 doi:10.20944/preprints202002.0194.v1

AGTR2, one possible novel key gene for the entry of 2019-nCoV into human cells

Chunmei Cui[1], Chuanbo Huang[1], Wanlu Zhou[2], Xiangwen Ji[1], Fenghong Zhang[2], Liang Wang[2], Yuan Zhou[1], Qinghua Cui[1*]

[1]Department of Biomedical Informatics, Department of Physiology and Pathophysiology, Center for Noncoding RNA Medicine, MOE Key Lab of Cardiovascular Sciences, School of Basic Medical Sciences, Peking University, 38 Xueyuan Rd, Beijing, 100191, China

[2]Co., Ltd of JeaMoon Technology, 6 Rd Middle Zuojiazhuang, Beijing, 100028, China.

*To whom the correspondence should be addressed:

Dr. Qinghua Cui

Tel: 8610-82801001

Email: cuiqinghua@hsc.pku.edu.cn

Keywords: 2019-nCoV, receptor, ACE2, AGTR2

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

Abstract

Recently, it was confirmed that ACE2 is the receptor of 2019-nCoV, the pathogen causing the recent outbreak of severe pneumonia in China. It is confused that ACE2 is widely expressed across a variety of organs and is expressed moderately but not highly in lung, which, however, is the major infected organ. It remains unclear why it is the lung but not other tissues among which ACE2 highly expressed is mainly infected. We hypothesized that there could be some other genes playing key roles in the entry of 2019-nCoV into human cells. Here we found that AGTR2 (angiotensin II receptor type 2), a G-protein coupled receptor, has interaction with ACE2 and is highly expressed in lung with a high tissue specificity. More importantly, simulation of 3D structure based protein-protein interaction reveals that AGTR2 shows a higher binding affinity with the Spike protein of 2019-nCov than ACE2 (energy score: -15.7 vs. -6.9 [kcal/mol]). Given these observations, we suggest that *AGTR2* could be a putative novel gene for the the entry of 2019-nCoV into human cells but need further confirmation by biological experiments. Finally, a number of compounds, biologics and traditional Chinese medicine that could decrease the expression level of AGTR2 were predicted.

Introduction

Recently, it was revealed that a novel coronavirus, 2019-nCoV, caused an outbreak of severe febrile respiratory illness in China [1]. As of February 4, 2020, more than 24000 cases and 492 deaths have been reported (https://news.163.com/special/epidemic/). Sequence analysis revealed that 2019-nCoV also belongs to the lineage B of the genus betacoronavirus, which suggests that it is highly similar with SARS virus. Therefore, it seems that the human angiotensin-converting enzyme 2 (ACE2) is also the receptor for the entry of 2019-nCoV into

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

human cells [2], as it works in the process of SARS virus infection. Recently, Letko and Munster confirmed this [3]. More recently, Hoffmann et al. found that 2019-nCoV spike protein uses cellular protease TMPRSS2 for its priming[4]. These studies suggest that ACE2 and TMPRSS2 are two putative targets for anti-2019-nCoV. Neither ACE2 nor TMPRSS2 showed a high or specific expression in lung, which, however, is the major infected tissue in human body. It still remains unclear why it is the lung but not other tissues which expresse high level of ACE2 is mainly infected. In this study, we found that AGTR2 can interact with ACE2. Moreover, AGTR2 has a high expression level and specificity in lung. In addition, AGTR2 is a G-protein coupled receptor. Given these observations, we suggest that *AGTR2* could be a putative novel gene for the the entry of 2019-nCoV into human cells.

Methods

Expression and protein functional association network data

We obtained the expression profiles of the genes from the dataset of HPA RNA-seq normal tissues at NCBI (https://www.ncbi.nlm.nih.gov/). We also obtained gene expression data from the Genotype-Tissue Expression (GTEx) database[5]. For screening human proteins interacting with ACE2, we searched ACE2 in STRING[6], a tool for visual analysis of protein-protein interactions, and selected homo sapiens. Then we constructed a PPI network centering on ACE2 with the default parameters of website. The functional protein association network of ACE2 was created using String. From Connectivity Map[7] (CMap, version 2.0), we obtained the gene expression profiles induced by ~1000 compounds. We obtained nearly 2000 gene expression profiles induced by a variety of agents including compounds, biologics, and traditional Chinese medicine from JeaMoon Map (JMap), which was developed by Co.,

 doi:10.20944/preprints202002.0194.v1

Ltd of JeaMoon Technology.

