

## Supplementary Material

### ***ELF5* is a respiratory epithelial cell-specific risk gene for severe COVID-19**

Maik Pietzner<sup>1,2,a</sup>, Robert Lorenz Chua<sup>3,a</sup>, Eleanor Wheeler<sup>2</sup>, Katharina Jechow<sup>3</sup>, Helena Radbruch<sup>4</sup>, Saskia Trump<sup>5</sup>, Bettina Heidecker<sup>6</sup>, Frank L. Heppner<sup>4,7,8</sup>, Roland Eils<sup>3,9,10</sup>, Marcus A. Mall<sup>10,11,12</sup>, Leif-Erik Sander<sup>13</sup>, Irina Lehmann<sup>5,10</sup>, Sören Lukassen<sup>3</sup>, Nick Wareham<sup>2</sup>, Christian Conrad<sup>3,b</sup>, Claudia Langenberg<sup>1,2,b\*</sup>

#### **Affiliations**

<sup>1</sup>Computational Medicine, Berlin Institute of Health (BIH) at Charité – Universitätsmedizin Berlin, Germany

<sup>2</sup>MRC Epidemiology Unit, University of Cambridge, Cambridge, UK

<sup>3</sup>Center for Digital Health, Berlin Institute of Health (BIH) at Charité – Universitätsmedizin Berlin, Germany

<sup>4</sup>Department of Neuropathology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin und Humboldt-Universität zu Berlin, Berlin, Germany

<sup>5</sup>Molecular Epidemiology Unit, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin and Berlin Institute of Health (BIH), Berlin, Germany

<sup>6</sup>Department of Cardiology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin und Humboldt-Universität zu Berlin, Berlin, Germany

<sup>7</sup>Cluster of Excellence, NeuroCure, Berlin, Germany

<sup>8</sup>German Center for Neurodegenerative Diseases (DZNE) Berlin, Berlin, Germany

<sup>9</sup>Health Data Science Unit, Heidelberg University Hospital and BioQuant, Heidelberg, Germany

<sup>10</sup>German Center for Lung Research (DZL), associated partner site, Augustenburger Platz 1, 13353 Berlin, Germany

<sup>11</sup>Department of Pediatric Respiratory Medicine, Immunology and Critical Care Medicine, Charité-Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

<sup>12</sup>Berlin Institute of Health at Charité – Universitätsmedizin Berlin, Berlin, Germany

<sup>13</sup>Department of Infectious Diseases and Respiratory Medicine, Charité - Universitätsmedizin Berlin, corporate member of Freie Universität Berlin, Humboldt-Universität zu Berlin, and Berlin Institute of Health (BIH), Berlin, Germany

<sup>a</sup>these authors contributed equally to this work

<sup>b</sup>these authors jointly supervised this work

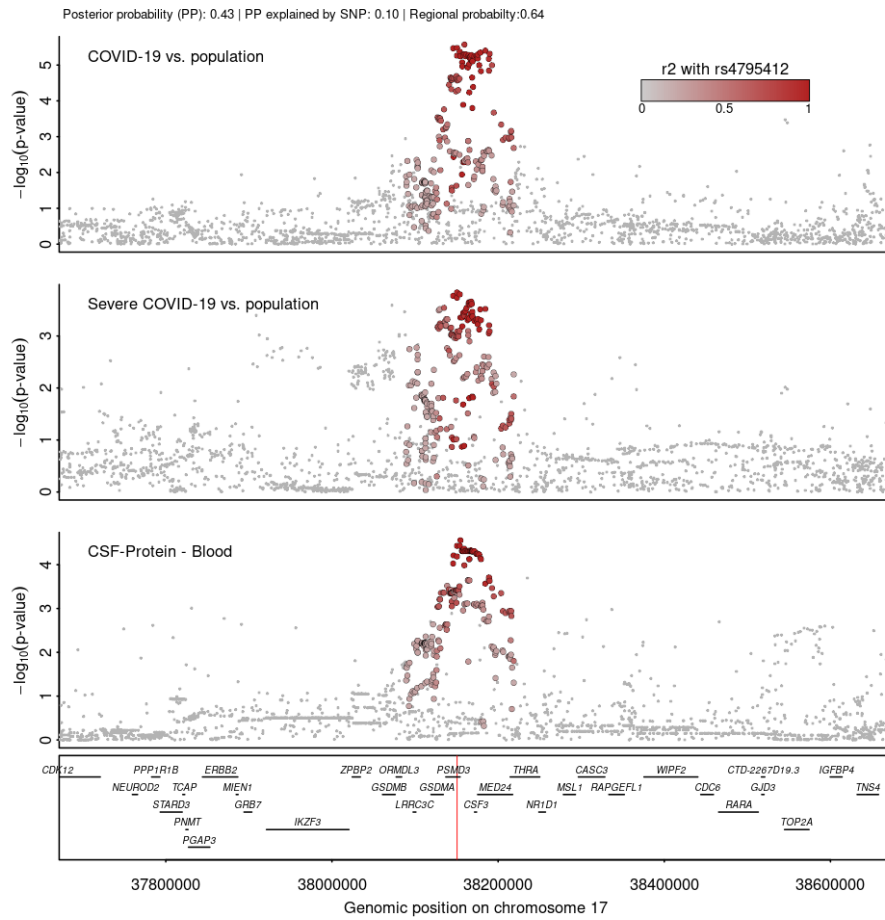
#### **Correspondence**

Prof Claudia Langenberg ([claudia.langenberg@bih-charite.de](mailto:claudia.langenberg@bih-charite.de))

