cytokine storm
660 contributed by PBMCs) may support a SASP phenotype, which is a pathological feature of aging and IPF lung
661 (97, 98).

662 It is possible that fibroblasts/myofibroblasts also contribute to this vicious cycle of inflammation and AT2
663 dysfunction. These findings show that these two distinct clinical syndromes, IPF, which predates the current
664 pandemic by many decades, and the novel COVID-19, share a similar profile of host immune response, both in

the lung microenvironment (in AT2 cells, to be specific) as well as in the circulating blood/PBMCs. We not only formally define the nature of that host response and provide computational tools to measure the extent of such response, but also chart the cascade of cytopathic changes in the alveoli that are critical for the profibrogenic state. It is noteworthy that the same host immune response is seen also in MIS-C, a new disease that co-emerged with COVID-19, and in KD (which shares overlapping features with MIS-C in clinical presentation (35)).

There are three major impacts and/or implications of the findings reported in this study.

First, our finding that COVID-19 and IPF share fundamental host immune response and alveolar cytopathic features is in keeping with the fact that both diseases share epidemiologic similarities—they primarily impact older adults, males more than females, and are characterized by progressive worsening of dyspnea and lung function (99, 100). The induction of ViP/sViP signatures in IPF is consistent with gathering consensus in the past decade that IPF may be a multi-trigger infection-driven chronic inflammatory condition (36, 101-103). Finally, patients with existing ILDs have greater odds of death due to COVID-19 compared with adults without ILD, even after controlling for age, sex, and comorbidities(104-106). On day 30 of COVID-19, 35% of patients with fibrotic idiopathic ILD had died compared with 19% of those with other ILDs(107). Whether the shared molecular or pathophysiologic features we report here synergize to increase fatality remains unknown; however, these shared features support the rationalization of clinical trials in COVID-19 using FDA-approved drugs for IPF, nintedanib (Ofev®), and pirfenidone (Esbriet®) (NCT04856111 and NCT04653831), and anecdotal case reports and case series have already chronicled the benefits of their use in COVID-19 lung disease (108-110).

Second, most severe COVID-19 patients develop pneumonia and hyperinflammation likely related to a macrophage activation syndrome (111) commonly named “cytokine storm”. Although this storm has been implicated, and it is comprised of the generic mix of all the typical cytokines(10), which exact component is the pro-fibrogenic driver remains unclear. By showing that an IL15/IL15RA-centric cytokine storm is the key shared phenomenon between CoV and IPF lungs, this study provides a link between hyperinflammation and the sequelae of fibrosis. Findings are consistent with prior reports of circulating IL15 as a biomarker for prognostication in IPF and other ILDs(112-114). As for what may be contributing to such a cytokine response, myeloid cells are likely to be major culprits, but dysfunctional AT2 cells cannot be ruled out. Prior studies using bleomycin-challenged transgenic mice that lack the *Telomeric Repeat Binding Factor 2* in the AT2 cells (*Trf2*$^{Fl/Fl}$*;Sftpc-CreER*) have shown that AT2 cells with telomere dysfunction upregulate an IL15-centric pathways(115). Mutations in the two main components of the telomerase holoenzyme complex that is responsible forcreating new telomeric DNA, *Telomerase reverse transcriptase (TERT)* or *telomerase RNA (TERC),* are major monogenic

698 causes of pulmonary fibrosis(116). Patients with IPF have shortened telomeres; short telomeres and TERT/TERC-
699 disease associated variants were associated with specific clinical and biological features and reduced transplant-
700 free survival (116, 117). Similarly, short telomeres increase the risk of severe COVID-19(95, 118, 119),
701 pulmonary fibrosis, and poorer outcomes(119, 120). It is possible that injury, DNA damage and IL15 signaling
702 in AT2 cells are one component in the profibrogenic cascade and suggest that IL15-targeted therapeutics may be
703 beneficial in the most severe cases of COVID-19 to prevent fibrotic sequelae.
704
705 Third, our finding that shared alveolar cytopathic changes (e.g., DNA damage, progenitor state arrest,
706 SASP, and ER stress) fuel fibrogenic programs in both COVID-19 and IPF are insightful because AT2 cells are
707 known to contain an elegant quality control feedback loop to respond to intrinsic or extrinsic stress; a failure of
708 such quality control results in diverse cellular phenotypes (reviewed in (88)): ER stress (121), defective
709 autophagy, mitochondrial dysfunction, apoptosis, inflammatory cell recruitment, profibrotic signaling, and
710 altered progenitor function (96, 122, 123) that ultimately converge to drive downstream fibrotic remodeling in
711 the lung. Prior work (70), which led to the discovery of the AT2-senescence signature (which we used here), had
712 demonstrated that senescence of AT2 cells (not loss of AT2) is sufficient to drive progressive pulmonary fibrosis.
713 Others agree, and this is now an established pathophysiologic trigger in lung fibrosis(64, 124-126). Our findings
714 are in keeping with published work (127) that suggests that AT2 senescence may be a targetable disease driver of
715 lung injury in COVID-19. Although AT2 senescence is a shared phenomenon, our PPI network analyses --which
716 integrated AT2 processes with the immune responses (ViP signatures) -- provided valuable clues into how platelet
717 activation and thromboinflammation may be uniquely seen in the setting of COVID-19. These findings are in
718 keeping with recent publications (128-130) which suggest that SARS-CoV-2 could induce epithelial senescence;
719 senescent AT2 cells would then assume a SASP phenotype, which in turn led to neutrophil and platelet activation,
720 and activation of the clotting cascade (128).
721
722 As for the limitations of this study, a direct comparison of lung samples from patients who survived
723 COVID-19 but went on to develop (or not) restrictive (fibrotic) lung disease was not possible due to the lack of
724 such datasets at this stage of the pandemic. It also remains unclear how to model post-COVID-19 progressive
725 lung fibrosis *in vitro* and hence, no attempt was made to do so. Our results suggest that AT2 cells alone may be
726 insufficient for such modeling because a specific host immune response that is carried in the PBMCs is a clear
727 determinant as to who progresses and who does not. Although AT2-specific modulation of ER-stress pathway
728 and SARS-CoV-challenged (treated vs untreated) hamsters were used to go beyond association and establish
729 causation, our study did not attempt to inhibit/reverse fibrosis in COVID-19 by acting on any profibrogenic
730 cellular pathway/process. Development of novel chemical matter/biologicals and validation of the therapeutic

efficacy of such agents will take time, but if successful, our findings show that their benefits will likely transcend beyond PCLD into IPF and other fibrotic lung conditions such as IPF.

In conclusion, this transdisciplinary work provides insights into the pathogenesis of PCLD, formally defines the fibrogenic processes in the lung, and rigorously validates high value gene signatures or even targets (i.e., IL15, senescence pathways, etc.) to track and manipulate the same. The insights, tools, computationally vetted disease models, and biomarkers (prognostic gene signatures) identified here are likely to spur the development of therapies for patients with fibrotic interstitial lung disease of diverse causes, including IPF, all of whom have limited or no good treatment options.

740 ## CONTRIBUTORS

741 DS, SS and PG conceptualized the project; SS and DS carried out all computational modeling and analyses and
742 DS contributed all software used in this work; GDK carried out the human serum cytokine analysis; VC and SS
743 carried out the immunohistochemical studies and their quantitative analysis under the supervision of PG; AGF
744 carried out the immunofluorescence studies on COVID-19 lung tissues; SS carried out the PPI network creation
745 and analysis; JD provided access to human subjects for cytokine analyses. CT, CRE and SD were responsible for
746 the COVID-19 lung models. SS, DS and PG prepared figures for data visualization; SS and PG wrote the original
747 draft of the manuscript; DS, SD, GK edited and revised the manuscript. DS, SD, and PG supervised various parts
748 of the project and secured funding; DS and PG administered the project. SS, DS, and PG have accessed and
749 verified the underlying data. All co-authors have approved the final version of the manuscript for submission.
750

751 ## DECLARATION OF INTERESTS

752 The authors declare no competing interests.
753

754 ## DATA SHARING STATEMENT

755 **Data availability**

756 Source data are provided with this paper. All data is available in the main text or the supplementary materials.
757 Publicly available datasets used are enlisted in **Supplemental Information 1**.
758

759 **Code availability**

760 The software codes are publicly available at the following links:
761 https://github.com/sahoo00/BoNE(131) and https://github.com/sahoo00/Hegemon (132). The PPI network
762 analysis can be found at: https://github.com/sinha7290/PPIN.

763

## ACKNOWLEDGMENTS

This work was supported by the National Institutes for Health (NIH) grants R00-CA151673 and R01-GM138385 (to DS) and R01-AI141630, CA100768 and CA160911 (to P.G), R01DK107585 (SD) and R01-AI155696 (to P.G, D.S and S.D), UCOP-RGPO (R00RG2628 & R00RG2642 to P.G, D.S and S.D). GDK was supported through the American Association of Immunologists Intersect Fellowship Program for Computational Scientists and Immunologists. JMD was supported by U19AI142742. We thank Zea Borok (UCSD) for helpful discussions during the preparation of this manuscript.

