EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Original research article

Profiling of lung SARS-CoV-2 and influenza virus infection dissects virus-specific host responses and gene signatures

Arutha Kulasinghe, Chin Wee Tan, Anna Flavia Ribeiro dos Santos Miggiolaro, James Monkman, Habib SadeghiRad, Dharmesh D. Bhuva, Jarbas da Silva Motta Junior, Caroline Busatta Vaz de Paula, Seigo Nagashima, Cristina Pellegrino Baena, Paulo Souza-Fonseca-Guimaraes, Lucia de Noronha, Timothy McCulloch, Gustavo Rodrigues Rossi, Caroline Cooper, Benjamin Tang, Kirsty R. Short, Melissa J. Davis, Fernando Souza-Fonseca-Guimaraes, Gabrielle T. Belz, Ken O'Byrne

Please cite this article as: Kulasinghe A, Tan CW, dos Santos Miggiolaro AFR, *et al*. Profiling of lung SARS-CoV-2 and influenza virus infection dissects virus-specific host responses and gene signatures. *Eur Respir J* 2021; in press (https://doi.org/10.1183/13993003.01881-2021).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

Profiling of Lung SARS-CoV-2 and Influenza Virus Infection Dissects Virus-Specific Host Responses and Gene Signatures

Arutha Kulasinghe[1,2,8*], Chin Wee Tan[3,4*], Anna Flavia Ribeiro dos Santos Miggiolaro[5*], James Monkman[1,2], Habib SadeghiRad[1,2], Dharmesh D Bhuva[3,4], Jarbas da Silva Motta Junior[5], Caroline Busatta Vaz de Paula[5], Seigo Nagashima[5], Cristina Pellegrino Baena[6], Paulo Souza-Fonseca-Guimaraes[4], Lucia de Noronha[7], Timothy McCulloch[2,8], Gustavo Rodrigues Rossi[2,8], Caroline Cooper[9,10], Benjamin Tang[11], Kirsty R. Short[12,13#], Melissa J Davis[3,4,14#], Fernando Souza-Fonseca-Guimaraes[2,8#], Gabrielle T. Belz[2,3,8,13#], Ken O'Byrne[1,2#]

[1]*Queensland University of Technology, School of Biomedical Sciences, Faculty of Health, Brisbane, Queensland 4000, Australia.*
[2]*Translational Research Institute, Brisbane, Queensland 4102, Australia.*
[3]*The Walter and Eliza Hall Institute of Medical Research, Parkville, Melbourne, Victoria 3052, Australia.*
[4]*Department of Medical Biology, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, VIC 3010, Australia.*
[5]*Postgraduate Program of Health Sciences - School of Medicine - Hospital Marcelino Champagnat - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, PR, Brazil.*
[6]*School of Medicine & Center of Education, Research and Innovation – Hospital Marcelino Champagnat - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Brazil.*
[7]*Laboratory of Experimental Pathology - School of Medicine - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Brazil.*
[8]*University of Queensland Diamantina Institute, University of Queensland, Woollongabba, Queensland, 4102 Australia.*
[9]*Pathology Queensland, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia.*
[10]*University of Queensland, Faculty of Medicine, Woolloongabba, Queensland, Australia.*
[11]*Centre for Immunology and Allergy Research, the Westmead Institute for Medical Research, Sydney, NSW, Australia.*
[12]*The University of Queensland, School of Chemistry and Molecular Biosciences, St Lucia, Brisbane, Queensland 4072, Australia.*
[13]*Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane, Queensland, Australia.*
[14]*Department of Clinical Pathology, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, VIC, Australia.*

*co-first authors
#co-senior authors

Corresponding Author: Dr Arutha Kulasinghe, School of Biomedical Sciences, Queensland University of Technology, 37 Kent Street, Woolloongabba, Queensland 4102, Australia.
Email: Arutha.kulasinghe@qut.edu.au

Keywords: respiratory disease, spatial transcriptomics, lung pathology, immune response.

Abstract

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) that emerged in late 2019 has spread globally, causing a pandemic of respiratory illness designated coronavirus disease 2019 (COVID-19). A better definition of the pulmonary host response to SARS-CoV-2 infection is required to understand viral pathogenesis and to validate putative COVID-19 biomarkers that have been proposed in clinical studies. Here, we use targeted transcriptomics of FFPE tissue using the Nanostring GeoMX^{TM} platform to generate an in-depth picture of the pulmonary transcriptional landscape of COVID-19, pandemic H1N1 influenza and uninfected control patients. Host transcriptomics showed a significant upregulation of genes associated with inflammation, type I interferon production, coagulation and angiogenesis in the lungs of COVID-19 patients compared to non-infected controls. SARS-CoV-2 was non-uniformly distributed in lungs (emphasising the advantages of spatial transcriptomics) with the areas of high viral load associated with an increased type I interferon response. Once the dominant cell type present in the sample, within patient correlations and patient-patient variation had been controlled for, only a very limited number of genes were differentially expressed between the lungs of fatal influenza and COVID-19 patients. Strikingly, the interferon-associated gene *IFI27*, previously identified as a useful blood biomarker to differentiate bacterial and viral lung infections, was significantly upregulated in the lungs of COVID-19 patients compared to patients with influenza. Collectively, these data demonstrate that spatial transcriptomics is a powerful tool to identify novel gene signatures within tissues, offering new insights into the pathogenesis of SARS-COV-2 to aid in patient triage and treatment.

Introduction

Since its emergence in 2019, severe acute respiratory syndrome-associated coronavirus-2 (SARS-CoV-2) has caused a broad spectrum of disease, ranging from asymptomatic infections to acute respiratory distress syndrome (ARDS). Despite a lower fatality rate than other related coronaviruses, such as SARS-CoV-1 and middle east respiratory syndrome coronavirus (MERS-CoV), SARS-CoV-2 is highly transmissible and to date has resulted in over >175M infections and >3.5M deaths worldwide and rising [1]. COVID-19 is the clinical manifestation of SARS-CoV-2 infection.

ARDS develops in 42% of COVID-19 patients presenting with pneumonia and accounts for a significant number of deaths associated with COVID-19 [2, 3]. ARDS is a form of hypoxemic respiratory failure defined by the presence of diffuse alveolar damage (DAD) commonly associated with bacterial pneumonia, sepsis, pancreatitis or trauma. The pathogenesis of ARDS is typically attributed to inflammatory injury to the alveolar-capillary membrane which leads to pulmonary oedema and respiratory insufficiency. In response to injury, alveolar macrophages orchestrate inflammation of the lung by recruiting neutrophils and circulating macrophages to the site of injury leading to the death of alveolar epithelial cells and further impairing lung function [4]. Severe COVID-19 is often also characterised by concurrent vascular disease and coagulopathy. This includes breakdown of the vascular barrier, oedema, endothelialitis and thrombosis. Thrombotic and microvascular complications are frequently recorded in deceased patients, suggesting that vascular pathology is a major driver of severe disease [5]. While these mechanisms contribute to disease progression, it is unclear which mechanisms are the predominant factors in causing fatality.

Transcriptomic analysis of patient tissue offers a unique opportunity to not only understand the pathogenesis of SARS-CoV-2 but also to assist in biomarker validation. To date, the major search for clinical biomarkers, including transcriptomic analyses on COVID-19 patient samples, have been performed on whole blood. Blood samples from infected patients are highly accessible but it is not clear whether blood parameters accurately reflect either the viral load or the level of tissue damage that clinical treatments aim to control [6]. The lung is the primary site of viral replication, yet only a limited number of transcriptomic studies have been performed on the lungs of COVID-19 patients. Such studies have used transcriptomic analyses on only a small patient cohort (1-2 patients) [7, 8] or relied on 'bulk' sequencing of whole lung tissue, where any insight into tissue heterogeneity in this complex organ is lost [5, 9, 10]. In contrast, Nanostring GeoMX^{TM} Digital Spatial Profiling (DSP) gene expression panels applied directly on tissue sections take into account the spatial location of transcriptomic features and are able to directly sample intra-organ heterogeneity [11, 12]. This provides a much deeper picture of the cellular changes driven by viral infection, and is a powerful way to define the characteristics of the host response to the virus and the spatial relationships between them. Indeed, such preliminary transcriptomic analysis of COVID-19 patient lungs appears promising [13]. However, such analyses require the application of advanced bioinformatics techniques to control for the numerous potential confounders present in gene expression data sampled in such a complex experimental design. In addition, there is a clear need to distinguish the host SARS-CoV-2 transcriptomic response (and associated spatial heterogeneity) from other infectious triggers of ARDS such as influenza virus, which may be circulating concurrently with SARS-CoV-2.

To address these questions, we investigated the pathogenesis of SARS-CoV-2 in lung tissue using spatial transcriptomics approaches from rapid autopsy samples taken from 10 COVID-

19 patients, five 2009 pandemic H1N1 (pH1N1) influenza virus patients and four control patient samples collected at a single institution. The Nanostring GeoMX^{TM} DSP platform combined with bioinformatic modelling allowed deconvolution of individual variability from viral and host factors. These data showed that viral load has a heterogenous impact within a patient on the pulmonary transcriptomic response, and there is significant overlap in the gene expression profiles between influenza virus and SARS-CoV-2 infected samples. However, several unique transcriptomic signatures were detected in the lungs of COVID-19 patients. This study demonstrates the strengths of spatial transcriptomics coupled with powerful linear modelling bioinformatic techniques to broadly identify key pathways involved in viral infections and to clearly differentiate the involvement of distinct cellular and molecular pathways in different infections.

Materials and methods

Study design.

Tissue microarray (TMA) cores were prepared from autopsied pulmonary tissue from 10 SARS-CoV-2 and 5 pH1N1 patients who died from respiratory failure (ARDS)(Supplementary Table 1-3). Control material was obtained from 4 uninfected patients (Supplementary Table 3). All SARS-CoV-2 and pH1N1 patients were confirmed for infection through RTqPCR of nasopharyngeal swab specimens, and imaging with computed tomography (CT) showed diffuse and bilateral opacities with ground-glass attenuation, suggestive of viral pulmonary infection. Autopsy and biopsy materials were obtained from the Pontificia Universidade Catolica do Parana PUCPR the National Commission for Research Ethics (CONEP) under ethics approval numbers: protocol number 3.944.734/2020 (for COVID-19 group), and 2.550.445/2018 (for H1N1 and Control group). All methods were carried out following relevant guidelines and regulations. Families permitted the post-mortem biopsy of COVID-19 and H1N1pdm09 samples and conventional autopsy for the cases of the Control group. The

study was also approved under University of Queensland and Queensland University of Technology Human Research Ethics Committee (HREC) ratification.

For both groups, initially, the imaging exams such as X-rays and CTs were analyzed to identify the pulmonary segments with more severe lung injury. Once we confirmed radiographically evident and representative lesions (especially the left lobes pulmonary segments, since they can be removed in a more agile and practical mini-thoracotomy technique), we collected the sample through an anterior mini-thoracotomy on the fourth or fifth intercostal space. The time elapsed between the patient's death, obtaining the consent form by the relatives, and removing the fragments through the mini-thoracotomy did not exceed 4 hours (COVID-19 and H1N1 groups). In all these cases (COVID-19 and H1N1 group), the samples were similarly sized (3x3 cm) and were delicately handled and resected using surgical scissors. The samples were then fixed into a 10% concentration formalin solution and kept in it for at least 24 hours until blocking and slicing for microscopic analysis. The lung formalin-fixed paraffin-embedded (FFPE) samples were stained with hematoxylin and eosin (H&E) to observe the histopathological aspects and find the appropriate and representative areas for punch and construct the TMA. Clinical data were obtained from medical records during hospitalization in the ICU (at Hospital Marcelino Champagnat in Curitiba, Brazil, for the COVID-19 group and at Hospital de Clínicas in Curitiba, Brazil for the H1N1 group).

Control group samples:

Also included in this study was a Control group of lung samples from patients who died from neoplastic or cardiovascular diseases (not involving lung lesions). The samples of the Control group were obtained by the conventional autopsy technique.

The time elapsed between the patient's death, obtaining the consent form by the relatives, and removing the fragments through the conventional autopsy did not exceed 6 hours.

These Control group samples were then fixed into a 10% concentration formalin solution and kept in it for at least 24 hours until blocking and slicing for microscopic analysis.

The lung formalin-fixed paraffin-embedded (FFPE) samples were stained with hematoxylin and eosin (H&E) to observe the histopathological aspects and find the proper punch areas to construct the TMA.

Clinical data were obtained from medical records during hospitalization at Hospital de Clínicas in Curitiba, Brazil.

Tissue Preparation and Histopathology. Thirty 5μm thick serial sections were cut from the TMA blocks onto positively charged slides (Bond Apex) and sections were stained with hematoxylin and eosin (H&E) and Masson's trichrome stain. Brightfield images were obtained using an Aperio (Leica Biosystems, US) slide scanner for assessment of histopathology by a pathologist. TMA cores were assessed for overall histological pattern. Regions of interest (ROIs) were semi-quantitatively analysed for alveolar haemorrhage and oedema, hyaline membrane formation, acute inflammation, type 2 pneumocyte hyperplasia, squamous metaplasia, capillary congestion, fibroblastic foci, and interstitial inflammation.

Immunohistochemistry and RNAscope®. Immunohistochemistry was performed on a Leica Bond-RX autostainer (Leica Biosystems, US) with antibody targeting SARS-CoV-2 spike protein (Abcam, ab272504) at 2 μg/ml. Heat induced epitope retrieval was performed in buffer ER1 at 100°C for 20 minutes, and signal visualized with 3,3'-Diaminobenzidine (DAB) substrate. Slides were imaged using a Zeiss Axioscanner (Carl Zeiss, Germany) and IHC was scored by a pathologist for bronchiolar epithelium, type 2 pneumocytes, interstitial lymphocytes and alveolar macrophage compartments.

RNAscope® probes (ACDbio, US) targeting SARS-CoV-2 spike mRNA (nCoV2019, #848561-C3), ACE2 host receptor mRNA (#848151-C2), and host serine protease TMPRSS2 mRNA (#470341-C1) were used as per manufacturer instructions for automation on Leica Bond RX. DNA was visualised with Syto13 (Thermofisher Scientific), channel 1 with Opal 570 (1:500), channel 2 with Opal 620 (1:1500), and channel 3 with Opal 690 (1:1500)(PerkinElmer). Fluorescent images were acquired with Nanostring Mars prototype DSP at 20x.

Nanostring GeoMX DSP Covid-19 Immune Atlas Panel. Freshly cut sections of each TMA were processed according to the Nanostring GeoMX Digital Spatial Profiler (DSP) slide preparation manual by the technology access program. Briefly, slides were baked 1 hour at 60^{o}C and then processed by Leica Bond RX autostainer. Slides were pre-treated with Proteinase K and then hybridised with mRNA probes contained in the Covid-19 Immune Atlas panel with additional SARS-CoV-2 spike-in panel. After incubation, slides were washed and then stained with CD68, CD3, PanCK, and Syto83 for 1 hour then loaded into GeoMX DSP instrument for scanning and ROI selection. Selected ROIs were guided by RNAscope® and IHC positivity on SARS-CoV-2 tissues with similar tissue structures captured on H1N1 and uninfected tissue cores. Oligonucleotides linked to hybridised mRNA targets were cleaved and collected for counting using Illumina i5 and i7 dual indexing. PCR reactions were performed with 4 μl of a GeoMx DSP sample. AMPure XP beads (Beckman Coulter) were used at 1.2× bead-to-sample ratio for PCR product purification. Paired-end sequencing (2×75) was performed using NextSeq550 up to 400M total aligned reads. Fastq files were processed by DND system and uploaded to GeoMX DSP system where raw and Q3 normalized counts of all targets were aligned with regions of interest (ROIs).

Transcriptomic Data. Data used in this study result from an mRNA assay conducted with the NanoString's GeoMx Covid-19 Immune atlas panel using the GeoMX Digital Spatial Profiler (DSP). The data were measurements of RNA abundance of over 1800 genes, including 22 add-in COVID-19 related genes, 4 SARS-CoV-2 specific genes and 2 negative control (SARS-CoV-2 Neg, NegProbe) genes and 32 internal reference genes. Samples were acquired from 3 TMA slides each containing 10 tissue cores from 10 patients. The three slides correspond to either COVID-19, H1N1 or Control Tissue biopsies. Transcriptomic measurements were made on regions of interests within each core. Control samples are COVID-19-free and H1N1-free,

but they are not necessarily disease-free samples; they originate from patients with other non-viral diseases. In total, 60 ROIs were analysed (47 COVID-19, 8 H1N1 and 5 "Control"). Factors considered in this dataset include disease type (COVID-19, H1N1 and Control), patient of origin, dominant tissue type (Type 2 pneumocytes, bronchiolar epithelium, hyaline membranes, macrophages) and viral load (based on RNAscope scores).

Bioinformatic analyses. Data exploration and quality checks were conducted on the Q3 normalised count data generated from the CTA DSP mRNA assay. Relative log expression (RLE) plots were used assess the presence of unwanted variation in the form of batch effects [14]. Data were normalised using the trimmed mean of M-values (TMM) method [15] using all genes in the panel. Specifically, $\log_2$-transformed transcript abundance data were median-centred for each gene, and then within each sample the difference between the observed and population median of each gene was calculated. Principal components analysis (PCA) of the samples was conducted to identify variability related to specific factors in the dataset and experimental design.

Differential expression (DE) analysis, Gene Ontology and KEGG pathway enrichment analysis were carried out using R/Bioconductor packages edgeR (v3.30.3) [16, 17] and limma (v3.44.3) [18]. Briefly, differential expression was modelled using linear models with various experimental factors as predictors. Variation in gene expression was modelled as the combination of a common dispersion that applies to all genes and a gene-specific dispersion. Limma was used to model and estimate the variation of each gene by borrowing information from all other genes using an empirical Bayes approach, thereby allowing estimation of the common and gene-wise variation with as few as 2 replicates per condition. Once the linear model was fit to a given experimental design, various contrasts of interest were used to query

the data for differential expression. The resulting statistic was an empirical Bayes moderated t-statistic which was more robust than a t-statistic from a classic t-test. Based on the observed differences between cores and tissue structures from the PCA analyses, considerations were made to allow for similarity that exists for regions originating from the same core using *duplicationCorrelations* in limma [19]. The flexible modelling framework afforded by linear models was used to take into account differences between heterogenous tissue structures (i.e. bronchiolar epithelial samples *vs* the rest of the samples) by including them as covariates in the models.

Two main factors were investigated in this analysis, namely disease type (COVID-19, H1N1 and Control) and viral load. For disease type, two comparisons were investigated: COVID-19 against *Control* with *cores* and *dominant tissue type* as covariates, and COVID-19 against H1N1 samples, with *cores* and *dominant tissue type* as covariates. For these two contrasts, the voom-limma with *duplicationCorrelations* pipeline [20] was used to fit linear models. The TREAT criteria was then applied [21] (p-value <0.05) to perform statistical tests and subsequently calculate the t-statistics, log-fold change (logFC), and adjusted *p*-values for all genes. For viral load, the voom-limma with *duplicationCorrelations* pipeline [20] (p-value <0.05) was used to fit linear models for the comparison between high and low viral load regions originating from COVID-19 patient cores (disease type as a covariate). The contrast was applied on two resolutions, firstly by grouping ROIs from the same cores with the same degree of viral load (cores/patients-based approach); secondly by treating each regions/ROIs sample as an independent observation (ROI-based approach). Gene set enrichment analysis (GSEA) were performed using the *fry* approach from the limma package. Gene sets from the Molecular Signatures Database (MsigDB) Hallmarks [22, 23] , C2 (curated gene sets) and C7 (immunologic signature gene sets) categories were tested using *fry*. Pathways from the KEGG

pathway database were tested using the *kegga* function in the limma package and gene ontology enrichment was assessed using *goana* from the limma package [19].

Results

Patient Characteristics

Lung tissues were obtained by rapid autopsy from 10 SARS-CoV-2 infected patients, five pH1N1 influenza patients and four non-virally infected control patients. The SARS-CoV-2 cohort was made up of four females and six males with mean age of 76 years (ranging from 46-93 years). These patients exhibited multiple comorbidities including obesity, diabetes and heart disease and had received varying treatment following hospital admission. Patient survival after admission ranged from 8 to 38 days, with time on mechanical ventilation ranging from 3 to 21 days (Supplementary Table 1). Blood profiling performed on admission and again immediately prior to death showed elevated D-dimer levels for 7 of 10 patients consistent with vascular coagulopathy (Supplementary Table 2).

The pH1N1 influenza virus patient cohort comprised five males with a mean age of 45.8 years (ranging from 31-53 years). All patients had received mechanical ventilation from the time of admission until death. The uninfected control cohort consisted of four males with a mean age of 36.25 (1 ranging from 18-60 years) whose tissues were donated following death from non-viral and non-respiratory diseases (Extended Data Table 3).

Histopathological Characterisation

Tissue microarrays (TMAs) were prepared from pulmonary formalin-fixed paraffin-embedded cores of each of the three cohorts (one core per patient) and blinded histological examination was performed by a pathologist (Supplementary Table 4). Acute phase diffuse alveolar damage

(DAD) characterised by hyaline membranes, type 2 pneumocyte hyperplasia and alveolar septal thickening was a prominent feature of SARS-CoV-2 cores. pH1N1 tissues also showed features of DAD, with alveolar haemorrhage evident in two of the cores. These findings were not seen in control cores. Fibrosis was not a prominent feature, apart from one SARS-CoV-2 core that exhibited moderate interstitial and alveolar fibrosis with concordant histological organising pneumonia. Two pH1N1 cores showed mild interstitial fibrosis that were characterised by acute phase or late organising phase DAD.

RNAscope® and ROI Selection for Digital Spatial Transcriptional Profiling

Transcriptional profiling of target cores was guided by the detection of SARS-CoV-2 spike mRNA. Two cores (Figure 2; LN1 and LN3) exhibited strong signals for viral mRNA (Fig. 2a) and these cores were comprehensively evaluated using the GeoMx DSP platform (shaded regions, Fig. 2b). Twenty-five regions of interest (ROIs) were selected from these two RNAscope® positive cores and 21 ROIs were selected from remaining eight cores of SARS-CoV-2 infected patients (minimum of one ROI per patient) for which viral mRNA was below detection by RNAscope. Eight ROIs were selected from the lungs of five pH1N1 patients (minimum of one ROI per patient), and five ROIs were selected from the lungs of control patients (minimum of one ROI per patient).

SARS-CoV-2 RNAscope® abundance was semi-quantitatively assessed for each ROI. Representative scoring criteria from 0 to 3 are shown in Supplementary Fig. 2 where positive scores were allocated only to regions within LN1 and LN3 (Supplementary Table 5). In addition, concordant spike protein immunohistochemistry (IHC) appeared to be specific for type 2 pneumocytes and bronchiolar epithelial regions, however, some staining was also observed within interstitial lymphocytes and alveolar macrophages (Supplementary Fig. 1).

Interestingly, spike protein IHC indicated virus in cores other than LN1 and LN3 (Supplementary Table 5) that did not produce a detectable signal using RNAscope®. Low level non-specific staining was observed in some regions of pH1N1 and control tissues Supplementary Table 5. We thus focussed our ROI selection on cores that indicated high mRNA integrity and viral detection by RNAscope®.

Histopathology and Dominant Features of ROIs

Assessment of histology by a pathologist allowed ROIs to be grouped by tissue architecture prior to further analysis. The predominant cell types found within each ROI were annotated in addition to scoring of their histopathology and spike protein IHC (Supplementary Table 5). ROIs could be broadly categorised into regions dominated by bronchiolar epithelial structures, type 2 pneumocytes, hyaline membranes and macrophages (Supplementary Fig. 2). Alveolar oedema, acute inflammation and squamous metaplasia were not observed in diseased or control cores, however, type 2 pneumocyte hyperplasia and capillary congestion were present in SARS-CoV-2 pulmonary tissues.