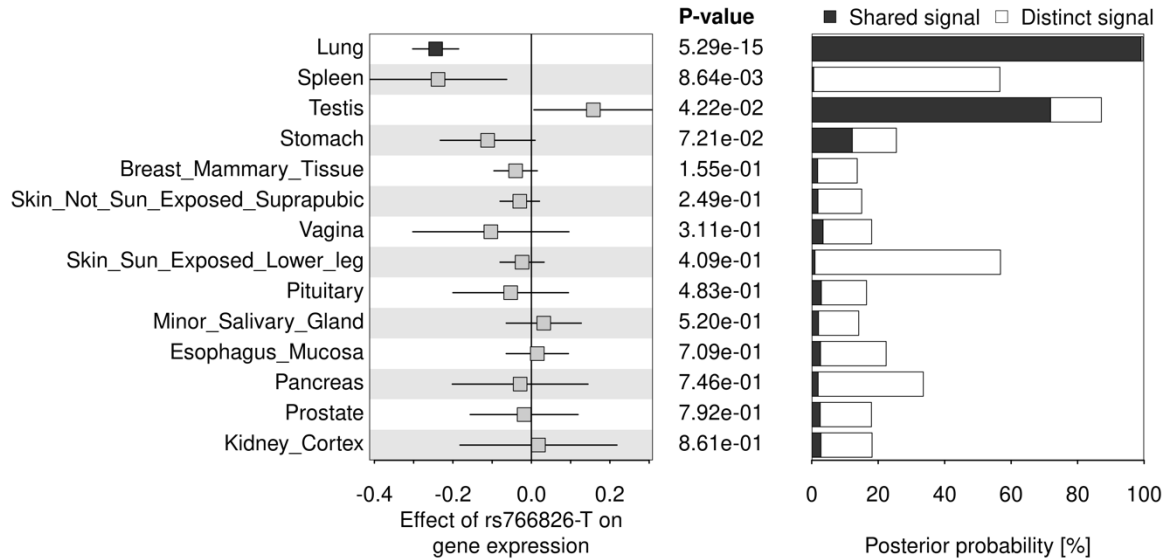
Prof Christian Conrad ([christian.conrad@bih-charite.de](mailto:christian.conrad@bih-charite.de))

Dr Maik Pietzner ([maik.pietzner@bih-charite.de](mailto:maik.pietzner@bih-charite.de))