## REFERENCES


1. Spagnolo P, Balestro E, Aliberti S, Cocconcelli E, Biondini D, Casa GD, et al. Pulmonary fibrosis secondary to COVID-19: a call to arms? Lancet Respir Med. 2020;8(8):750-2.
2. Myall KJ, Mukherjee B, Castanheira AM, Lam JL, Benedetti G, Mak SM, et al. Persistent Post-COVID-19 Interstitial Lung Disease. An Observational Study of Corticosteroid Treatment. Ann Am Thorac Soc. 2021;18(5):799-806.
3. Wells AU, Devaraj A, Desai SR. Interstitial Lung Disease after COVID-19 Infection: A Catalog of Uncertainties. Radiology. 2021;299(1):E216-e8.
4. Aronson KI, Podolanczuk AJ. Lungs after COVID-19: Evolving Knowledge of Post-COVID-19 Interstitial Lung Disease. Ann Am Thorac Soc. 2021;18(5):773-4.
5. Dadhwal R, Sharma M, Surani S. Restrictive Lung Disease in Patients With Subclinical Coronavirus Infection: Are We Bracing Ourselves for Devastating Sequelae? Cureus. 2021;13(1):e12501.
6. Atabati E, Dehghani-Samani A, Mortazavimoghaddam SG. Association of COVID-19 and other viral infections with interstitial lung diseases, pulmonary fibrosis, and pulmonary hypertension: A narrative review. Can J Respir Ther. 2020;56:1-9.
7. Ahmad Alhiyari M, Ata F, Islam Alghizzawi M, Bint IBA, Salih Abdulhadi A, Yousaf Z. Post COVID-19 fibrosis, an emerging complicationof SARS-CoV-2 infection. IDCases. 2021;23:e01041.
8. Alharthy A, Abuhamdah M, Balhamar A, Faqihi F, Nasim N, Ahmad S, et al. Residual Lung Injury in Patients Recovering From COVID-19 Critical Illness: A Prospective Longitudinal Point-of-Care Lung Ultrasound Study. J Ultrasound Med. 2020.
9. Gagiannis D, Steinestel J, Hackenbroch C, Schreiner B, Hannemann M, Bloch W, et al. Clinical, Serological, and Histopathological Similarities Between Severe COVID-19 and Acute Exacerbation of Connective Tissue Disease-Associated Interstitial Lung Disease (CTD-ILD). Front Immunol. 2020;11:587517.
10. George PM, Wells AU, Jenkins RG. Pulmonary fibrosis and COVID-19: the potential role for antifibrotic therapy. Lancet Respir Med. 2020;8(8):807-15.
11. van Gassel RJJ, Bels J, Remij L, van Bussel BCT, Posthuma R, Gietema HA, et al. Functional Outcomes and Their Association With Physical Performance in Mechanically Ventilated Coronavirus Disease 2019 Survivors at 3 Months Following Hospital Discharge: A Cohort Study. Crit Care Med. 2021.
12. Kaarteenaho R, Kinnula VL. Diffuse alveolar damage: a common phenomenon in progressive interstitial lung disorders. Pulm Med. 2011;2011:531302.
13. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. N Engl J Med. 2020;383(2):120-8.
14. Cheon IS, Li C, Son YM, Goplen NP, Wu Y, Cassmann T, et al. Immune signatures underlying post-acute COVID-19 lung sequelae. Sci Immunol. 2021:eabk1741.
15. Bharat A, Querrey M, Markov NS, Kim S, Kurihara C, Garza-Castillon R, et al. Lung transplantation for patients with severe COVID-19. Sci Transl Med. 2020;12(574).
16. Croci GA, Vaira V, Trabattoni D, Biasin M, Valenti L, Baselli G, et al. Emergency Lung Transplantation after COVID-19: Immunopathological Insights on Two Affected Patients. Cells. 2021;10(3).
17. Bharat A, Machuca TN, Querrey M, Kurihara C, Garza-Castillon R, Jr., Kim S, et al. Early outcomes after lung transplantation for severe COVID-19: a series of the first consecutive cases from four countries. Lancet Respir Med. 2021;9(5):487-97.
18. Sahoo D, Katkar GD, Khandelwal S, Behroozikhah M, Claire A, Castillo V, et al. AI-guided discovery of the invariant host response to viral pandemics. EBioMedicine. 2021;68:103390.
19. Sahoo D, Dill DL, Tibshirani R, Plevritis SK. Extracting binary signals from microarray time-course data. Nucleic Acids Res. 2007;35(11):3705-12.
20. Fabregat A, Jupe S, Matthews L, Sidiropoulos K, Gillespie M, Garapati P, et al. The Reactome Pathway Knowledgebase. Nucleic Acids Res. 2018;46(D1):D649-D55.

21. Szklarczyk D, Gable AL, Nastou KC, Lyon D, Kirsch R, Pyysalo S, et al. The STRING database in 2021: customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Res. 2021;49(D1):D605-D12.
22. Fink G, Orlova-Fink N, Schindler T, Grisi S, Ferrer APS, Daubenberger C, et al. Inactivated trivalent influenza vaccination is associated with lower mortality among patients with COVID-19 in Brazil. BMJ Evid Based Med. 2020.
23. Wagstaffe HR, Pickering H, Houghton J, Mooney JP, Wolf AS, Prevatt N, et al. Influenza Vaccination Primes Human Myeloid Cell Cytokine Secretion and NK Cell Function. J Immunol. 2019;203(6):1609-18.
24. Hoek KL, Samir P, Howard LM, Niu X, Prasad N, Galassie A, et al. A cell-based systems biology assessment of human blood to monitor immune responses after influenza vaccination. PLoS One. 2015;10(2):e0118528.
25. Carson WE, Giri JG, Lindemann MJ, Linett ML, Ahdieh M, Paxton R, et al. Interleukin (IL) 15 is a novel cytokine that activates human natural killer cells via components of the IL-2 receptor. J Exp Med. 1994;180(4):1395-403.
26. Carson WE, Ross ME, Baiocchi RA, Marien MJ, Boiani N, Grabstein K, et al. Endogenous production of interleukin 15 by activated human monocytes is critical for optimal production of interferon-gamma by natural killer cells in vitro. J Clin Invest. 1995;96(6):2578-82.
27. Herazo-Maya JD, Sun J, Molyneaux PL, Li Q, Villalba JA, Tzouvelekis A, et al. Validation of a 52-gene risk profile for outcome prediction in patients with idiopathic pulmonary fibrosis: an international, multicentre, cohort study. Lancet Respir Med. 2017;5(11):857-68.
28. Tindle C, Fuller M, Fonseca A, Taheri S, Ibeawuchi SR, Beutler N, et al. Adult stem cell-derived complete lung organoid models emulate lung disease in COVID-19. Elife. 2021;10.
29. Lopez-Sanchez I, Dunkel Y, Roh YS, Mittal Y, De Minicis S, Muranyi A, et al. GIV/Girdin is a central hub for profibrogenic signalling networks during liver fibrosis. Nat Commun. 2014;5:4451.
30. Bakowski MA, Beutler N, Wolff KC, Kirkpatrick MG, Chen E, Nguyen TH, et al. Drug repurposing screens identify chemical entities for the development of COVID-19 interventions. Nat Commun. 2021;12(1):3309.
31. Rogers TF, Zhao F, Huang D, Beutler N, Burns A, He WT, et al. Isolation of potent SARS-CoV-2 neutralizing antibodies and protection from disease in a small animal model. Science. 2020;369(6506):956-63.
32. Varghese F, Bukhari AB, Malhotra R, De A. IHC Profiler: an open source plugin for the quantitative evaluation and automated scoring of immunohistochemistry images of human tissue samples. PLoS One. 2014;9(5):e96801.
33. Chamberlain CS, Crowley EM, Kobayashi H, Eliceiri KW, Vanderby R. Quantification of collagen organization and extracellular matrix factors within the healing ligament. Microsc Microanal. 2011;17(5):779-87.
34. Rieppo L, Janssen L, Rahunen K, Lehenkari P, Finnilä MAJ, Saarakkala S. Histochemical quantification of collagen content in articular cartilage. PLoS One. 2019;14(11):e0224839.
35. Ghosh P, Katkar GD, Shimizu C, Kim J, Khandelwal S, Tremoulet AH, et al. An Artificial Intelligence-guided signature reveals the shared host immune response in MIS-C and Kawasaki disease. Nat Commun. 2022;13(1):2687.
36. Imai Y, Kuba K, Rao S, Huan Y, Guo F, Guan B, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. Nature. 2005;436(7047):112-6.
37. Bonham CA, Strek ME, Patterson KC. From granuloma to fibrosis: sarcoidosis associated pulmonary fibrosis. Curr Opin Pulm Med. 2016;22(5):484-91.
38. Patterson KC, Strek ME. Pulmonary fibrosis in sarcoidosis. Clinical features and outcomes. Ann Am Thorac Soc. 2013;10(4):362-70.
39. Meltzer EB, Barry WT, D'Amico TA, Davis RD, Lin SS, Onaitis MW, et al. Bayesian probit regression model for the diagnosis of pulmonary fibrosis: proof-of-principle. BMC Med Genomics. 2011;4:70.

40. Bauer Y, Tedrow J, de Bernard S, Birker-Robaczewska M, Gibson KF, Guardela BJ, et al. A novel genomic signature with translational significance for human idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol. 2015;52(2):217-31.
41. Gangwar I, Kumar Sharma N, Panzade G, Awasthi S, Agrawal A, Shankar R. Detecting the Molecular System Signatures of Idiopathic Pulmonary Fibrosis through Integrated Genomic Analysis. Sci Rep. 2017;7(1):1554.
42. Ammar R, Sivakumar P, Jarai G, Thompson JR. A robust data-driven genomic signature for idiopathic pulmonary fibrosis with applications for translational model selection. PLoS One. 2019;14(4):e0215565.
43. Lederer DJ, Martinez FJ. Idiopathic Pulmonary Fibrosis. N Engl J Med. 2018;378(19):1811-23.
44. Reyfman PA, Walter JM, Joshi N, Anekalla KR, McQuattie-Pimentel AC, Chiu S, et al. Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. Am J Respir Crit Care Med. 2019;199(12):1517-36.
45. Adams TS, Schupp JC, Poli S, Ayaub EA, Neumark N, Ahangari F, et al. Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. Sci Adv. 2020;6(28):eaba1983.
46. Xu G, Liu Y, Liao M, Gou J, Wang X, Yuan J, et al. Persistent viral activity, cytokine storm, and lung fibrosis in a case of severe COVID-19. Clin Transl Med. 2020;10(7):e224.
47. Wu X, Liu X, Zhou Y, Yu H, Li R, Zhan Q, et al. 3-month, 6-month, 9-month, and 12-month respiratory outcomes in patients following COVID-19-related hospitalisation: a prospective study. Lancet Respir Med. 2021;9(7):747-54.
48. Marshall M. The lasting misery of coronavirus long-haulers. Nature. 2020;585(7825):339-41.
49. Kølner-Augustson L, Prior TS, Skivild V, Aalestrup A, Bendstrup E. Fatigue in idiopathic pulmonary fibrosis measured by the Fatigue Assessment Scale during antifibrotic treatment. Eur Clin Respir J. 2020;8(1):1853658.
50. Bloem AEM, Mostard RLM, Stoot N, Vercoulen JH, Peters JB, Janssen DJA, et al. Severe Fatigue is Highly Prevalent in Patients with IPF or Sarcoidosis. J Clin Med. 2020;9(4).
51. Bloem AEM, Mostard RLM, Stoot N, Vercoulen JH, Peters JB, Spruit MA. Perceptions of fatigue in patients with idiopathic pulmonary fibrosis or sarcoidosis. J Thorac Dis. 2021;13(8):4872-84.
52. van der Meer FJ, Orsel K, Barkema HW. The new influenza A H1N1 virus: balancing on the interface of humans and animals. Can Vet J. 2010;51(1):56-62.
53. Bermejo-Martin JF, Martin-Loeches I, Rello J, Antón A, Almansa R, Xu L, et al. Host adaptive immunity deficiency in severe pandemic influenza. Crit Care. 2010;14(5):R167.
54. Baden LR, Drazen JM, Kritek PA, Curfman GD, Morrissey S, Campion EW. H1N1 influenza A disease--information for health professionals. N Engl J Med. 2009;360(25):2666-7.
55. Zhong NS, Zheng BJ, Li YM, Poon, Xie ZH, Chan KH, et al. Epidemiology and cause of severe acute respiratory syndrome (SARS) in Guangdong, People's Republic of China, in February, 2003. Lancet. 2003;362(9393):1353-8.
56. Beigel JH, Farrar J, Han AM, Hayden FG, Hyer R, de Jong MD, et al. Avian influenza A (H5N1) infection in humans. N Engl J Med. 2005;353(13):1374-85.
57. Zaki AM, van Boheemen S, Bestebroer TM, Osterhaus AD, Fouchier RA. Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia. N Engl J Med. 2012;367(19):1814-20.
58. Zhai Y, Franco LM, Atmar RL, Quarles JM, Arden N, Bucasas KL, et al. Host Transcriptional Response to Influenza and Other Acute Respiratory Viral Infections--A Prospective Cohort Study. PLoS Pathog. 2015;11(6):e1004869.
59. Liao M, Liu Y, Yuan J, Wen Y, Xu G, Zhao J, et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. Nat Med. 2020;26(6):842-4.
60. Wu K, Byers DE, Jin X, Agapov E, Alexander-Brett J, Patel AC, et al. TREM-2 promotes macrophage survival and lung disease after respiratory viral infection. J Exp Med. 2015;212(5):681-97.