Data normalization strategy and exploratory analysis of spatial transcriptomic data

Q3 housekeeping normalised data was corrected for systematic bias by TMM to allow for comparisons between SARS-CoV-2, pH1N1 influenza virus and control samples (Supplementary Fig. 3). Principal components analysis (PCA) plots were then used to explore the variability in SARS-CoV-2, pH1N1 influenza virus and control lung samples and to determine whether they could be clearly separated simply based on the type of infecting pathogen and to identify factors that might confound differential expression analysis. Principal components (PCs) capture orthogonal dimensions of variability in the data in descending order of contribution. Non-infected samples clearly separate from other samples on PCs 1 and 3,

while SARS-CoV-2 and pH1N1 influenza samples remained associated on these axes (Fig. 3a). Indeed, across PCs 1-4, SARS-CoV-2 and pH1N1 influenza samples substantially overlapped. This suggests that the transcriptomic profile of SARS-CoV-2 and pH1N1 influenza virus tissues were not highly variable at low viral load, and that a common transcriptome was induced by these two viruses in fatal cases.

To further examine the major sources of variation in the dataset, we labelled PCA plots by multiple different parameters including sample core (Fig. 3b), dominant cell type (Fig. 3c) and viral load (Fig. 3d). When samples were labelled by core, PC 3 and PC 4, were able to distinguish patient specific effects, as shown by a clear separation of cores LN1 and LN3, with clustering of other cores with multiple ROIs (LN10) (Fig. 3b). This suggests that within patient correlations and patient-patient variation has an important underlying effect on the observed transcriptome and that this needs to be taken into account in any subsequent analysis Additionally, samples separating on PC2, and clearly clustered from other samples when plotting PC1 and 2, were grouped based on the dominant cell type found within the sample core (Fig. 3c). Specifically, ROIs where bronchiolar epithelial cells were the dominant cell type clustered together independent of whether the patient had pH1N1 influenza virus or SARS-CoV-2 (Fig. 3c), or belonged to the non-infected controls. This implied that the tissue heterogeneity captured by dominant cell type was a substantial contributor to the variability in our experiment. While epithelial cells are the primary cellular target of both SARS-CoV-2 and influenza virus[24], our data did not identify viral epithelial tropism as a key regulator of signal as the cores dominated by bronchial epithelial cells did not exhibit the highest viral load (as defined by RNA Scope for SARS-CoV-2)(Fig. 3d). However, the high SARS-CoV-2 viral load cores separate from other virally-infected samples and control samples on PC3 (Fig. 3d). Together, these analyses suggest that unbiased and accurate comparative transcriptomics analyses between uninfected, SARS-CoV-2 and pH1N1 influenza lung samples must take

account of the dominant cell type present in samples, inter-patient variations and within-patient correlations. The ability to include structural and spatial covariates demonstrates the key advantage in using a spatial transcriptomics as compared with bulk RNA sequencing, but also necessitates careful construction of the model used to compare expression across the ROIs sampled.

Severe COVID-19 infection results in a global pro-inflammatory response and downregulation of immune cell effector and regeneration pathways

Differential gene expression analysis was subsequently performed on all sample cores, taking into account the effect of the dominant cell type and the appropriate within and inter-patient variations. 286 genes were identified as significantly upregulated and 20 genes were significantly downregulated in the lungs of SARS-CoV-2 patients compared with non-virally infected controls (Fig. 4, Supplementary Table 6, 7).

Down regulated genes included immune-related genes such as B lymphocyte antigen *CD19*, cytokine genes such as *IL7*, *IL9* and *IL13* and components of the complement cascade (e.g. C9) together with *CSF3*, colony stimulating factor, normally associated with mobilization of cells from the bone marrow (Supplementary Table 6). A number of tumor-associated genes including *GAGE1* and *MAGEC2* were downregulated, with these genes being associated with cell survival. Interestingly, *PCK1*, which is required for tissue gluconeogenesis, and *MPL*, that regulates the generation of megakaryocytes and thus platelet formation were also highly reduced [25].

A more diverse array of upregulated genes was found to be differentially expressed in the lungs of SARS-CoV-2 patients compared with those from patients who died of non-viral causes (Supplementary Table 7). These included genes associated with antigen presentation (e.g. *B2M, HLA-DRA, HLA-C* and *HLA-E),* the Type I interferon (IFN) response (e.g. *LY6E* and *IFI27*), fibrosis and epithelial cell growth (e.g. *COL3A1, FN1, KRT7, COL1A1, KRT18 and KRT19)* and the complement cascade (*C1QA* and *C1R*). Consistent with these observations, gene set enrichment analysis showed that hallmark genes of biological processes such as reactive oxygen species, complement, IL6/JAK/STAT3 signalling, apoptosis, p53 signalling IFN-γ response, hypoxia, IFN-α response, coagulation, TGF-β signalling, IL-2/STAT5 signalling, angiogenesis, TNF-α signalling via NFκ-B, and inflammatory response were all significantly upregulated in SARS-CoV-2 infected lungs (Table 1). Gene set enrichment using blood coagulation, hypoxia responses and angiogenesis related genes from nanoString's nCounter® PanCancer Progression Panel further confirmed the upregulation of pathways associated with blood coagulation and angiogenesis responses (Table 2). These analyses show that SARS-CoV-2 samples significantly upregulate genes associated with blood coagulation, angiogenesis, the IFN-α and IFN-γ response and positively regulate cytokine production and cytokine stimulus during the immune response (Fig. 5). These data are consistent with previous descriptions of a 'cytokine storm' response in patients with severe SARS-CoV-2, often accompanied by various coagulopathies [26]. They are also in agreement with the notion that a late stage or delayed type I IFN response may be observed in concert with a pro-inflammatory response [27].

Limited differential gene expression associated with SARS-CoV-2 viral load

The pathogenesis of SARS-CoV-2 is a complex interplay of host immunopathology and viral load. To determine how viral load influenced the host transcriptomic signal, we performed a differential expression analysis of SARS-CoV-2 samples stratifying based on viral load. Patient sample cores were classified as either 'high' or 'low' based on the presence of viral mRNA by RNAscope®. Thus, two cores which were scored positively by RNAscope®, LN1 and LN3, were designated as 'high' while the remaining 8 cores from SARS-CoV-2 patients were scored as 'low'. This analysis showed that only 17 up-regulated and 7 down-regulated genes were differentially expressed between the high viral load SARS-CoV-2 infected cores and the other SARS-CoV-2 cores (Fig. 6, Table 3). Consistent with our RNAScope analyses, SAR-CoV-2 specific genes *S* and *ORF1ab* were the most strongly upregulated genes in the high viral load group (1.916 and 1.861 log2 fold change, respectively). This is also consistent with previous observations that in at least some COVID-19 patients, these are the most highly expressed viral genes in the lung [28]. All other upregulated genes with >1 log2 fold change in expression were associated with the type I IFN response (*CXCL10, IFIT3, ISG15, MX1, GBP1* and *IFI6*) (Fig. 6a, Table 3).

The above analysis assumed that every ROI derived from the same core would have the same degree of viral infection (i.e. high or low). However, viral infection is unlikely to be uniform across tissues. Indeed, none of the bronchiolar epithelial ROIs in the two viral high cores were classified as high (Fig. 3d, Supplementary Table 5). To examine the impact of this, we considered each ROI as an independent sample which was classified as either 'high' (score > 0) or 'low' (score = 0) by RNAscope positivity. By stratifying cores with this criterion and accounting for the dominant cell type as well as correlations between ROIs from the same core, we identified 33 differentially upregulated genes and 132 downregulated genes in high viral load ROIs compared to low viral load ROIs (Supplementary Table 8). Once again, *S* and

ORF1ab were the most strongly upregulated genes in the high viral load group (2.415 and 2.365 log2 fold change respectively). Other upregulated genes of note were those associated with the type I IFN response (e.g. *RSAD2, OASL, TLR8, IL1RN* and *IKBKE*), chemokines associated with IFN–γ (e.g. *CXCL11, CCL8* and *CCL7*) and the SARS-CoV-2 receptor *ACE2*. A more diverse array of genes were downregulated in high viral load samples including those associated with the complement cascade (e.g. *C3*), ribosomal proteins (e.g. *RPS6* and *RPL23*) and antioxidants (e.g. *PRDX5*) (Figure 6B). Taken together, these data suggest that a pronounced type I interferon gene signature is associated with SARS-CoV-2 RNA expression and the spatial context provided by this technology enables finer details of viral load associations to be determined.

Severe COVID-19 is associated with a limited, differential transcriptome when compared with severe influenza

Only a limited number of genes were identified as differentially expressed between COVID-19 and pH1N1 influenza samples (2 downregulated and 4 upregulated genes, Fig. 7). Of the six genes significantly differentially regulated in COVID-19 samples, three were associated with the type I IFN response (*LY6E, IFI27* & *IFI6*), two were heat shock protein family members (*HSPA6* & *HSPA1A*) and one, *CT45A1,* when combined with various growth factors, is associated with cell survival and tumorigenesis. Interestingly, all three of the interferon-inducible genes were also significantly upregulated in COVID-19 samples compared to those without a viral infection (Supplementary Table 7). Two heat shock protein family members (*HSPA1A* and *HSPA*) were also identified as significantly down-regulated in COVID-19 tissues. These two genes were also significantly upregulated in pH1N1 tissues compared to uninfected tissues (data not shown), suggesting a potential influenza specific signature. To understand the relevance of these genes to COVID-19 and pH1N1, we have compared these

differentially expressed genes to other published spatial transcriptomics approaches both experimentally and analytically and that is shown by the varying sizes of the gene sets (Figure 8). It is clear that our approach is able to identify a focussed set of gene signatures specific to COVID-19 infection as compared to other studies (*IFI27* and *LY6E* common across all studies, and *IFI6* common in 4 out of the 5 studies). Together, these data indicate that fatal pH1N1 influenza virus and SARS-CoV-2 lung samples have only subtly different host transcriptomes but may be potentially distinguished by specific gene signatures. This is further clarified in the pH1N1 vs Control (Supplementary Fig 6) and GSEA analysis (Supplementary Fig 7) which shows many common activation pathways for both COVID-19 and pH1N1 infected tissue when compared to control. These data suggest that many DE pathways identified in the COVID-19 sample may be indicative of a more generic viral-induced ARDS, whilst only a limited number of DE genes may be unique to COVID-19 itself.

Discussion

Understanding the biological functions, networks, and host-pathogen interactions that impact organ and disease development requires both cellular information and a spatial context. Analyses of bulk RNA sequencing or scRNA sequencing provides a global overview of an organ's response to a pathogen, typically identifying broad inflammatory pathways. However, such approaches fail to identify subtle individual cellular changes that are spatially distinct. This has particular impact when considering innate and adaptive immune cells that may be crucial for understanding pathogen-specific responses, and to distinguish the profile of one pathogen from another. In contrast, spatially resolved transcriptomes of virally-infected tissues offers the possibility of disentangling the individual infected cells, contributions of viral load, cellular responses and hence patient-to-patient variability.

The clinical spectrum of COVID-19 is highly variable. Association of viral load with disease severity would provide a mechanism for early stratification of patients for treatment options. Indeed, analyses of nasopharyngeal swabs suggests that nasopharyngeal SARS-CoV-2 RNA is independently correlated with disease severity [29], similar to earlier observations with SARS-CoV-1[30]. However, the association between viral replication in the lung and severe disease remains more complicated. Here, we observed that 8 out of 10 patients who died because of COVID-19 had no detectable viral RNA in lung biopsies as determined by RNAscope. Whilst this rate of viral RNA detection may be lower than expected, it is important to recognise that viral antigen deposition in the lung is heterogeneous, and therefore these data may simply reflect random sample selection. There was no notable difference in time from admission to death in patients who were positive vs. those who were negative by viral RNA-Scope. However, we do not have information for these patients on time between symptom onset and hospital admission which may have influenced viral clearance. Nevertheless, it is important to note that all patients in this study were PCR positive for SARS-CoV-2 RNA via nasopharyngeal swab. Therefore, if the absence of viral RNA in the lungs of 8 patients reflects viral clearance this was restricted to the lower respiratory tract.

Transcriptomic analysis showed that areas of high viral load were associated with a pronounced type I IFN response, consistent with other preliminary spatial analysis of the lungs of COVID-19 patients [13] and assumedly due to increased viral pathogen-associated molecular patterns (PAMPs) available for type I IFN stimulation. Nevertheless, even in samples where no viral RNA was detected, a pronounced type I IFN gene signature could still be observed in the pulmonary tissue. These observations add weight to the growing understanding of the role type I IFNs in SARS-CoV-2 infections [31]. Recent findings in a COVID-19 tissue atlas study found upregulation of IFN-α and IFN-γ response genes and oxidative phosphorylation pathways in

the lungs of COVID-19 patients vs control [32]. Current evidence suggests that an early and robust induction of type I interferons can help control viral replication and help ensure a mild infection. In contrast, a delayed induction of type I IFN (i.e. after the peak of viral replication) is largely irrelevant for viral control as viral titres have already declined at this point in the infection but may help perpetuate the detrimental pro-inflammatory response and lung damage [31, 33]. The transcriptomic data presented here thus speak to the potential dual-role of type I IFNs in SARS-CoV-2 infection.

Pulmonary transcriptomic analysis is a powerful tool to delineate the pathogenesis of SARS-CoV-2 and identify how it differs compared to other respiratory pathogens such as influenza virus. Previous studies have suggested that the lungs of COVID-19 patients display an increased incidence of alveolar capillary microthrombi and thrombosis with microangiopathy compared to those of influenza patients [5]. Consistent with these observations, an earlier report using bulk lung RNA analyses to identify 69 differentially-expressed angiogenesis-related genes in COVID-19 patients, but not in influenza patients [13]. In the present study, we observed that both COVID-19 and influenza patients had an upregulation of genes associated with coagulation and angiogenesis, these genes were not differentially expressed between the two virus-infected patient groups. These contrasting data most likely reflects the fact that all samples in the present study were collected at the end stage of disease and accordingly, earlier differences in influenza and COVID-19 disease pathogenesis and severity were unable to be examined. However, these data do highlight both the difficulties and importance of finding specific biomarkers that are sensitive enough to differentiate COVID-19 from other respiratory viral infections even at end-stage disease.

Several potential clinical and transcriptional biomarkers for triaging patients with COVID-19 have been explored. Clinical biomarkers include C-reactive protein, serum amyloid A, interleukin-6, lactate dehydrogenase, neutrophil-to-lymphocyte ratio, D-dimer, lymphocytes and platelet count [34, 35]. However, the majority of these studies have focussed on biomarkers in patient blood. While peripheral blood samples are a convenient and highly accessible site for clinical sampling, a biomarker that could be rapidly identified in the respiratory tissue (e.g. via a nasal swab) may be more accessible in low resource settings. *IFI27* levels in the blood have been previously identified to have a 88% diagnostic accuracy (AUC) and 90% specificity in discriminating between influenza and bacterial infections in patients [36]. Interestingly, previous analyses of *IFI27* in the blood of COVID-19 patients revealed that *IFI27* was upregulated in SARS-CoV-2 infection [37-39]. Here, we identified *IFI27* as differentiated upregulated in the lungs of both COVID-19 patients *vs* control patients and in the highly restricted set of genes differentially expressed between COVID-19 and influenza patients. These data raise the exciting possibility that *IFI27* may not only represent a biomarker for severe COVID-19 but that it may also help differentiate this disease from other clinically similar viral infections. This becomes particularly important as the 'second wave' of COVID-19 in the US and Europe is set to overlap with the winter influenza season. Validation of this gene in nasal specimens, as well associations with disease severity will be required to confirm *IFI27* as a gene signature that is useful in stratifying COVID-19 patients.

This study was subject to several important limitations. Firstly, all data was derived from a small sample cohort derived from a single study site, including unbalanced patient vs control groups, and it remains to be determined how much these data can be extrapolated to other patient populations. Importantly, future studies with a larger number of patients should include an equal number of males and females as male sex is independently associated with increased

COVID-19 severity [40]. Furthermore, additional studies are required across a broader range of patients (i.e. those with mild and moderate disease) to determine the therapeutic value of measuring *IFI27* levels during infection. Moreover, there have been rapid advances using antibody guided ROI capture, whole transcriptome assays, and cellular deconvolution to infer cell types. However, despite these limitations these data reveal the unprecedented power of spatial profiling combined with detailed multiparameter bioinformatic analyses to dissect the key variables that contribute to differential gene expression across highly variable patient cohorts and the heterogeneous distribution of virus and immune responsiveness within tissues.

Data and materials availability: Nanostring data that support the findings of this study will be deposited with the Gene Expression Omnibus (GEO) repository and made available on publication. All the data to evaluate the conclusions in this paper are present either in the main text or in supplementary materials.

Acknowledgments. We thank S. Weaver, M. Haraguchi (Fred Hutch Experimental Histopathology Department, USA), L. Pan, A. Nam (Nanostring Technologies Seattle, USA), I. Stefani (University of Queensland Diamantina Institute Translational Innate Immunology Group) and G. Smyth (WEHI). We thank the patients and their families that made this study possible.

Funding. This work was supported by grants and fellowships from the National Health and Medical Research Council (NHMRC) of Australia (1157741 AK; 1135898 GTB, 1140406 FSFG), Priority driven Collaborative Cancer Research Scheme, funded by Cure Cancer Australia with the assistance of Cancer Australia and the Can Too Foundation (1182179 AK; 1158085 FSFG), Princess Alexandra Research Foundation (KOB), University of Queensland

(GTB, FSFG), Queensland University of Technology (AK), The Garnett Passe and Rodney Williams Memorial Foundation (AK), Walter and Eliza Hall Institute of Medical Research (CT, DB, and MJD). MJD is supported by the Betty Smyth Centenary Fellowship in Bioinformatics. DB is supported by a Chan Zuckerberg Initiative Program grant awarded to G. Smyth.

Competing interests. Fernando Souza-Fonseca-Guimaraes is a consultant for Biotheus Inc. However, all opinions and reviews presented in this manuscript belong to the authors alone and are independent of the relationships with Biotheus. Other authors declare no competing interests.

Author Contributions

AK, FSFG, KRS, GTB, KOB conceived of the study. AK, AFRDSM, JM, JDSMJ, CBVDP, SN, CPB, PSFG, LDN, CC, TM, GRR performed the experimentation. AK, CWT, DB, JM, CC, BT, KRS, MJD, FSFG, GTB performed the data analysis and interpretation. All authors critically reviewed and approved the manuscript prior to submission.

References

1. World Health Organization Coronavirus Disease (COVID-19) Dashboard. 2020 [cited; Available from: https://covid19.who.int/
2. Grasselli G, Tonetti T, Protti A, Langer T, Girardis M, Bellani G, Laffey J, Carrafiello G, Carsana L, Rizzuto C, Zanella A, Scaravilli V, Pizzilli G, Grieco DL, Di Meglio L, de Pascale G, Lanza E, Monteduro F, Zompatori M, Filippini C, Locatelli F, Cecconi M, Fumagalli R, Nava S, Vincent JL, Antonelli M, Slutsky AS, Pesenti A, Ranieri VM, collaborators. Pathophysiology of COVID-19-associated acute respiratory distress syndrome: a multicentre prospective observational study. *Lancet Respir Med* 2020.
3. Gibson PG, Qin L, Puah SH. COVID-19 acute respiratory distress syndrome (ARDS): clinical features and differences from typical pre-COVID-19 ARDS. *Med J Aust* 2020: 213(2): 54-56 e51.
4. Short KR, Kroeze EJBV, Fouchier RAM, Kuiken T. Pathogenesis of influenza-induced acute respiratory distress syndrome. *The Lancet Infectious Diseases* 2014: 14(1): 57-69.
5. Ackermann M, Verleden SE, Kuehnel M, Haverich A, Welte T, Laenger F, Vanstapel A, Werlein C, Stark H, Tzankov A, Li WW, Li VW, Mentzer SJ, Jonigk D. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in Covid-19. 2020: 383(2): 120-128.
6. Xiong Y, Liu Y, Cao L, Wang D, Guo M, Jiang A, Guo D, Hu W, Yang J, Tang Z, Wu H, Lin Y, Zhang M, Zhang Q, Shi M, Liu Y, Zhou Y, Lan K, Chen Y. Transcriptomic characteristics of bronchoalveolar lavage fluid and peripheral blood mononuclear cells in COVID-19 patients. *Emerging Microbes & Infections* 2020: 9(1): 761-770.
7. Blanco-Melo D, Nilsson-Payant BE, Liu W-C, Uhl S, Hoagland D, Møller R, Jordan TX, Oishi K, Panis M, Sachs D, Wang TT, Schwartz RE, Lim JK, Albrecht RA, tenOever BR. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* 2020: 181(5): 1036-1045.e1039.
8. Margaroli C, Benson P, Sharma N, Madison M, Robison S, Arora N, Zhang L, Patel R, Gaggar A. Spatial mapping of SARS-CoV-2 and H1N1 Influenza Lung Injury Identifies Differential Transcriptional Signatures. *Cell* 2020.
9. Nienhold R, Ciani Y, Koelzer VH, Tzankov A, Haslbauer JD, Menter T, Schwab N, Henkel M, Frank A, Zsikla V, Willi N, Kempf W, Hoyler T, Barbareschi M, Moch H, Tolnay M, Cathomas G, Demichelis F, Junt T, Mertz KD. Two distinct immunopathological profiles in autopsy lungs of COVID-19. *medRxiv* 2020.
10. Staines HM, Kirwan DE, Clark DJ, Adams ER, Augustin Y, Byrne RL, Cocozza M, Cubas-Atienza AI, Cuevas LE, Cusinato M, Davies BMO, Davis M, Davis P, Duvoix A, Eckersley NM, Forton D, Fraser A, Garrod G, Hadcocks L, Hu Q, Johnson M, Kay GA, Klekotko K, Lewis Z, Mensah-Kane J, Menzies S, Monahan I, Moore C, Nebe-von-Caron G, Owen SI, Sainter C, Sall AA, Schouten J, Williams C, Wilkins J, Woolston K, Fitchett JRA, Krishna S, Planche T. Dynamics of IgG seroconversion and pathophysiology of COVID-19 infections. *medRxiv* 2020.
11. Merritt CR, Ong GT, Church SE, Barker K, Danaher P, Geiss G, Hoang M, Jung J, Liang Y, McKay-Fleisch J, Nguyen K, Norgaard Z, Sorg K, Sprague I, Warren C, Warren S, Webster PJ, Zhou Z, Zollinger DR, Dunaway DL, Mills GB, Beechem JM. Multiplex digital spatial profiling of proteins and RNA in fixed tissue. *Nature biotechnology* 2020: 38(5): 586-599.
12. Monkman J, Taheri T, Ebrahimi Warkiani M, O'Leary C, Ladwa R, Richard D, O'Byrne K, Kulasinghe A. High-Plex and High-Throughput Digital Spatial Profiling of Non-Small-Cell Lung Cancer (NSCLC). *Cancers* 2020: 12(12).

