Supplementary Data Description file

Integrated histopathology, spatial and single cell transcriptomics resolve cellular drivers of early and late alveolar damage in COVID-19

Jimmy Tsz Hang Lee[1,*], Sam N. Barnett[2,3*], Kenny Roberts[1], Helen Ashwin[4], Luke Milross[5], Jae-Won Cho[6], Alik Huseynov[2], Benjamin Woodhams[1,7], Alexander Aivazidis[1], Tong Li[1], Joaquim Majo[8], Patricia Chaves[2], Michael Lee[2], Antonio M. A. Miranda[2], Zuzanna Jablonska[2], Vincenzo Arena[9], Brian Hanley[10], Michael Osborn[10], Virginie Uhlmann[7], Xiao-Ning Xu[11], Gary R. McLean[2,12], Sarah A. Teichmann[1,13], Anna M. Randi[2,3], Andrew Filby[14], Paul M. Kaye[4], Andrew J. Fisher[5,15,**,#], Martin Hemberg[6,**,#], Michela Noseda[2,3**,#], Omer Ali Bayraktar[1,**,#]

* These authors contributed equally
** These authors jointly supervised this work
# corresponding authors: a.j.fisher@newcastle.ac.uk, mhemberg@bwh.harvard.edu, m.noseda@imperial.ac.uk, ob5@sanger.ac.uk

1. Cellular Genetics Programme, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK.
2. National Heart and Lung Institute, Imperial College London, London, UK.
3. British Heart Foundation Centre of Research Excellence, Imperial College London, London, UK.
4. York Biomedical Research Institute, Hull York Medical School, University of York, York, UK.
5. Newcastle University Translational and Clinical Research Institute, Newcastle upon Tyne, UK.
6. The Gene Lay Institute of Immunology and Inflammation, Brigham and Women's Hospital, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA.
7. European Bioinformatics Institute, European Molecular Biology Laboratory (EMBL), Cambridge, UK.
8. Department of Cellular Pathology, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.
9. Department of Woman and Child Health and Public Health, Fondazione Policlinico Universitario A. Gemelli IRCCS, Istituto di Anatomia Patologica, Università Cattolica Del Sacro Cuore, Rome, Italy.
10. Department of Cellular Pathology, Northwest London Pathology, Imperial College London NHS Trust, London, UK
11. Department of Infectious Disease, Imperial College London, London, UK.
12. London Metropolitan University, London, UK.
13. Cambridge Stem Cell Institute & Department of Medicine, University of Cambridge, Cambridge, UK.
14. Biosciences Institute and Innovation, Methodology and Application Research Theme, Newcastle University, Newcastle upon Tyne, UK.
15. Institute of Transplantation, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.

Supplementary Data 1: COVID-19 patient metadata used for snRNA-seq / RNAscope.

Supplementary Data 2: Datasets used for generation of COVID-19 sc-/snRNA-seq object.

Supplementary Data 3: Cell state abbreviations.

Supplementary Data 4: Cell state differential gene expression analysis. Differentially expressed genes were calculated using the Wilcoxon Rank Sum Test with Benjamini-Hochberg adjustment.

Supplementary Data 5: EdgeR Pseudobulk analysis of sc-/snRNA-seq data (COVID-19 vs. Healthy / cell state). Differentially expressed genes were calculated using QLF (Quasi-Likelihood F-test).

Supplementary Data 6: MSigDB Pathway enrichment analysis in EP.AT1 (COVID-19 vs. Healthy / cell state).

Supplementary Data 7: MSigDB Pathway enrichment analysis in EP.AT2 (COVID-19 vs. Healthy / cell state).

Supplementary Data 8: COVID-19 patient metadata used for WTA profiling.

Supplementary Data 9: EdgeR Pseudobulk analysis of EDAD vs ODAD spatial WTA. Differentially expressed genes were calculated using QLF (Quasi-Likelihood F-test). P values were adjusted for multiple comparisons using the Benjamini-Hochberg method.

Supplementary Data 10: EdgeR Pseudobulk analysis of EDAD vs PRES spatial WTA. Differentially expressed genes were calculated using QLF (Quasi-Likelihood F-test). P values were adjusted for multiple comparisons using the Benjamini-Hochberg method.

Supplementary Data 11: EdgeR Pseudobulk analysis of ODAD vs PRES spatial WTA. Differentially expressed genes were calculated using QLF (Quasi-Likelihood F-test). P values were adjusted for multiple comparisons using the Benjamini-Hochberg method.

Supplementary Data 12: EdgeR Pseudobulk analysis of MDAD vs PRES spatial WTA. Differentially expressed genes were calculated using QLF (Quasi-Likelihood F-test). P values were adjusted for multiple comparisons using the Benjamini-Hochberg method.

Supplementary Data 13: SVG format of Supplementary Figure 1D.