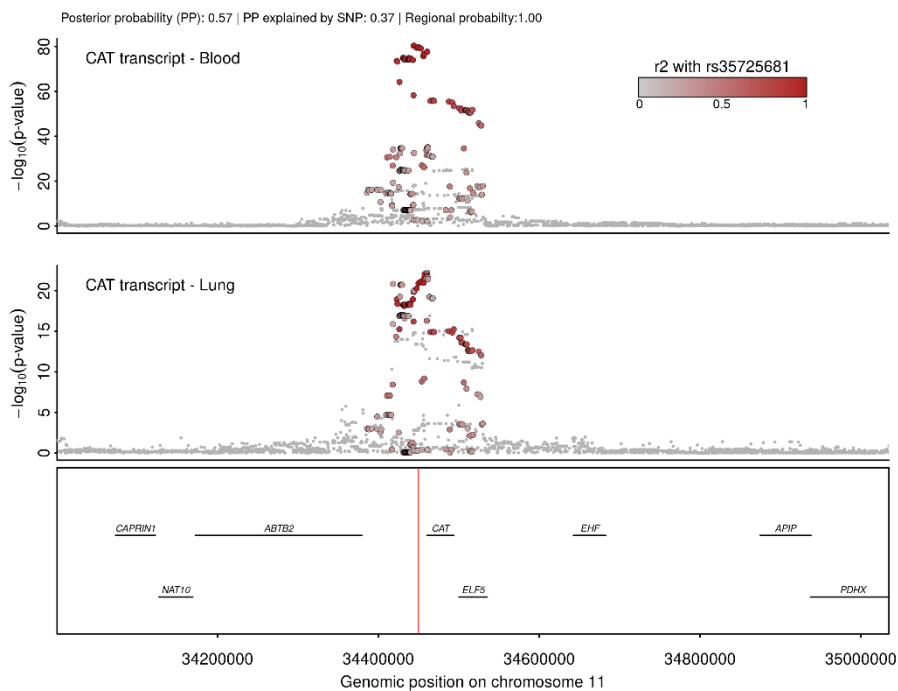
## FIGURES



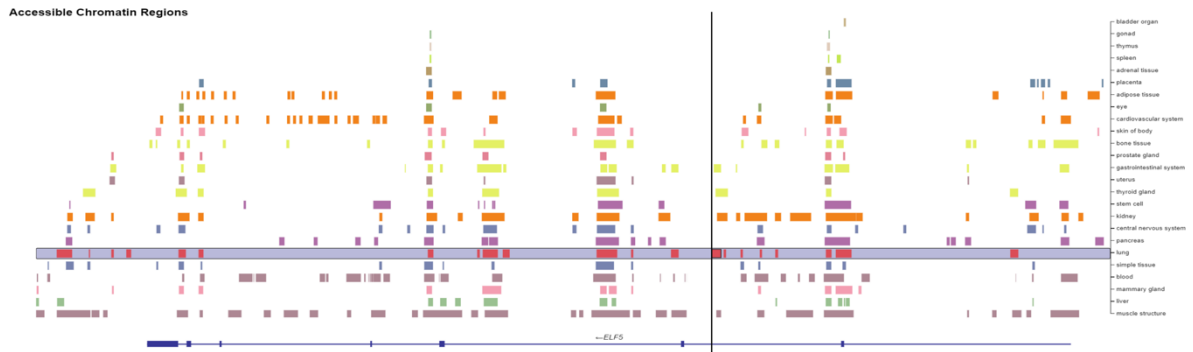
**Supplementary Figure 1 Stacked regional association plots at *CSF3*.** Each panel contains regional association statistics (p-values) for the trait listed in the upper left corner along genomic coordinates. Each dot represents a single nucleotide polymorphisms and colours indicate linkage disequilibrium (LD;  $r^2$ ) with the most likely causative variant (rs4795412) at this locus (darker colours stronger LD).



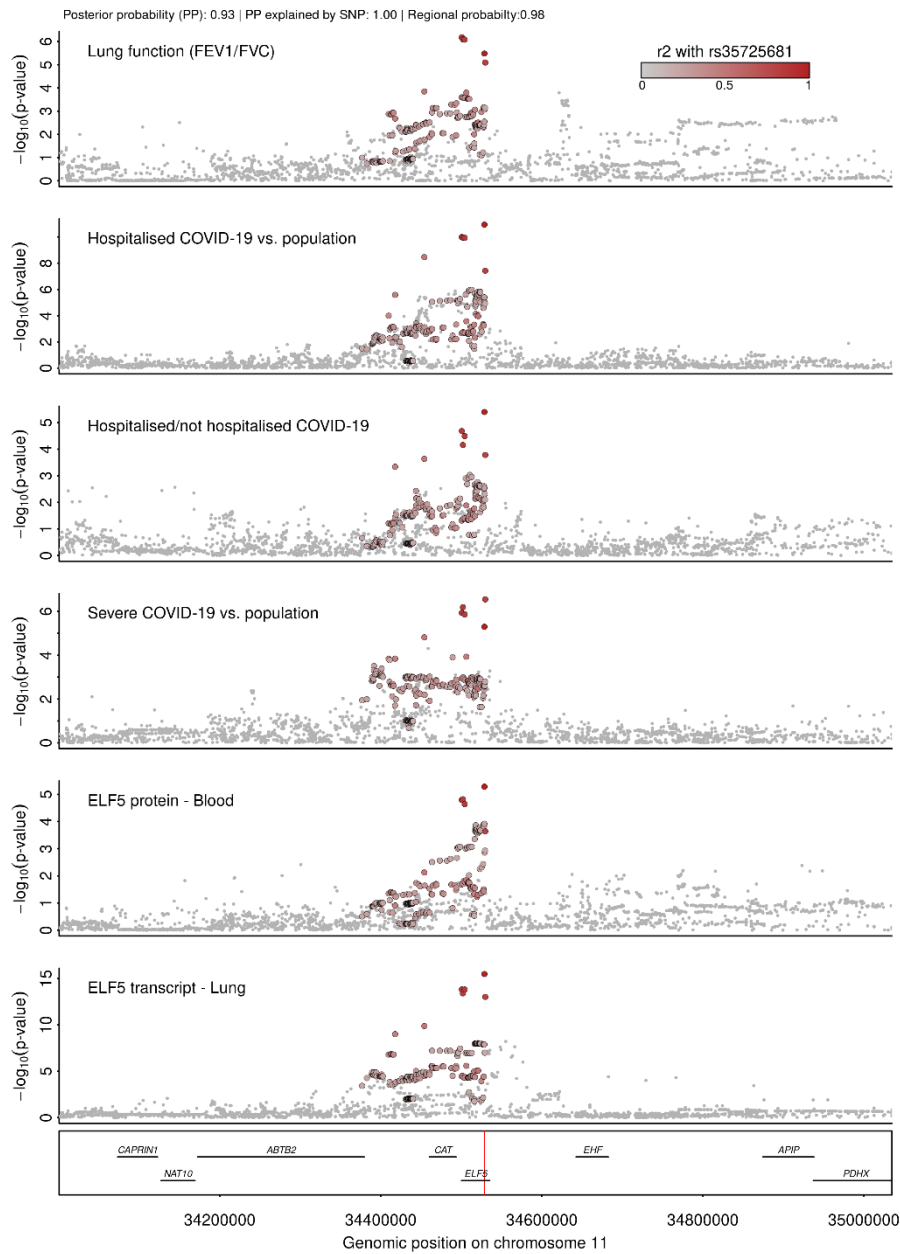
**Supplementary Figure 2** Summary of cross-tissue colocalisation for *ELF5*. The left panel shows effect estimates for rs766826 on *ELF5* expression in all tissues with detectable levels in the GTEx v8 resource. Significant effects ( $p < 0.001$ ) are highlighted in black. The left hand side shows posterior probabilities for testing for a shared genetic signal between *ELF5* abundance in plasma and *ELF5* expression in each tissue.



