European Review for Medical and Pharmacological Sciences 2021; 25: 2409-2414

Gene expression analysis identifies genes associated with SARS-COV-2 receptors, ACE2 and TMPRSS2, in normal human lung tissues

R.-W. HU[1,2], C. LIU[3], J.-P. GONG[1,2], Z.-X. CAO[1,2]

[1]Tongji Cancer Research Institute, Tongji Hospital, Tongji Medical College in Huazhong University of Science and Technology, Wuhan, Hubei, China
[2]Department of Gastrointestinal Surgery, Tongji Hospital, Tongji Medical College in Huazhong University of Science and Technology, Wuhan, Hubei, China
[3]Department of Radiology, The First Affiliated Hospital of Anhui Medical University, Hefei, Anhui, China

Abstract. – The COVID-19 (Corona Virus Disease 2019) outbreak, which seriously affected people's lives across the world, has not been effectively controlled. Previous studies have demonstrated that SARS-COV-2 (Severe acute respiratory syndrome coronavirus 2) infecting host cells mainly rely on binding to receptor proteins, namely ACE2 and TMPRSS2. COVID-19 transmission is faster than the severe acute respiratory syndrome (SARS) pneumonia outbreak in 2002. This is mainly attributed to the different pathways of virus-infected host cells, coupled with patients' atypical clinical characteristics. SARS-CoV-2 is mainly transmitted through respiratory droplets and contact, infecting lung tissues before damaging other body organs, such as the liver, brain, kidney and heart. The present study identified potential target genes for SARS-COV-2 receptors, ACE2 and TMPRSS2, in normal human lung tissue. The findings provide novel insights that will guide future drug development approaches for treatment of COVID-19.

Key Words:
SARS-COV-2, COVID-19, TMPRSS2, ACE2, Pneumonia.

Abbreviations

COVID-19 = Corona Virus Disease 2019; SARS-COV-2 = Severe acute respiratory syndrome coronavirus 2; GTEx = Genotype-Tissue Expression; ACE2 = Angiotensin Converting Enzyme 2, Protein Coding; TMPRSS2 = Transmembrane Serine Protease 2, Protein Coding; ZNF385B = Zinc Finger Protein 385B, Protein Coding; AHCYL2 = Adenosylhomocysteine Like 2, Protein Coding; SDR16C5 = Short Chain Dehydrogenase/Reductase Family 16C Member 5, Protein Coding; SNX30 = Sorting Nexin Family Member 30, Protein Coding; RHOBTB2 = Rho Related BTB Domain Containing 2, Protein Coding; TMEM163 = Transmembrane Protein 163, Protein Coding; TACC2 = Transforming Acidic Coiled-Coil Containing Protein 2, Protein Coding; SFTPB = Surfactant Protein B, Protein Coding; SNX25 = Sorting Nexin 25, Protein Coding; FASN = Fatty Acid Synthase, Protein Coding; CTSH = Cathepsin H, Protein Coding; DUOXA1 = Dual Oxidase Maturation Factor 1, Protein Coding; FREM2 = FRAS1 Related Extracellular Matrix 2, Protein Coding; DUOX1 = Dual Oxidase 1, Protein Coding; TFCP2L1 = Transcription Factor CP2 Like 1, Protein Coding)..

Introduction

The COVID-19 outbreak, that turned into a global pandemic, currently threatens public health worldwide[1]. SARS-COV-2, which mainly infects[2] host cells through the assistance of ACE2 and TMPRSS2[3-6], is transmitted through respiratory droplets and contact. Consequently, the virus infects human lung tissue before damaging other organs[7-10]. COVID-19 easily spreads throughout the world, owing to a lack of typical clinical symptoms in some infected patients during early stages of infection[11-13]. This is one of the main reasons why the disease became a global pandemic[14,15]. The Genotype-Tissue Expression (GTEx) database[16-18] contains whole genome sequence information of normal human tissues, hence represents a useful tool for identifying genes related to ACE2 and TMPRSS2 in normal human lung

Corresponding Authors: Zhixin Cao, MD; e-mail: zxcao@tjh.tjmu.edu.cn
Jianping Gong, MD; e-mail: jpgong@tjh.tjmu.edu.cn