13. Desai N, Neyaz A, Szabolcs A, Shih AR, Chen JH, Thapar V, Nieman LT, Solovyov A, Mehta A, Lieb DJ, Kulkarni AS, Jaicks C, Pinto CJ, Juric D, Chebib I, Colvin RB, Kim AY, Monroe R, Warren SE, Danaher P, Reeves JW, Gong J, Rueckert EH, Greenbaum BD, Hacohen N, Lagana SM, Rivera MN, Sholl LM, Stone JR, Ting DT, Deshpande V. Temporal and Spatial Heterogeneity of Host Response to SARS-CoV-2 Pulmonary Infection. *medRxiv* 2020.
14. Gandolfo LC, Speed TP. RLE plots: Visualizing unwanted variation in high dimensional data. *PLOS ONE* 2018: 13(2): e0191629.
15. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biology* 2010: 11(3): R25.
16. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research* 2012: 40(10): 4288-4297.
17. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2009: 26(1): 139-140.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* 2015: 43(7): e47-e47.
19. Smyth GK, Michaud J, Scott HS. Use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* 2005: 21(9): 2067-2075.
20. Law CW, Chen Y, Shi W, Smyth GK. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biology* 2014: 15(2): R29.
21. McCarthy DJ, Smyth GK. Testing significance relative to a fold-change threshold is a TREAT. *Bioinformatics* 2009: 25(6): 765-771.
22. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. 2005: 102(43): 15545-15550.
23. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The Molecular Signatures Database (MSigDB) hallmark gene set collection. *Cell systems* 2015: 1(6): 417-425.
24. Radzikowska U, Ding M, Tan G, Zhakparov D, Peng Y, Wawrzyniak P, Wang M, Li S, Morita H, Altunbulakli C, Reiger M, Neumann AU, Lunjani N, Traidl-Hoffmann C, Nadeau KC, O'Mahony L, Akdis C, Sokolowska M. Distribution of ACE2, CD147, CD26, and other SARS-CoV-2 associated molecules in tissues and immune cells in health and in asthma, COPD, obesity, hypertension, and COVID-19 risk factors. *Allergy* 2020.
25. Carvelli J, Demaria O, Vély F, Batista L, Benmansour NC, Fares J, Carpentier S, Thibult M-L, Morel A, Remark R, André P, Represa A, Piperoglou C, Miranda LA, Baron W, Belaid N, Caillet C, Caraguel F, Carrette B, Carrette F, Chanuc F, Courtois R, Fenis A, Giordano M, Girard-Madoux M, Giraudon-Paoli M, Gourdin N, Grondin G, Guillot F, Habif G, Jaubert S, Lopez J, Le Van M, Lovera N, Mansuy M, Bonnet E, Sansaloni A, Reboul A, Mitry E, Nekkar-Constant C, Péri V, Ricaut P, Simon L, Vallier J-B, Vétizou M, Zerbib R, Ugolini S, Etiennot M, Galluso J, Lyonnet L, Forel J-M, Papazian L, Velly L, André B, Briantais A, Faucher B, Jean E, Seguier J, Veit V, Harlé J-R, Pastorino B, Delteil C, Daniel L, Boudsocq J-P, Clerc A, Delmond E, Vidal P-O, Savini H, Coutard B, Cordier PY, Le Dault E, Guervilly C, Simeone P, Gainnier M, Morel Y, Ebbo M, Schleinitz N, Vivier E, the Explore C-IPHg, the Explore C-MIg. Association of COVID-19 inflammation with activation of the C5a–C5aR1 axis. *Nature* 2020.

26. Shi W, Lv J, Lin L. Coagulopathy in COVID-19: Focus on vascular thrombotic events. *J Mol Cell Cardiol* 2020: 146: 32-40.
27. Israelow B, Song E, Mao T, Lu P, Meir A, Liu F, Alfajaro MM, Wei J, Dong H, Homer RJ, Ring A, Wilen CB, Iwasaki A. Mouse model of SARS-CoV-2 reveals inflammatory role of type I interferon signaling. *Journal of Experimental Medicine* 2020: 217(12).
28. Liu T, Jia P, Fang B, Zhao Z. Differential Expression of Viral Transcripts From Single-Cell RNA Sequencing of Moderate and Severe COVID-19 Patients and Its Implications for Case Severity. 2020: 11(2568).
29. Pujadas E, Chaudhry F, McBride R, Richter F, Zhao S, Wajnberg A, Nadkarni G, Glicksberg BS, Houldsworth J, Cordon-Cardo C. SARS-CoV-2 viral load predicts COVID-19 mortality. *The Lancet Respiratory Medicine* 2020: 8(9): e70.
30. Giannis D, Ziogas IA, Gianni P. Coagulation disorders in coronavirus infected patients: COVID-19, SARS-CoV-1, MERS-CoV and lessons from the past. *Journal of Clinical Virology* 2020: 127: 104362.
31. Park A, Iwasaki A. Type I and Type III Interferons - Induction, Signaling, Evasion, and Application to Combat COVID-19. *Cell host & microbe* 2020: 27(6): 870-878.
32. Delorey TM, Ziegler CGK, Heimberg G, Normand R, Yang Y, Segerstolpe Å, Abbondanza D, Fleming SJ, Subramanian A, Montoro DT, Jagadeesh KA, Dey KK, Sen P, Slyper M, Pita-Juárez YH, Phillips D, Biermann J, Bloom-Ackermann Z, Barkas N, Ganna A, Gomez J, Melms JC, Katsyv I, Normandin E, Naderi P, Popov YV, Raju SS, Niezen S, Tsai LTY, Siddle KJ, Sud M, Tran VM, Vellarikkal SK, Wang Y, Amir-Zilberstein L, Atri DS, Beechem J, Brook OR, Chen J, Divakar P, Dorceus P, Engreitz JM, Essene A, Fitzgerald DM, Fropf R, Gazal S, Gould J, Grzyb J, Harvey T, Hecht J, Hether T, Jané-Valbuena J, Leney-Greene M, Ma H, McCabe C, McLoughlin DE, Miller EM, Muus C, Niemi M, Padera R, Pan L, Pant D, Pe'er C, Pfiffner-Borges J, Pinto CJ, Plaisted J, Reeves J, Ross M, Rudy M, Rueckert EH, Siciliano M, Sturm A, Todres E, Waghray A, Warren S, Zhang S, Zollinger DR, Cosimi L, Gupta RM, Hacohen N, Hibshoosh H, Hide W, Price AL, Rajagopal J, Tata PR, Riedel S, Szabo G, Tickle TL, Ellinor PT, Hung D, Sabeti PC, Novak R, Rogers R, Ingber DE, Jiang ZG, Juric D, Babadi M, Farhi SL, Izar B, Stone JR, Vlachos IS, Solomon IH, Ashenberg O, Porter CBM, Li B, Shalek AK, Villani A-C, Rozenblatt-Rosen O, Regev A. COVID-19 tissue atlases reveal SARS-CoV-2 pathology and cellular targets. *Nature* 2021: 595(7865): 107-113.
33. Lee JS, Shin EC. The type I interferon response in COVID-19: implications for treatment. *Nature reviews Immunology* 2020: 20(10): 585-586.
34. Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A. The role of biomarkers in diagnosis of COVID-19 - A systematic review. *Life Sci* 2020: 254: 117788-117788.
35. Malik P, Patel U, Mehta D, Patel N, Kelkar R, Akrmah M, Gabrilove JL, Sacks H. Biomarkers and outcomes of COVID-19 hospitalisations: systematic review and meta-analysis. *BMJ Evidence-Based Medicine* 2020.
36. Tang BM, Shojaei M, Parnell GP, Huang S, Nalos M, Teoh S, O'Connor K, Schibeci S, Phu AL, Kumar A, Ho J, Meyers AFA, Keynan Y, Ball T, Pisipati A, Kumar A, Moore E, Eisen D, Lai K, Gillett M, Geffers R, Luo H, Gul F, Schreiber J, Riedel S, Booth D, McLean A, Schughart K. A novel immune biomarker IFI27 discriminates between influenza and bacteria in patients with suspected respiratory infection. 2017: 49(6): 1602098.
37. Xu G, Qi F, Li H, Yang Q, Wang H, Wang X, Liu X, Zhao J, Liao X, Liu Y, Liu L, Zhang S, Zhang Z. The differential immune responses to COVID-19 in peripheral and lung revealed by single-cell RNA sequencing. *Cell Discovery* 2020: 6(1): 73.
38. Wilk AJ, Rustagi A, Zhao NQ, Roque J, Martínez-Colón GJ, McKechnie JL, Ivison GT, Ranganath T, Vergara R, Hollis T, Simpson LJ, Grant P, Subramanian A, Rogers AJ,

Blish CA. A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nature Medicine* 2020: 26(7): 1070-1076.
39. Arunachalam PS, Wimmers F, Mok CKP, Perera RAPM, Scott M, Hagan T, Sigal N, Feng Y, Bristow L, Tak-Yin Tsang O, Wagh D, Coller J, Pellegrini KL, Kazmin D, Alaaeddine G, Leung WS, Chan JMC, Chik TSH, Choi CYC, Huerta C, Paine McCullough M, Lv H, Anderson E, Edupuganti S, Upadhyay AA, Bosinger SE, Maecker HT, Khatri P, Rouphael N, Peiris M, Pulendran B. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. 2020: 369(6508): 1210-1220.
40. Longmore DK, Miller JE, Bekkering S, Saner C, Mifsud E, Zhu Y, Saffery R, Nichol A, Colditz G, Short KR, Burgner DP. Diabetes and Overweight/Obesity Are Independent, Nonadditive Risk Factors for In-Hospital Severity of COVID-19: An International, Multicenter Retrospective Meta-analysis. 2021: dc202676.
41. Margaroli C, Benson P, Sharma NS, Madison MC, Robison SW, Arora N, Ton K, Liang Y, Zhang L, Patel RP, Gaggar A. Spatial mapping of SARS-CoV-2 and H1N1 lung injury identifies differential transcriptional signatures. *Cell reports Medicine* 2021: 2(4): 100242.
42. Grant RA, Morales-Nebreda L, Markov NS, Swaminathan S, Querrey M, Guzman ER, Abbott DA, Donnelly HK, Donayre A, Goldberg IA, Klug ZM, Borkowski N, Lu Z, Kihshen H, Politanska Y, Sichizya L, Kang M, Shilatifard A, Qi C, Lomasney JW, Argento AC, Kruser JM, Malsin ES, Pickens CO, Smith SB, Walter JM, Pawlowski AE, Schneider D, Nannapaneni P, Abdala-Valencia H, Bharat A, Gottardi CJ, Budinger GRS, Misharin AV, Singer BD, Wunderink RG, Grant RA, Morales-Nebreda L, Markov NS, Swaminathan S, Querrey M, Guzman ER, Abbott DA, Donnelly HK, Donayre A, Goldberg IA, Klug ZM, Borkowski N, Lu Z, Kihshen H, Politanska Y, Sichizya L, Kang M, Shilatifard A, Qi C, Lomasney JW, Argento AC, Kruser JM, Malsin ES, Pickens CO, Smith SB, Walter JM, Pawlowski AE, Schneider D, Nannapaneni P, Abdala-Valencia H, Bharat A, Gottardi CJ, Budinger GRS, Misharin AV, Singer BD, Wunderink RG, Wagh AA, Hauser AR, Wolfe AR, Thakrar A, Yeldandi AV, Wang AA, Levenson AR, Joudi AM, Tran B, Gao CA, Kurihara C, Schroedl CJ, Horvath CM, Meza D, Odell DD, Kamp DW, Winter DR, Ozer EA, Shanes ED, Bartom ET, Rendleman EJ, Leibenguth EM, Wehbe F, Liu GY, Gadhvi GT, Navarro HT, Sznajder JI, Dematte JE, Le J, Arnold JM, Du JC, Coleman J, Bailey JI, Deters JS, Fiala JA, Starren J, Ridge KM, Secunda K, Aren K, Gates KL, Todd K, Gradone LD, Textor LN, Wolfe LF, Pesce LL, Nunes Amaral LA, Rosenbaum ML, Kandpal M, Jain M, Sala MA, Saine M, Carns M, Alexander MJ, Cuttica MJ, Prickett MH, Khan NH, Chandel NS, Soulakis ND, Rivas OR, Seed PC, Reyfman PA, Go PD, Sporn PHS, Cooper PR, Tomic R, Patel R, Garza-Castillon R, Kalhan R, Morimoto RI, Mylvaganam RJ, Kim SS, Gatesy SWM, Thakkar S, Ben Maamar S, Han S, Rosenberg SR, Nozick S, Green SJ, Russell SR, Poor TA, Zak TJ, Lombardo TA, Stoeger T, Shamaly T, Ren Z, The NUSSI. Circuits between infected macrophages and T cells in SARS-CoV-2 pneumonia. *Nature* 2021: 590(7847): 635-641.
43. Desai N, Neyaz A, Szabolcs A, Shih AR, Chen JH, Thapar V, Nieman LT, Solovyov A, Mehta A, Lieb DJ, Kulkarni AS, Jaicks C, Xu KH, Raabe MJ, Pinto CJ, Juric D, Chebib I, Colvin RB, Kim AY, Monroe R, Warren SE, Danaher P, Reeves JW, Gong J, Rueckert EH, Greenbaum BD, Hacohen N, Lagana SM, Rivera MN, Sholl LM, Stone JR, Ting DT, Deshpande V. Temporal and spatial heterogeneity of host response to SARS-CoV-2 pulmonary infection. *Nature Communications* 2020: 11(1): 6319.
44. Butler D, Mozsary C, Meydan C, Foox J, Rosiene J, Shaiber A, Danko D, Afshinnekoo E, MacKay M, Sedlazeck FJ, Ivanov NA, Sierra M, Pohle D, Zietz M, Gisladottir U, Ramlall V, Sholle ET, Schenck EJ, Westover CD, Hassan C, Ryon K, Young B, Bhattacharya C, Ng DL, Granados AC, Santos YA, Servellita V, Federman S, Ruggiero P, Fungtammasan A, Chin C-S, Pearson NM, Langhorst BW, Tanner NA, Kim Y, Reeves JW,

Hether TD, Warren SE, Bailey M, Gawrys J, Meleshko D, Xu D, Couto-Rodriguez M, Nagy-Szakal D, Barrows J, Wells H, O'Hara NB, Rosenfeld JA, Chen Y, Steel PAD, Shemesh AJ, Xiang J, Thierry-Mieg J, Thierry-Mieg D, Iftner A, Bezdan D, Sanchez E, Campion TR, Sipley J, Cong L, Craney A, Velu P, Melnick AM, Shapira S, Hajirasouliha I, Borczuk A, Iftner T, Salvatore M, Loda M, Westblade LF, Cushing M, Wu S, Levy S, Chiu C, Schwartz RE, Tatonetti N, Rennert H, Imielinski M, Mason CE. Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nature Communications* 2021: 12(1): 1660.

Figure Legends

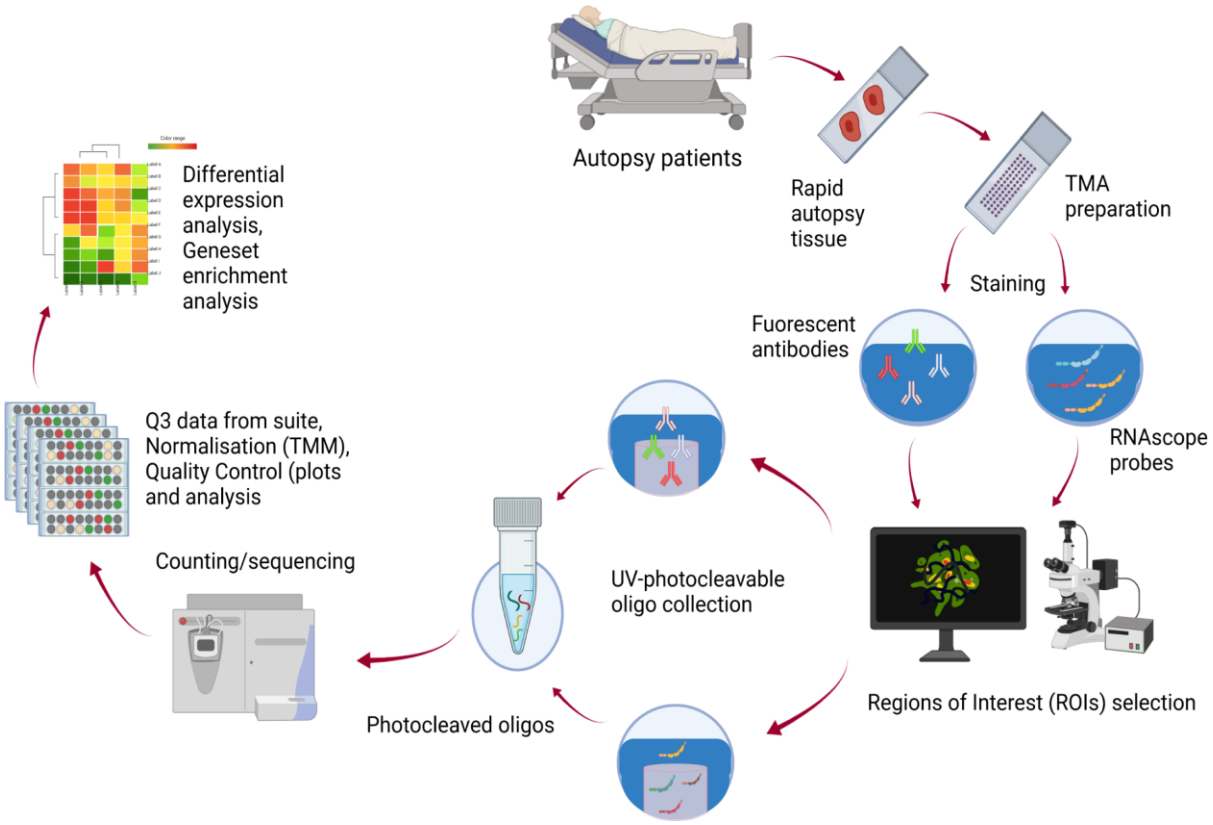
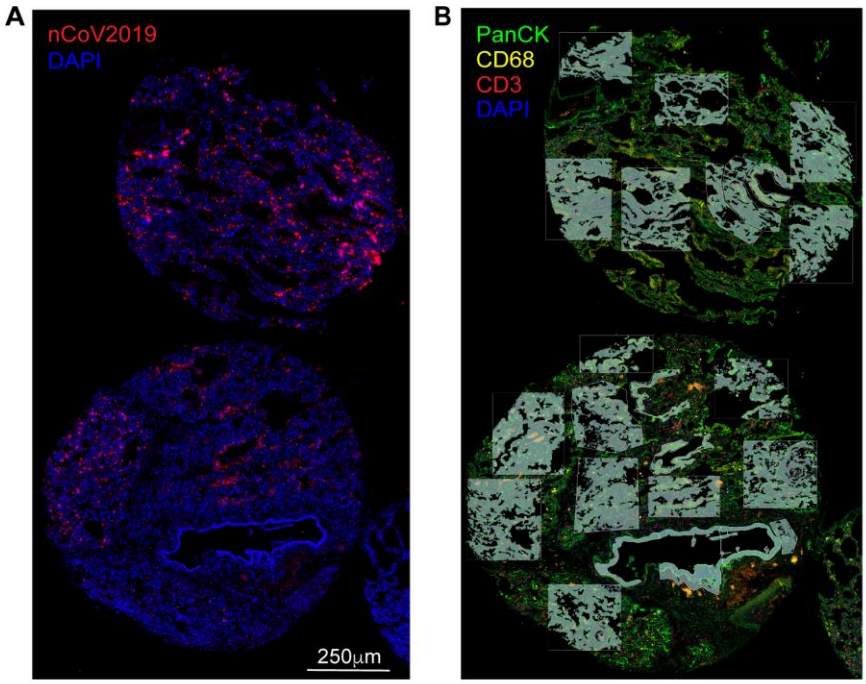
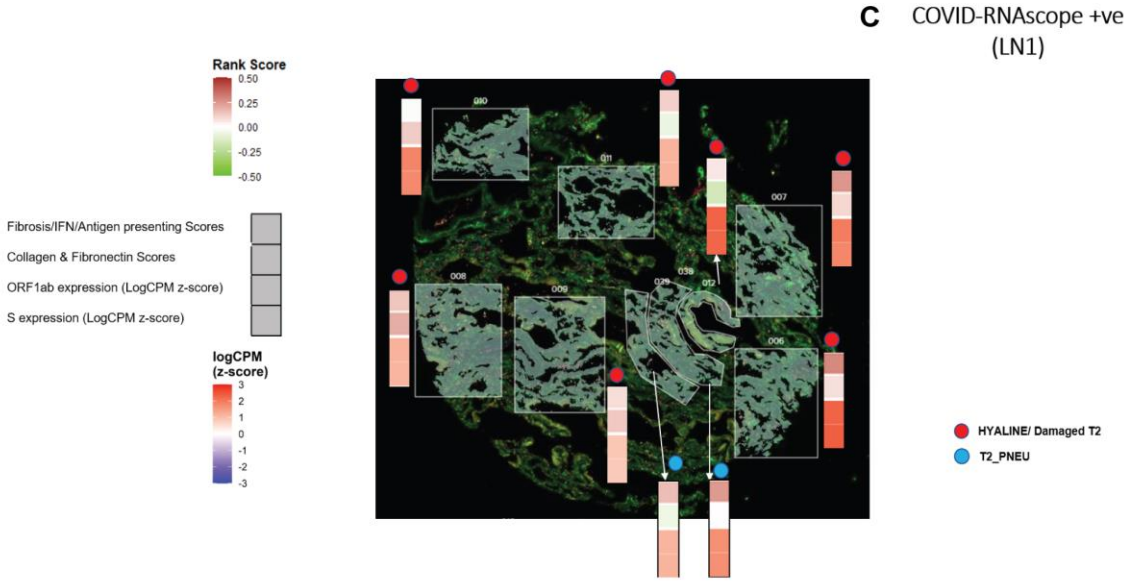


Figure 1. Schematic of the study


A
nCoV2019
DAPI
250μm
B
PanCK
CD68
CD3
DAPI


C
COVID-RNAscope +ve
(LN1)
Rank Score
0.50
0.25
0.00
-0.25
-0.50
Fibrosis/IFN/Antigen presenting Scores
Collagen & Fibronectin Scores
ORF1ab expression (LogCPM z-score)
S expression (LogCPM z-score)
logCPM
(z-score)
3
2
1
0
-1
-2
-3
HYALINE/ Damaged T2
T2_PNEU

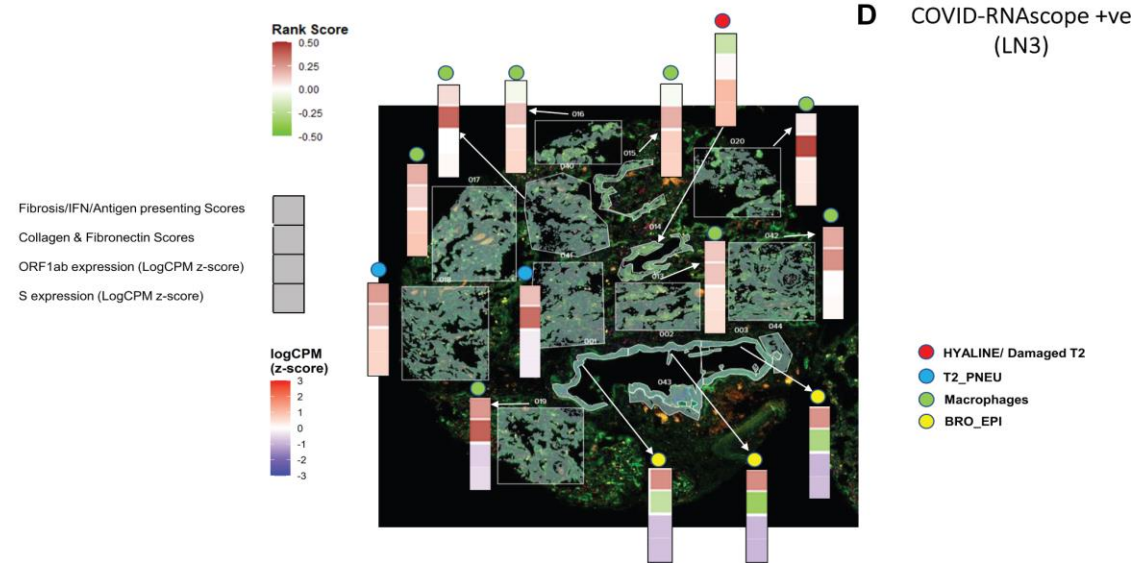

D
COVID-RNAscope +ve
(LN3)
Rank Score
0.50
0.25
0.00
-0.25
-0.50
Fibrosis/IFN/Antigen presenting Scores
Collagen & Fibronectin Scores
ORF1ab expression (LogCPM z-score)
S expression (LogCPM z-score)
logCPM
(z-score)
3
2
1
0
-1
-2
-3
HYALINE/ Damaged T2
T2_PNEU
Macrophages
BRO_EPI

Figure 2. RNA-FISH and digital spatial profiling morphology marker visualisation of nCoV2019 mRNA positive cores. ROI selection was guided by RNAscope-FISH staining for nCoV2019 spike mRNA. **(a)** RNA-FISH staining on COVID TMA showing the cores LN1 and LN3 which were highly positive for nCoV2019 spike mRNA (red). Nuclei are shown in blue (DAPI staining). Scale bar, 1000 µm. **(b)** Representative ROIs selected across LN1 and LN3 SARS-CoV-2 virus positive cores on immunohistological paraffin embedded sections stained for monocytes (CD68, yellow), T cells (CD3ε, red), lung epithelial cells (PanCK, green) and nuclei (DAPI, blue). Scale bar, 500 µm. **C.** High resolution image of high viral core LN1 where each region of interest has been determined for the majority cell type (Hyaline membrane/T2 pneumocyte). D. High resolution image of high viral core LN3 where each region of interest has been determined for the majority cell type (Hyaline membrane, T2 pneumocyte, macrophages, bronchiolar epithelial cells).