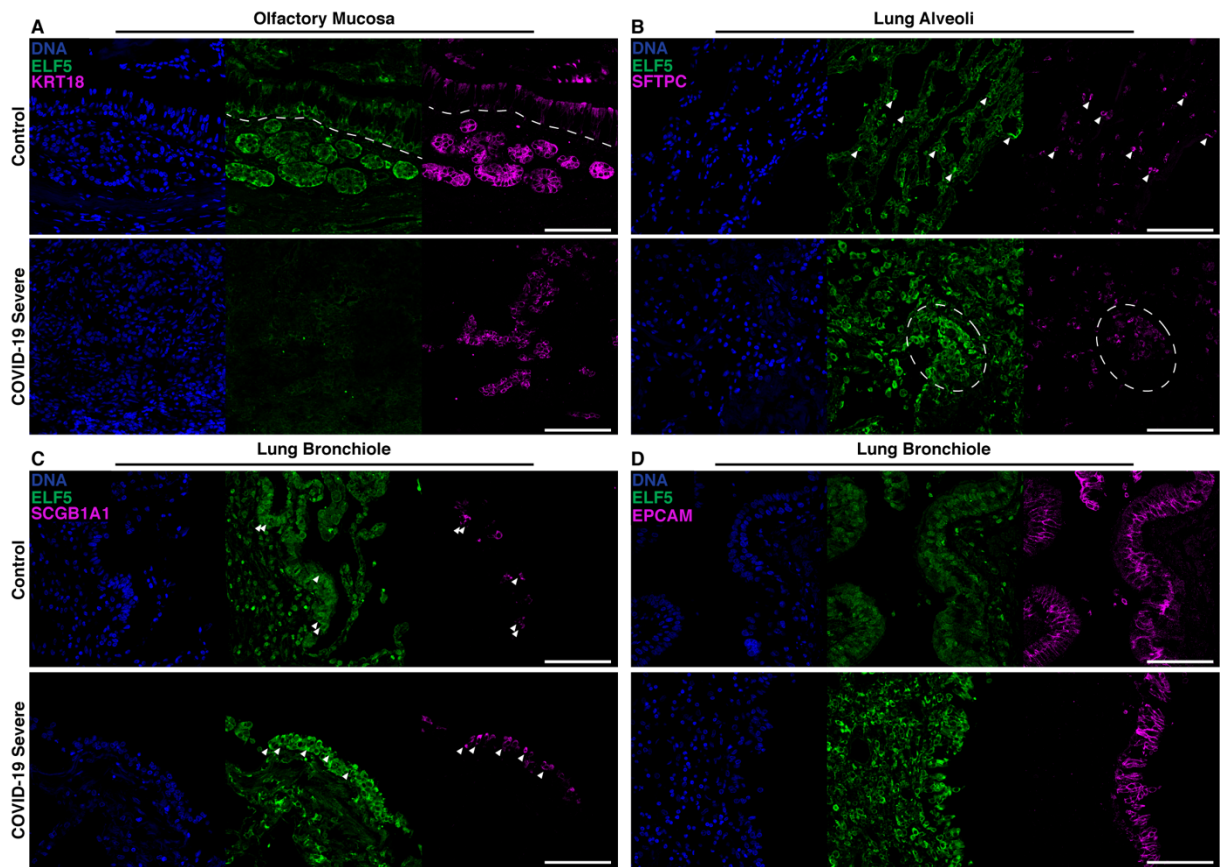
**Supplementary Figure 3** Stacked regional association plots at *CAT*. Each panel contains regional association statistics (p-values) for the trait listed in the upper left corner along genomic coordinates. Each dot represents a single nucleotide polymorphisms and colours indicate linkage disequilibrium (LD;  $r^2$ ) with the most likely causative variant (rs35725681) at this locus (darker colours stronger LD).