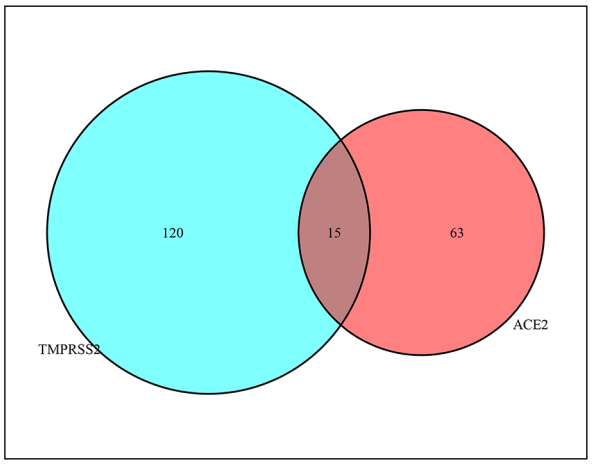

Figure 1. Screening the intersection related genes of ACE2 and TMPRSS2.

tissues. Identification of such homologs could reveal potential therapeutic targets of SARS-COV-2, thereby set up a platform for future development of drugs to treat the disease. To obtain whole genome sequencing information of normal lung tissues, we downloaded data from UCSC Xena (https://xenabrowser.net/datapages/) website[19], then used the Pearson correlation coefficient to understand the relationship between the genes. All data analyses were performed using packages implemented in R software (version: 3.6.0), with data followed by $p<0.05$ considered statistically significant.

Finally, 288 patients with normal lung tissue were included. Pearson correlation allowed identification of 78 genes (cor≥0.45) that were statistically significant (**Supplementary Table I**). Genes related to TMPRSS2 expression, with statistical significance, as well as 135 genes (cor≥0.65) were also screened (**Supplementary Table II**). Analysis of genes related to ACE2 and TMPRSS2 resulted in 15 genes, that were finally included (Figure 1). Pearson correlation coefficients and *p*-values for the 15 genes, as well as ACE2, and TMPRSS2 are shown in Table I. We also generated scatter plots of ACE2, TMPRSS2 and the 15-related genes using R software (Figures 2 and 3). The 15 genes were also subjected to enrichment analysis, using Gene Ontology (GO)[20,21] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[22,23], and results are as shown in Figure 4. GO enrichment analysis included Biological Processes (BP), Cellular Components (CC) and Molecular Function (MF), which focus on the biological process of gene pathway, cytological components of gene mechanism, and molecular biological function of the genes, respectively. BP and CC enrichment analyses revealed five, and three statistically significant pathways and cytological components, respectively, whereas MF and KEGG enrichment analyses resulted in no statistically significant pathways. Moreover, BP enrichment analysis revealed that the genes were significantly enriched in the hormone metabolic process, thyroid hormone generation, hydrogen peroxide biosynthetic process, thyroid hormone metabolic process and antibiotic biosynthetic pathways. On the other hand,

Table I. The Pearson correlation coefficient and p-value between these 15 genes and ACE2, TMPRSS2.

Parameters	ACE2		TMPRSS2	
	cor	*p*-value	cor	*p*-value
AHCYL2	0.450	9.52E-16	0.651	3.68E-36
CTSH	0.480	4.93E-18	0.667	2.23E-38
DUOX1	0.488	1.26E-18	0.724	5.69E-48
DUOXA1	0.485	2.30E-18	0.705	1.43E-44
FASN	0.479	5.97E-18	0.678	4.44E-40
FREM2	0.487	1.36E-18	0.672	3.46E-39
RHOBTB2	0.457	3.00E-16	0.748	6.95E-53
SDR16C5	0.450	9.08E-16	0.707	6.21E-45
SFTPB	0.471	2.83E-17	0.700	9.92E-44
SNX25	0.474	1.42E-17	0.724	5.49E-48
SNX30	0.450	9.15E-16	0.652	2.89E-36
TACC2	0.471	2.78E-17	0.714	3.78E-46
TFCP2L1	0.493	4.88E-19	0.656	7.65E-37
TMEM163	0.468	4.32E-17	0.680	2.00E-40
ZNF385B	0.450	8.67E-16	0.701	7.56E-44