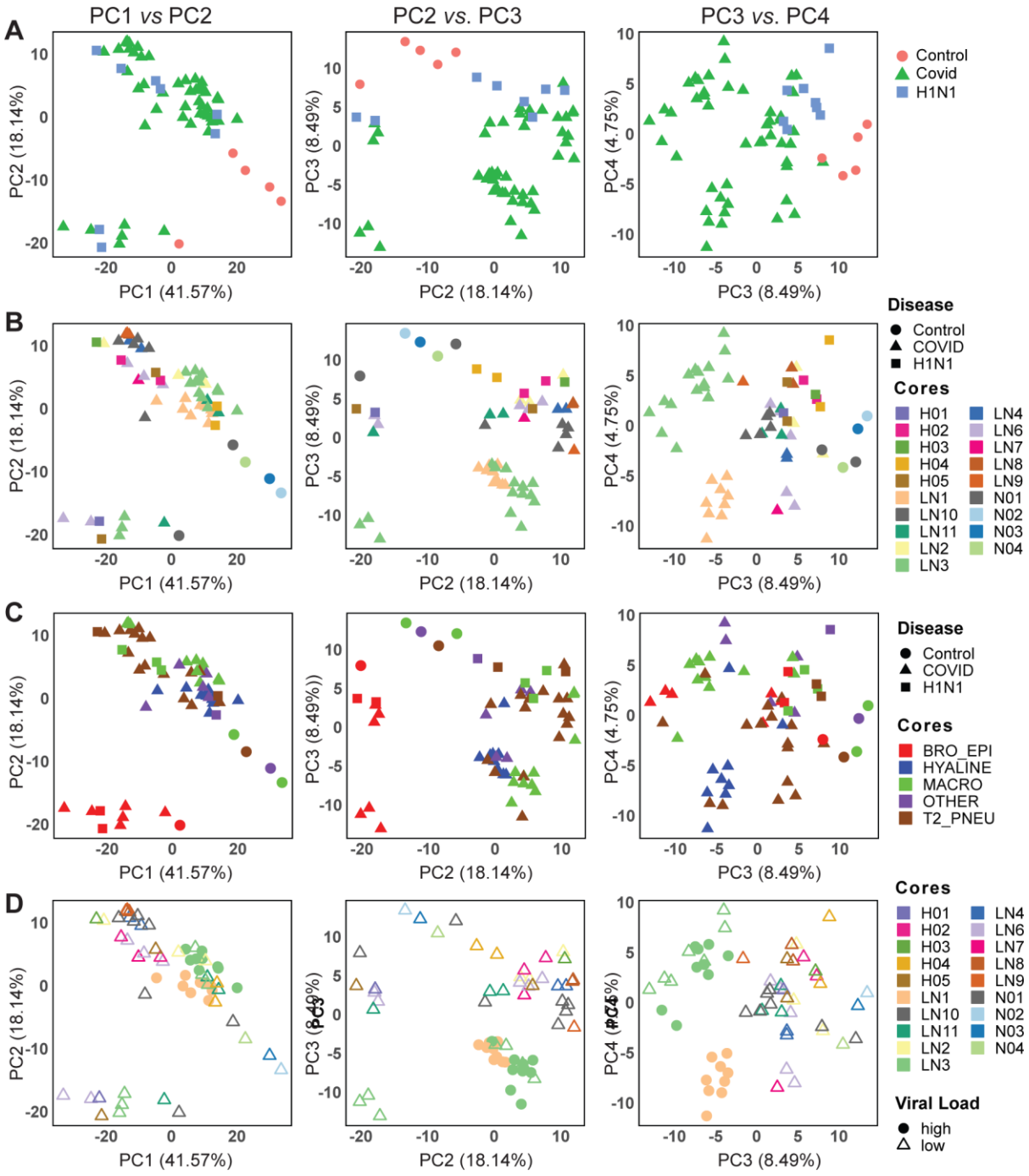


Figure 3. Principal components identify variability and factors in the transcriptomic data. Principal components (PCs) capturing orthogonal dimensions of variability in the transcriptomic data in descending order of contribution (i.e. PC1-PC2, PC3-PC2 and PC3-PC4) were plotted stratifying based on the following factors in the experimental data: (**a**) Disease types, (**b**) cores/patients with disease types, (**c**) dominant tissue types with disease types and (**d**) Viral Load in cores/patients. Disease type in panels **a**, **b** and **c** labelled as shapes with Control as circles, COVID-19 as triangles and H1N1 as squares. Viral load labelled in panel **b**

as shapes with high viral load as solid circle and low viral low as unfilled triangle. Cores and patients labelled in panel **b** and **c** using differing colors. Dominant tissue types in panel **c** labelled as differing fill colors (BRO_EPI, bronchiolar epithelium; HYALINE, hyaline membranes; MACRO, macrophages; T2_PNEU, Type2 pneumocytes; OTHERS, other classifications).

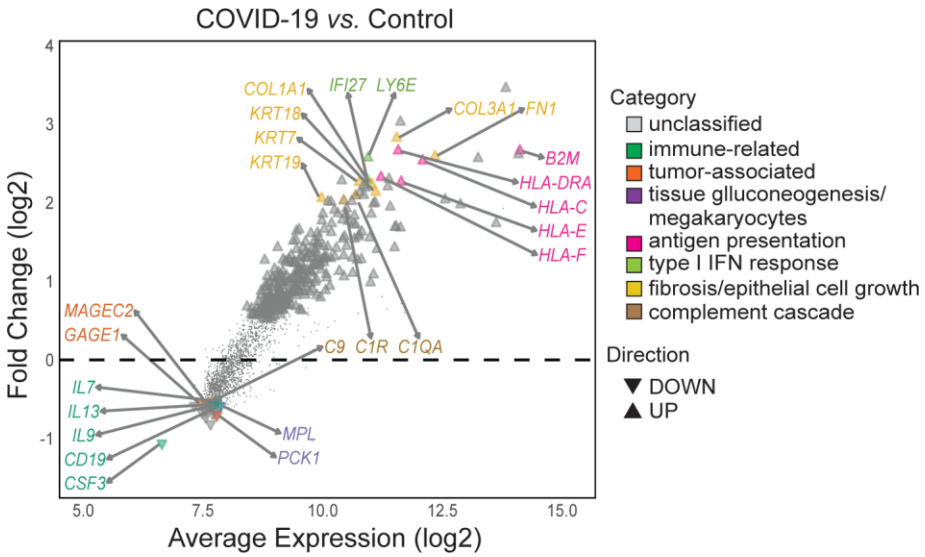


Figure 4. COVID-19 infection drives pro-inflammatory response and suppresses immune cell effector and regeneration. Distribution of differentially expressed genes as a function of the average transcript expression and fold change (log2) identified in COVID-19 samples *vs* uninfected (control) samples. Down-regulated genes included immune related, cytokine, tumor/cell-survival associated genes and cell regeneration genes, inferring a suppression of immune cell effect and regeneration. Up-regulated genes are enriched with pro-inflammatory genes including Type I interferon (IFN) response and fibrosis genes. Representative genes from these processes are highlighted. Differential expression genes generated using the *voom-limma* pipeline with *limma:duplicationCorrelations* and applying TREAT criteria with absolute log fold change > 1.2 with p-value <0.05.

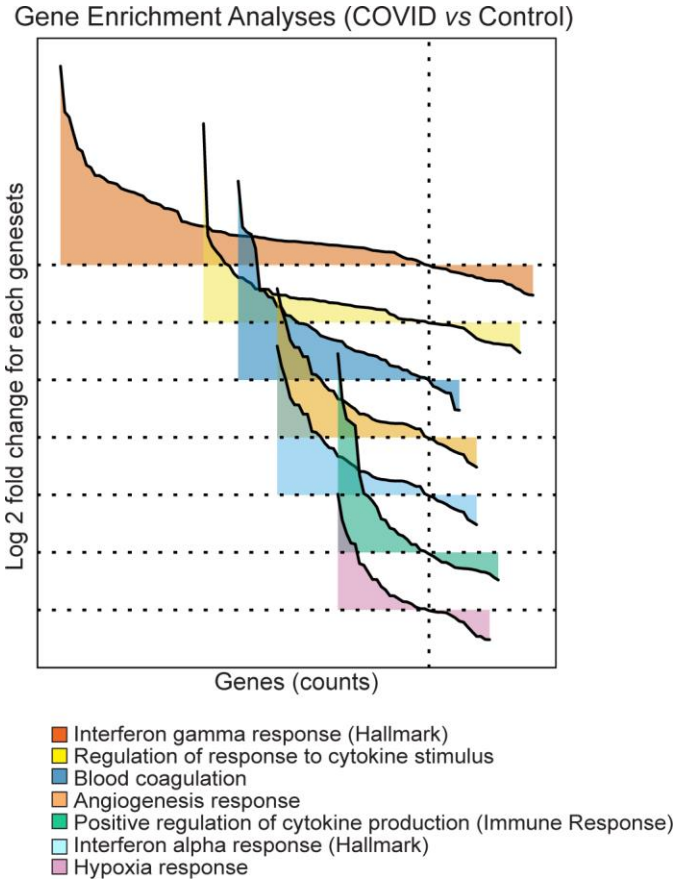


Figure 5. Gene set enrichment analyses (GSEA) reveals upregulated cytokine responses with accompanying coagulopathies in COVID-19 infection. GSEA for differentially expressed genes comparing COVID-19 and uninfected samples were conducted using the estimated log2 fold change and p-values with the distribution of significantly differentiated genesets visualised as a function of the log2 fold change (logFC) and genes sorted by the logFC (waterfall plot). The genesets visualised are custom genesets for 'angiogenesis response', 'blood coagulation', 'hypoxia' responses based on nanoString's nCounter® PanCancer Progression Panel were identified by GSEA to be differentially upregulated in COVID-19 samples. (see *Table* 2), MSigDB Hallmark gene set for IFN-α and IFN-γ and MSigDB Gene Ontology (GO) gene sets for 'Regulation of response to cytokine stimulus' and 'Positive regulation of cytokine production involved in immune response'. Genesets ordered by gene counts. Refer to comprehensive list in Table 1. GSEA conducted using *limma-fry* with FDR <0.05. The direction and relative change in expression are shown. Details of each geneset are provided in Supplementary Fig 4 and Supplementary Fig 5.

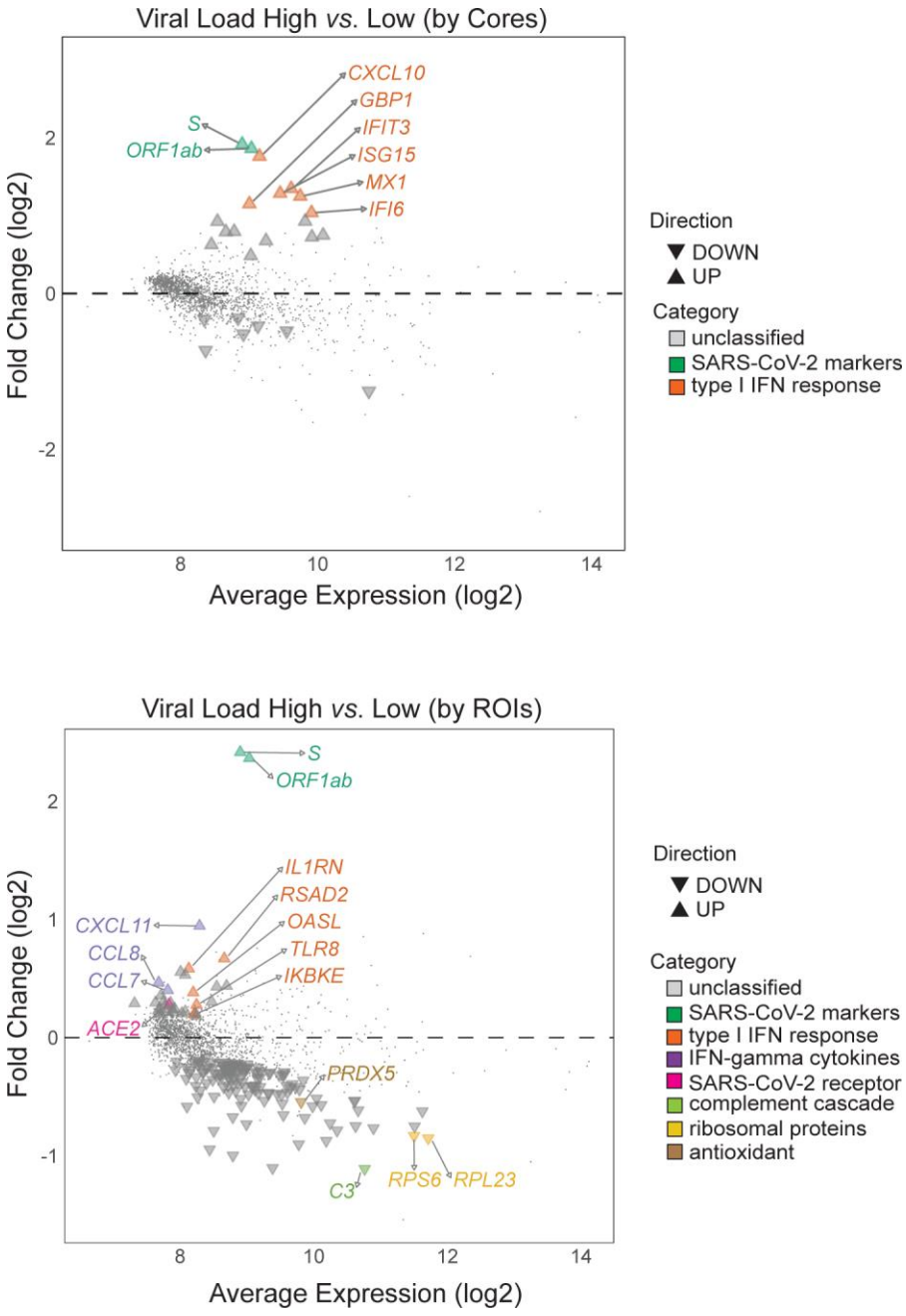


Figure 6. Resolution provided by spatial data reveals type I interferon gene signature associates with COVID-19 viral load. Distribution of differentially expressed genes as a function of the average transcript expression and fold change (log2) identified in high viral load vs low viral load COVID-19 samples. Analyses were conducted at two resolutions, firstly by (**top**) grouping ROIs from the same cores with the same degree of viral load (core patients-based approach); secondly by (**bottom**) treating each regions/ROIs sample as an independent observation (ROI-based approach). In both approaches, SAR-CoV-2 specific genes *S* and *ORF1ab* were strongly upregulated in the high viral load group, consistent with results of the RNAscope. Consistent with previous reports, type I IFN response genes were associated with

SAR-CoV-2 RNA expression. Notably the core-based approach reveals limited differential gene expression with the finer resolution ROI-based approach revealing additional differential changes in complement cascade, ribosomal protein and antioxidants genes. Representative genes are highlighted. Differential expression genes derived using *voom-limma* pipeline with *limma:duplicationCorrelations* and applying absolute log2 fold change > 1.0 with p-value <0.05.

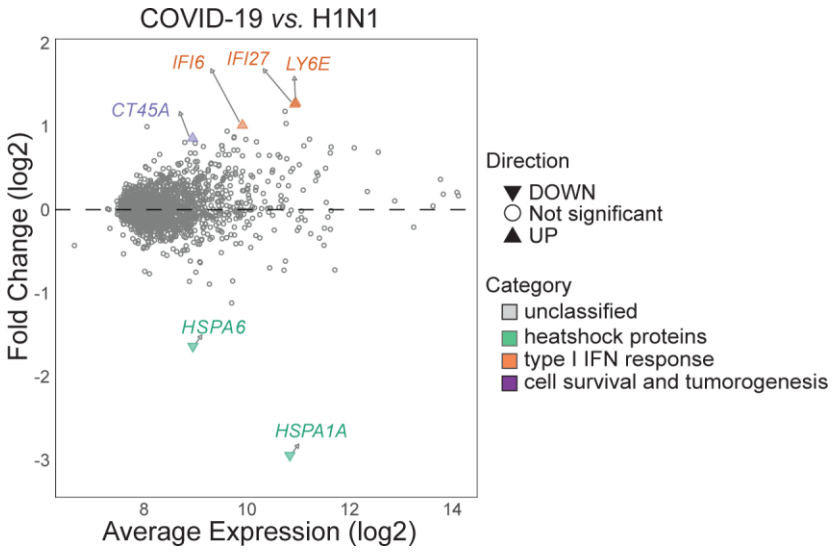


Figure 7. Limited transcriptomic differences associated with COVID-19 infection compared with pH1N1. Distribution of differentially expressed genes as a function of the average transcript expression and fold change (log2) identified in COVID-19 samples vs pH1N1 samples were visualised with the 6 differentially expressed genes labelled and classified based on associated biological processes. The genes associated with the type I Interferon response, heat shock protein family members and (associated) with cell survival and tumorigenesis are shown. This exclusive list of genes reveals the subtle differences between COVID-19 and pH1N1 infected transcriptome and presents a potential disease specific transcriptomic signature for distinguishing the two disease types. Differential expression genes derived using *voom-limma* pipeline with *limma:duplicationCorrelations* and applying *TREAT* criteria with absolute log fold change > 1.2 with p-value <0.05.

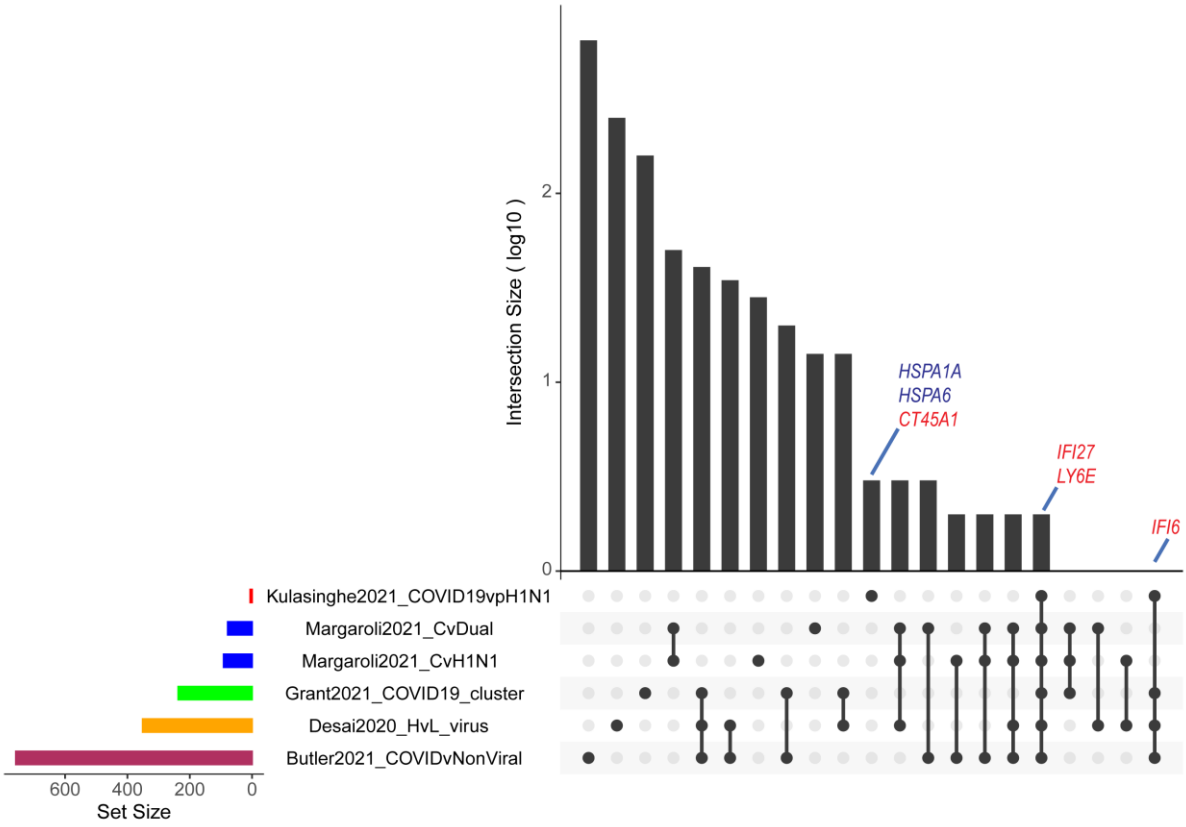


Figure 8. Comparison of COVID specific gene sets identified across multiple studies. Upset plot describing the overlaps between gene sets across 5 different studies. The size of the gene sets varies across the studies with COVID specific gene set identified in this study (COVID19 vs pH1N1) having the smallest number of genes (6). Notably *IFI27* and *LY6E* are common across all the studies while IFI6 is common amongst 4 out of the 5 studies. Note: red font: upregulation, blue font: downregulation in this study. Studies compared include Margaroli *et al* 2021 [41], Grant *et al* 2021 [42], Desai *et al* 2020 [43] and Butler *et al* 2021 [44].

Supplementary Figure 1. **Representative assessment of virus load in DSP ROIs using RNAscope and IHC for SARS-CoV-2 spike mRNA and protein, respectively, on serial sections.** Fluorescent IHC staining was performed for DSP acquisition using monocytes (CD68, yellow), T cells (CD3ε, red), lung epithelial cells (PanCK, green) and nuclei (DAPI, blue), (left panels, **A**, **D**, **G**, **J**). RNAscope staining was performed as per manufacturer's instructions using nCoV2019-s probe for SARS-CoV-2 spike mRNA (center panels, **B**, **E**, **H**, **K**). IHC was performed with antibody against SARS-CoV-2 spike protein (right panels, **C**, **F**, **I**, **L**). ROIs were semi-quantitatively scored from 0 to 3 based on RNAscope signal abundance. **A-C**. Representative ROI with RNAscope score of 3 (high). **D-F**. Representative ROI with RNAscope score of 2 (intermediate). **G-I**. Representative ROI with RNAscope score of 1 (low). **J-L**. Representative ROI with RNAscope score of 0 (below detection). Scale bar = 100μm.