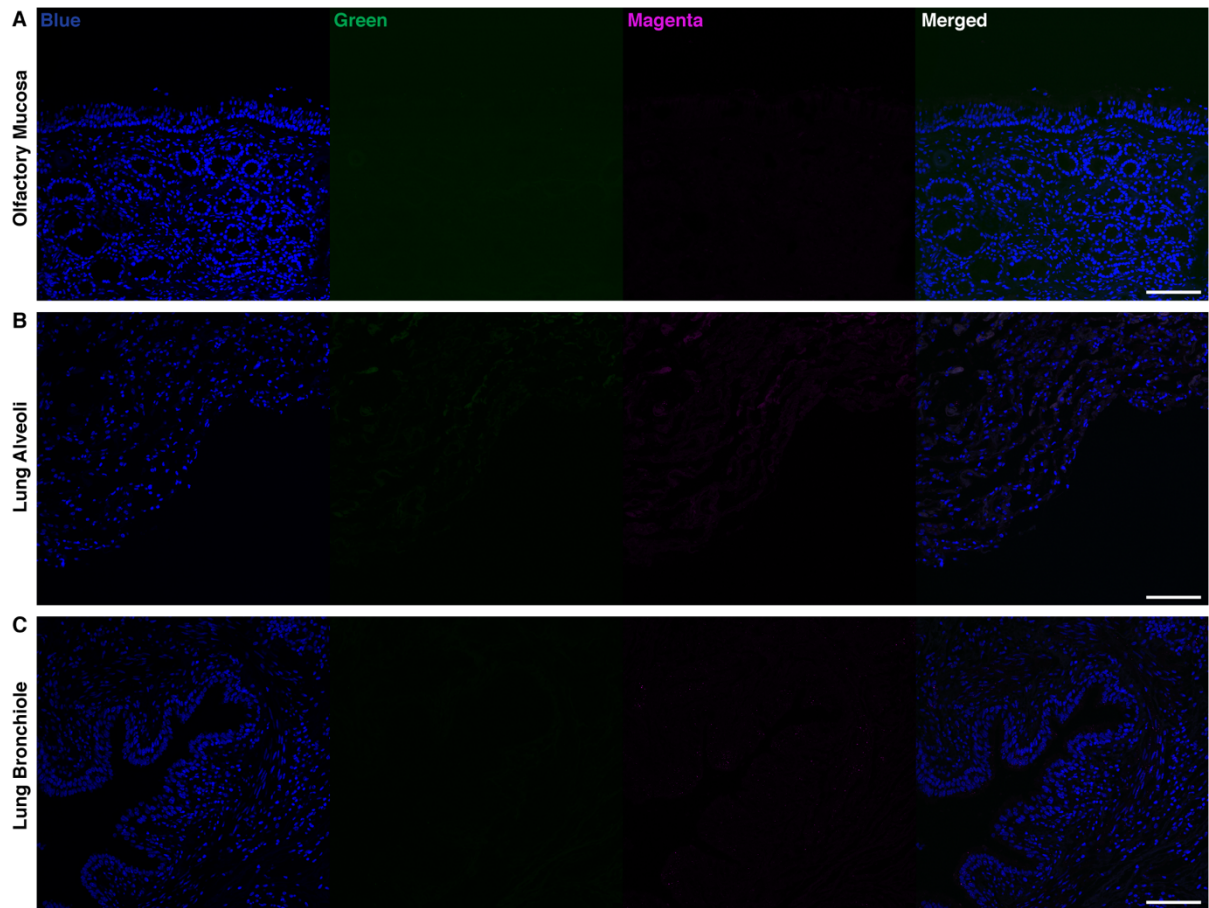
**Supplementary Figure 4 Open chromatin regions at ELF5 across various tissues.** Each colour bar indicates open chromatin regions identified using ATAC-seq experiments. The relevant lung-specific region is highlighted. Data was obtained from <https://t2d.hugeamp.org/variant.html?variant=rs766826> and open chromatin regions in lung are based on experiments in alveolar type 2 cells (<https://cmdga.org/annotations/DSR063NOE/>). The position of rs766826 is highlighted by a black bar.



**Supplementary Figure 5 Stacked regional association plots at *ELF5*.** Each panel contains regional association statistics (p-values) for the trait listed in the upper left corner along genomic coordinates. Each dot represents a single nucleotide polymorphisms and colours indicate linkage disequilibrium (LD;  $r^2$ ) with the most likely causative variant (rs766826) at this locus (darker colours stronger LD). This figure is similar to Figure 2 in the main text, but now also including summary statistics for the FEV1/FCV ratio as a measure of lung function taken from Shrine et al.<sup>1</sup>.