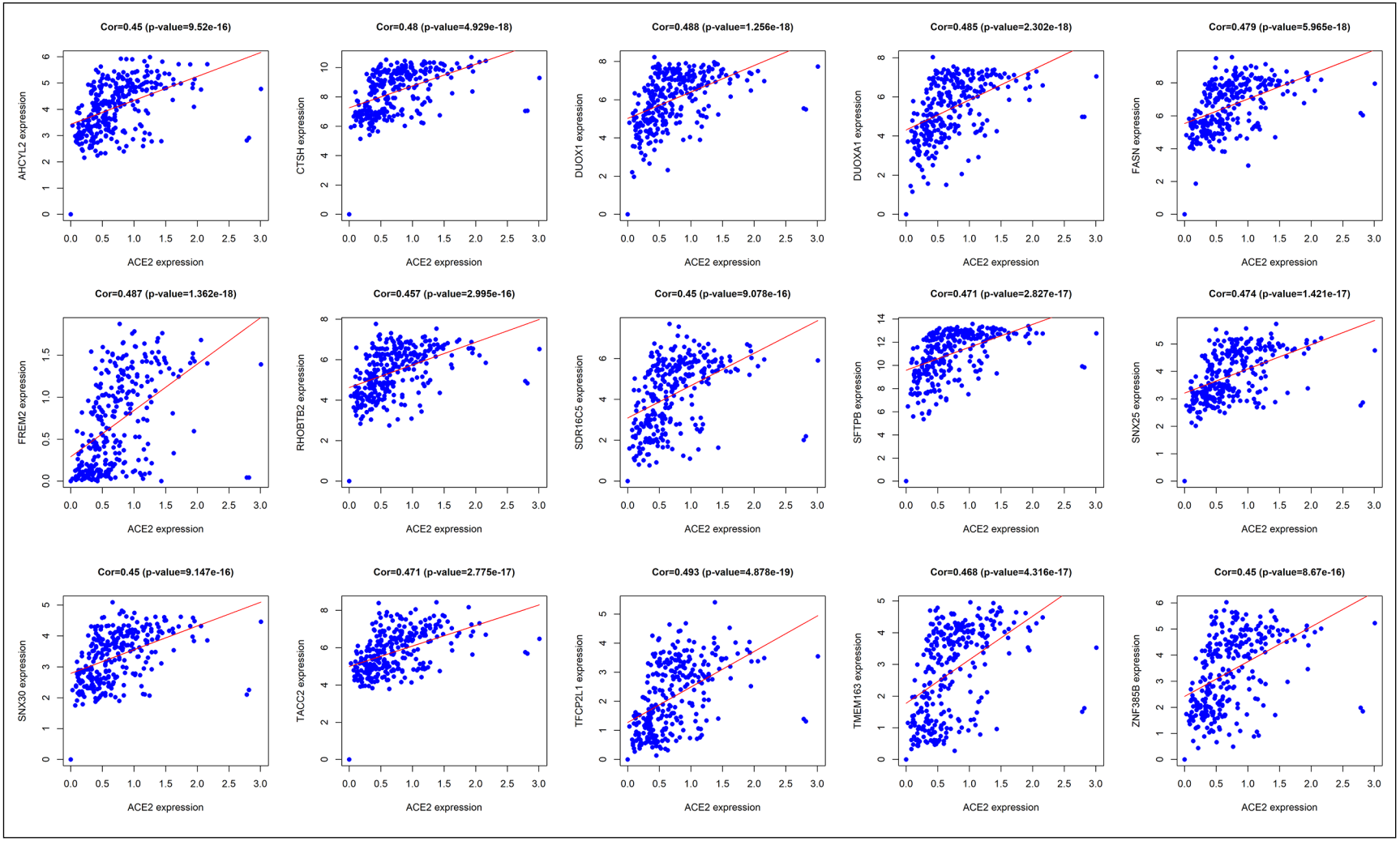

Figure 2. A scatter plot indicating the correlation between ACE2 and the 15 related genes.