914 61. Evren E, Ringqvist E, Tripathi KP, Sleiers N, Rives IC, Alisjahbana A, et al. Distinct developmental
915 pathways from blood monocytes generate human lung macrophage diversity. Immunity. 2021;54(2):259-75.e7.
916 62. Leahy TR, McManus R, Doherty DG, Grealy R, Coulter T, Smyth P, et al. Interleukin-15 is associated
917 with disease severity in viral bronchiolitis. Eur Respir J. 2016;47(1):212-22.
918 63. Zhang Z, Luo D, Zhong X, Choi JH, Ma Y, Wang S, et al. SCINA: A Semi-Supervised Subtyping
919 Algorithm of Single Cells and Bulk Samples. Genes (Basel). 2019;10(7).
920 64. Parimon T, Yao C, Stripp BR, Noble PW, Chen P. Alveolar Epithelial Type II Cells as Drivers of Lung
921 Fibrosis in Idiopathic Pulmonary Fibrosis. Int J Mol Sci. 2020;21(7).
922 65. Garcia CK. Insights from human genetic studies of lung and organ fibrosis. J Clin Invest.
923 2018;128(1):36-44.
924 66. Pardo A, Selman M. Molecular mechanisms of pulmonary fibrosis. Front Biosci. 2002;7:d1743-61.
925 67. Choi J, Park JE, Tsagkogeorga G, Yanagita M, Koo BK, Han N, et al. Inflammatory Signals Induce AT2
926 Cell-Derived Damage-Associated Transient Progenitors that Mediate Alveolar Regeneration. Cell Stem Cell.
927 2020;27(3):366-82.e7.
928 68. Strunz M, Simon LM, Ansari M, Kathiriya JJ, Angelidis I, Mayr CH, et al. Alveolar regeneration
929 through a Krt8+ transitional stem cell state that persists in human lung fibrosis. Nat Commun. 2020;11(1):3559.
930 69. Kobayashi Y, Tata A, Konkimalla A, Katsura H, Lee RF, Ou J, et al. Persistence of a regeneration-
931 associated, transitional alveolar epithelial cell state in pulmonary fibrosis. Nat Cell Biol. 2020;22(8):934-46.
932 70. Yao C, Guan X, Carraro G, Parimon T, Liu X, Huang G, et al. Senescence of Alveolar Type 2 Cells
933 Drives Progressive Pulmonary Fibrosis. Am J Respir Crit Care Med. 2021;203(6):707-17.
934 71. Abdelnabi R, Foo CS, Kaptein SJF, Zhang X, Do TND, Langendries L, et al. The combined treatment of
935 Molnupiravir and Favipiravir results in a potentiation of antiviral efficacy in a SARS-CoV-2 hamster infection
936 model. EBioMedicine. 2021;72:103595.
937 72. Cox RM, Wolf JD, Plemper RK. Therapeutically administered ribonucleoside analogue MK-
938 4482/EIDD-2801 blocks SARS-CoV-2 transmission in ferrets. Nat Microbiol. 2021;6(1):11-8.
939 73. Toots M, Yoon JJ, Hart M, Natchus MG, Painter GR, Plemper RK. Quantitative efficacy paradigms of
940 the influenza clinical drug candidate EIDD-2801 in the ferret model. Transl Res. 2020;218:16-28.
941 74. Toots M, Yoon JJ, Cox RM, Hart M, Sticher ZM, Makhsous N, et al. Characterization of orally
942 efficacious influenza drug with high resistance barrier in ferrets and human airway epithelia. Sci Transl Med.
943 2019;11(515).
944 75. Sheahan TP, Sims AC, Zhou S, Graham RL, Pruijssers AJ, Agostini ML, et al. An orally bioavailable
945 broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple
946 coronaviruses in mice. Sci Transl Med. 2020;12(541).
947 76. Urakova N, Kuznetsova V, Crossman DK, Sokratian A, Guthrie DB, Kolykhalov AA, et al. β-d-N (4)-
948 Hydroxycytidine Is a Potent Anti-alphavirus Compound That Induces a High Level of Mutations in the Viral
949 Genome. J Virol. 2018;92(3).
950 77. Jayk Bernal A, Gomes da Silva MM, Musungaie DB, Kovalchuk E, Gonzalez A, Delos Reyes V, et al.
951 Molnupiravir for Oral Treatment of Covid-19 in Nonhospitalized Patients. N Engl J Med. 2022;386(6):509-20.
952 78. Yang P, Gu H, Zhao Z, Wang W, Cao B, Lai C, et al. Angiotensin-converting enzyme 2 (ACE2)
953 mediates influenza H7N9 virus-induced acute lung injury. Sci Rep. 2014;4:7027.
954 79. Vasarmidi E, Tsitoura E, Spandidos DA, Tzanakis N, Antoniou KM. Pulmonary fibrosis in the
955 aftermath of the COVID-19 era (Review). Exp Ther Med. 2020;20(3):2557-60.
956 80. Pang J, Xu F, Aondio G, Li Y, Fumagalli A, Lu M, et al. Efficacy and tolerability of bevacizumab in
957 patients with severe Covid-19. Nat Commun. 2021;12(1):814.
958 81. Lamers MM, van der Vaart J, Knoops K, Riesebosch S, Breugem TI, Mykytyn AZ, et al. An organoid-
959 derived bronchioalveolar model for SARS-CoV-2 infection of human alveolar type II-like cells. EMBO J.
960 2021;40(5):e105912.
961 82. In-Yee Lee and Jan-Ming Ho and Ming-Syan C. CLUGO: a clustering algorithm for automated
962 functional annotations based on gene ontology. 2005. p. 4 -.

963 83. Barrett TJ, Cornwell M, Myndzar K, Rolling CC, Xia Y, Drenkova K, et al. Platelets amplify
964 endotheliopathy in COVID-19. Sci Adv. 2021;7(37):eabh2434.
965 84. Bonaventura A, Vecchié A, Dagna L, Martinod K, Dixon DL, Van Tassell BW, et al. Endothelial
966 dysfunction and immunothrombosis as key pathogenic mechanisms in COVID-19. Nat Rev Immunol.
967 2021;21(5):319-29.
968 85. Maruhashi T, Higashi Y. Pathophysiological Association of Endothelial Dysfunction with Fatal
969 Outcome in COVID-19. Int J Mol Sci. 2021;22(10).
970 86. Senchenkova EY, Komoto S, Russell J, Almeida-Paula LD, Yan LS, Zhang S, et al. Interleukin-6
971 mediates the platelet abnormalities and thrombogenesis associated with experimental colitis. Am J Pathol.
972 2013;183(1):173-81.
973 87. Nosaka M, Ishida Y, Kimura A, Kuninaka Y, Taruya A, Ozaki M, et al. Crucial Involvement of IL-6 in
974 Thrombus Resolution in Mice via Macrophage Recruitment and the Induction of Proteolytic Enzymes. Front
975 Immunol. 2019;10:3150.
976 88. Katzen J, Beers MF. Contributions of alveolar epithelial cell quality control to pulmonary fibrosis. J Clin
977 Invest. 2020;130(10):5088-99.
978 89. Borok Z, Horie M, Flodby P, Wang H, Liu Y, Ganesh S, et al. Loss in Epithelial Progenitors Reveals an
979 Age-linked Role for Endoplasmic Reticulum Stress in Pulmonary Fibrosis. Am J Respir Crit Care Med.
980 2020;201(2):198-211.
981 90. Rock JR, Barkauskas CE, Cronce MJ, Xue Y, Harris JR, Liang J, et al. Multiple stromal populations
982 contribute to pulmonary fibrosis without evidence for epithelial to mesenchymal transition. Proc Natl Acad Sci
983 U S A. 2011;108(52):E1475-83.
984 91. Barkauskas CE, Cronce MJ, Rackley CR, Bowie EJ, Keene DR, Stripp BR, et al. Type 2 alveolar cells
985 are stem cells in adult lung. J Clin Invest. 2013;123(7):3025-36.
986 92. Frank DB, Peng T, Zepp JA, Snitow M, Vincent TL, Penkala IJ, et al. Emergence of a Wave of Wnt
987 Signaling that Regulates Lung Alveologenesis by Controlling Epithelial Self-Renewal and Differentiation. Cell
988 Rep. 2016;17(9):2312-25.
989 93. Zhou J, Mao B, Zhou Q, Ding D, Wang M, Guo P, et al. Endoplasmic reticulum stress activates
990 telomerase. Aging Cell. 2014;13(1):197-200.
991 94. Sanchez-Vazquez R, Guío-Carrión A, Zapatero-Gaviria A, Martínez P, Blasco MA. Shorter telomere
992 lengths in patients with severe COVID-19 disease. Aging (Albany NY). 2021;13(1):1-15.
993 95. Aviv A. Short telomeres and severe COVID-19: The connection conundrum. EBioMedicine.
994 2021;70:103513.
995 96. Verheyden JM, Sun X. A transitional stem cell state in the lung. Nat Cell Biol. 2020;22(9):1025-6.
996 97. Liu RM, Liu G. Cell senescence and fibrotic lung diseases. Exp Gerontol. 2020;132:110836.
997 98. Parimon T, Hohmann MS, Yao C. Cellular Senescence: Pathogenic Mechanisms in Lung Fibrosis. Int J
998 Mol Sci. 2021;22(12).
999 99. Gaddam RR, Chambers S, Bhatia M. ACE and ACE2 in inflammation: a tale of two enzymes. Inflamm
1000 Allergy Drug Targets. 2014;13(4):224-34.
1001 100. Inlay MA, Bhattacharya D, Sahoo D, Serwold T, Seita J, Karsunky H, et al. Ly6d marks the earliest
1002 stage of B-cell specification and identifies the branchpoint between B-cell and T-cell development. Genes Dev.
1003 2009;23(20):2376-81.
1004 101. Sodhi CP, Nguyen J, Yamaguchi Y, Werts AD, Lu P, Ladd MR, et al. A Dynamic Variation of
1005 Pulmonary ACE2 Is Required to Modulate Neutrophilic Inflammation in Response to. J Immunol.
1006 2019;203(11):3000-12.
1007 102. Jia H. Pulmonary Angiotensin-Converting Enzyme 2 (ACE2) and Inflammatory Lung Disease. Shock.
1008 2016;46(3):239-48.
1009 103. Kuba K, Imai Y, Penninger JM. Multiple functions of angiotensin-converting enzyme 2 and its
1010 relevance in cardiovascular diseases. Circ J. 2013;77(2):301-8.