Simulating interaction between 2019-nCoV Spike protein and AGTR2

The experimental protein structures of AGTR2 was downloaded from the PDB database (entry number 5UNG). Note that PDB:5UNG is a chimera protein merging AGTR2 and soluble cytochrome b562, therefore only the structure part which belongs to residues 35-335 of AGTR2 was retained for further analysis. The experimental protein structure of ACE2 was downloaded from the PDB database (entry number 1R42). The monomer structure of Spike protein (Spro, NCBI seq ID: QHR63290) was predicted by SWISS-MODEL, using the template PDB structure 6ACD, chain C (the subunit bounds to ACE2 from SARS-Cov Spro homo-trimer, covering residues 23-1146 of nCoV-2019 spike glycoprotein, with sequence identity = 76.28%). Then the structure of ACE2 and nCoV-2019 Spro was submitted to the template-based docking server PRISM (http://cosbi.ku.edu.tr/prism/) to predict their potential interaction interface, using the known interaction interface between SARS-Cov Spro and ACE2 as the template (PDB:6ACG, chain C and D). For other proteins, because there was no available template complex between these human and virus proteins, these proteins were instead submitted to the template-free docking server ZDOCK (http://zdock.umassmed.edu/) to predict their potential interaction interface. The residues participate in subunit self-assembly of virus proteins and the residues of transmembrane helices plus cytoplasmic regions of AGTR2 were blocked during the calculation, since these residues are unlikely to participate in the inter-protein interactions. The best-scored interface without prominent overlaps with the transmembrane helices of AGTR2 was considered. Finally, the predicted protein complex were analyzed by PISA

 doi:10.20944/preprints202002.0194.v1

(https://www.ebi.ac.uk/msd-srv/prot_int/cgi-bin/piserver) and InterProSurf (http://curie.utmb.edu/prosurf.html) to obtain the interacting energy and interface residues, respectively. The protein complex structures were visualized by the PyMol tool (http://pymol.org).

Screening agents that could decrease the expression level of AGTR2 and TMPRSS2

We screen CMap database and JMap database to identify the potential agents that decrease the expression level of AGTR2 using a cutoff of fold change (FC) 2.0. For JMap, only small molecules, biologics, and traditional Chinese medicine (TCM) are considered.

Results

ACE2 functional protein association network and expression analysis

Expression analysis revealed that both ACE2 and TMPRSS2 have a relative low expression level in lung. Thus, to find more potential genes that play key roles in the entry of 2019-nCoV into human lung cells, we first generated the ACE2 functional protein association network (**Figure 1**) using String. After a quick annotation of the associated proteins, we found that only AGTR1 (angiotensin II receptor type 1) and AGTR2 are G-protein coupled receptors (GPCRs). We think the putative key gene should be also a receptor. Thus, we focused on the two GPCRs. Expression analysis further revealed that AGTR1 is dominantly expressed in plancenta and show a quite low expression in lung (**Figure 2A**), whereas AGTR2 is dominantly expressed in lung and show a high specificity in lung (**Figure 2B**). AGTR2 also shows a most high expression in the GTEx dataset. Taken together, the result suggest AGTR2 could be a possible key gene in the entry of 2019-nCoV into human lung cells.

Simulation of interaction between AGTR2 and 2019-nCoV Spike protein

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

According to the ZDOCK prediction result, AGTR2 and Spro would form a protein complex with energy score -15.7 kcal/mol (**Figure 3**). According to the PRISM prediction result, ACE2 and Spro would form a protein complex with energy score -6.9 kcal/mol (while the interacting energy score between SARS-nCOV's Spro and ACE2 is calculated as -8.8 kcal/mol based on the experimental complex structure PDB:6ACG). The cartoon representation of the complex structure is shown below, where ACE2 and Spro are colored purple and red (**Figure 4**), respectively. In this predicted complex structure, Spro exploits the same domain which is also used by the SARS-nCOV's Spro to interact with its receptor ACE2.

Agents that could decrease the expression level of AGTR2

We previously identified the agents that decrease the expression level of ACE2 (doi: 10.20944/preprints202002.0047.v1) using CMap and JMap. Here, by screening CMap and JMap, we identified the potential agents to decrease the expression level of AGTR2 (**Figure 5**). Interestingly, a number of agents that decreased the expression of ACE2 also down-regulated the expression of AGTR2, for example, Sambucus, Urtica, Andrographis, lipopolysaccharide, emetine, Aza-dC, valproic acid, oxLDL etc.