Supplementary Figure 2. **Representative assessment of dominant tissue histopathology in DSP ROIs**. TMAs were stained for DSP acquisition using monocytes (CD68, yellow), T cells (CD3ε, red), lung epithelial cells (PanCK, green) and nuclei (DAPI, blue)(left panels). H&E (middle panels) and Masson's trichrome (right panels) histological staining was performed on serial sections and ROIs overlaid for assessment by a pathologist. **A-C**. ROIs characterised by alveolar type 2 pneumocytes. **D-F**. ROIs characterised by hyaline membranes. **G-I**. ROIs characterised by macrophage infiltration. **J-L**. ROIs specifically drawn around bronchiolar epithelium lining greater airways. Scale bar = 100μm.

Supplementary Figure 3. TMM normalisation of the data removed systematic bias in the uninfected samples.

Relative log expression (RLE) plots for the Q3-normalised counts identified a systematic higher gene expression in the (**top**) *uninfected control* samples. The Q3-data was normalised using the (**bottom**) trimmed mean of M-values (TMM) method [15] using the all the genes in the gene panel. The normalisation removed the systematic bias in the regions corresponding to cores from control (uninfected) patients, allowing for cross-region analyses between the different disease types. TMM normalisation conducted using the *calcNormFactors* function in *edgeR*.

Supplementary Figure 4. Gene Set Enrichment Analysis (GSEA) for COVID-19 vs control (Custom Gene Sets). The waterfall plot shows the direction and relative change in gene expression profiles of custom gene sets for 'angiogenesis response', 'blood coagulation', 'hypoxia' responses based on nanoString's nCounter® PanCancer Progression Panel.

Supplementary Figure 5. Gene Set Enrichment Analysis (GSEA) for COVID-19 vs control (MsigDB). The waterfall plot shows the direction and relative change in gene expression profiles of relevant gene sets identified via GSEA from Molecular Signature Database (MsigDB) collections Hallmark and Gene Ontology.

Supplementary Figure 6. Mean-difference plot for pH1N1 vs Control.
Distribution of differentially expressed genes as a function of the average transcript expression and fold change (log2) identified in pH1N1 samples vs uninfected (control) samples. Representative genes from relevant processes are highlighted. Categories and processes visualised as per corresponding figure for COVID vs Control in Figure 4. Differential expression genes generated using the *voom-limma* pipeline with *limma:duplicationCorrelations* and applying TREAT criteria with absolute log fold change > 1.2 with p-value <0.05.

Supplementary Figure 7. Gene Set Enrichment Analysis (GSEA) for pH1n1 vs control. GSEA for differentially expressed genes comparing pH1N1 and uninfected samples were conducted using the estimated log2 fold change and p-values with the distribution of significantly differentiated genesets visualised as a function of the log2 fold change (logFC) and genes sorted by the logFC (waterfall plot). The genesets visualised are custom genesets for **(A)** 'angiogenesis response', **(B)** 'blood coagulation', **(G)** 'hypoxia' responses based on nanoString's nCounter® PanCancer Progression Panel were identified by GSEA to be differentially upregulated in COVID-19 samples, MSigDB Hallmark gene set for **(E)** IFN-α and **(F)** IFN-γ and MSigDB Gene Ontology (GO) gene sets for **(C)** 'Positive regulation of cytokine production involved in immune response' and **(D)** 'Regulation of response to cytokine stimulus'. Genesets ordered by gene counts. GSEA conducted using *limma-fry* with FDR <0.05. The direction and relative change in expression are shown. Details of each geneset are provided in Supplementary Fig 4 *and* Supplementary *Fig 5*.

Table 1. Geneset Enrichment Analysis for SARS-CoV-2 vs uninfected DE genes using Molecular Signatures Database (MSigDB) Hallmark gene sets.

Geneset Name [a,b]	N Genes [c]	Direction [d]	two-sided directional		non-directional	
			p-value [e]	FDR [f]	p-value [e]	FDR [f]
HALLMARK_REACTIVE_OXYGEN_SPECIES_PATHWAY	13	Up	6.92	2.74	2.61	6.92
HALLMARK_COMPLEMENT	76	Up	6.60	23.44	21.74	6.60
HALLMARK_KRAS_SIGNALING_DN	40	Down	6.60	11.03	10.67	6.60
HALLMARK_IL6_JAK_STAT3_SIGNALING	73	Up	6.46	19.71	18.62	6.46
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	16	Up	6.37	2.47	2.35	6.37
HALLMARK_APOPTOSIS	79	Up	6.37	17.81	17.02	6.37
HALLMARK_OXIDATIVE_PHOSPHORYLATION	33	Up	5.25	5.28	5.05	5.25
HALLMARK_P53_PATHWAY	56	Up	4.98	16.44	15.69	4.98
HALLMARK_MTORC1_SIGNALING	47	Up	4.84	12.27	11.85	4.84
HALLMARK_INTERFERON_GAMMA_RESPONSE	110	Up	4.84	13.77	13.25	4.84
HALLMARK_PROTEIN_SECRETION	10	Up	4.78	1.14	1.11	4.78
HALLMARK_GLYCOLYSIS	48	Up	4.75	14.00	13.44	4.75
HALLMARK_PANCREAS_BETA_CELLS	5	Down	4.59	0.09	0.09	4.59
HALLMARK_HYPOXIA	60	Up	4.55	18.84	17.84	4.55
HALLMARK_MYC_TARGETS_V2	8	Up	4.49	0.25	0.24	4.49
HALLMARK_CHOLESTEROL_HOMEOSTASIS	15	Up	4.33	1.83	1.77	4.33
HALLMARK_INTERFERON_ALPHA_RESPONSE	52	Up	4.33	8.74	8.41	4.33
HALLMARK_SPERMATOGENESIS	19	Down	3.68	1.59	1.54	3.68
HALLMARK_MYC_TARGETS_V1	28	Up	3.60	4.40	4.22	3.60
HALLMARK_KRAS_SIGNALING_UP	77	Up	3.38	16.13	15.43	3.38
HALLMARK_ADIPOGENESIS	30	Up	3.30	4.88	4.67	3.30
HALLMARK_COAGULATION	56	Up	3.05	15.14	14.49	3.05
HALLMARK_XENOBIOTIC_METABOLISM	36	Up	2.98	11.63	11.25	2.98
HALLMARK_ANDROGEN_RESPONSE	16	Up	2.86	2.24	2.14	2.86
HALLMARK_FATTY_ACID_METABOLISM	20	Up	2.79	4.63	4.44	2.79
HALLMARK_APICAL_JUNCTION	61	Up	2.69	18.58	17.70	2.69

HALLMARK_UV_RESPONSE_UP	46	Up	2.68	12.07	11.67	2.68
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	81	Up	2.64	14.94	14.32	2.64
HALLMARK_ESTROGEN_RESPONSE_EARLY	40	Up	2.60	8.46	8.16	2.60
HALLMARK_TGF_BETA_SIGNALING	20	Up	2.53	3.35	3.20	2.53
HALLMARK_ESTROGEN_RESPONSE_LATE	51	Up	2.51	12.47	12.01	2.51
HALLMARK_DNA_REPAIR	27	Up	1.99	5.96	5.71	1.99
HALLMARK_MITOTIC_SPINDLE	27	Up	1.94	3.12	2.98	1.94
HALLMARK_IL2_STAT5_SIGNALING	85	Up	1.81	14.68	14.10	1.81
HALLMARK_ANGIOGENESIS	19	Up	1.81	2.23	2.13	1.81
HALLMARK_UV_RESPONSE_DN	44	Up	1.78	8.73	8.41	1.78
HALLMARK_HEDGEHOG_SIGNALING	12	Up	1.75	0.98	0.96	1.75
HALLMARK_G2M_CHECKPOINT	67	Up	1.70	12.43	11.99	1.70
HALLMARK_APICAL_SURFACE	13	Up	1.68	1.32	1.28	1.68
HALLMARK_TNFA_SIGNALING_VIA_NFKB	107	Up	1.57	13.21	12.71	1.57
HALLMARK_INFLAMMATORY_RESPONSE	115	Up	1.49	20.51	19.29	1.49
HALLMARK_E2F_TARGETS	61	Up	1.46	8.04	7.76	1.46
HALLMARK_NOTCH_SIGNALING	21	Up	1.26	3.71	3.54	1.26
HALLMARK_MYOGENESIS	35	Up	1.11	7.37	7.11	1.11
HALLMARK_PI3K_AKT_MTOR_SIGNALING	56	Up	1.11	18.55	17.70	1.11
HALLMARK_PEROXISOME	17	Up	0.80	2.12	2.04	0.80
HALLMARK_ALLOGRAFT_REJECTION	141	Up	0.42	22.32	20.92	0.42
HALLMARK_WNT_BETA_CATENIN_SIGNALING	27	Up	0.38	5.58	5.34	0.38
HALLMARK_BILE_ACID_METABOLISM	14	Down	0.25	1.79	1.73	0.25
HALLMARK_HEME_METABOLISM	19	Up	0.04	2.19	2.10	0.04

[a] Geneset signatures in green are enriched in SARS-CoV-2 compared to control (i.e. Direction is Up) while red is the opposite direction (i.e. Down). [b] Genesets highlighted in yellow indicates biological processed upregulated in SARS-CoV-2 infected lungs. List ordered by descending FDR. [c] N-Genes, number of genes in the geneset. [d] Direction, direction of change (i.e. up or down). [e] p-value, statistical significance of the difference. [f] FDR, false discovery rate. All p-value and FDR reported as -log10 value.

Table 2. Geneset Enrichment Analysis for SARS-CoV-2 vs uninfected DE genes using nanoString's nCounter® PanCancer Progression Panel custom gene sets.

Geneset Name[a]	N Genes[b]	Direction[c]	two-sided directional		non-directional	
			p-value[d]	FDR[e]	p-value[d]	FDR[e]
Blood Coagulation	47	Up	4.30	**3.56**	11.62	11.06
Angiogenesis Response	56	Up	2.30	**1.74**	22.35	21.31
Hypoxia Response	36	Up	1.89	**1.43**	7.35	6.97
Positive Regulation of Angiogenesis	35	Down	0.98	0.79	8.38	7.94
Negative Regulation of Angiogenesis	17	Up	0.51	0.42	1.62	1.35
Fibrosis	9	Up	0.27	0.23	1.05	0.96
Regulation of Angiogenesis	11	Down	0.08	0.08	1.12	0.96

[a] Genesets highlighted in green indicate biological processes identified by GSEA to be differentially upregulated in SARS-CoV-2 infected lungs versus uninfected samples with FDR < 0.05. [b] N-Genes, number of genes in the geneset. [c] Direction, direction of change (i.e. up or down). [d] *p-v*alue, statistical significance of the difference. [e] FDR, false discovery rate. All p-value and FDR reported as -log10 value.

Table 3. Genes differentially expressed in the lungs of SARS-CoV-2 patients with high viral load compared with patients with low viral load (patient-based analysis).

Gene name [a]	Fold Change (log2)	Average Expression	p-value [b]	Adjusted p-value [c]
S	1.916	8.900	4.54	2.01
ORF1ab	1.861	9.036	4.41	2.01
CXCL10	1.761	9.153	5.03	2.24
IFIT3	1.346	9.616	4.85	2.19
ISG15	1.286	9.454	4.09	1.78
MX1	1.249	9.748	5.45	2.48
GBP1	1.151	9.005	8.88	5.61
IFI6	1.037	9.913	3.45	1.43
STAT1	0.927	9.820	4.37	2.01
IDO1	0.925	8.539	3.73	1.58
OAS3	0.795	8.783	4.45	2.01
RSAD2	0.792	8.662	3.28	1.36
TAP1	0.748	10.082	3.34	1.39
PSMB9	0.725	9.919	3.78	1.59
GBP4	0.677	9.248	3.84	1.61
HERC6	0.627	8.451	3.20	1.31
IFI35	0.485	9.026	3.34	1.39
NR3C1	-0.313	8.839	3.44	1.43
ATF2	-0.330	8.348	3.52	1.46
SERINC3	-0.415	9.130	3.24	1.33
JAK1	-0.479	9.550	3.86	1.61
FYN	-0.518	8.918	3.60	1.51
C4BPA	-0.728	8.366	3.61	1.51
SERPINA1	-1.251	10.750	3.41	1.42

[a] Up-regulated genes in high viral load compared to low vial load SARS-CoV-2 lung samples are shown in green, downregulated genes shown in red. **[b]** *p*-value, statistical significance of the difference. **[c]** *adjusted p-v*alue, *the smallest familywise significance level for multiple comparison testing. All p-value and adjusted p-value reported as -log10 value.*

Supplementary Data

Spatial Profiling of Lung SARS-CoV-2 and Influenza Virus Infection Dissects Virus-Specific Host Responses and Gene Signatures

Arutha Kulasinghe[1,2*], Chin Wee Tan[3,4*], Anna Flavia Ribeiro dos Santos Miggiolaro[5*], James Monkman[1,2], Dharmesh Bhuva[3,4], Jarbas da Silva Motta Junior[5], Caroline Busatta Vaz de Paula[5], Seigo Nagashima[5], Cristina Pellegrino Baena[6], Paulo Souza-Fonseca-Guimaraes[4], Lucia de Noronha[7], Tim McCulloch[2,8], Gustavo Rodrigues Rossi[2,8], Caroline Cooper[9,10], Benjamin Tang[11], Kirsty R. Short[12,13#], Melissa J Davis[3,4,14#], Fernando Souza-Fonseca-Guimaraes[2,8#], Gabrielle T. Belz[2,3,8,13#], Ken O'Byrne[1,2#]

[1]*Queensland University of Technology, School of Biomedical Sciences, Faculty of Health, Brisbane, Queensland 4000, Australia.*
[2]*Translational Research Institute, Brisbane, Queensland 4102, Australia.*
[3]*The Walter and Eliza Hall Institute of Medical Research, Parkville, Melbourne, Victoria 3052, Australia.*
[4]*Department of Medical Biology, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, VIC 3010, Australia.*
[5]*Postgraduate Program of Health Sciences - School of Medicine - Hospital Marcelino Champagnat - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, PR, Brazil.*
[6]*School of Medicine & Center of Education, Research and Innovation – Hospital Marcelino Champagnat - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Brazil.*
[7]*Laboratory of Experimental Pathology - School of Medicine - Pontifícia Universidade Católica do Paraná (PUCPR), Curitiba, Brazil.*
[8]*University of Queensland Diamantina Institute, University of Queensland, Woollongabba, Queensland, 4102 Australia.*
[9]*Pathology Queensland, Princess Alexandra Hospital, Woolloongabba, Queensland, Australia.*
[10]*University of Queensland, Faculty of Medicine, Woolloongabba, Queensland, Australia.*
[11]*Centre for Immunology and Allergy Research, the Westmead Institute for Medical Research, Sydney, NSW, Australia.*
[12]*The University of Queensland, School of Chemistry and Molecular Biosciences, St Lucia, Brisbane, Queensland 4072, Australia.*
[13]*Australian Infectious Diseases Research Centre, The University of Queensland, Brisbane, Queensland, Australia.*
[14]*Department of Clinical Pathology, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, VIC, Australia.*

*co-first authors
#co-senior authors

Corresponding Author: Dr Arutha Kulasinghe, Email: Arutha.kulasinghe@qut.edu.au

Supplementary Table 1. Patient Case History Parameters

Case	Sex	Age	Mechanical ventilation (Days)	Admission to death (days)	Co-morbidities	Existing medication	COVID treatment
LN1	F	87	8	8	Systemic Arterial Hypertension, Dyslipidemia, Hypothyroidism, Senile dementia	Losartan, Rivaroxabana, Acetylsalicylic Acid, Levotiroxina, Ferrous Sulphate, Quetiapine, Galantamine Rosuvastatin, Zolpidem	Hydroxychloroquine, Azithromycin, Oseltamivir, Piperacillin, Tazobactam
LN2	M	53	8	13	Class II Obesity	None	Hydroxychloroquine, Azithromycin, Oseltamivir, Ceftriaxone
LN3	F	85	0	23	Systemic Arterial Hypertension, Brain Stroke, Vascular Dementia	Losartana, Amlodipine, Acetylsalicylic Acid, Clopidogrel	Piperacilin, tazobactam
LN4	M	73	10	38	Type 2 Diabetes Mellitus, Chronic Kidney Disease Dialysis, Atrial Fibrillation, Coronary Disease, Heart Failure, Peripheral Obstructive Artery Disease	Insulin NPH, Losartana, Hydrochlorotiazide, Metoprolol, Acetylsalicylic Acid, Clopidogrel, Rosuvastatin, Cilostazol, Eritropoyetina	Hydroxychloroquine, Azithromycin, Oseltamivir, Metronidazol, Meropenen, Linezolida
LN6	M	80	21	23	Systemic Arterial Hypertension, Coronary Disease, Heart Failure, Class III obesity	Acetylsalicylic Acid, Metoprolol, Rivaroxabana, Ezetimibe, Pitavastatina, Trimetazidima, Carbamazepine, Trazodona, Inhaled Beclomethasone, Inhaled Formoterol	Hydroxychloroquine, Azithromycin, Oseltamivir, Ceftriaxone
LN7	M	81	8	8	Systemic Arterial Hypertension, Chronic Kidney Disease Dialysis, Brain Stroke	Acetylsalicylic Acid, Clopidogrel, Rosuvastatin, Allopurinol, Clobazam	Hydroxychloroquine, Azithromycin, Oseltamivir, Meropenen, Linezolida
LN8	F	70	14	20	Type 2 Diabetes Mellitus Systemic Arterial Hypertension Liver transplant (at 2013) – viral hepatitis Bladder cancer	Ciclosporin, Mycophenolate Mofetil, Metformin, Atenolol, Omeprazole	Azithromycin, Oseltamivir, Enoxaparin, Prednisone. Piperacillin + Tazobactam
LN9	M	86	3	6	Prostate cancer, Abdominal Aortic Aneurysm, Giant Cells Arteritis	Acetylsalicylic Acid, Prednisone	Hydroxychloroquine, Azithromycin, Oseltamivir, Meropenen

LN10	M	46	5	8	Dyslipidemia	Sinvastatin	Azithromycin, Dexamethasone
LN11	F	93	6	6	Type 2 Diabetes Mellitus, Systemic Arterial Hypertension, Dyslipidemia, Senile dementia	Atenolol, Acetylsalicylic Acid, Sertraline, Rosuvastatin	Azithromycin, Oseltamivir, Meropenen, Enoxaparin, Dexamethasone

Supplementary Table 2. Patient Clinical Laboratory Parameters

Case	CRP (mg/L)	D-Dimer (µg/mL)	Leukocytes	Neutrophils	Band Cells	Lymphocytes	Hemo-globin (g/dL)	Platelets	CRP	D-Dimer (µg/mL	Leukocytes	Neutrophils	Band Cells	Lymphocytes	Hemoglobin (g/dL)	Platelets
	Ref: <5mg/L	Ref: <500µg/mL	Ref: 3600 to 12000/mm³	Ref: 1440 to 9600/ mm³	Ref: 0 to 600/ mm³	Ref: 720 to 5040/ mm³	Ref: 12.5-17 g/dL	Ref: 150000 to 450000 / mm³	Ref: <5mg/L	Ref: <500µg/ml	Ref: 3600 to 12000/mm³	Ref: 1440 to 900/mm³	Ref: 0 to 600/ mm³	Ref: 720 to 5040/ mm³	Ref:12.5-17 g/dL	Ref: 150000 to 450000/ mm³
	Initial tests								Final tests							
LN1	313	2014	15100	13892	1208	604	8.7	366000	201	NA	15000	12750	2400	1200	9.5	239000
LN2	146	425	11000	9790	550	990	13.5	325000	133	7394	50200	40160	3012	3514	7.2	318000
LN3	393	NA	19000	7623	396	570	9.9	282000	139	NA	9900	8118	693	891	10	481000
LN4	83	3436	9200	7176	1932	552	8.6	38000	270	NA	22000	4268	2420	440	8	356000
LN6	52	816	4700	3807	188	423	12.8	112000	407	4507	9400	7802	1034	1316	9.7	142000
LN7	301	13662	11800	11210	1770	354	10.7	252000	291	NA	24900	23406	1743	498	9.1	175000
LN8	16	4160	7380	5850	N/A	820	12.5	218000	16	1129	15880	11910	N/A	1260	7.6	266000
LN9	105	11184	8300	5133	708	249	11.5	110000	307	13535	8900	7921	445	712	9	99000
LN10	155	594	4500	3285	315	1035	15.2	138000	27.5	765	19500	17355	1365	975	10.2	175000
LN11	107	6544	7100	6319	355	639	9.5	111000	330	N/A	15600	14820	1092	780	8.1	76000

Supplementary Table 3. H1N1 and Control Patient Case History Parameters

Group	Case	Cause of death	Sex	Age	Mechanical ventilation (Days)	Days from admission to death	Comorbidities
H1N1	99-6999	H1N1	M	51	19	19	Haemophilus influenza
H1N1	99-7000	H1N1	M	57	12	12	Haemophilus influenza
H1N1	99-7223	H1N1	M	31	1	1	Mycoplasma pneumoniae
H1N1	99-7888	H1N1	M	37	1	1	Haemophilus influenza
H1N1	99-8157	H1N1	M	53	3	3	None
Control	N07-09	Neuroendocrine CA	M	45			
Control	N08-21	Hepatic CA	M	60			
Control	N09-19	Surgery	M	22			
Control	N09-57	Lymphoma	M	18			

Supplementary Table 4. Histopathological Parameters

Group	Case	Fibrosis	Histology
COVID	LN1	Anatomical fibrosis	Acute phase DAD
COVID	LN2	Anatomical fibrosis	Acute phase DAD
COVID	LN3	Anatomical fibrosis	Acute phase DAD
COVID	LN4	Anatomical fibrosis	Reactive Type 2 pneumocyte hyperplasia
COVID	LN6	Anatomical fibrosis	Reactive Type 2 pneumocyte hyperplasia
COVID	LN7	Anatomical fibrosis	Reactive Type 2 pneumocyte hyperplasia
COVID	LN8	Alveolar, moderate interstitial fibrosis	Organising pneumonia
COVID	LN9	Anatomical fibrosis	Organising pneumonia
COVID	LN10	Anatomical fibrosis	Acute phase DAD
COVID	LN11	Anatomical fibrosis	Acute phase DAD
H1N1	99-6999	Anatomical fibrosis	Alveolar haemorrhage
H1N1	99-7000	Anatomical fibrosis	Alveolar haemorrhage
H1N1	99-7223	Anatomical fibrosis	Type 2 pneumocyte hyperplasia
H1N1	99-7888	Mild interstitial fibrosis	Acute phase DAD
H1N1	99-8157	Mild interstitial fibrosis	Late organising phase DAD
Control	N07-09	Anatomical fibrosis	Nil
Control	N08-21	Anatomical fibrosis	Carcinoma primary or secondary
Control	N09-19	Anatomical fibrosis	Pulmonary oedema
Control	N09-57	Anatomical fibrosis	Pulmonary oedema

* Histology of FFPE cores was assessed by pathologist for levels of fibrosis, organisation and tissue injury. DAD; Diffuse Alveolar Damage