1011 104. Esposito AJ, Menon AA, Ghosh AJ, Putman RK, Fredenburgh LE, El-Chemaly SY, et al. Increased
1012 Odds of Death for Patients with Interstitial Lung Disease and COVID-19: A Case-Control Study. Am J Respir
1013 Crit Care Med. 2020;202(12):1710-3.
1014 105. Drake TM, Docherty AB, Harrison EM, Quint JK, Adamali H, Agnew S, et al. Outcome of
1015 Hospitalization for COVID-19 in Patients with Interstitial Lung Disease. An International Multicenter Study.
1016 Am J Respir Crit Care Med. 2020;202(12):1656-65.
1017 106. Podolanczuk AJ, Richeldi L. COVID-19 and Interstitial Lung Disease: Keep Them Separate. Am J
1018 Respir Crit Care Med. 2020;202(12):1614-6.
1019 107. Gallay L, Uzunhan Y, Borie R, Lazor R, Rigaud P, Marchand-Adam S, et al. Risk Factors for Mortality
1020 after COVID-19 in Patients with Preexisting Interstitial Lung Disease. Am J Respir Crit Care Med.
1021 2021;203(2):245-9.
1022 108. FIRDAUS H, RAZAK NA, REZA M, KHAN NA, SHAMEEM M. Clinical Improvement with
1023 Pirfenidone in Two Patients of COVID 19. Journal of Clinical & Diagnostic Research. 2020;14(12).
1024 109. Momen ABI, Khan F, Saber S, Sultana A, Alam RF, Raihan SG, et al. Usefulness of Pirfenidone in
1025 Covid Lung: A Case Series. European Journal of Medical and Health Sciences. 2021;3(1):24-6.
1026 110. Xi Z, Zhigang Z, Ting L. Post-inflammatory pulmonary fibrosis in a discharged COVID-19 patient:
1027 Effectively treated with Pirfenidone. Archives of Pulmonology and Respiratory Care. 2020;6(1):051-3.
1028 111. McGonagle D, Sharif K, O'Regan A, Bridgewood C. The Role of Cytokines including Interleukin-6 in
1029 COVID-19 induced Pneumonia and Macrophage Activation Syndrome-Like Disease. Autoimmun Rev.
1030 2020;19(6):102537.
1031 112. Muro S, Taha R, Tsicopoulos A, Olivenstein R, Tonnel AB, Christodoulopoulos P, et al. Expression of
1032 IL-15 in inflammatory pulmonary diseases. J Allergy Clin Immunol. 2001;108(6):970-5.
1033 113. Takada T, Ohashi K, Hayashi M, Asakawa K, Sakagami T, Kikuchi T, et al. Role of IL-15 in interstitial
1034 lung diseases in amyopathic dermatomyositis with anti-MDA-5 antibody. Respir Med. 2018;141:7-13.
1035 114. Wuttge DM, Wildt M, Geborek P, Wollheim FA, Scheja A, Akesson A. Serum IL-15 in patients with
1036 early systemic sclerosis: a potential novel marker of lung disease. Arthritis Res Ther. 2007;9(5):R85.
1037 115. Alder JK, Barkauskas CE, Limjunyawong N, Stanley SE, Kembou F, Tuder RM, et al. Telomere
1038 dysfunction causes alveolar stem cell failure. Proc Natl Acad Sci U S A. 2015;112(16):5099-104.
1039 116. Abdul-Careem MF, Mian MF, Yue G, Gillgrass A, Chenoweth MJ, Barra NG, et al. Critical role of
1040 natural killer cells in lung immunopathology during influenza infection in mice. J Infect Dis. 2012;206(2):167-
1041 77.
1042 117. Schneider WM, Chevillotte MD, Rice CM. Interferon-stimulated genes: a complex web of host
1043 defenses. Annu Rev Immunol. 2014;32:513-45.
1044 118. Froidure A, Mahieu M, Hoton D, Laterre PF, Yombi JC, Koenig S, et al. Short telomeres increase the
1045 risk of severe COVID-19. Aging (Albany NY). 2020;12(20):19911-22.
1046 119. McGroder CF, Zhang D, Choudhury MA, Salvatore MM, D'Souza BM, Hoffman EA, et al. Pulmonary
1047 fibrosis 4 months after COVID-19 is associated with severity of illness and blood leucocyte telomere length.
1048 Thorax. 2021.
1049 120. Mongelli A, Barbi V, Gottardi Zamperla M, Atlante S, Forleo L, Nesta M, et al. Evidence for Biological
1050 Age Acceleration and Telomere Shortening in COVID-19 Survivors. Int J Mol Sci. 2021;22(11).
1051 121. Borok Z, Horie M, Flodby P, Wang H, Liu Y, Ganesh S, et al. Grp78 Loss in Epithelial Progenitors
1052 Reveals an Age-linked Role for Endoplasmic Reticulum Stress in Pulmonary Fibrosis. Am J Respir Crit Care
1053 Med. 2020;201(2):198-211.
1054 122. Liang J, Zhang Y, Xie T, Liu N, Chen H, Geng Y, et al. Hyaluronan and TLR4 promote surfactant-
1055 protein-C-positive alveolar progenitor cell renewal and prevent severe pulmonary fibrosis in mice. Nat Med.
1056 2016;22(11):1285-93.
1057 123. Xu Y, Mizuno T, Sridharan A, Du Y, Guo M, Tang J, et al. Single-cell RNA sequencing identifies
1058 diverse roles of epithelial cells in idiopathic pulmonary fibrosis. JCI Insight. 2016;1(20):e90558.

1059 124. Chilosi M, Carloni A, Rossi A, Poletti V. Premature lung aging and cellular senescence in the
1060 pathogenesis of idiopathic pulmonary fibrosis and COPD/emphysema. Transl Res. 2013;162(3):156-73.
1061 125. Chen H, Chen H, Liang J, Gu X, Zhou J, Xie C, et al. TGF-β1/IL-11/MEK/ERK signaling mediates
1062 senescence-associated pulmonary fibrosis in a stress-induced premature senescence model of Bmi-1 deficiency.
1063 Exp Mol Med. 2020;52(1):130-51.
1064 126. Rana T, Jiang C, Liu G, Miyata T, Antony V, Thannickal VJ, et al. PAI-1 Regulation of TGF-β1-
1065 induced Alveolar Type II Cell Senescence, SASP Secretion, and SASP-mediated Activation of Alveolar
1066 Macrophages. Am J Respir Cell Mol Biol. 2020;62(3):319-30.
1067 127. Evangelou K, Veroutis D, Paschalaki K, Foukas PG, Lagopati N, Dimitriou M, et al. Pulmonary
1068 infection by SARS-CoV-2 induces senescence accompanied by an inflammatory phenotype in severe COVID-
1069 19: possible implications for viral mutagenesis. Eur Respir J. 2022.
1070 128. Lee S, Yu Y, Trimpert J, Benthani F, Mairhofer M, Richter-Pechanska P, et al. Virus-induced
1071 senescence is a driver and therapeutic target in COVID-19. Nature. 2021;599(7884):283-9.
1072 129. Lipskaia L, Maisonnasse P, Fouillade C, Sencio V, Pascal Q, Flaman JM, et al. Evidence that SARS-
1073 CoV-2 Induces Lung-Cell Senescence: Potential Impact on COVID-19 Lung Disease. Am J Respir Cell Mol
1074 Biol. 2021.
1075 130. D'Agnillo F, Walters KA, Xiao Y, Sheng ZM, Scherler K, Park J, et al. Lung epithelial and endothelial
1076 damage, loss of tissue repair, inhibition of fibrinolysis, and cellular senescence in fatal COVID-19. Sci Transl
1077 Med. 2021;13(620):eabj7790.
1078 131. Sahoo D, Dill DL, Gentles AJ, Tibshirani R, Plevritis SK. Boolean implication networks derived from
1079 large scale, whole genome microarray datasets. Genome Biol. 2008;9(10):R157.
1080 132. Sahoo D, Seita J, Bhattacharya D, Inlay MA, Weissman IL, Plevritis SK, et al. MiDReG: a method of
1081 mining developmentally regulated genes using Boolean implications. Proc Natl Acad Sci U S A.
1082 2010;107(13):5732-7.
1083 133. Dabydeen SA, Desai A, Sahoo D. Unbiased Boolean analysis of public gene expression data for cell
1084 cycle gene identification. Mol Biol Cell. 2019;30(14):1770-9.
1085 134. Ju Z, Jiang H, Jaworski M, Rathinam C, Gompf A, Klein C, et al. Telomere dysfunction induces
1086 environmental alterations limiting hematopoietic stem cell function and engraftment. Nat Med. 2007;13(6):742-
1087 7.