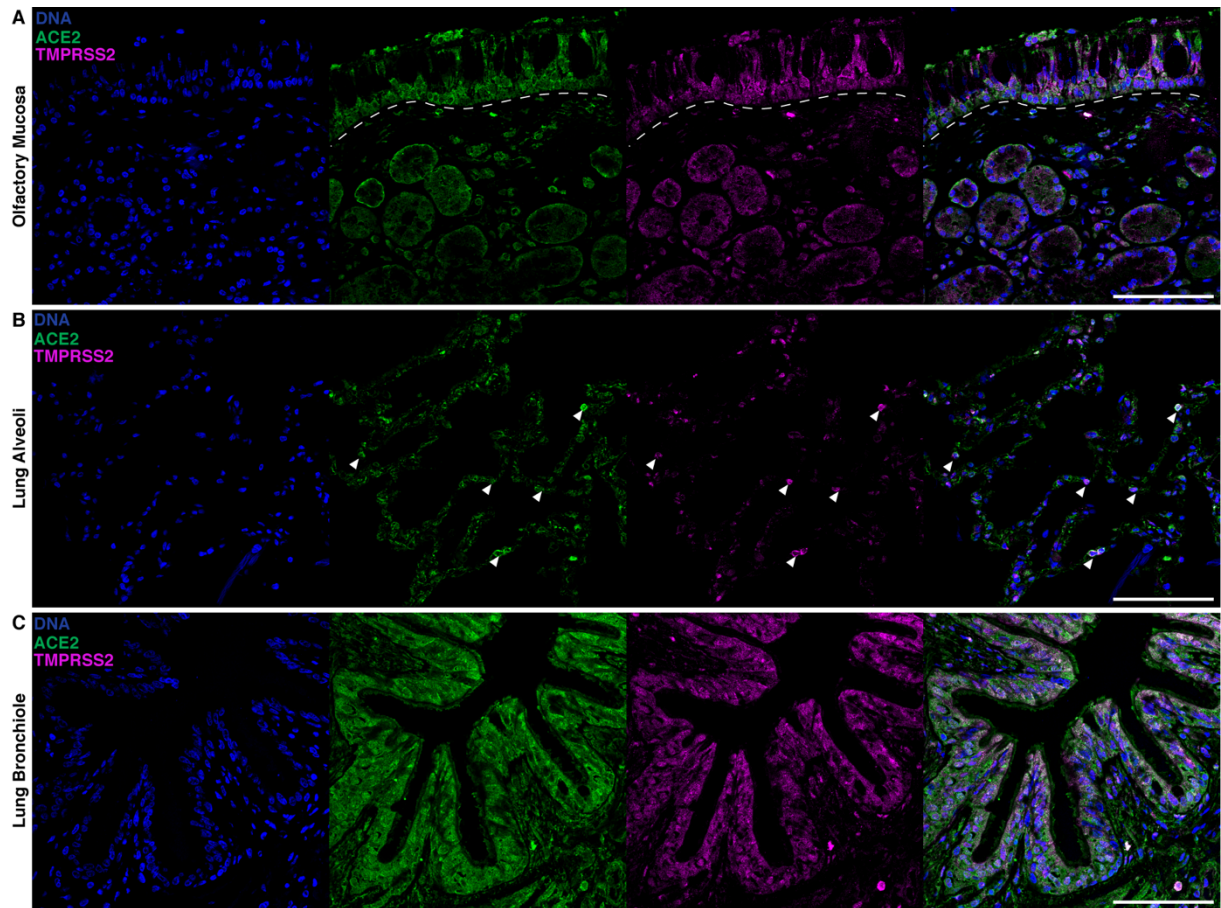


**Supplementary Figure 6 ELF5 expression by epithelial cells of the olfactory mucosa and lung.** Immunofluorescent staining of ELF5 in separate channels in control and COVID-19 patients in the **A** olfactory mucosa, **B** lung alveoli, and **C** lung bronchiole. **A** Dashed lines separate the olfactory epithelium and the lamina propria. **B** Arrowheads highlight AT2 cells expressing ELF5; dashed outline highlights clusters of AT2 cells expressing ELF5. **C** left: Arrowheads highlight secretory cells expressing ELF5; right: arrowheads highlight airway epithelial cells expressing ELF5. Marker genes for sustentacular and Bowman gland cells (**A**, KRT18), alveolar type II cells (**B**, SFTPC), secretory cells (**C**, SCGB1A1), and epithelial cells (**C**, EPCAM) are shown in purple. Validation staining for each tissue: control (n = 2); COVID-19 (n = 2). Scale bar = 100 $\mu$ m.