Supplementary Table 5. Histopathological Scoring of ROIs

ROI#	Group	Case	Dominant Tissue Type	H&E	H&E	H&E	H&E	H&E	H&E	Spike IHC	Spike IHC	Spike IHC	Spike IHC	RNAscope Score
				Alveolar Haemorrhage	Hyaline Membranes	Type 2 Pneumocyte Hyperplasia	Capillary Congestion	Fibroblastic Foci	Interstitial Inflammation	Bronchiolar Epithelium	Type 2 Pneumocytes	Interstitial Lymphocytes	Alveolar Macrophages	
10	COVID	LN1	Hyaline membrane	1	3	2	2	0	0	0	1	0	0	1
11	COVID	LN1	Hyaline membrane	0	3	3	3	0	0	0	2	0	3	1
12	COVID	LN1	Hyaline membrane	1	3	3	3	0	0	0	2	0	0	1
6	COVID	LN1	Hyaline membranes	0	3	3	3	0	0	0	2	0	3	3
7	COVID	LN1	Hyaline membranes, type 2 pneumocytes	0	3	3	2	0	0	0	2	1	3	3
8	COVID	LN1	Hyaline membranes, type 2 pneumocytes	0	3	3	2	0	0	0	2	0	3	2
9	COVID	LN1	Hyaline membranes, type 2 pneumocytes	1	3	2	2	0	0	0	0	0	0	1
38	COVID	LN1	Type 2 pneumocytes	1	0	3	3	0	0	0	2	0	3	2
39	COVID	LN1	Type 2 pneumocytes	1	3	2	2	0	0	0	2	0	0	1
22	COVID	LN2	Lymphocytes	0	0	0	0	0	3	0	0	2	0	0
23	COVID	LN2	Lymphocytes and type 2 pneumocytes	0	0	2	2	0	2	0	2	2	0	0
21	COVID	LN2	Type 2 pneumocytes	0	0	3	3	0	0	0	2	0	0	0
1	COVID	LN3	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
2	COVID	LN3	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
3	COVID	LN3	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
14	COVID	LN3	Hyaline membranes	0	3	0	0	0	0	0	1	1	3	2
13	COVID	LN3	Macrophages	0	3	2	3	0	1	0	0	1	3	1
15	COVID	LN3	Macrophages	2	3	2	3	0	2	0	2	1	2	1
16	COVID	LN3	Macrophages	0	3	0	3	0	0	0	0	0	3	1
17	COVID	LN3	Macrophages	0	0	3	3	0	0	0	2	0	3	2
19	COVID	LN3	Macrophages	0	3	3	3	0	1	0	2	1	2	1
20	COVID	LN3	Macrophages	0	3	2	3	0	0	0	1	1	3	1
40	COVID	LN3	Macrophages	0	3	0	3	0	1	0	2	1	2	0
42	COVID	LN3	Macrophages	0	3	2	3	0	0	0	2	1	3	1
43	COVID	LN3	Subepithelial lymphocytes	0	0	0	3	0	1	0	0	1	0	0

44	COVID	LN3	Subepithelial lymphocytes	0	0	0	3	0	1	0	0	2	0	0
18	COVID	LN3	Type 2 pneumocytes	0	0	2	3	0	1	0	2	1	3	2
41	COVID	LN3	Type 2 pneumocytes	0	3	2	3	0	1	0	2	2	3	0
24	COVID	LN4	Type 2 pneumocytes	0	0	3	3	0	0	0	3	0	0	0
25	COVID	LN4	Type 2 pneumocytes	0	0	3	3	0	0	0	3	0	3	0
4	COVID	LN6	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
5	COVID	LN6	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
27	COVID	LN6	Type 2 pneumocytes	0	0	1	0	0	1	0	1	1	0	0
28	COVID	LN6	Type 2 pneumocytes	0	0	2	2	0	0	0	2	0	0	0
45	COVID	LN6	Type 2 pneumocytes	0	0	3	2	0	1	0	1	1	0	0
26	COVID	LN7	Type 2 pneumocytes	0	0	1	2	0	1	0	2	2	0	0
29	COVID	LN8	Macrophages	3	0	3	3	3	3	0	2	2	3	0
30	COVID	LN8	Macrophages	3	0	3	3	3	3	0	1	2	3	0
37	COVID	LN9	Macrophages	1	0	2	3	3	2	0	2	2	3	0
31	COVID	LN10	Type 2 pneumocytes and Bronchiolar epithelium	0	0	2	3	0	3	3	2	2	2	0
32	COVID	LN10	Type 2 pneumocytes	2	0	3	3	1	2	0	3	2	2	0
33	COVID	LN10	Type 2 pneumocytes	1	3	3	3	0	0	0	3	1	3	0
46	COVID	LN10	Type 2 pneumocytes	1	1	3	1	0	1	0	3	2	2	0
47	COVID	LN10	Type 2 pneumocytes	1	0	3	2	0	1	0	3	2	3	0
34	COVID	LN11	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	0
36	COVID	LN11	Hyaline membranes	0	3	2	0	0	2	0	3	2	2	0
35	COVID	LN11	Type 2 pneumocytes	0	3	3	0	0	1	0	2	2	3	0
8	H1N1	N09-6999	Bronchiolar epithelium	0	0	0	0	0	0	3	0	0	0	
7	H1N1	N09-6999	Macrophages	3	0	3	0	0	3	0	2	0	1	
1	H1N1	N09-7000	Bronchiolar epithelium	0	0	0	0	0	0	2	0	0	0	
2	H1N1	N09-7223	Macrophages	2	0	3	3	0	0	0	0	0	1	
3	H1N1	N09-7223	Macrophages	2	0	3	3	0	0	0	1	0	2	
5	H1N1	N09-7888	Intravascular lymphocytes	0	0	0	0	0	0	0	0	0	0	

6	H1N1	N09-7888	Type 2 pneumocytes	1	3	3	3	0	0	0	0	0	1	
4	H1N1	N09-8157	Type 2 pneumocytes	1	0	2	0	0	3	0	1	1	0	
5	NORMAL	N07-009	Type 2 pneumocytes	1	0	0	3	0	0	0	0	0	0	
3	NORMAL	N08-021	Carcinoma	0	0	0	0	0	0	0	0	0	0	
4	NORMAL	N09-019	Bronchiolar epithelium	0	0	0	0	0	0	2	0	0	1	
1	NORMAL	N09-019	Macrophages	0	0	1	0	0	0	0	0	0	0	
2	NORMAL	N09-57	Macrophages	0	0	1	0	0	0	0	0	0	0	

Supplementary Table 6. Genes downregulated in the lungs of SARS-CoV-2 patients compared with non-virally infected controls.

Gene name	Fold Change (log2)	Average Expression	t-statistics	p-value	Adjusted p-value
CSF3	-1.060	6.642	-2.740	0.00401	0.02794
PPP2R1A	-0.817	7.657	-3.234	0.00098	0.00819
IFNA8	-0.728	7.554	-2.695	0.00452	0.03075
MAGEC2	-0.704	7.792	-3.019	0.00184	0.01440
PCK1	-0.694	7.777	-3.596	0.00032	0.00313
SARS-CoV-2-Neg	-0.683	7.714	-2.842	0.00302	0.02212
NegProbe	-0.672	7.483	-4.324	0.00003	0.00035
MLANA	-0.658	7.558	-3.325	0.00074	0.00639
NAALAD2	-0.596	7.650	-2.722	0.00420	0.02883
CD19	-0.593	7.831	-2.685	0.00464	0.03149
FGF17	-0.582	7.318	-2.496	0.00761	0.04729
MPL	-0.582	7.875	-2.810	0.00330	0.02351
IL9	-0.575	7.721	-2.662	0.00493	0.03323
GAGE1	-0.567	7.469	-2.651	0.00507	0.03380
IL13	-0.567	7.716	-2.488	0.00778	0.04784
PRAME	-0.561	7.564	-2.517	0.00721	0.04529
IFNL1	-0.558	7.762	-2.473	0.00808	0.04905
LDHA	-0.542	7.791	-2.479	0.00794	0.04852
C9	-0.530	7.670	-2.511	0.00733	0.04586
IL7	-0.526	7.794	-2.553	0.00656	0.04192

Supplementary Table 7. Genes upregulated in the lungs of SARS-CoV-2 patients compared with non-virally infected controls.

Gene name	Fold Change (log2)	Average Expression	t-statistics	p-value	Adjusted p-value
CD74	3.456	13.824	10.579	7.54E-16	2.79E-13
SSX1	3.026	11.624	17.734	2.38E-26	4.41E-23
COL3A1	2.822	11.546	5.737	1.53E-07	3.83E-06
B2M	2.660	14.118	8.550	2.11E-12	1.78E-10
HLA-DRA	2.659	11.577	9.218	1.49E-13	1.85E-11
ACTB	2.612	14.090	7.752	5.15E-11	2.89E-09
FN1	2.587	12.345	6.410	1.10E-08	3.23E-07
LY6E	2.571	10.949	11.983	3.96E-18	3.67E-15
SFTPA1	2.563	13.247	2.945	2.57E-03	1.91E-02
HLA-C	2.532	12.090	10.324	2.00E-15	6.16E-13
HLA-F	2.323	11.219	10.984	1.62E-16	7.48E-14
CD63	2.298	11.385	9.479	5.33E-14	8.97E-12
IFI27	2.289	10.932	7.896	2.88E-11	1.72E-09
ANXA1	2.280	10.400	11.337	4.30E-17	2.65E-14
HLA-E	2.262	11.640	8.574	1.92E-12	1.69E-10
KRT7	2.259	10.763	4.324	2.85E-05	3.52E-04
COL1A1	2.234	11.022	4.423	2.02E-05	2.59E-04
CD163	2.216	10.854	7.565	1.09E-10	5.64E-09
CD81	2.203	11.362	8.971	3.95E-13	4.30E-11
FLNA	2.186	10.894	7.563	1.10E-10	5.64E-09
KRT18	2.133	11.108	5.014	2.36E-06	3.87E-05
CD68	2.127	10.776	9.258	1.27E-13	1.82E-11
TYMP	2.093	10.862	8.715	1.09E-12	1.07E-10
C1QA	2.087	10.690	8.302	5.68E-12	4.05E-10
KRT19	2.055	9.979	4.589	1.11E-05	1.56E-04
A2M	2.038	12.558	5.212	1.13E-06	2.04E-05
C1R	2.025	10.440	7.179	5.11E-10	2.15E-08
APP	2.017	10.607	9.011	3.38E-13	3.91E-11
HSP90B1	2.012	10.113	9.217	1.50E-13	1.85E-11
CTSS	2.008	10.175	9.804	1.50E-14	3.46E-12
COL1A2	1.998	11.096	4.787	5.43E-06	8.04E-05
GPX1	1.985	11.060	8.944	4.41E-13	4.54E-11
HLA-A	1.975	12.878	7.191	4.87E-10	2.10E-08
OAZ1	1.946	11.412	8.360	4.50E-12	3.47E-10
CTSH	1.925	10.501	4.983	2.64E-06	4.24E-05
CD9	1.905	10.245	6.956	1.25E-09	4.71E-08
LGALS3	1.861	10.319	8.577	1.89E-12	1.69E-10
CCL18	1.857	10.005	5.360	6.42E-07	1.28E-05

UBB	1.856	11.016	8.325	5.19E-12	3.84E-10
DDIT4	1.853	10.493	5.811	1.14E-07	2.89E-06
HLA-DRB3	1.833	10.665	6.021	5.04E-08	1.39E-06
LAMP1	1.781	10.410	9.314	1.02E-13	1.57E-11
CTNNB1	1.770	9.983	9.668	2.54E-14	5.23E-12
LYZ	1.764	10.533	5.294	8.25E-07	1.57E-05
HLA-B	1.735	13.623	5.849	9.83E-08	2.57E-06
H3C10	1.734	10.187	5.658	2.06E-07	4.92E-06
HLA-DRB4	1.727	10.644	4.854	4.23E-06	6.42E-05
PKM	1.716	10.692	9.615	3.13E-14	5.80E-12
SERPING1	1.693	11.532	4.841	4.43E-06	6.67E-05
GNAS	1.686	11.510	5.717	1.64E-07	4.05E-06
RPL7A	1.679	11.626	5.656	2.07E-07	4.92E-06
CAPN2	1.663	10.033	6.528	6.84E-09	2.11E-07
CD14	1.658	9.507	6.540	6.53E-09	2.08E-07
COL5A1	1.637	9.452	4.485	1.60E-05	2.13E-04
COL6A3	1.628	10.820	4.432	1.93E-05	2.53E-04
CEACAM6	1.623	10.178	3.871	1.32E-04	1.45E-03
FCGRT	1.613	9.893	8.366	4.40E-12	3.47E-10
ADH1A	1.609	9.753	5.614	2.44E-07	5.41E-06
C3	1.589	10.763	3.336	7.27E-04	6.35E-03
CLU	1.567	10.458	3.590	3.28E-04	3.20E-03
SERPINA1	1.564	10.750	3.803	1.65E-04	1.76E-03
C1S	1.556	9.915	5.640	2.20E-07	5.17E-06
TXNIP	1.553	10.982	4.240	3.76E-05	4.53E-04
CD44	1.539	10.424	6.465	8.77E-09	2.66E-07
RPS27A	1.537	10.665	5.344	6.82E-07	1.33E-05
CCND1	1.520	9.444	4.602	1.05E-05	1.50E-04
ACTA2	1.511	9.385	3.619	2.98E-04	2.98E-03
STAT1	1.510	9.820	6.549	6.31E-09	2.08E-07
C1QB	1.498	9.642	7.111	6.70E-10	2.76E-08
RAB7A	1.482	9.890	10.081	5.11E-15	1.35E-12
IFI6	1.476	9.913	5.091	1.77E-06	3.06E-05
IFITM1	1.447	10.502	5.016	2.33E-06	3.86E-05
STAT3	1.444	10.017	6.934	1.36E-09	5.04E-08
MUC1	1.437	10.106	3.188	1.13E-03	9.35E-03
BAX	1.430	9.060	8.024	1.73E-11	1.07E-09
RPS6	1.428	11.502	3.915	1.13E-04	1.27E-03
STAT2	1.427	9.587	8.199	8.58E-12	5.67E-10
LRP1	1.421	10.286	6.209	2.41E-08	6.86E-07
C7	1.411	9.782	3.621	2.96E-04	2.98E-03
APOL6	1.406	9.456	6.537	6.60E-09	2.08E-07

ALDOA	1.405	11.133	5.569	2.89E-07	6.15E-06
ITGB2	1.404	9.841	5.884	8.59E-08	2.27E-06
TPSB2	1.382	9.653	3.640	2.78E-04	2.81E-03
CD97	1.360	9.880	7.046	8.71E-10	3.43E-08
MIF	1.348	9.717	6.914	1.47E-09	5.35E-08
HIF1A	1.331	9.341	7.236	4.06E-10	1.84E-08
COX6A1	1.326	10.440	6.442	9.62E-09	2.87E-07
BRD2	1.324	9.546	8.252	6.94E-12	4.76E-10
CD164	1.316	9.799	7.197	4.75E-10	2.09E-08
CTSL	1.315	10.252	4.860	4.14E-06	6.34E-05
H3-3A	1.310	10.042	6.200	2.50E-08	7.01E-07
CD55	1.302	10.089	3.305	7.92E-04	6.76E-03
HLA-DMA	1.302	9.545	5.506	3.69E-07	7.58E-06
SF3B1	1.301	9.651	7.236	4.07E-10	1.84E-08
NDUFA13	1.301	10.040	7.072	7.84E-10	3.16E-08
TPM4	1.297	10.618	4.747	6.25E-06	9.18E-05
LDHB	1.295	9.866	4.482	1.61E-05	2.14E-04
JUNB	1.291	10.170	4.907	3.48E-06	5.50E-05
CD99	1.288	9.679	7.587	9.99E-11	5.44E-09
CD4	1.281	9.295	7.296	3.20E-10	1.56E-08
EWSR1	1.279	10.008	7.851	3.46E-11	2.00E-09
IFITM2	1.278	10.807	4.428	1.96E-05	2.55E-04
MDM2	1.275	9.172	5.665	2.00E-07	4.88E-06
DUSP1	1.268	10.724	3.429	5.40E-04	5.02E-03
H3C2	1.267	8.791	3.406	5.80E-04	5.34E-03
PSMB9	1.260	9.919	5.612	2.46E-07	5.41E-06
CXCL16	1.253	9.219	5.272	8.93E-07	1.67E-05
SLC1A5	1.241	9.260	5.592	2.65E-07	5.78E-06
ENG	1.239	10.026	4.293	3.13E-05	3.79E-04
HLA-DPA1	1.234	9.398	6.573	5.73E-09	1.93E-07
IFNGR1	1.232	9.428	6.229	2.23E-08	6.45E-07
IDH1	1.232	9.349	5.575	2.83E-07	6.10E-06
GRB2	1.229	9.422	8.060	1.49E-11	9.54E-10
CDKN1A	1.226	9.848	4.181	4.62E-05	5.52E-04
H3-5	1.226	9.942	5.021	2.29E-06	3.85E-05
SREBF1	1.224	9.546	4.050	7.19E-05	8.33E-04
PRDX5	1.201	9.811	4.870	3.99E-06	6.19E-05
CFB	1.199	9.777	3.790	1.71E-04	1.82E-03
COX5B	1.194	9.909	7.000	1.05E-09	4.04E-08
VEGFA	1.163	9.514	3.966	9.57E-05	1.09E-03
NDUFB4	1.156	9.611	7.478	1.54E-10	7.72E-09
GDF15	1.156	8.995	3.007	1.91E-03	1.47E-02

BCL2L1	1.145	9.540	4.368	2.41E-05	3.04E-04
TNFRSF1A	1.145	9.966	5.291	8.33E-07	1.57E-05
CD46	1.143	9.585	4.053	7.12E-05	8.29E-04
AKT1	1.141	9.673	4.991	2.55E-06	4.15E-05
VCAN	1.138	9.483	3.485	4.53E-04	4.28E-03
TNC	1.138	9.525	2.639	5.29E-03	3.51E-02
CEBPB	1.131	9.756	5.357	6.49E-07	1.28E-05
GPI	1.129	9.164	6.537	6.61E-09	2.08E-07
RHOB	1.128	10.349	2.769	3.71E-03	2.60E-02
MX1	1.126	9.748	2.983	2.04E-03	1.55E-02
ITGAV	1.116	9.074	5.240	1.01E-06	1.85E-05
CD47	1.114	9.427	4.416	2.04E-05	2.61E-04
ITGB1	1.104	10.627	3.489	4.48E-04	4.25E-03
TPM1	1.102	9.776	2.983	2.04E-03	1.55E-02
MAP2K2	1.101	9.453	6.643	4.34E-09	1.49E-07
SOD1	1.087	10.075	4.533	1.35E-05	1.83E-04
S100A9	1.085	11.062	2.568	6.37E-03	4.09E-02
POLR2A	1.075	9.117	7.245	3.93E-10	1.84E-08
STAT6	1.074	9.514	6.699	3.48E-09	1.24E-07
TNFSF10	1.071	9.246	3.820	1.55E-04	1.68E-03
NDUFA11	1.069	9.594	5.562	2.98E-07	6.27E-06
COL5A2	1.064	8.843	3.013	1.87E-03	1.45E-02
MCL1	1.060	10.445	3.783	1.75E-04	1.85E-03
ANP32B	1.056	9.170	5.163	1.35E-06	2.42E-05
ADH1B	1.052	9.578	2.629	5.40E-03	3.57E-02
CYBB	1.050	8.950	4.316	2.89E-05	3.54E-04
TAP1	1.045	10.082	3.427	5.43E-04	5.03E-03
DAB2	1.044	9.054	4.704	7.31E-06	1.05E-04
ILF3	1.043	9.518	5.512	3.60E-07	7.49E-06
NDUFS8	1.039	8.971	5.898	8.14E-08	2.18E-06
MRC1	1.037	9.450	2.894	2.62E-03	1.94E-02
SBNO2	1.034	9.398	5.367	6.26E-07	1.26E-05
IL6ST	1.029	9.659	4.523	1.40E-05	1.89E-04
IL1R1	1.024	9.869	3.332	7.27E-04	6.35E-03
PDGFRB	1.020	9.320	3.964	9.62E-05	1.09E-03
TUBB	1.020	9.530	3.945	1.02E-04	1.15E-03
ATP5F1D	1.015	9.477	5.317	7.56E-07	1.46E-05
ITGA3	1.014	9.674	2.578	6.19E-03	4.02E-02
SFN	1.012	9.517	2.468	8.22E-03	4.97E-02
CSF1R	1.005	9.096	3.566	3.51E-04	3.40E-03
COX6B1	0.990	9.843	4.581	1.13E-05	1.58E-04
CDKN1B	0.988	9.159	5.503	3.73E-07	7.59E-06

TECR	0.988	9.326	5.257	9.47E-07	1.75E-05
ALCAM	0.985	9.264	3.354	6.79E-04	6.08E-03
JAK1	0.983	9.550	5.616	2.42E-07	5.41E-06
NDUFB7	0.982	11.000	4.080	6.50E-05	7.62E-04
CD276	0.979	8.998	5.952	6.60E-08	1.80E-06
FCER1G	0.977	9.521	3.459	4.92E-04	4.63E-03
ADH1C	0.972	9.747	3.024	1.81E-03	1.43E-02
MLPH	0.968	8.818	2.566	6.37E-03	4.09E-02
SDHA	0.961	8.997	5.613	2.45E-07	5.41E-06
G6PD	0.948	9.144	5.156	1.38E-06	2.46E-05
HLA-DMB	0.942	9.205	3.565	3.52E-04	3.40E-03
CDC25B	0.941	9.435	2.949	2.24E-03	1.68E-02
ETS2	0.937	10.148	2.776	3.63E-03	2.57E-02
CD53	0.929	9.639	3.953	9.96E-05	1.13E-03
PSMB10	0.927	9.059	5.831	1.05E-07	2.71E-06
C2	0.927	8.818	3.342	7.06E-04	6.24E-03
DDIT3	0.926	8.847	3.141	1.29E-03	1.04E-02
ZEB2	0.926	8.916	4.334	2.71E-05	3.40E-04
LAMA5	0.924	9.451	3.138	1.30E-03	1.05E-02
ATP2A2	0.923	9.483	5.114	1.62E-06	2.85E-05
LCP1	0.919	9.461	3.604	3.11E-04	3.08E-03
SIRPA	0.910	9.089	5.091	1.76E-06	3.06E-05
ITGA5	0.900	9.093	3.507	4.22E-04	4.03E-03
RHOA	0.897	9.787	3.646	2.72E-04	2.77E-03
HK1	0.892	9.384	6.681	3.74E-09	1.31E-07
TGFBR2	0.883	10.060	3.387	6.13E-04	5.57E-03
BRD4	0.881	9.493	4.735	6.51E-06	9.42E-05
LTBP1	0.875	8.716	2.736	4.04E-03	2.80E-02
PEBP1	0.872	9.606	2.805	3.35E-03	2.38E-02
U2AF1	0.871	9.414	4.423	1.99E-05	2.57E-04
NDUFA1	0.862	9.731	3.599	3.16E-04	3.11E-03
C4A-B	0.860	9.518	2.502	7.51E-03	4.68E-02
FLNB	0.860	9.313	3.668	2.54E-04	2.62E-03
FUBP1	0.856	9.347	5.016	2.33E-06	3.86E-05
CCL15	0.855	9.385	4.396	2.18E-05	2.77E-04
PARP9	0.852	9.343	4.921	3.31E-06	5.29E-05
PRKACA	0.851	8.852	4.797	5.20E-06	7.76E-05
DTX3L	0.846	9.100	5.046	2.08E-06	3.57E-05
SLC25A1	0.832	9.134	3.812	1.59E-04	1.71E-03
LGALS9	0.832	9.140	3.185	1.13E-03	9.35E-03
SMARCA4	0.831	9.063	5.027	2.24E-06	3.80E-05
JUP	0.826	9.245	3.102	1.44E-03	1.15E-02