1088
1089

1090 **FIGURE LEGENDS**

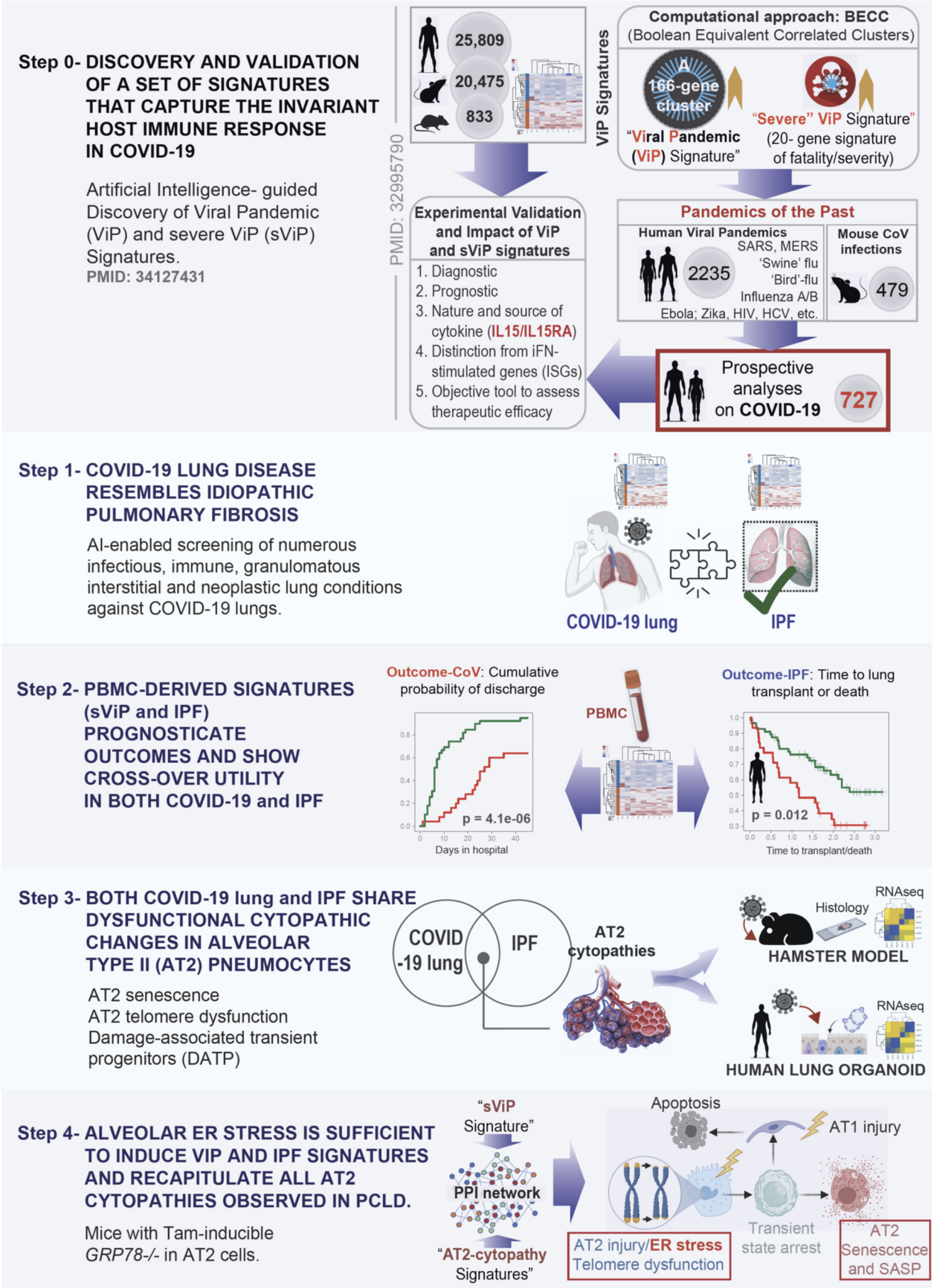


**Figure 1. Study design: Artificial Intelligence-guided navigation of COVID-19 lung disease.**

(*From top to bottom*) *Step 0*: Over 45,000 human, mouse, and rat gene expression databases were mined using machine learning tools called Boolean Equivalent Correlated Clusters (BECC(133)) to identify invariant host response to viral pandemics (ViP). In the absence of a sufficiently large number of COVID-19 datasets at the onset of the COVID-19 pandemic, these ViP signatures were trained on only two datasets from the pandemics of the past (Influenza and avian flu; GSE47963, n = 438; GSE113211, n = 118) and used without further training to prospectively analyze the samples from the current pandemic (i.e., COVID-19; n = 727 samples from diverse datasets). A subset of 20-genes classified disease severity called severe-ViP (sViP) signature. The ViP signatures appeared to capture the 'invariant' host response, i.e., the shared fundamental nature of the host immune response induced by all viral pandemics, including COVID-19. *Step1*: The set of ViP/sViP signatures and a CoV-lung specific(13) gene signature was analyzed on diverse transcriptomic datasets representing a plethora of lung diseases; these efforts identified COVID-19 lung disease to be the closest to Idiopathic pulmonary fibrosis (IPF); both conditions induced a common array of gene signatures. *Step 2*: Clinically useful whole-blood and PBMC-derived prognostic signatures previously validated in IPF(27) showed crossover efficacy in COVID-19, and *vice versa*. *Step 3*: Gene signatures of alveolar type II (AT2) cytopathic changes that are known to fuel IPF were analyzed in COVID-19 lung, and predicted shared features were validated in human and hamster lungs and lung-organoid derived models. *Step 4:* Protein- protein interaction (PPI) network built using sViP and AT2 cytopathy-related signatures was analyzed to pinpoint ER stress as a major shared feature in COVID-19 lung disease and IPF, which was subsequently validated in human and hamster lungs.

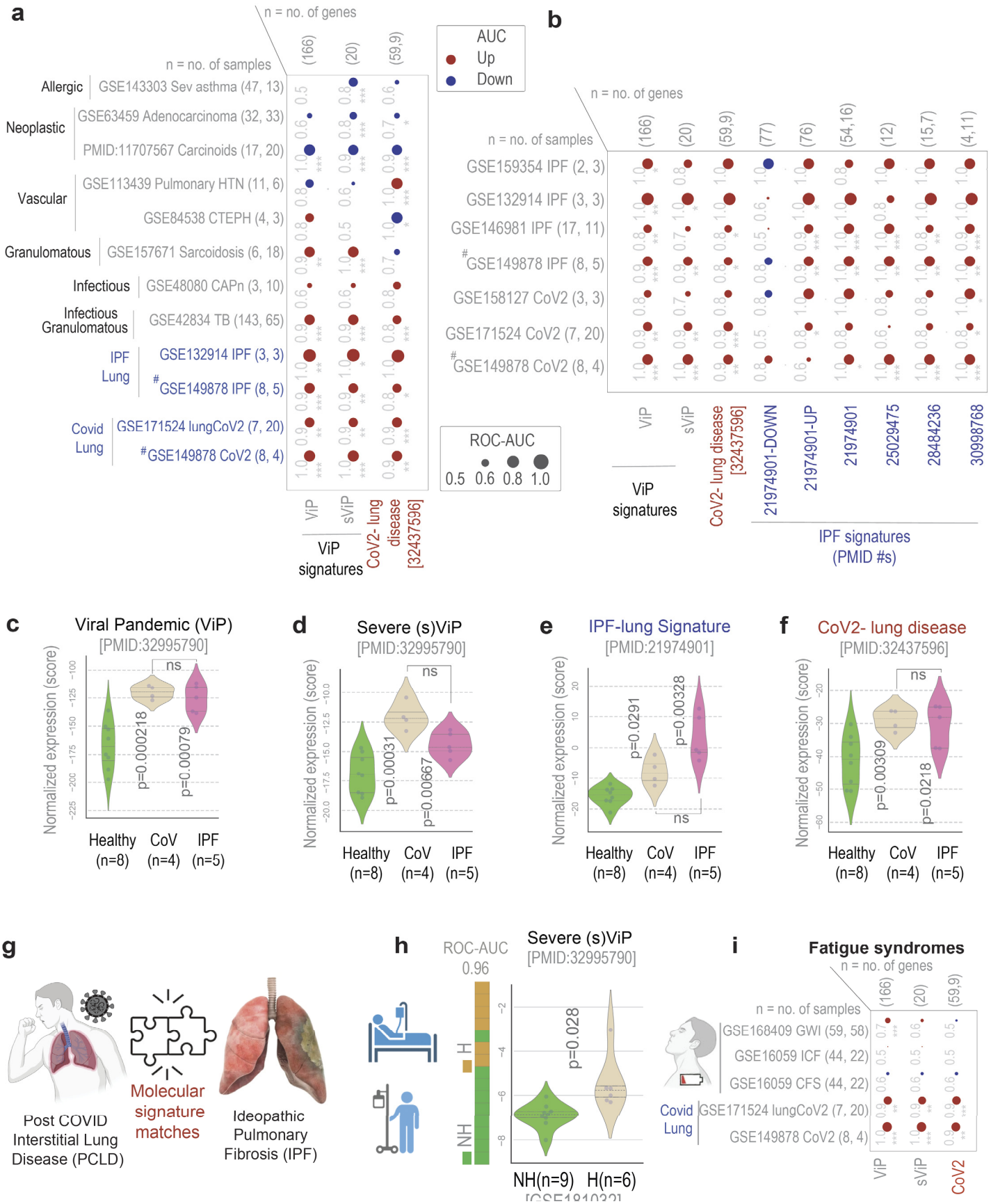


**Figure 2. Identification of lung condition(s) that resemble COVID-19 lung disease.**

**a.** Bubble plots show the ROC-AUC values (radii of circles are based on the ROC-AUC) show the direction of gene regulation (Up, red; Down, blue) for sample classification based on the induction of ViP/sViP(18) and a previously published human CoV-lung signatures(13) in diverse datasets representing diverse lung pathologies, including COVID-19. Tuberculosis, TB; community acquired pneumonia, CAPn; pulmonary hypertension, PHT; chronic thromboembolic pulmonary hypertension, CTEPH; idiopathic pulmonary fibrosis, IPF; CoV, COVID-19. **b.** Bubble plots in B show the ROC-AUC values for sample classification based on the induction of ViP signatures and several previously published IPF signatures in COVID-19 infected lungs and sc-RNA seq datasets from IPF explants. **c-f.** Violin plots display the extent of induction of ViP (E) or sViP (F), IPF (G) or CoV lung (H) signatures in a pooled cohort of healthy, CoV and IPF lung samples (GSE149878; GSE122960). Welch's two sample unpaired t-test is performed on the composite gene signature score (z-score of normalized tpm count) to compute the *p* values. In multi-group setting each group is compared to the first control group and only significant *p* values are displayed. **g.** Schematic summarizes the predicted molecular match between COVID-19 lung and IPF based on the crossover validation of gene signatures showcased in B-F. **h.** Bar plots show the ability of sViP signature to classify COVID-19 patients based on hospitalization status and violin plots show the degree of induction of the signature within each group (H, hospitalized; NH, non-hospitalized). Welch's two sample unpaired t-test is performed on the composite gene signature score (z-score of normalized tpm count) to compute the *p values*. **i.** Bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC) show the direction of gene regulation (Up, red; Down, blue) for the classification of healthy vs diseased samples based on the ViP/sViP and the CoV-lung signatures on three well known fatigue conditions (blood/PBMC samples from GWI, gulf war illness; CFS, chronic fatigue syndrome). The significance of the AUC was determined by Welch's t-test; * = $p < 0.05$, ** = $p < 0.01$, and *** = $p < 0.001$. (# denotes a pooled cohort of GSE149878; GSE122960).