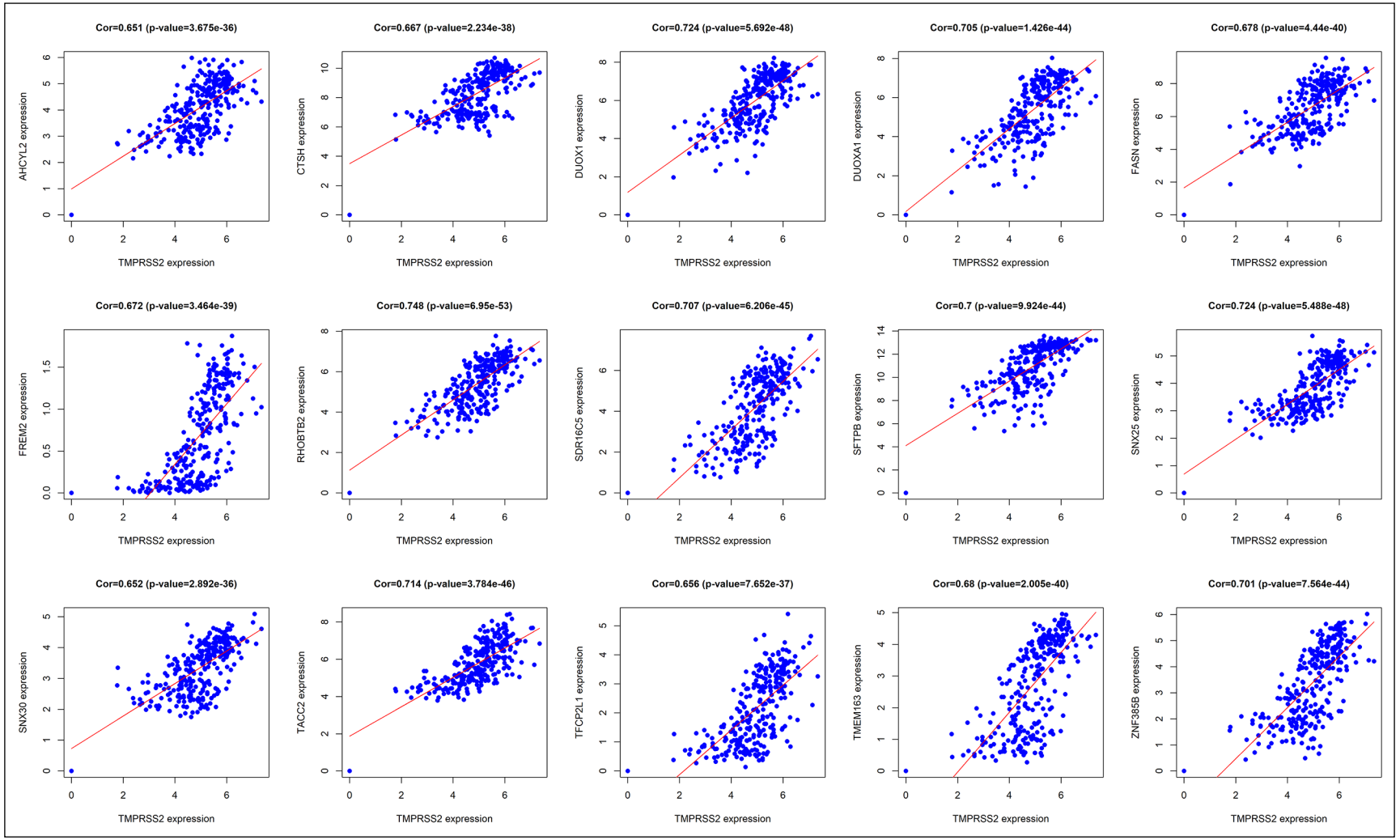

Figure 3. A scatter plot indicating a correlation between TMPRSS2 and the 15 related genes.

