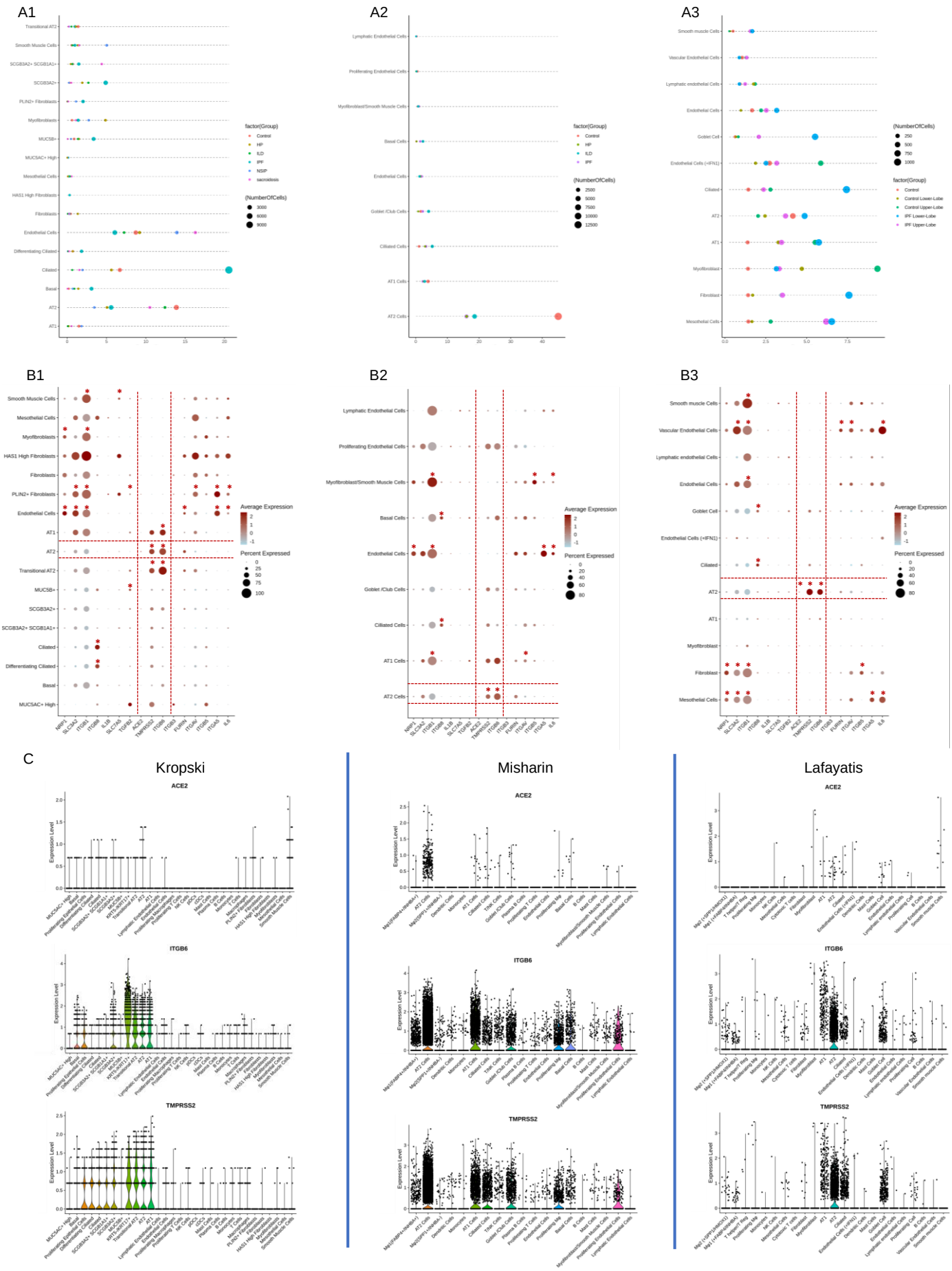
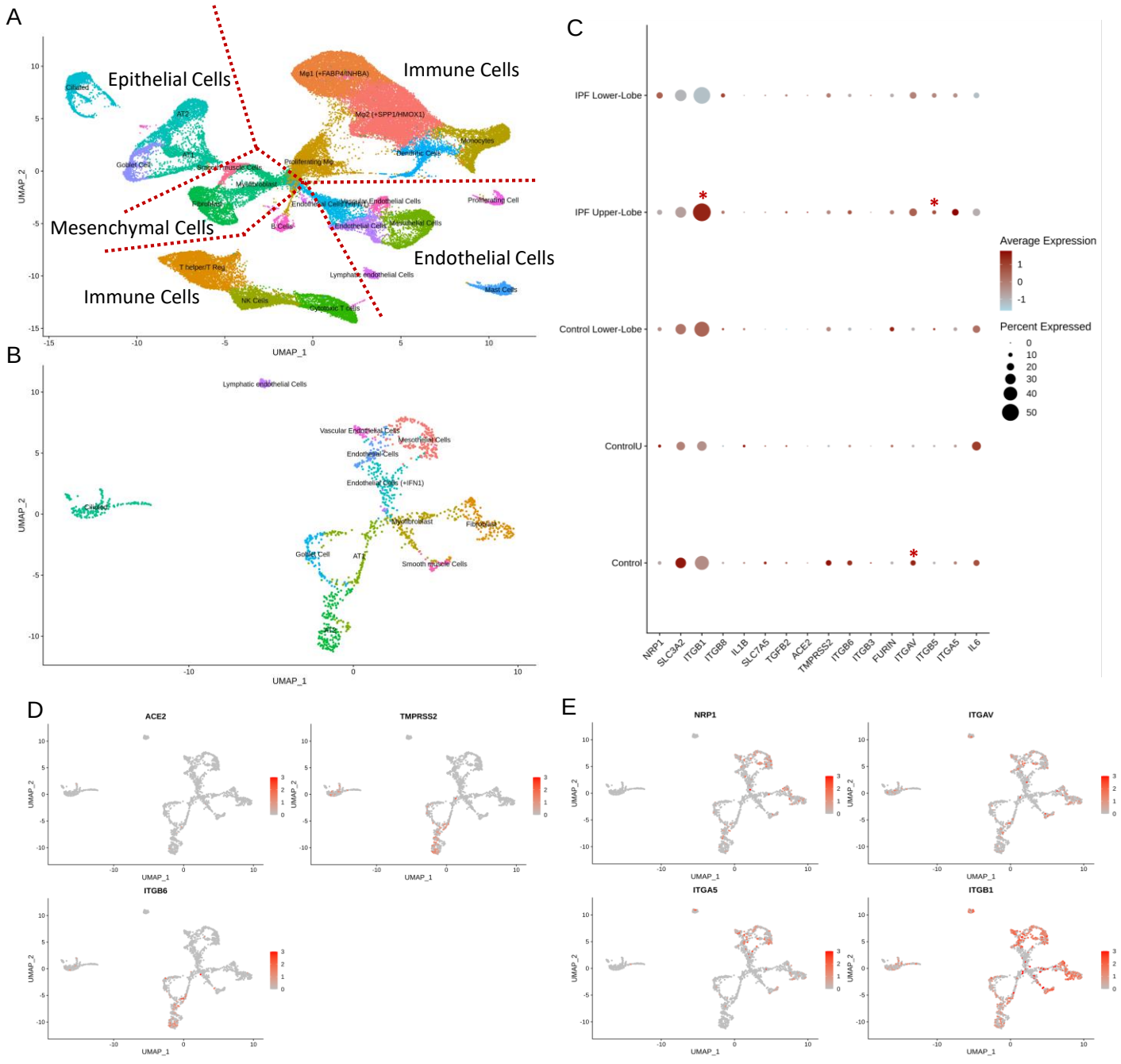
Gene expression and cellular distribution across datasets