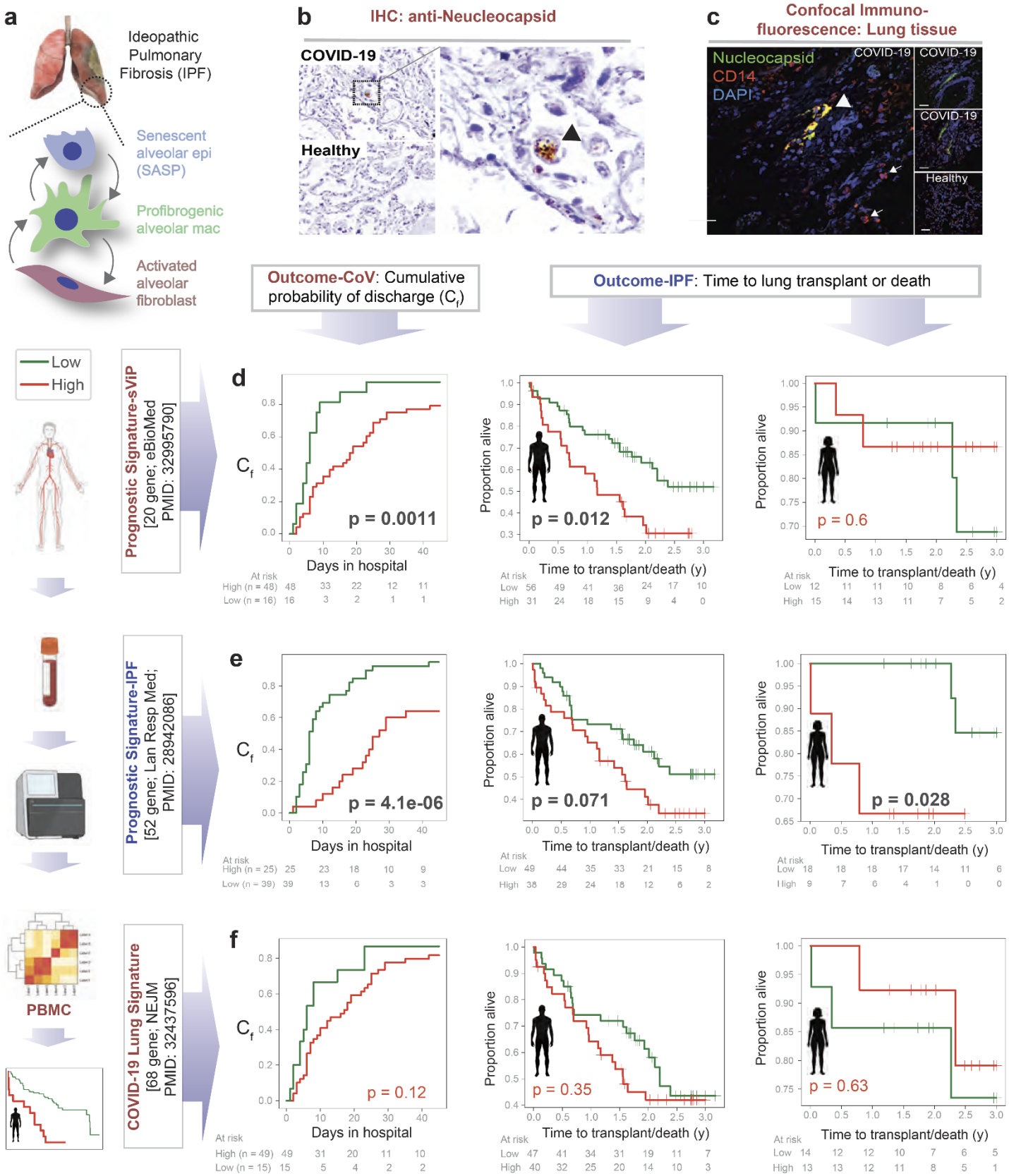


**Figure 3. PBMC-derived prognostic signatures in IPF also prognosticate outcome in COVID-19.**

**a.** Schematic showing the currentunderstanding of the complex interplay between senescent alveolar pneumocyte (AT2), reactive alveolar macrophage, and fibroblasts in the pathogenesis of pulmonary fibrosis. **b-c**. FFPE tissues from normal lungs and from the lungs of subjects deceased due to COVID-19 were analyzed for the SARS-CoV2 viral nucleocapsid protein either by IHC (B) or by confocal immunofluorescence (C). Lungs were co-stained also with CD14 (a marker for activated alveolar macrophage(61)) in C. **d.** Kaplan-Meier plots show the stratification of patients based on the degree of induction of the sViP signature in PBMCs (low-green and high-red groups; determined using the *StepMiner* algorithm(19) and its relationship to either the cumulative probability of discharge of COVID-19 patients (D-*left*; GSE157103; derived from the hospital-free days during a 45-day follow-up), or the time to lung transplant or death in male (D-*middle*) and female (D-*right*) IPF patients (GSE28221). Statistical significances were estimated using the log rank test. **e.** Kaplan-Meier plots on the same cohorts as in D, using a PBMC-derived 52-gene IPF-specific prognostic signature that was previously validated on numerous independent IPF cohorts(27). **f**. Kaplan-Meier plots on the same cohorts as in D-E, using a CoV-lung derived signature.

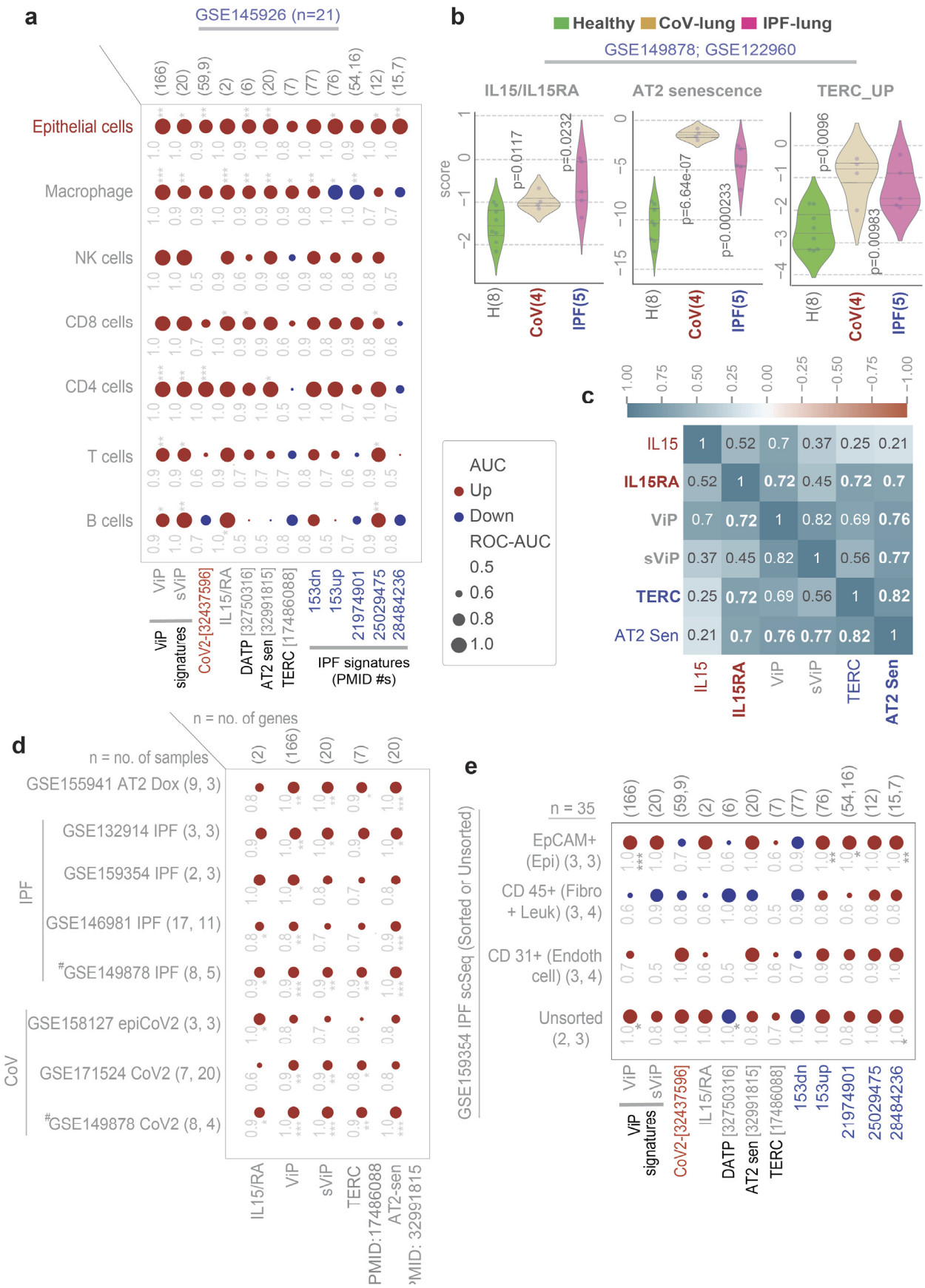


**Figure 4. Identification of shared AT2 cytopathies in COVID-19 and IPF.**

**a**. Bubble plot of ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification of various cell types between healthy and CoV lung based on various ViP/sViP and IPF gene signatures in **Fig 2a-b** alongside *IL15/IL15RA* and additional signatures of AT2 cytopathies that are encountered and implicated in IPF (DATP, damage associated transient progenitors; Sens, senescence; TERC, *Te*lomerase *R*NA *C*omponent). Numbers in parenthesis indicate PMIDs. **b.** Violin plots display the extent of induction of a *IL15/IL15RA* composite score (B-*left*), or a signature of AT2 senescence(70) (B-*middle*), or a gene set for telomere dysfunction (JU_AGING_TERC(134)) in a pooled cohort of healthy, CoV and IPF lung samples (GSE149878; GSE122960). Welch's two sample unpaired t-test is performed on the composite gene signature score (z-score of normalized tpm count) to compute the *p* values. In multi-group setting each group is compared to the first control group and only significant p values are displayed. **C.** Correlation matrix showing the Pearson's correlation coefficient between different gene signatures in the pooled healthy, CoV and IPF lung dataset (GSE149878; GSE122960). **d.** Bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification based on AT2 senescence-related signatures on other publicly available CoV and IPF bulk RNA seq datasets. **e.** Bubble plots of ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for different types of cells in an independent sc-RNA seq datasets of IPF explants and healthy lung tissues. The significance of the AUC was determined by Welch's t-test; * = $p < 0.05$, ** = $p < 0.01$, and *** = $p < 0.001$. (# denotes a pooled cohort of GSE149878; GSE122960).

