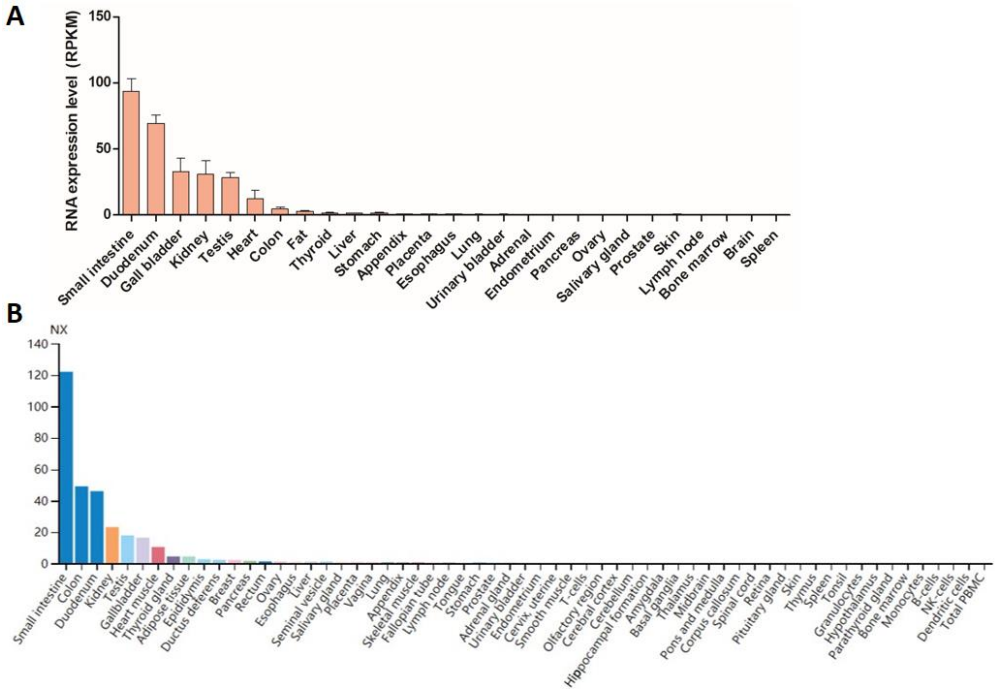

Supplementary Figure 1 The expression profiles of ACE2 in normal human tissues. (A) The expression profile of ACE2 from NCBI. (B) The expression profile of ACE2 from The Human Protein Atlas.

Methods

Inter-group difference test

The original information of patients infected by 2019-nCoV infection was retrieved from three recent publications[1-3]. The 2019-nCoV infection was laboratory-confirmed by real-time RT-PCR and next-generation sequencing. Categorical variables were expressed as number with percentage in parenthesis. All the variables were compared by Fisher's exact test for the three studies as one study has a relatively small sample size[1]. Continuous variables were expressed as estimated mean with inter-quartile range (IQR) in parenthesis. A two-sided α of less than 0.05 was considered statistically significant. Statistical analyses were performed using the R software, version 3.5.3.

ACE2 expression profile

The expression profiles of ACE2 from normal human tissues were obtained from public database NCBI (https://www.ncbi.nlm.nih.gov/) with accession number PRJEB4337. The original expression data was collected and then plotted by GraphPad Prism 5.

Single-cell sequencing analysis

The public single-cell RNA-seq sequencing data (GSE92332) was downloaded from the GEO database. The 10X matrix file of GSE92332_atlas_UMIcounts.txt was used for subsequent analysis based on R package Seurat (Version 3.1.2). In order to filter out low-quality cells and low-quality genes, strict parameters, "min.featrue=1000" and "min.cell=20", were used in the function CreateSeuratObject. The data was subsequently log-normalized by the function NormalizeData with the default parameters. The genes with highly variable expression were identified by the function FindVariableGenes. After the data were processed by the function ScaleData, PCA dimensionality reduction was performed utilizing the function RunPCA. Based on the analysis by the function of JackStraw and ScoreJackStraw, the first 15 PCA components were selected for further two-dimensional t-distributed stochastic neighbor embedding (tSNE). The setting of k.param was 30 in the function FindNeighbors, and the setting of the resolution was 1.5 in the function FindClusters. In the downloaded original gene-barcode matrix, Adam L.Haber and his coworkers had annotated the cell identities of each barcode. In our analysis, the cell types were identified by the expression of marker genes and the annotation by the original gene-barcode matrix. Gene expression of different cell types was illustrated by the functions of FeaturePlot and VlnPlot.

Reference:

1 Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. Lancet 2020.

2 Chen N ZM, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y, Xia J, Yu T, Zhang X, Zhang L. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 2020.
3 Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020.