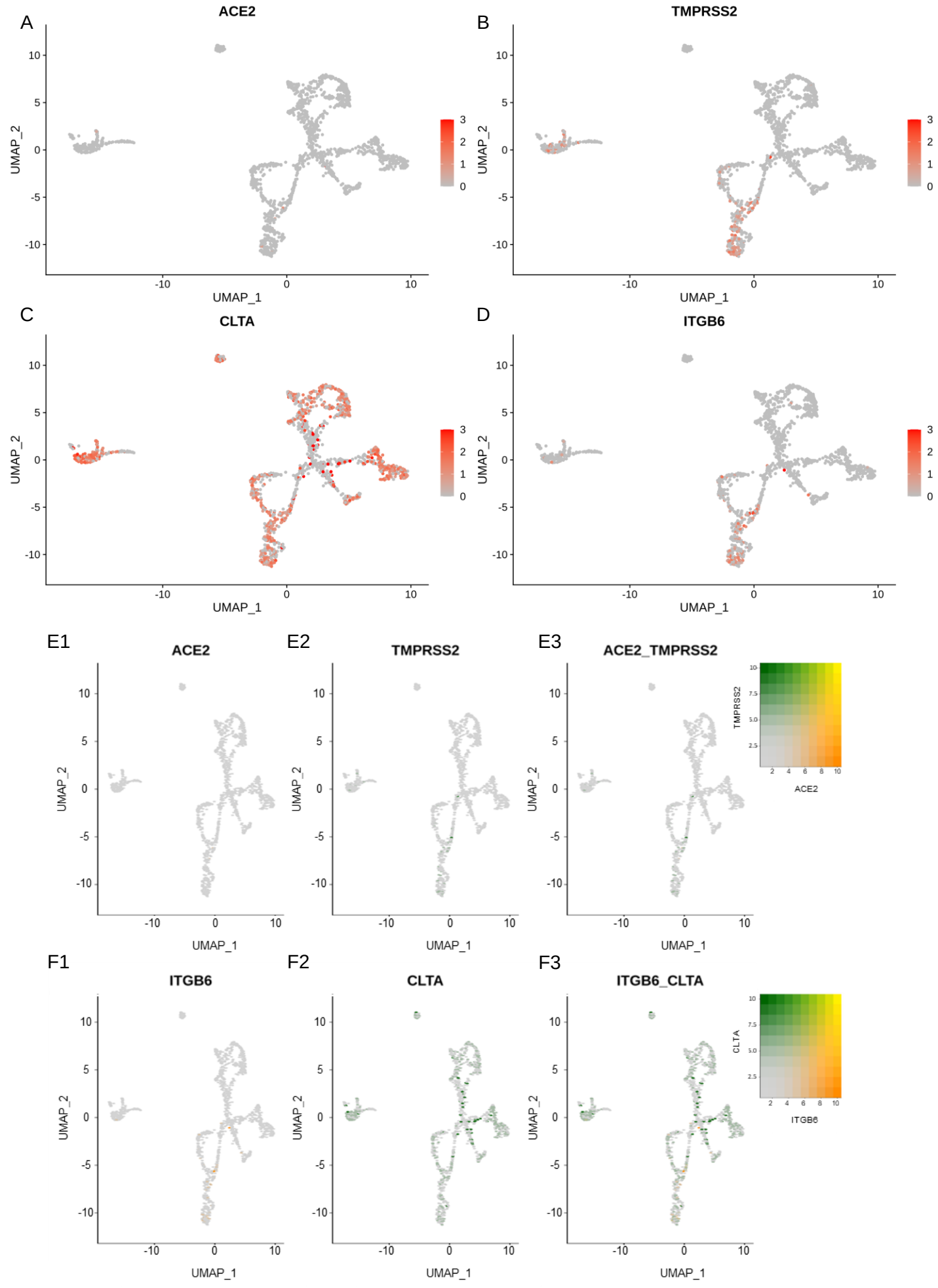
A) Dot plot of the quantification of cells as a percentage of all the endothelial, epithelial and mesenchymal cells located on the donor biopsies defined in different colours by lung location, disease status and study 1) Kropski, 2) Misharin 3) Lafayatis. The size of the dot represents the number of each cell type found in each donor. B) Dot plot of defining potential genes able to mediate on the SARS-CoV-2 virus cellular entry such as integrins, ACE2 and TMPRSS2 and receptors involved in cellular inflammatory responses, such as IL-6 and IL-1. Dot size represents fraction of cells within cell type expressing a given gene, and colour intensity represents count-based expression amounts (log(scaled UMI+1)) among expressing cells. Red stars indicates statically significantly largest proportion of target genes ($p < 0.05$ FDR correction) in relation to the other cellular subsets. The parallel red dot lines indicate the cell type with largest proportion of ACE2+ (1.5%- 3%) TMPRSS2+ (36% - 56%) in each dataset 1) Kropski, 2) Misharin, 3) Lafayatis (Bonferroni-adjusted p-value < 0.05). F) Violin plots represent the gene expression of cell membrane point of entry for SARS-CoV-2 ACE2 and TMPRSS2 across all the cell population and the integrin ITGB6, which is significantly upregulated in AT2 cells.