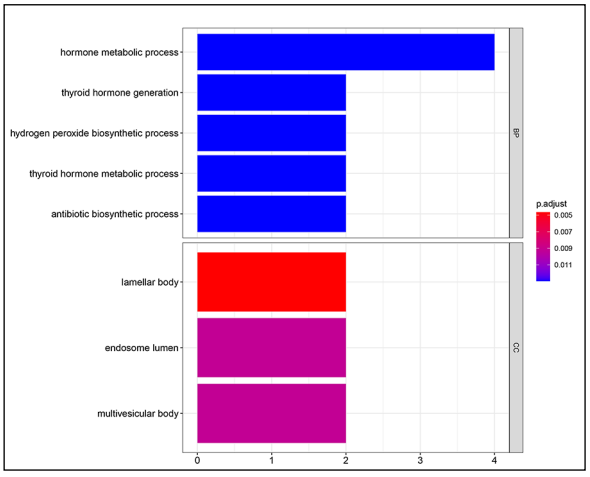

Figure 4. Gene Ontology enrichment analysis.

CC enrichment revealed that lamellar body, endosome lumen and multivesicular body may be regulating SARS-COV-2's ability to infect host cells (Table II).

Correlation between ACE2 and TMPRSS2 expression revealed a significant positive correlation among the following genes: ZNF385B, AHCYL2, SDR16C5, SNX30, RHOBTB2, TMEM163, TACC2, SFTPB, SNX25, FASN, CTSH, DUOXA1, FREM2, DUOX1 and TFCP2L1. This indicates that inhibiting expression of the 15 genes may downregulate ACE2 and TMPRSS2, thereby arrest SARS-COV-2's ability to infect cells and slow down its rate of transmission. Screening for co-expressed genes may help to identify potential gene therapy targets, and boost development of new targeted drugs as well as formulation of new clinical treatment plans[24]. This will be of great importance in saving people's lives and ending the COVID-19 pandemic.

Conclusions

Our results indicate that ZNF385B, AHCYL2, SDR16C5, SNX30, RHOBTB2, TMEM163, TACC2, SFTPB, SNX25, FASN, CTSH, DUOXA1, FREM2, DUOX1 and TFCP2L1 genes are significantly correlated with expression of ACE2 and TMPRSS2, in normal human lung tissues, hence may be potential targets for developing treatment therapies against SARS-COV-2.

Conflict of Interest

The Authors declare that they have no conflict of interests.

Table II. Results of GO enrichment analysis.

Methods	ID	Enrichment analysis	*p*	Gene name
BP	GO:0042445	Hormone Metabolic Process	0.013	ACE2/SDR16C5/DUOXA1/DUOX1
BP	GO:0006590	Thyroid Hormone Generation	0.013	DUOXA1/DUOX1
BP	GO:0050665	Hydrogen Peroxide Biosynthetic Process	0.013	DUOXA1/DUOX1
BP	GO:0042403	Thyroid Hormone Metabolic Process	0.013	DUOXA1/DUOX1
BP	GO:0017000	Antibiotic Biosynthetic Process	0.013	DUOXA1/DUOX1
CC	GO:0042599	Lamellar Body	0.005	SFTPB/CTSH
CC	GO:0031904	Endosome Lumen	0.009	SFTPB/CTSH
CC	GO:0005771	Multivesicular Body	0.009	SFTPB/CTSH

Authors' Contribution

Renwang Hu and Can Liu draft the manuscript. Jianping Gong and Zhixin Cao conceived the idea and recommended this magazine. All the authors participated in the revision of this manuscript. All authors read and approved the final manuscript.