IGF2R	0.826	9.243	4.152	5.09E-05	6.01E-04
DST	0.826	8.718	3.365	6.58E-04	5.91E-03
OAT	0.826	8.929	4.741	6.37E-06	9.30E-05
CSF3R	0.824	8.898	2.734	4.07E-03	2.81E-02
TAP2	0.823	9.623	2.847	2.98E-03	2.19E-02
ATP5ME	0.821	9.880	3.857	1.37E-04	1.50E-03
PTEN	0.820	8.906	3.899	1.19E-04	1.32E-03
SHC1	0.817	9.448	3.201	1.08E-03	9.00E-03
IDH2	0.816	9.175	3.729	2.08E-04	2.19E-03
SGK1	0.804	9.027	2.572	6.25E-03	4.05E-02
C1QBP	0.803	9.061	3.829	1.51E-04	1.64E-03
ATOX1	0.800	9.071	5.632	2.27E-07	5.25E-06
GPX4	0.791	9.956	2.815	3.26E-03	2.34E-02
MAPKAPK2	0.788	9.282	3.667	2.54E-04	2.62E-03
MAPK3	0.786	8.784	4.598	1.07E-05	1.51E-04
TYK2	0.785	9.072	4.890	3.71E-06	5.83E-05
FCGR2A	0.782	8.870	3.346	6.97E-04	6.21E-03
EIF4EBP1	0.779	8.822	3.148	1.26E-03	1.02E-02
RAD21	0.775	9.198	3.264	8.92E-04	7.55E-03
PRKCD	0.772	8.887	4.869	4.01E-06	6.19E-05
MS4A6A	0.771	9.283	2.659	4.97E-03	3.34E-02
UBA7	0.770	8.859	4.542	1.30E-05	1.79E-04
GNG12	0.765	9.043	3.341	7.07E-04	6.24E-03
TRAF7	0.764	8.947	4.551	1.26E-05	1.75E-04
PML	0.756	9.005	3.018	1.84E-03	1.44E-02
PSMB5	0.755	9.193	3.698	2.31E-04	2.40E-03
ITPK1	0.749	8.880	3.446	5.12E-04	4.79E-03
PSMB8	0.744	9.238	3.178	1.15E-03	9.49E-03
API5	0.737	8.811	4.508	1.47E-05	1.97E-04
IFI16	0.737	8.657	2.488	7.77E-03	4.78E-02
BCL6	0.737	9.083	3.242	9.53E-04	8.02E-03
CFI	0.733	8.797	2.547	6.68E-03	4.25E-02
ABL1	0.726	8.798	3.328	7.36E-04	6.38E-03
CSF1	0.719	8.808	2.697	4.50E-03	3.07E-02
NDUFB10	0.709	9.215	3.372	6.43E-04	5.81E-03
PARP4	0.706	8.960	3.616	3.00E-04	2.98E-03
PFKFB3	0.704	9.115	2.556	6.52E-03	4.18E-02
GLS	0.703	8.825	2.732	4.09E-03	2.81E-02
NFE2L2	0.699	9.332	3.314	7.67E-04	6.58E-03
PSMB7	0.695	9.198	3.703	2.27E-04	2.37E-03
FURIN	0.691	9.272	2.623	5.47E-03	3.61E-02
SIGLEC1	0.684	8.758	3.114	1.39E-03	1.12E-02

SERINC1	0.682	9.016	2.577	6.18E-03	4.02E-02
TWF1	0.681	8.924	4.175	4.71E-05	5.59E-04
IFI35	0.680	9.026	2.490	7.73E-03	4.78E-02
IFNGR2	0.679	9.439	3.327	7.38E-04	6.38E-03
DDB2	0.676	8.421	3.106	1.43E-03	1.14E-02
PCNA	0.674	8.884	2.474	8.05E-03	4.90E-02
IRF9	0.668	9.101	4.300	3.06E-05	3.73E-04
NOTCH1	0.668	8.847	2.967	2.13E-03	1.61E-02
LTBR	0.665	9.150	2.983	2.03E-03	1.55E-02
SMAD4	0.653	8.630	3.172	1.17E-03	9.62E-03
PCK2	0.652	8.514	2.953	2.22E-03	1.67E-02
ETS1	0.648	8.905	2.579	6.13E-03	4.01E-02
FH	0.646	8.682	3.545	3.76E-04	3.61E-03
IL13RA1	0.645	9.073	2.743	3.97E-03	2.77E-02
BAGE	0.640	9.522	2.821	3.21E-03	2.31E-02
RELA	0.638	9.181	3.658	2.62E-04	2.68E-03
CUL1	0.637	8.791	4.010	8.25E-05	9.49E-04
SF3A1	0.636	8.829	3.395	6.00E-04	5.47E-03
MBNL1	0.632	9.157	2.528	7.01E-03	4.41E-02
COX4I1	0.624	9.564	2.838	3.05E-03	2.23E-02
GNA11	0.624	9.045	3.014	1.86E-03	1.45E-02
E2F4	0.615	8.931	2.993	1.98E-03	1.52E-02
DNMT1	0.602	8.719	3.302	7.95E-04	6.76E-03
CDK4	0.599	8.653	2.603	5.76E-03	3.79E-02
KMT2D	0.599	8.883	3.162	1.21E-03	9.87E-03
SPOP	0.597	8.754	2.832	3.11E-03	2.26E-02
SERINC5	0.585	8.560	2.534	6.90E-03	4.38E-02
MAPK14	0.584	8.756	3.084	1.52E-03	1.20E-02
PTPN11	0.578	9.125	2.652	5.06E-03	3.38E-02
PYCR2	0.576	8.510	2.822	3.20E-03	2.31E-02
CKLF	0.575	8.820	2.811	3.29E-03	2.35E-02
ARID1A	0.570	8.545	2.874	2.77E-03	2.04E-02
TNFRSF14	0.559	8.985	3.397	5.96E-04	5.47E-03
UQCR11	0.558	8.619	2.951	2.23E-03	1.67E-02
SF3A3	0.554	8.864	2.671	4.82E-03	3.26E-02
NASP	0.530	8.950	2.484	7.86E-03	4.82E-02
NDUFS7	0.525	8.851	2.530	6.98E-03	4.41E-02
RBX1	0.515	8.778	2.768	3.71E-03	2.60E-02

*Highlighted genes (*LY6E, IFI27* & *IFI6)* also found to be upregulated in SARS-CoV-2 patients when compared to pH1N1 influenza samples.

Supplementary Table 8. Genes differentially expressed in the lungs of SARS-CoV-2 patients with high viral load compared with patients with low viral load (Region based analysis).

Gene name	Fold Change (log2)	Average Expression	t-statistics	p-value	Adjusted p-value
S	2.415	8.900	5.616	4.985E-07	2.350E-04
ORF1ab	2.365	9.036	5.611	5.076E-07	2.350E-04
CXCL11	0.946	8.291	4.533	2.738E-05	2.414E-03
RSAD2	0.669	8.662	3.014	3.737E-03	4.524E-02
IL1RN	0.584	8.130	4.057	1.421E-04	6.923E-03
ANGPTL4	0.555	8.008	3.216	2.073E-03	3.392E-02
CD274	0.531	8.076	6.292	3.621E-08	6.706E-05
CCL8	0.466	7.679	5.045	4.276E-06	1.320E-03
IL15RA	0.437	8.696	3.277	1.726E-03	3.043E-02
SOCS1	0.431	8.547	3.881	2.560E-04	9.676E-03
CCL7	0.401	7.816	4.008	1.679E-04	7.585E-03
OASL	0.382	8.196	2.989	4.020E-03	4.653E-02
SARS-CoV-2-Neg	0.357	7.714	4.643	1.845E-05	2.136E-03
IL19	0.305	7.750	3.648	5.454E-04	1.554E-02
CSF2RB	0.304	8.473	3.213	2.088E-03	3.392E-02
CALML5	0.296	7.632	3.022	3.655E-03	4.498E-02
WNT2	0.293	7.858	3.429	1.087E-03	2.314E-02
FGF17	0.290	7.318	3.552	7.402E-04	1.828E-02
ACE2	0.279	7.833	3.577	6.839E-04	1.759E-02
TLR8	0.277	8.246	3.190	2.238E-03	3.508E-02
CLECL1	0.269	7.774	3.056	3.312E-03	4.303E-02
MASP2	0.244	7.920	3.034	3.527E-03	4.414E-02
PPP3R2	0.241	7.704	3.107	2.861E-03	3.988E-02
RIMKLA	0.237	7.887	3.913	2.302E-04	9.546E-03
MAGEB2	0.227	7.673	3.039	3.479E-03	4.383E-02
RET	0.220	8.077	3.400	1.186E-03	2.392E-02
OTC	0.218	7.887	3.055	3.322E-03	4.303E-02
OTOA	0.216	7.789	3.100	2.915E-03	3.988E-02
SH2B2	0.207	8.230	3.101	2.905E-03	3.988E-02
MAPK8IP2	0.206	8.026	3.226	2.012E-03	3.357E-02
CCR3	0.201	7.703	2.971	4.222E-03	4.827E-02
IKBKE	0.197	8.207	3.117	2.774E-03	3.988E-02
MUC2	0.184	8.239	3.010	3.781E-03	4.548E-02
STAT5B	-0.169	8.393	-3.147	2.544E-03	3.831E-02
POLR2D	-0.195	8.264	-3.447	1.027E-03	2.239E-02
PIAS4	-0.196	8.294	-3.510	8.437E-04	1.930E-02
APC	-0.201	7.947	-3.306	1.581E-03	2.892E-02
PPP3CA	-0.208	8.486	-3.020	3.673E-03	4.498E-02

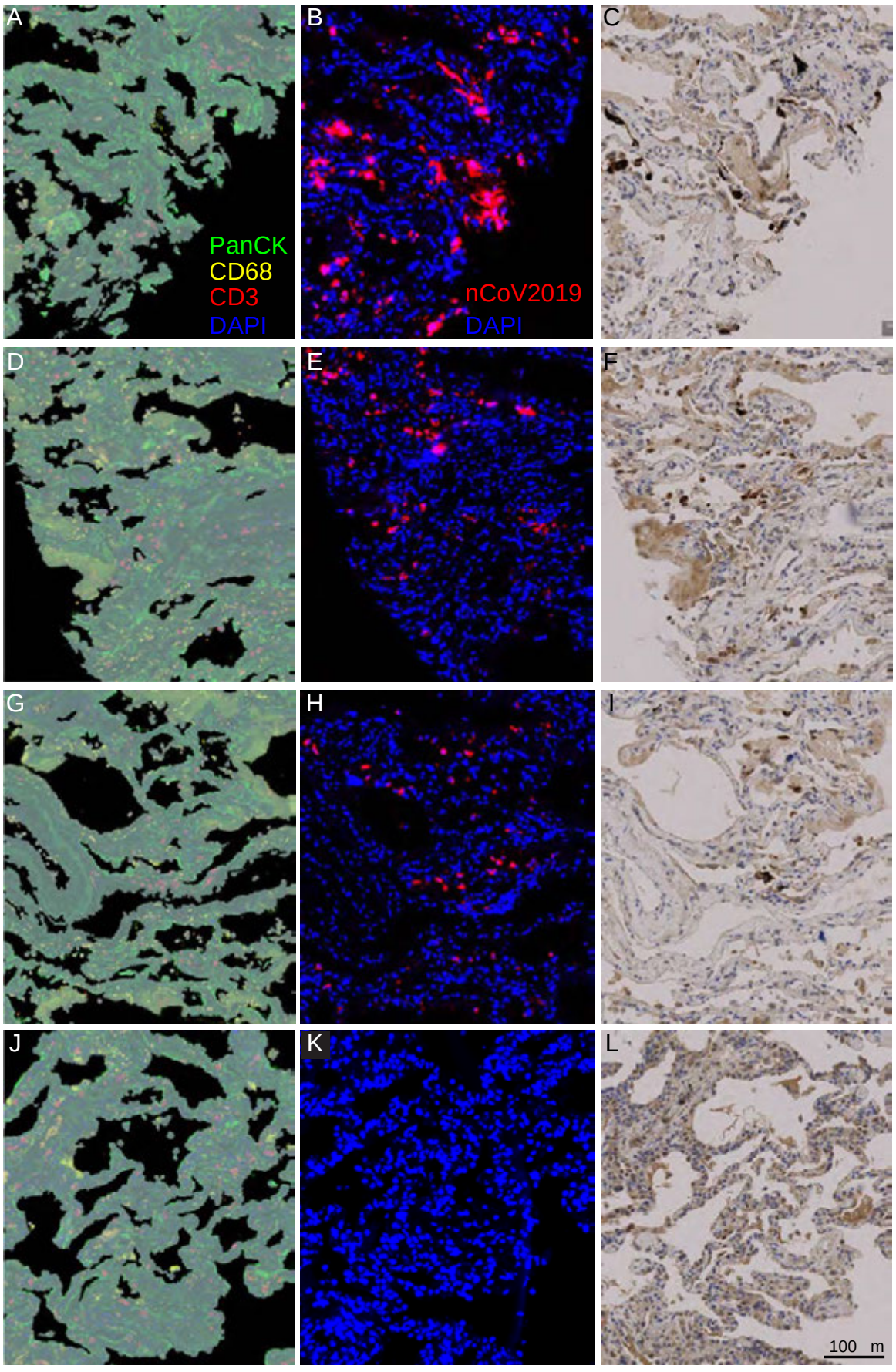
ITCH	-0.209	8.672	-3.158	2.460E-03	3.735E-02
ERCC3	-0.213	8.212	-3.049	3.379E-03	4.345E-02
IFNAR1	-0.213	8.297	-3.639	5.615E-04	1.575E-02
UQCR11	-0.214	8.619	-3.180	2.306E-03	3.558E-02
SF3A3	-0.218	8.864	-2.963	4.329E-03	4.888E-02
API5	-0.224	8.811	-3.000	3.889E-03	4.597E-02
PSMD7	-0.225	8.980	-3.162	2.431E-03	3.721E-02
TPR	-0.229	8.771	-3.120	2.752E-03	3.988E-02
DNMT1	-0.230	8.719	-3.424	1.102E-03	2.319E-02
RAF1	-0.232	8.798	-2.997	3.922E-03	4.597E-02
RIN1	-0.232	8.270	-3.313	1.549E-03	2.892E-02
CKLF	-0.237	8.820	-3.135	2.632E-03	3.899E-02
SRP54	-0.237	8.832	-2.975	4.177E-03	4.805E-02
KMT2D	-0.238	8.883	-3.317	1.531E-03	2.892E-02
HMGB1	-0.238	8.644	-3.096	2.950E-03	3.988E-02
MAPK1	-0.244	8.648	-3.229	1.996E-03	3.357E-02
PVRIG	-0.245	8.911	-3.006	3.824E-03	4.569E-02
TGFB3	-0.247	8.411	-2.997	3.922E-03	4.597E-02
KMT2C	-0.249	8.808	-3.543	7.620E-04	1.857E-02
FEN1	-0.250	8.094	-3.416	1.129E-03	2.349E-02
TBL1XR1	-0.252	9.164	-3.105	2.875E-03	3.988E-02
NDUFA2	-0.253	9.007	-3.307	1.576E-03	2.892E-02
ROCK1	-0.254	8.777	-3.739	4.073E-04	1.245E-02
HDAC2	-0.257	8.493	-3.107	2.858E-03	3.988E-02
ATRX	-0.260	8.724	-3.835	2.979E-04	1.012E-02
KRT10	-0.262	8.395	-3.296	1.631E-03	2.905E-02
TXN2	-0.263	8.965	-3.385	1.244E-03	2.477E-02
GOT2	-0.263	8.994	-3.408	1.157E-03	2.381E-02
TGFBR1	-0.267	8.665	-3.784	3.518E-04	1.143E-02
IRF3	-0.273	8.749	-3.340	1.427E-03	2.781E-02
FH	-0.273	8.682	-3.737	4.100E-04	1.245E-02
CDKN2C	-0.274	7.931	-3.200	2.175E-03	3.502E-02
PALMD	-0.275	8.157	-3.274	1.742E-03	3.043E-02
NDUFA3	-0.283	9.207	-3.088	3.023E-03	4.011E-02
SMARCB1	-0.286	8.735	-4.142	1.068E-04	5.650E-03
ATF2	-0.288	8.348	-4.294	6.329E-05	4.042E-03
PARP4	-0.289	8.960	-3.380	1.262E-03	2.486E-02
COX4I1	-0.291	9.564	-3.430	1.082E-03	2.314E-02
PTPN11	-0.291	9.125	-3.770	3.686E-04	1.177E-02
NOTCH1	-0.292	8.847	-3.040	3.470E-03	4.383E-02
FUBP1	-0.296	9.347	-3.512	8.382E-04	1.930E-02
LRP6	-0.298	8.132	-3.477	9.349E-04	2.086E-02

SMAD4	-0.298	8.630	-3.694	4.710E-04	1.363E-02
DAXX	-0.298	8.341	-4.749	1.264E-05	1.672E-03
NFE2L2	-0.299	9.332	-3.304	1.593E-03	2.892E-02
CREBBP	-0.301	8.512	-4.209	8.489E-05	5.240E-03
NDUFB1	-0.301	9.165	-4.336	5.465E-05	3.771E-03
TRAF7	-0.306	8.947	-4.108	1.199E-04	6.167E-03
SPOP	-0.306	8.754	-3.973	1.886E-04	8.317E-03
JAK1	-0.307	9.550	-3.087	3.032E-03	4.011E-02
MAVS	-0.307	8.688	-3.491	8.948E-04	2.021E-02
PDGFA	-0.310	8.602	-3.089	3.013E-03	4.011E-02
ATP2A2	-0.317	9.483	-3.327	1.485E-03	2.835E-02
PRKACA	-0.321	8.852	-3.602	6.317E-04	1.696E-02
BRD2	-0.324	9.546	-2.991	3.995E-03	4.653E-02
H2AX	-0.329	8.721	-3.752	3.905E-04	1.226E-02
MCM7	-0.333	8.619	-3.869	2.664E-04	9.866E-03
LEPR	-0.334	8.243	-3.571	6.969E-04	1.768E-02
PSMB5	-0.337	9.193	-3.719	4.338E-04	1.296E-02
ABL1	-0.339	8.798	-3.534	7.819E-04	1.866E-02
TP53	-0.340	8.576	-4.612	2.063E-05	2.189E-03
PRKDC	-0.344	8.734	-4.303	6.126E-05	4.042E-03
CDKN1B	-0.346	9.159	-3.533	7.860E-04	1.866E-02
ITGA6	-0.348	8.091	-3.142	2.579E-03	3.853E-02
PPP3R1	-0.357	8.489	-4.851	8.717E-06	1.535E-03
COX6B1	-0.362	9.843	-3.018	3.692E-03	4.498E-02
OAT	-0.375	8.929	-4.838	9.146E-06	1.535E-03
P4HA2	-0.380	8.858	-3.271	1.759E-03	3.045E-02
STMN1	-0.385	8.393	-3.891	2.474E-04	9.546E-03
CD99	-0.393	9.679	-3.577	6.827E-04	1.759E-02
ITGA1	-0.396	8.799	-3.466	9.671E-04	2.132E-02
SERINC3	-0.404	9.130	-4.568	2.418E-05	2.239E-03
DST	-0.405	8.718	-3.336	1.445E-03	2.788E-02
THBD	-0.406	9.088	-3.019	3.685E-03	4.498E-02
NDUFA11	-0.408	9.594	-3.833	2.996E-04	1.012E-02
MKI67	-0.409	8.237	-3.298	1.620E-03	2.905E-02
ITGAV	-0.415	9.074	-3.188	2.254E-03	3.508E-02
NDUFS8	-0.415	8.971	-4.370	4.856E-05	3.597E-03
RAD21	-0.417	9.198	-3.918	2.262E-04	9.546E-03
NDUFA1	-0.417	9.731	-3.620	5.967E-04	1.625E-02
CDK4	-0.419	8.653	-5.278	1.793E-06	6.642E-04
SF3B1	-0.427	9.651	-3.799	3.353E-04	1.109E-02
RHOA	-0.439	9.787	-3.561	7.194E-04	1.801E-02
RANBP2	-0.441	9.123	-4.815	9.945E-06	1.535E-03