Gene expression and cellular distribution Lafyatis et.al, 2019



Computational analysis to identify cell types by using unbiased methods. Shown on the top left is a A) UMAP projection of ~53K cells across 9 donors (PF (n:3) and control lungs (n:6)). Cells represented by points coloured according to cell type and divided with dashed red lines on cellular ontogenies (immune, endothelial, epithelial and mesenchymal). B) UMAP clustering of selected cells for further analysis (Endothelial, Epithelial and Mesenchymal cells) represented by points coloured accordingly to cellular function. C) Dot-blot representing the disease status and lung location versus gene expression. The red stars represent a significant gene enrichment (Bonferroni-adjusted p-value < 0.05). Manifold UMAP projection of potential receptors facilitating the entry of SARS-CoV-2 in D) AT2 (epithelial cell) – ACE2 (coronavirus receptor –top left), TMPRSS2 (coronavirus S priming for entry –top right), and ITGB6 (RGD affinity receptor) E) In mesothelial cells (mesenchymal cells) are found the potential coronavirus receptors – top left NPR1, top right ITGAV and on the lower row ITGA5 (left) and ITGB1 (right) .Colour coding as follows: grey mRNA negative, red positive.

Gene expression and cellular distribution Lafyatis et.al, 2019



UMAP projections of epithelial, endothelial, mesenchymal cells (18609 cells) as coloured in different red intensities (representing gene expression amounts –log(scaled UMI+1)) or grey (as negative) of A) ACE2, B) TMPRSS2, C) CLTA and D) ITBG6.

E) UMAP representation of ACE2^{+} (E.1) coloured yellow and TMPRSS2^{+} (E.2 green) cells. E.3) Double positive ACE2^{+} TMPRSS2^{+} cells mRNA expressing both genes (coloured green-yellow) comprise around 0.05 % of type II pneumocytes. These double positive cells were omitted for this analysis since too few cells were observed. F) The lower UMAP projections represented in yellow dots are F.1) ITGB6+ cells (grey are negative cells) and F.2) the UMAP projection coloured with green dots represent CLTA+. The (F.3) expression of ITGB6+ CLTA+ is enriched in type II pneumocytes ($p<0.0001$, $n=22$). Statistical significance assessed by Fisher Exact Test.