Discussion

The relative low expression of ACE2 in lung suggest that there should be more genes that play key roles in the entry of 2019-nCoV into human cells. To address this issue, here using bioinformatics we identified a putative gene, *AGTR2*. Simulation of 3D based protein-protein interaction reveals that AGTR2 shows a higher binding affinity with Spro than ACE2. In addition, we also found another two receptor genes, *PIGR* and *ADGRF1*, shows relatively

 doi:10.20944/preprints202002.0194.v1

high and specific expression in lung (trachea and/or bronchial epitheial cell, data not shown), suggesting these two genes could also have the possibility to have roles in 2019-nCoV's entry into human cells. Moreover, a number of agents (compounds, biologics, and traditional Chinese medicine) that could decrease the expression of AGTR2 were identified. In addition, besides the above agent types, the JMap also contained a number of other agent types, for example, environmental factors, among which, heating/heat stress were identified in both AGTR2 and ACE2 predictions. Finally, we would suggest that AGTR2 could be a novel key gene for the entry of 2019-nCoV into human cells and the identified agents should be carefully considered in anti-2019-nCoV usage.

References

[1] N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G.F. Gao, W. Tan, I. China Novel Coronavirus, T. Research, A Novel Coronavirus from Patients with Pneumonia in China, 2019, N Engl J Med (2020).

[2] S. Perlman, Another Decade, Another Coronavirus, N Engl J Med (2020).

[3] M. Letko, V. Munster, Functional assessment of cell entry and receptor usage for lineage B β-coronaviruses, including 2019-nCoV, bioRxiv (2020) 2020.2001.2022.915660.

[4] M. Hoffmann, H. Kleine-Weber, N. Krüger, M. Müller, C. Drosten, S. Pöhlmann, The novel coronavirus 2019 (2019-nCoV) uses the SARS-coronavirus receptor ACE2 and the cellular protease TMPRSS2 for entry into target cells, bioRxiv (2020) 2020.2001.2031.929042.

[5] G.T. Consortium, The Genotype-Tissue Expression (GTEx) project, Nat Genet 45 (2013) 580-585.

[6] D. Szklarczyk, A.L. Gable, D. Lyon, A. Junge, S. Wyder, J. Huerta-Cepas, M. Simonovic, N.T. Doncheva, J.H. Morris, P. Bork, L.J. Jensen, C.V. Mering, STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets, Nucleic Acids Res 47 (2019) D607-D613.

[7] J. Lamb, E.D. Crawford, D. Peck, J.W. Modell, I.C. Blat, M.J. Wrobel, J. Lerner, J.P. Brunet, A. Subramanian, K.N. Ross, M. Reich, H. Hieronymus, G. Wei, S.A. Armstrong, S.J. Haggarty, P.A. Clemons, R. Wei, S.A. Carr, E.S. Lander, T.R. Golub, The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease, Science 313 (2006) 1929-1935.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

Table 1. The interacting interface residues between AGTR2 and Spro.

Amino Acid	Residue Number	Monomer Area	Complex Area	Protein	Delta Area
A	95	36.72	15.24	AGTR2	21.48
P	98	32.61	15.62	AGTR2	16.99
L	99	33.45	2.52	AGTR2	30.93
T	102	49.63	31.16	AGTR2	18.47
Y	106	110.05	98.13	AGTR2	11.92
L	111	118.35	69.36	AGTR2	48.99
F	112	79.05	4.74	AGTR2	74.31
G	113	29.61	8.65	AGTR2	20.96
V	115	116.65	14.85	AGTR2	101.8
M	116	85.71	0.53	AGTR2	85.18
K	118	49.96	19.23	AGTR2	30.73
V	119	72.56	22.72	AGTR2	49.84
F	120	39.64	2.27	AGTR2	37.37
L	176	99.9	31.4	AGTR2	68.5
F	179	107.62	15.44	AGTR2	92.18
Y	180	170.99	19.67	AGTR2	151.32
F	181	18.07	6.39	AGTR2	11.68
P	200	48.63	33.78	AGTR2	14.85
W	207	96.5	69.05	AGTR2	27.45
G	455	80.32	34.9	Spro	45.42
Y	458	143.84	3.33	Spro	140.51
N	459	119.97	59.48	Spro	60.49
L	461	50.59	34.61	Spro	15.98
L	464	52.06	36.89	Spro	15.17
G	491	58.34	41.09	Spro	17.25
V	492	119.63	58.04	Spro	61.59
E	493	156.5	0.07	Spro	156.43
G	494	29.68	19.51	Spro	10.17
F	495	204.17	89.08	Spro	115.09
N	496	106.08	91.33	Spro	14.75
Y	498	133.53	32.31	Spro	101.22
F	499	125.91	26.17	Spro	99.74
L	501	38.95	20.3	Spro	18.65
Q	502	87.83	15.04	Spro	72.79
S	503	28.5	5.37	Spro	23.13
Q	507	72.72	35.23	Spro	37.49