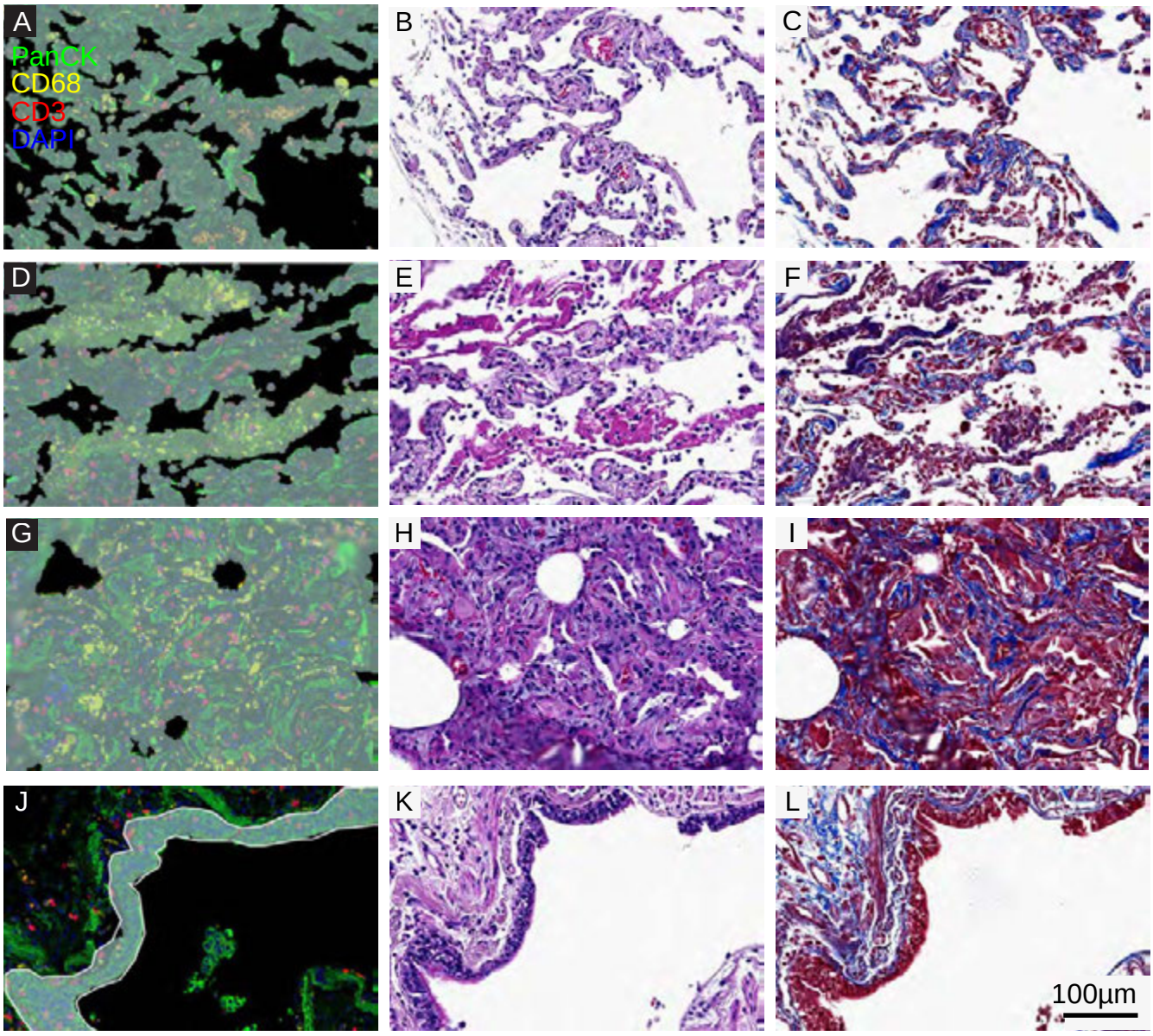
TYMS	-0.447	8.301	-3.116	2.785E-03	3.988E-02
THY1	-0.451	8.833	-3.122	2.736E-03	3.988E-02
PEBP1	-0.465	9.606	-3.042	3.453E-03	4.383E-02
CCND2	-0.466	8.729	-3.839	2.935E-04	1.012E-02
PDGFRB	-0.466	9.320	-2.954	4.442E-03	4.985E-02
ANP32B	-0.469	9.170	-4.150	1.039E-04	5.650E-03
LAMA5	-0.473	9.451	-3.097	2.944E-03	3.988E-02
TECR	-0.475	9.326	-4.975	5.523E-06	1.461E-03
ITGAX	-0.481	9.025	-3.080	3.094E-03	4.064E-02
C4A-B	-0.484	9.518	-2.969	4.251E-03	4.830E-02
ST6GAL1	-0.498	8.508	-4.435	3.872E-05	2.988E-03
SPRY1	-0.504	8.263	-4.574	2.367E-05	2.239E-03
H3-3A	-0.515	10.042	-4.083	1.303E-04	6.523E-03
PTEN	-0.518	8.906	-5.953	1.362E-07	1.261E-04
TPM4	-0.533	10.618	-3.101	2.904E-03	3.988E-02
SLC1A5	-0.534	9.260	-4.012	1.652E-04	7.585E-03
APP	-0.536	10.607	-3.189	2.247E-03	3.508E-02
PRDX5	-0.544	9.811	-3.624	5.894E-04	1.625E-02
TUBB	-0.558	9.530	-4.185	9.207E-05	5.501E-03
HSP90B1	-0.569	10.113	-3.219	2.054E-03	3.392E-02
NOTCH3	-0.576	8.919	-4.749	1.261E-05	1.672E-03
CTNNB1	-0.581	9.983	-4.456	3.585E-05	2.886E-03
FLNC	-0.583	8.103	-3.906	2.354E-04	9.546E-03
RPS27A	-0.615	10.665	-3.400	1.188E-03	2.392E-02
RPL7A	-0.623	11.626	-3.227	2.007E-03	3.357E-02
SREBF1	-0.655	9.546	-3.583	6.709E-04	1.759E-02
F13A1	-0.672	8.789	-3.190	2.240E-03	3.508E-02
LDHB	-0.679	9.866	-3.855	2.791E-04	1.012E-02
CD9	-0.725	10.245	-3.510	8.442E-04	1.930E-02
MCAM	-0.733	8.993	-4.464	3.495E-05	2.886E-03
ITGB1	-0.748	10.627	-4.604	2.128E-05	2.189E-03
GNAS	-0.751	11.510	-4.154	1.025E-04	5.650E-03
FLNA	-0.766	10.894	-3.266	1.782E-03	3.055E-02
RHOB	-0.785	10.349	-3.695	4.698E-04	1.363E-02
H3C8	-0.788	8.504	-4.706	1.476E-05	1.822E-03
RPS6	-0.831	11.502	-3.900	2.401E-04	9.546E-03
RPL23	-0.852	11.713	-4.159	1.005E-04	5.650E-03
H3C10	-0.873	10.187	-4.047	1.469E-04	6.975E-03
TPM1	-0.902	9.776	-4.869	8.144E-06	1.535E-03
SFRP2	-0.946	8.442	-3.895	2.442E-04	9.546E-03
H3C2	-0.998	8.791	-4.823	9.631E-06	1.535E-03
ACTA2	-1.103	9.385	-4.334	5.498E-05	3.771E-03

C3	-1.110	10.763	-3.832	3.006E-04	1.012E-02

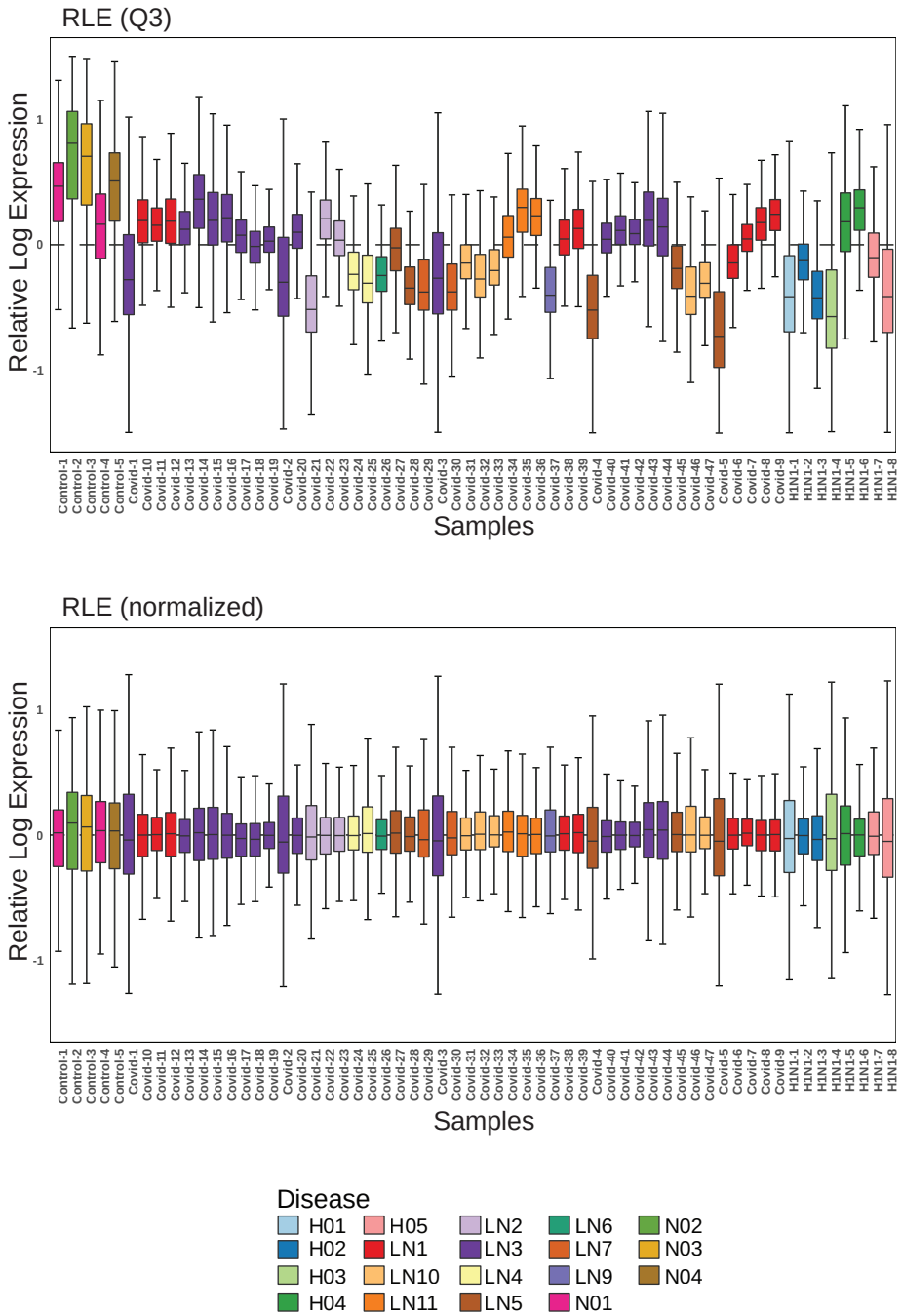
Genes upregulated genes in high viral load (green) compared with low vial load SARS-CoV-2 lung samples. Downregulated genes are shown in red.



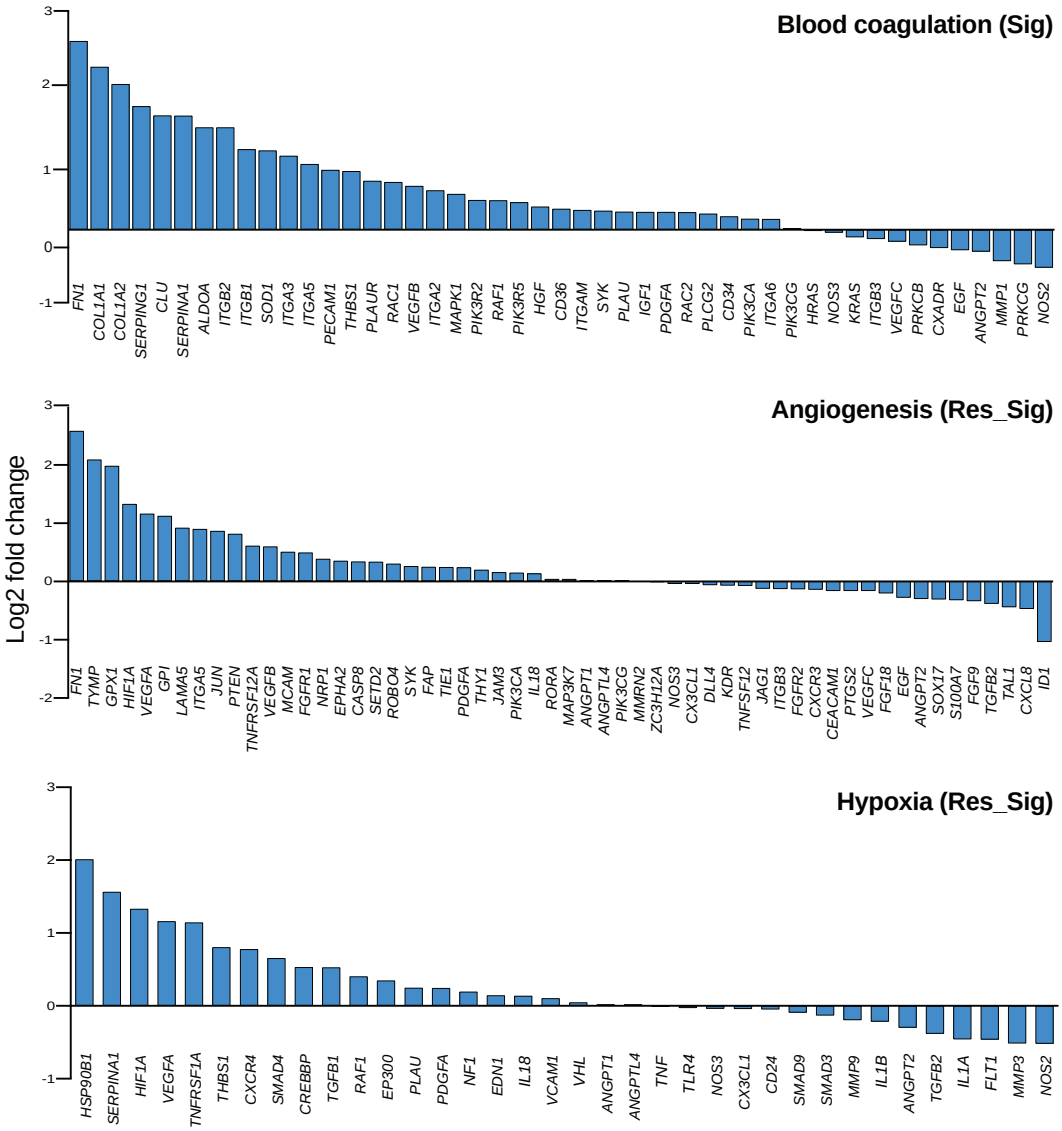
Kulasinghe *et al*, Supplementary Figure 1



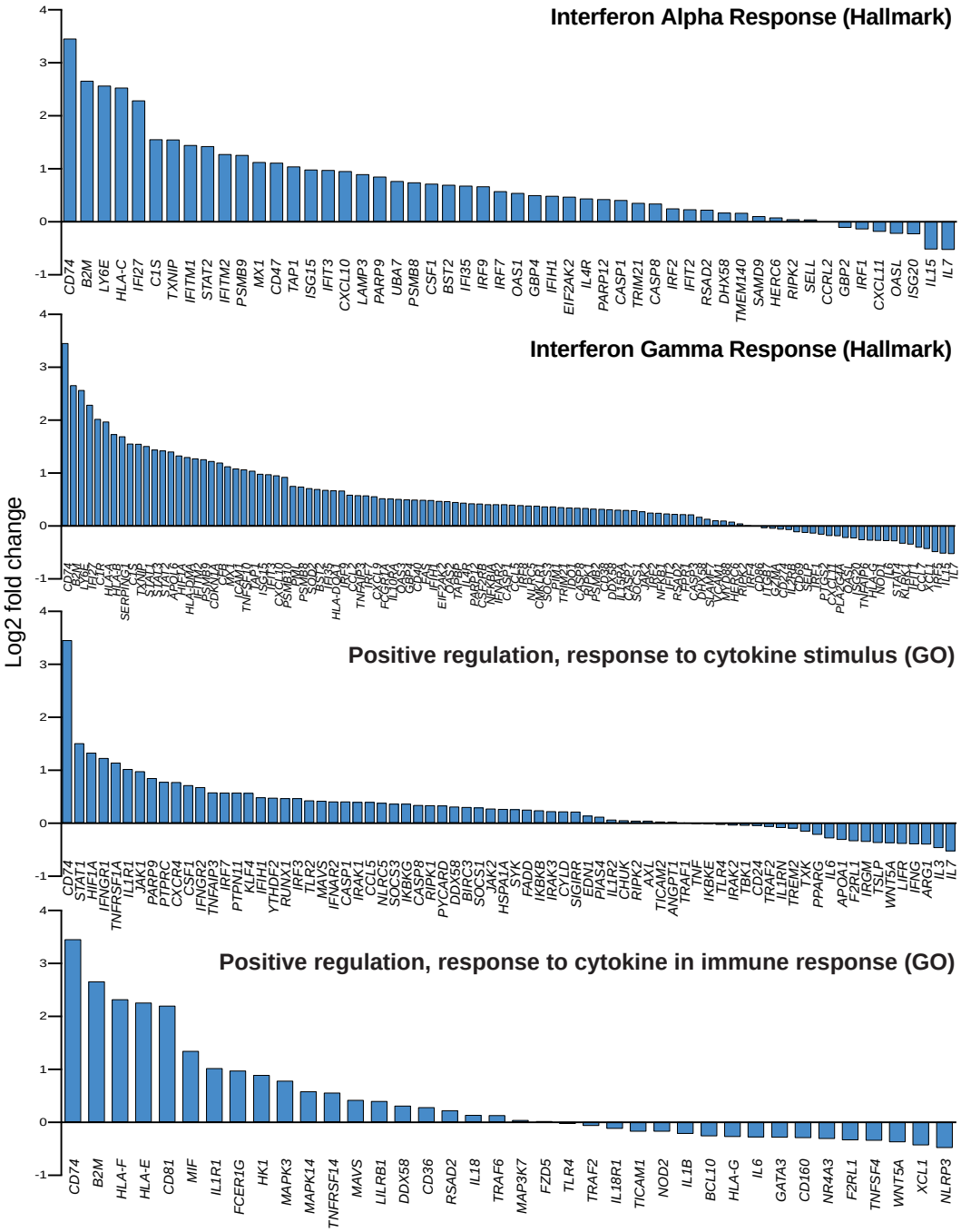
Kulasinghe *et al*, Supplementary Figure 2



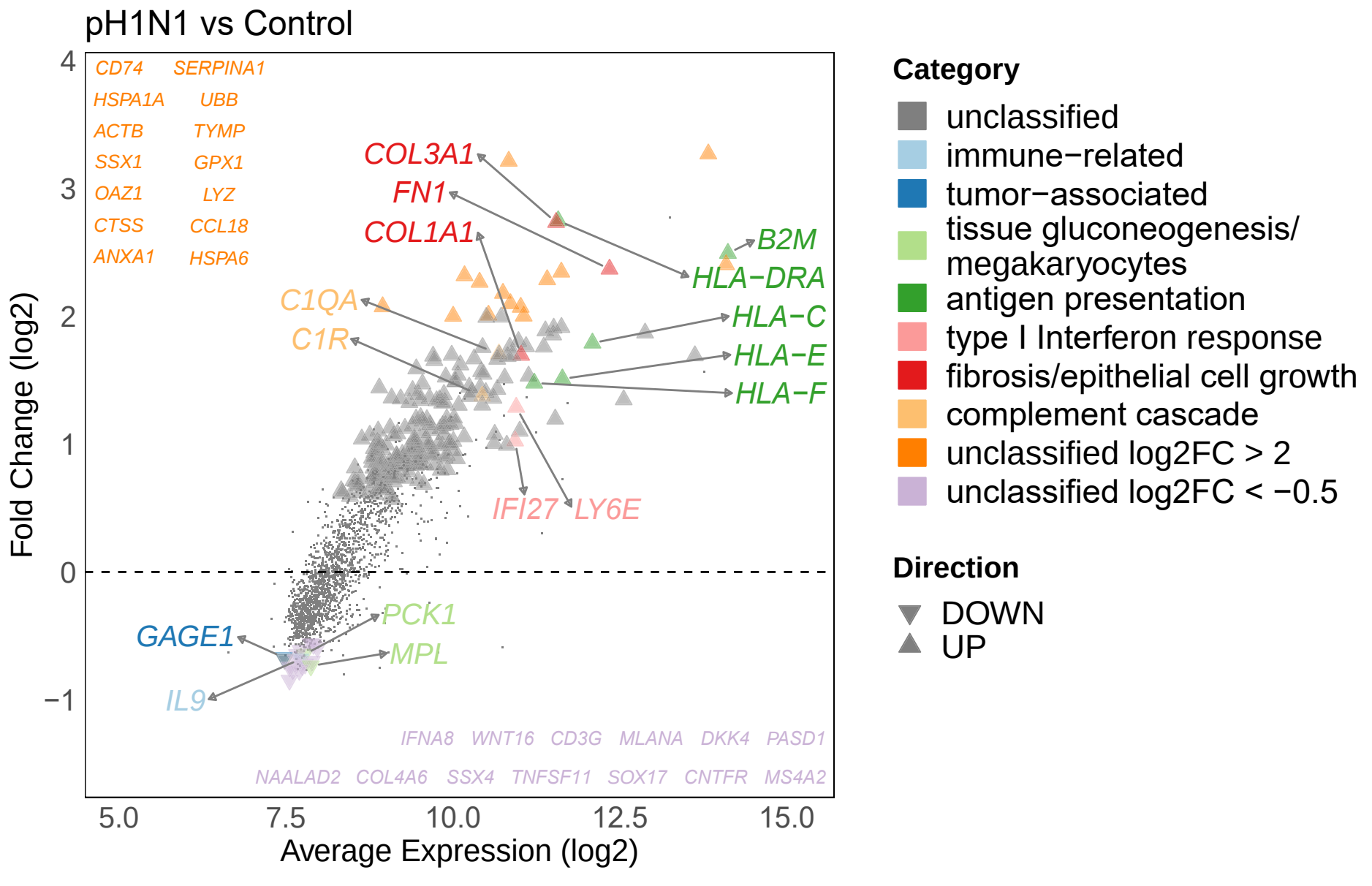
Kulasinghe *et al*, Supplementary Figure 3



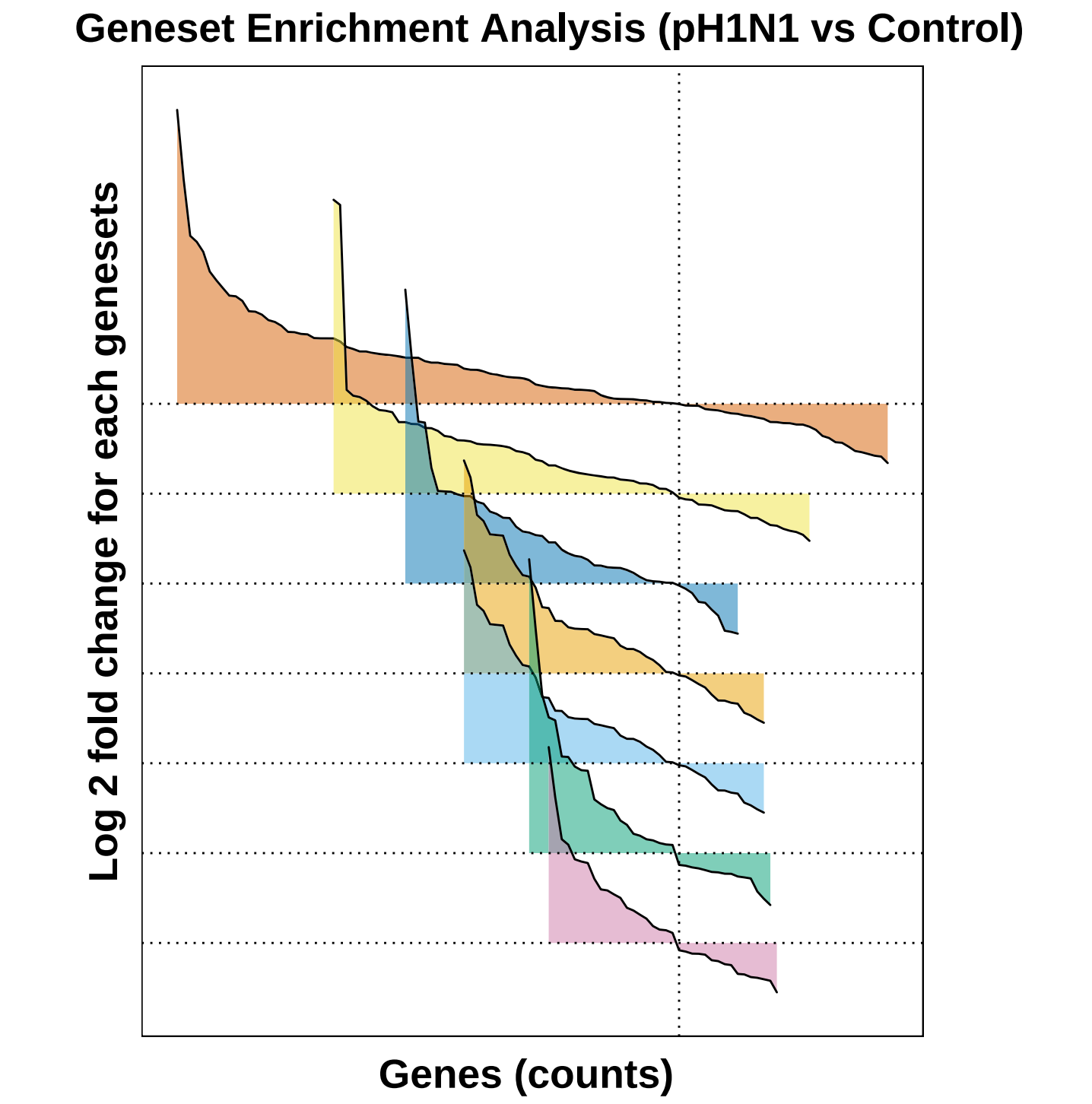
Kulasinghe *et al*, Supplementary Figure 4



Kulasinghe *et al*, Supplementary Figure 5



Kulasinghe et al, Supplementary figure 6



Kulasinghe et al, Supplementary figure 7