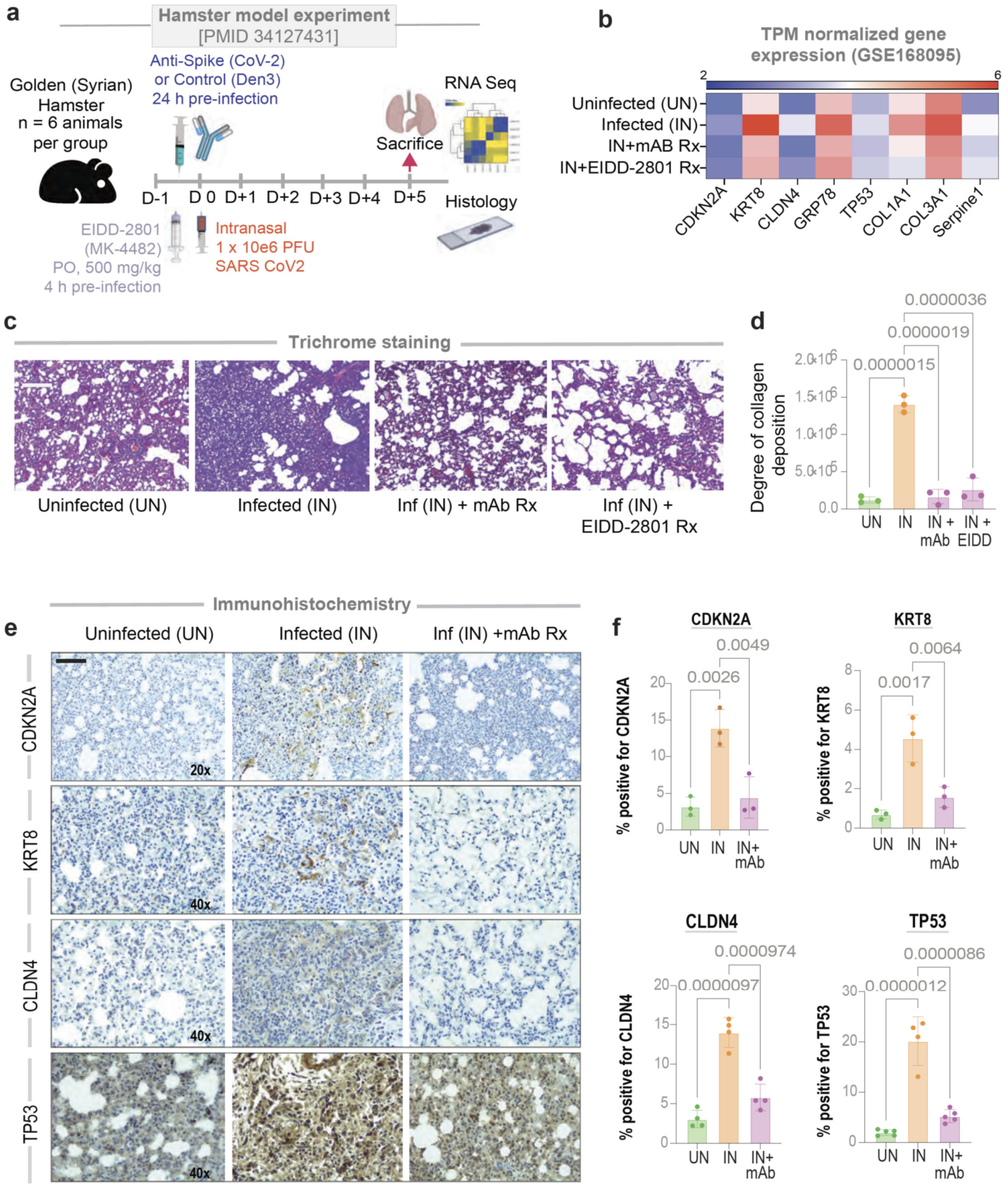


**Figure 5. SARS-CoV-2-challenged hamster lungs faithfully recapitulate AT2 cytopathies in COVID-19.**

**a.** Schematic showing the experimental design for COVID-19 modeling in golden Syrian hamsters. *Uninf*, uninfected; *Den3 and Anti-CoV-2* indicate SARS-CoV-2 challenged groups that received either a control mAb or the clone CC12.2 of anti-CoV-2 IgG(31), respectively. PFU, plaque-forming units. **b**. Heatmap showing transcripts/million (TPM)-normalized gene expression of specific genes associated with AT2 senescence (*CDKN2A*, *TP53*), damage associated transition state progenitor (DATP) arrested state (*CLDN4*, *KRT8*), ER stress (*HSPA5*/*GRP78*) and SASP-state (*SERPINE1*, *COL1A*, *COL3A*). **c-d**. Lungs harvested from the uninfected, infected, and infected but anti-CoV-2 IgG and infected but EIDD-2801 treated hamsters have been analyzed for the abundance of collagen deposition by Trichrome staining and quantified by ImageJ. Representative images are shown panel c. Scale bar = 200 µm. Results of the quantification are displayed as bar plot in D. Statistical significance was analyzed by one way ANOVA. Error bars represent S.E.M; n = 3. **e-f**. Lungs harvested from the uninfected, infected, and infected but anti-CoV-2 IgG groups were analyzed by IHC for the indicated proteins and quantified by ImageJ- IHC profiler. Representative images are shown panel E. Scale bar = 200 µm. Results of the quantification are displayed as violin plots in F. Statistical significance was analyzed by one way ANOVA; n = 4.

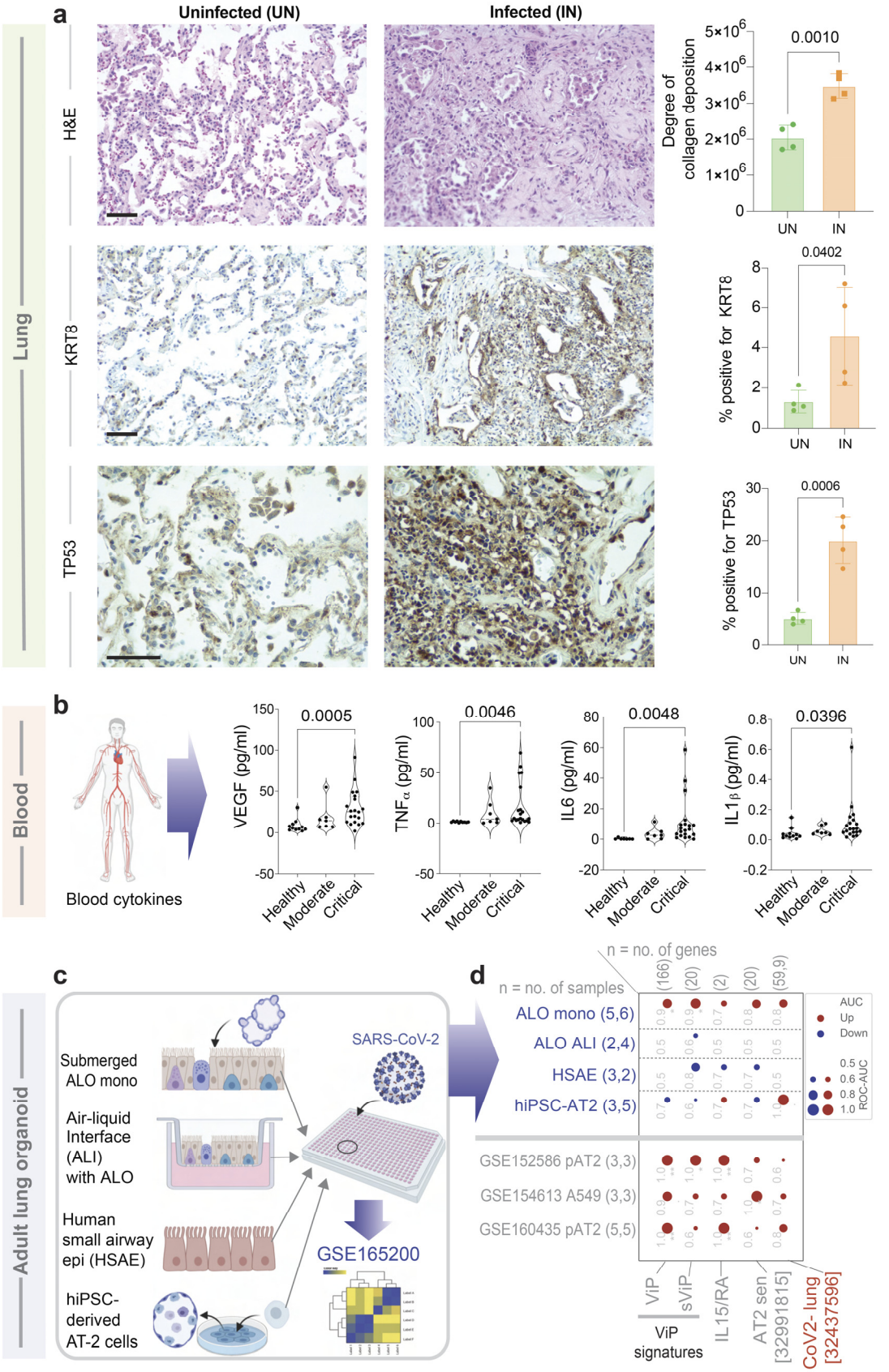


**Figure 6. AT2 cytopathic features of IPF are detected in COVID-19 lung tissues and human pre-clinical organoid models of COVID-19.**

**a**. FFPE lung tissues collected during rapid autopsy of deceased COVID-19-infected subjects and histologically normal surgically resected lung tissues were analyzed by H&E staining (*top*) and for the expression of KRT8 (a marker of DATP(67); *middle*) and TP53 (a marker of senescence; *bottom*) by IHC. Representative fields are shown on the left. Scale bar = 200 µm. Images were quantified by IHC profiler (ImageJ) and displayed as violin plots on the right. **b**. Violin plots display the abundance of key senescence-associated cytokines in the plasma from adults with COVID-19 and healthy volunteers. Statistical significance was analyzed by unpaired t-test. See also **Supplemental Information 3-4** for patient demographics and cytokine concentrations. **c-d**. Schematic in C summarizes the various pre-clinical *in vitro* models of the human lung that were challenged with SARS-CoV-2, followed by RNA sequencing(28) (*top* rows; blue font; D). Other independent SARS-CoV-2-challenged alveolar pneumocyte datasets are listed in grey font (D). Bubble plots in D show the ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification of uninfected vs infected samples based on the signatures (below).