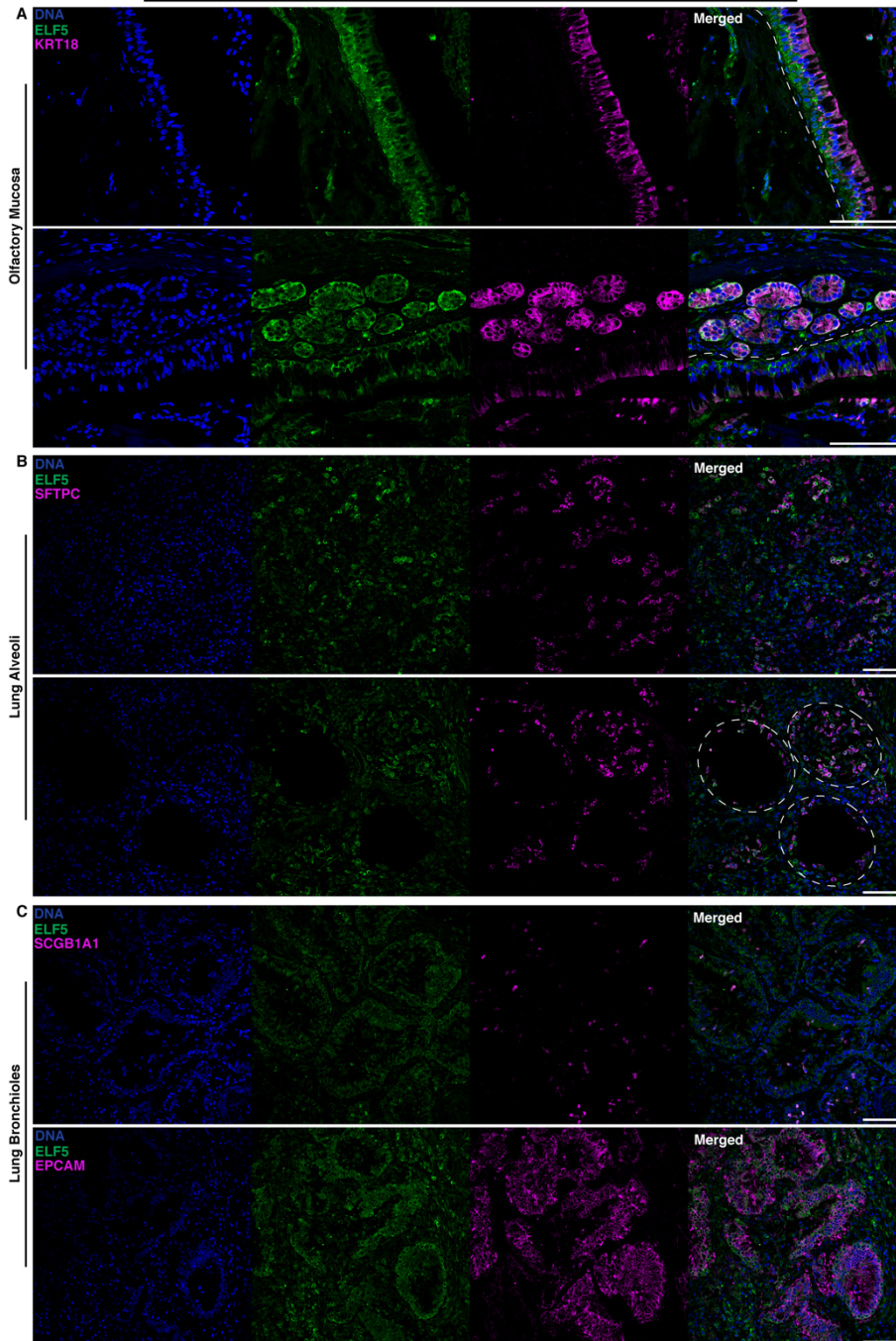
**Supplementary Figure 7. No primary antibody controls of the olfactory mucosa and lung.** Representative images of negative controls for the **A** olfactory mucosa, **B** lung alveoli, and **C** lung bronchiole under high laser power. Scale bar = 100 $\mu$ m.