References

1) Kannan S, Shaik Syed Ali P, Sheeza A, Hemalatha K. COVID-19 (Novel Coronavirus 2019) - recent trends. Eur Rev Med Pharmacol Sci 2020; 24: 2006-2011.
2) Hillen HS, Kokic G, Farnung L, Dienemann C, Tegunov D, Cramer P. Structure of replicating SARS-CoV-2 polymerase. Nature 2020; 584: 154-156.
3) Clausen TM, Sandoval DR, Spliid CB, Pihl J, Perrett HR, Painter CD, Narayanan A, Majowicz SA, Kwong EM, McVicar RN, Thacker BE, Glass CA, Yang Z, Torres JL, Golden GJ, Bartels PL, Porell RN, Garretson AF, Laubach L, Feldman J, Yin X, Pu Y, Hauser BM, Caradonna TM, Kellman BP, Martino C, Gordts PLSM, Chanda SK, Schmidt AG, Godula K, Leibel SL, Jose J, Corbett KD, Ward AB, Carlin AF, Esko JD. SARS-CoV-2 infection depends on cellular heparan sulfate and ACE2. Cell 2020; 183: 1043-1057.e15.
4) Dong M, Zhang J, Ma X, Tan J, Chen L, Liu S, Xin Y, Zhuang L. ACE2, TMPRSS2 distribution and extrapulmonary organ injury in patients with COVID-19. Biomed Pharmacother 2020; 131: 110678.
5) Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C, Pöhlmann S. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 2020; 181: 271-280.e8.
6) Lukassen S, Chua RL, Trefzer T, Kahn NC, Schneider MA, Muley T, Winter H, Meister M, Veith C, Boots AW, Hennig BP, Kreuter M, Conrad C, Eils R. SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. EMBO J 2020; 39: e105114.
7) Akhmerov A, Marbán E. COVID-19 and the Heart. Circ Res 2020; 126: 1443-1455.
8) Hassanein M, Radhakrishnan Y, Sedor J, Vachharajani T, Vachharajani VT, Augustine J, Demirjian S, Thomas G. COVID-19 and the kidney. Cleve Clin J Med 2020; 87: 619-631.
9) Jothimani D, Venugopal R, Abedin MF, Kaliamoorthy I, Rela M. COVID-19 and the liver. J Hepatol 2020; 73: 1231-1240.
10) 10.Khan S, Gomes J. Neuropathogenesis of SARS-CoV-2 infection. eLife 2020; 9: e59136.
11) Chang TH, Wu JL, Chang LY. Clinical characteristics and diagnostic challenges of pediatric COVID-19: a systematic review and meta-analysis. J Formos Med Assoc 2020; 119: 982-989.
12) Wilder-Smith A, Chiew CJ, Lee VJ. Can we contain the COVID-19 outbreak with the same measures as for SARS? Lancet Infect Dis 2020; 20: e102-e107.
13) Li C, Xu BH. The viral, epidemiologic, clinical characteristics and potential therapy options for COVID-19: a review. Eur Rev Med Pharmacol Sci 2020; 24: 4576-4584.
14) Wang Z, Fu Y, Guo Z, Li J, Li J, Cheng H, Lu B, Sun Q. Transmission and prevention of SARS-CoV-2. Biochem Soc Trans 2020; 48: 2307-2316.
15) Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. J Autoimmun 2020; 109: 102433.
16) GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013; 45: 580-585.
17) GTEx Consortium. Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 2015; 348: 648-660.
18) Gibson G. Human genetics. GTEx detects genetic effects. Science 2015; 348: 640-641.
19) Goldman MJ, Craft B, Hastie M, Repečka K, McDade F, Kamath A, Banerjee A, Luo Y, Rogers D, Brooks AN, Zhu J, Haussler D. Visualizing and interpreting cancer genomics data via the Xena platform. Nat Biotechnol 2020; 38: 675-678.
20) Kim HI, Kim JH, Park YJ. Transcriptome and Gene Ontology (GO) enrichment analysis reveals genes involved in biotin metabolism that affect L-lysine production in corynebacterium glutamicum. Int J Mol Sci 2016; 17: 353.
21) Peng J, Wang H, Lu J, Hui W, Wang Y, Shang X. Identifying term relations cross different gene ontology categories. BMC Bioinformatics. 2017; 18: 573.
22) Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 2000; 28: 27-30.
23) Du J, Yuan Z, Ma Z, Song J, Xie X, Chen Y. KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model. Mol Biosyst 2014; 10: 2441-2447.
24) Trivedi N, Verma A, Kumar D. Possible treatment and strategies for COVID-19: review and assessment. Eur Rev Med Pharmacol Sci 2020; 24: 12593-12608.