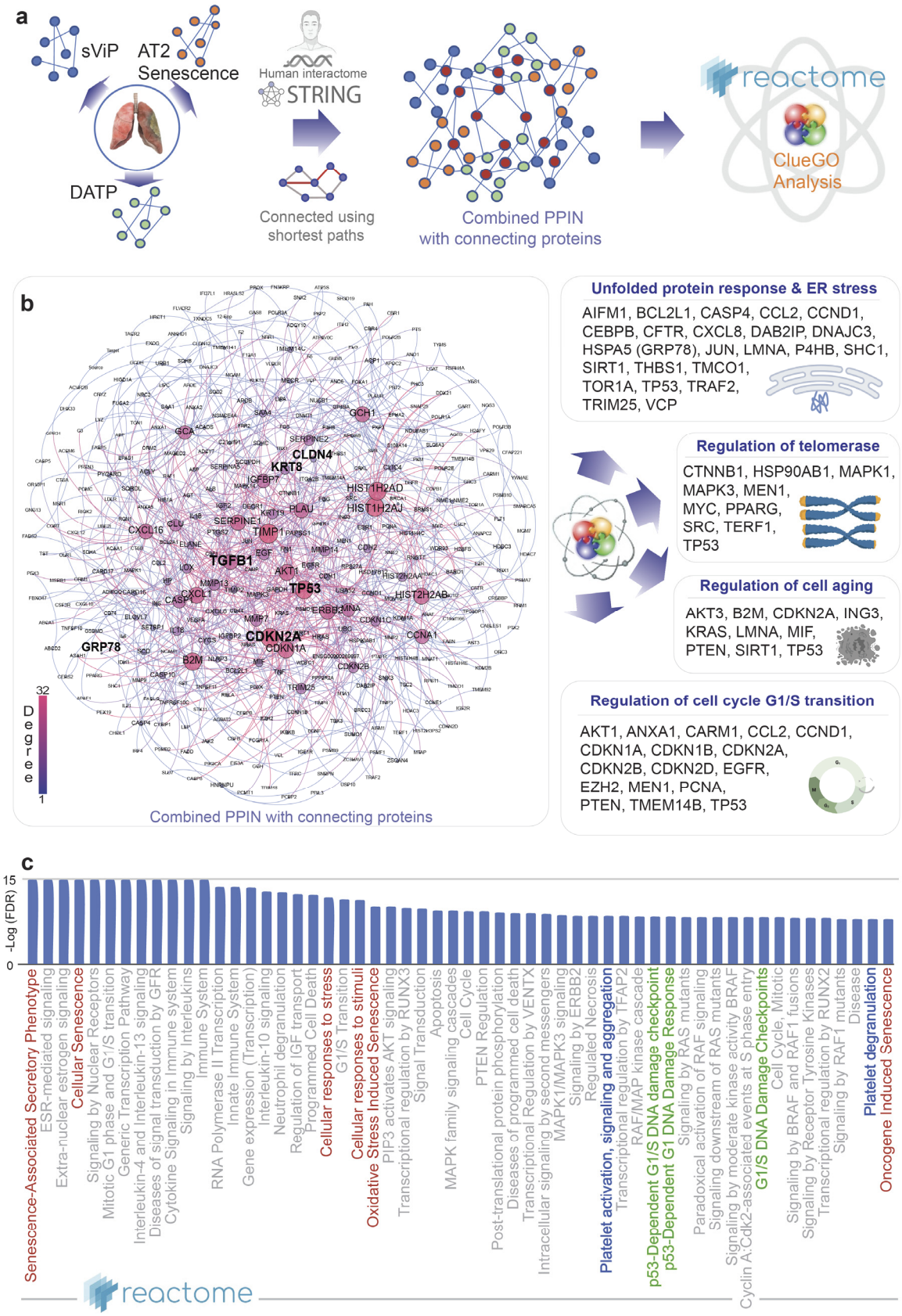


**Figure 7. Protein-protein interaction network (PPIN) analysis predicts ER stress in AT2 cells as a driver pathophysiology in COVID-19**

**a**. Schematic summarizes the key steps in building a PPI network using as 'seeds' the genes in the sVIP, AT2 senescence and transition state arrest (DATP) signatures and connecting them using shortest paths prior to carrying out network analyses. **b**. The combined PPI network is shown on the left and the degree of connectivity is indicated. See also **Supplemental Information 5** for the complete list of nodes sorted by their degrees of connectivity. Schematics on the right summarize the results of CluGo analysis, which identifies cellular processes and key genes that may perpetuate lung fibrosis. **c**. Reactome pathway analysis of the PPI network lists the most significant pathways that are associated with the node list in the network in B. Red = senescence related processes; Green = DNA damage response pathways; Blue = Platelet activation and degranulation.

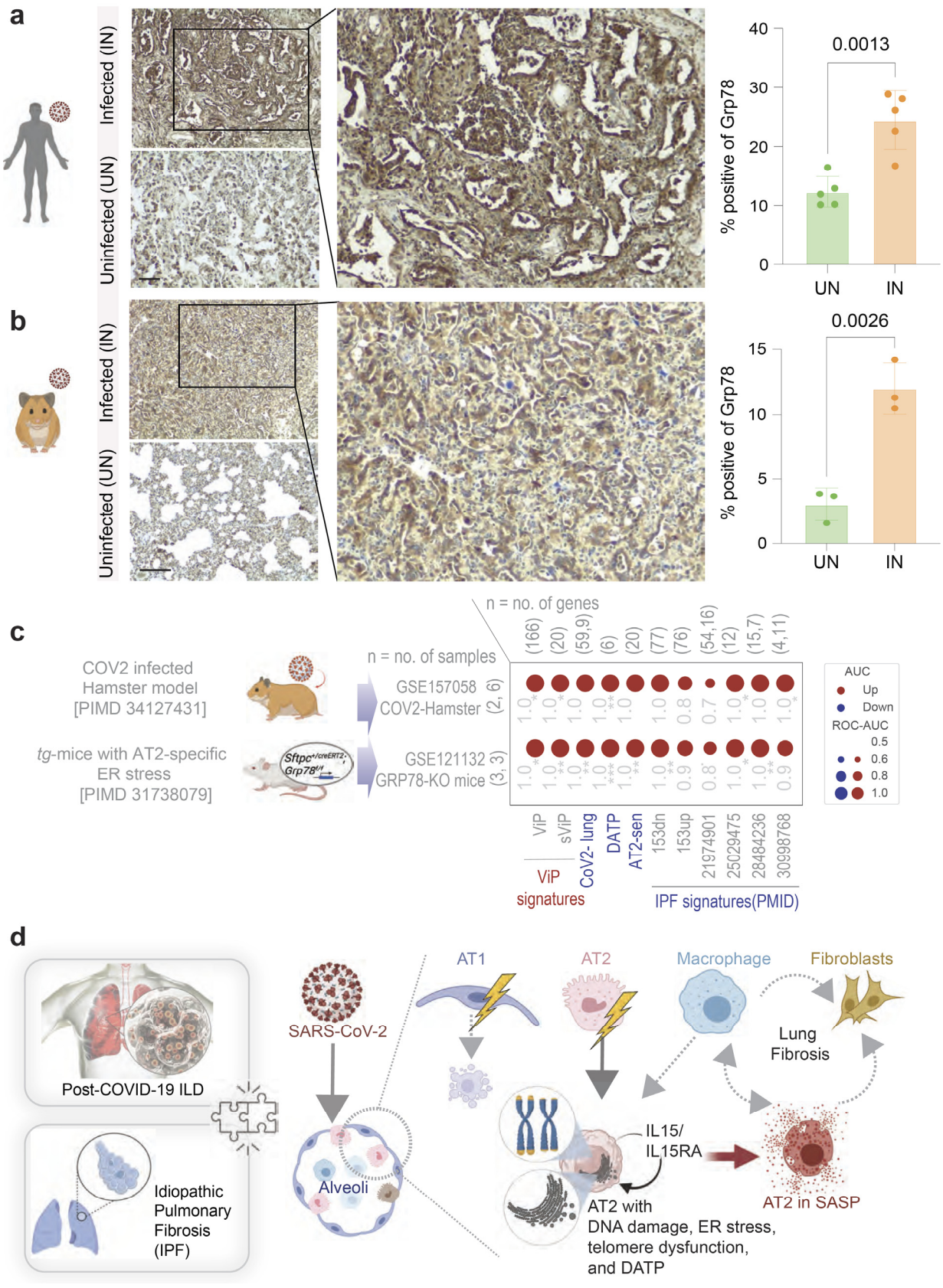


**Figure 8. Induction of ER stress in AT2 cells is observed in COVID-19 and is sufficient to mimic the host immune response in COVID-19 and IPF**

**a-b**. FFPE uninfected (control) and SARS-CoV-2-infected human (A) and hamster (B) lungs were analyzed for GRP78 expression by IHC. Representative fields are shown on the left. Scale bar = 200 µm. Images were quantified by IHC profiler (ImageJ) and displayed as violin plots on the right. **c**. Bubble plots show the ROC-AUC values (radii of circles are based on the ROC-AUC) demonstrating the direction of gene regulation (Up, red; Down, blue) for the classification of WT vs *GRP78*-KO murine lung (bottom row) and uninfected vs. infected hamster lung (top row) samples based on the signatures (below). **d**. Schematic summarizes the findings in this work and the proposed working model for the contributions of various alveolar cells in fueling the fibrotic progression in both IPF and COVID-19. The significance of the AUC was determined by Welch's t-test ; * = $p < 0.05$, ** = $p < 0.01$, and *** = $p < 0.001$.