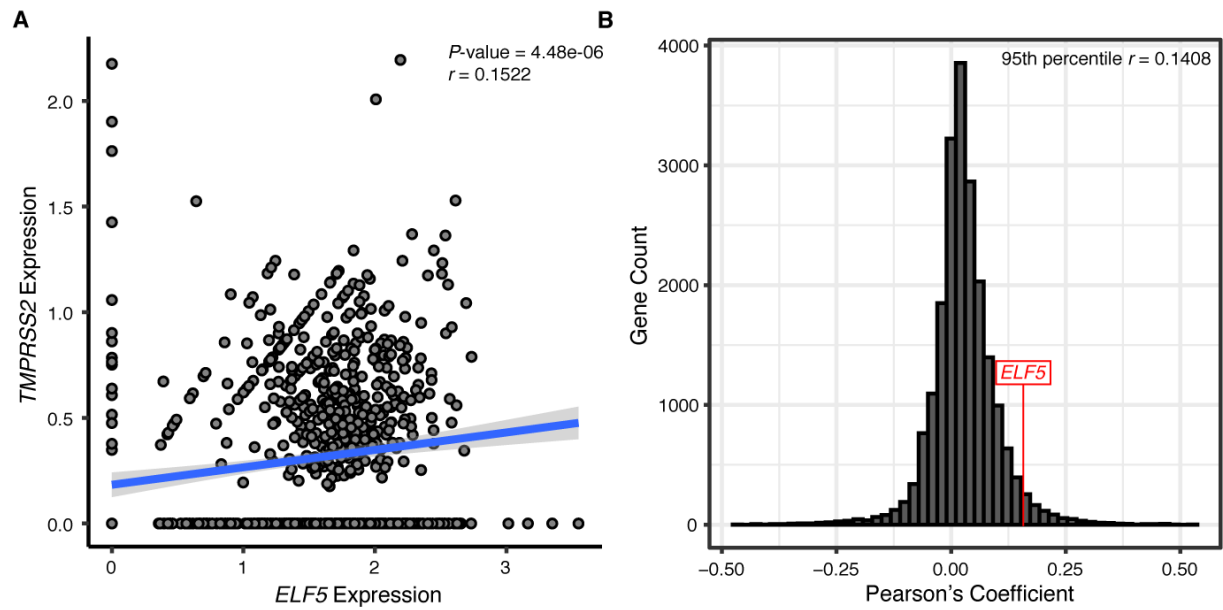


**Supplementary Figure 8. ACE2 and TMPRSS2 are expressed by different epithelial cells of the olfactory mucosa and the lung. A** Sustentacular cells, horizontal basal cells, and Bowman gland cells of the olfactory mucosa co-express ACE2 and TMPRSS2. **B** Arrows highlight punctuated expression of TMPRSS2 reflecting the distribution of AT2 cells together with ACE2. **C** Airway epithelial cells of the bronchioles expressing ACE2 and TMPRSS2. Validation staining for each tissue: control (n = 2). Scale bar = 100 $\mu$ m.



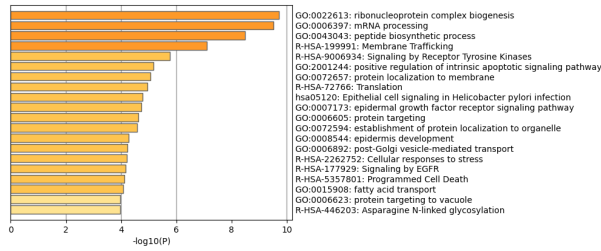


**Supplementary Figure 9. ELF5 expression by epithelial cells of the olfactory mucosa and lung in COVID-19 cases with longer disease duration.** Immunofluorescent staining of ELF5 in separate channels of COVID-19 patients that died after hospitalization after 14 days. ELF5 expression in the **A** olfactory mucosa, **B** lung alveoli, and **C** lung bronchiole. **A** Dashed lines separate the olfactory epithelium and the lamina propria. **B** dashed outline highlights clusters of AT2 cells expressing ELF5. Marker genes for sustentacular and Bowman gland cells (**A**, KRT18), alveolar type II cells (**B**, SFTPC), secretory cells (**C**, SCGB1A1), and epithelial cells (**C**, EPCAM) are shown in purple. Validation staining for each tissue: control (n = 2); COVID-19 (n = 2). Scale bar = 100 $\mu$ m.

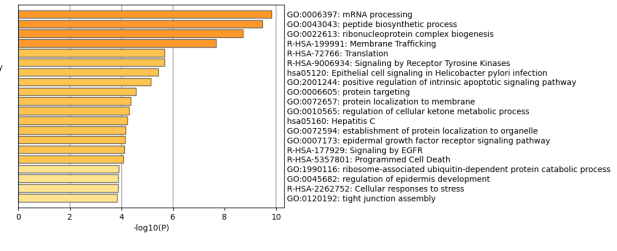


**Supplementary Figure 10. Co-expression of ELF5 and TMPRSS2.** A Scatterplot opposing normalised expression levels of ELF5 and TMPRSS2 in sustentacular cells of the olfactory mucosa. The blue line indicates a linear regression fit and correlation coefficient and p-value are given in the legend. B Distribution of correlation coefficients of pairwise gene expression across all genes detected in sustentacular cells. The red line indicates where the ELF5 – TMPRSS2 correlations is placed.

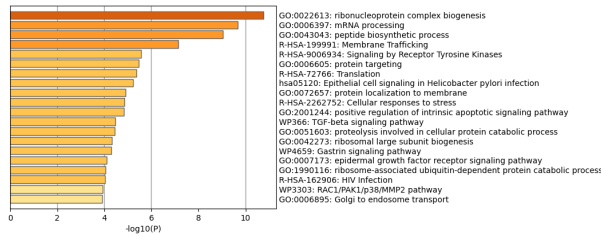
### Olfactory mucosa – Secretory cells



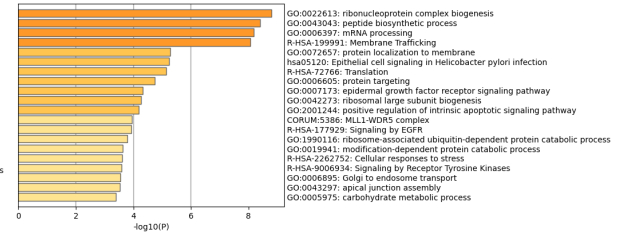
### Olfactory mucosa – Sustentacular cells



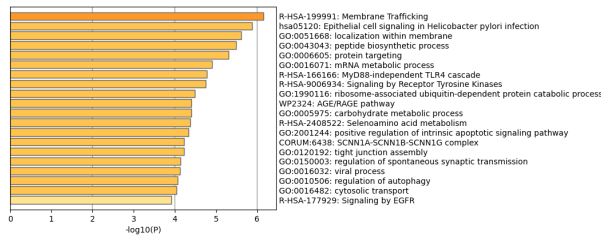
### Olfactory mucosa – Bowman cells



### Nasopharynx – Secretory cells



### Lung – Secretory cells



**Supplementary Figure 11.** Results from cell-type specific pathway enrichment analysis for predicted ELF5 targets.