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

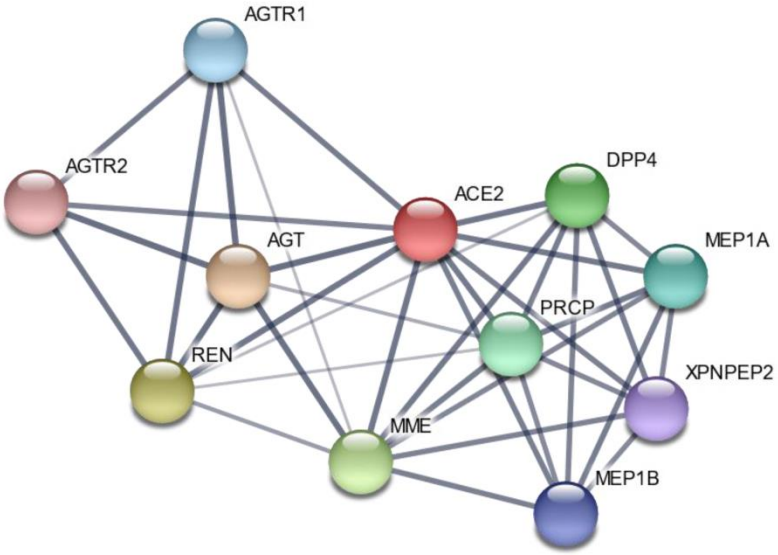

Figure 1. Protein functional association network built by String.

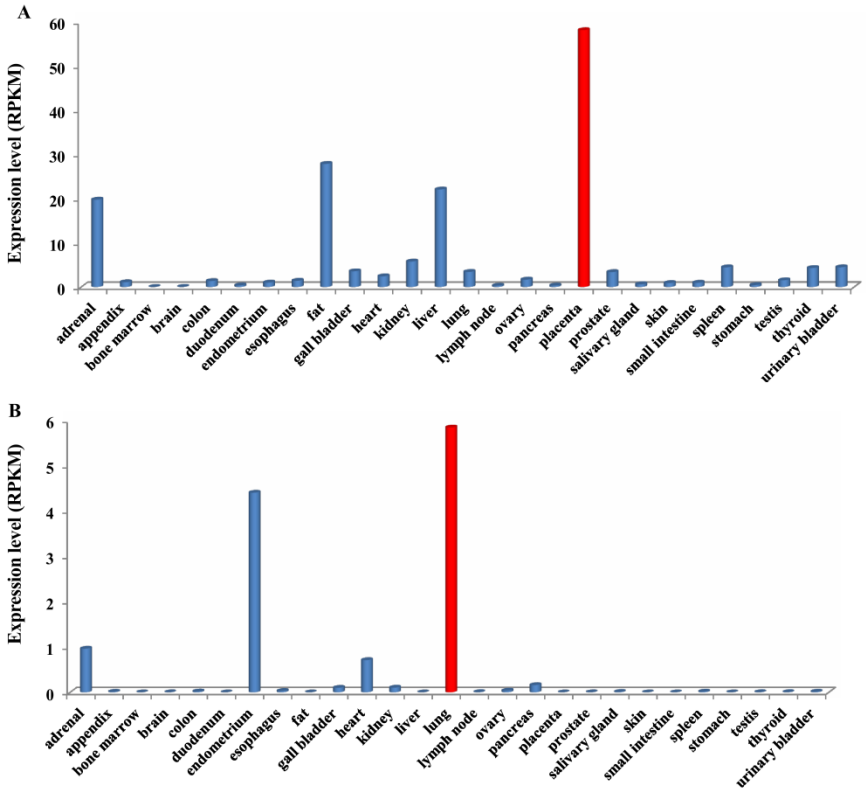

Figure 2. Expression profiles of AGTR1 (A) and AGTR2 (B).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

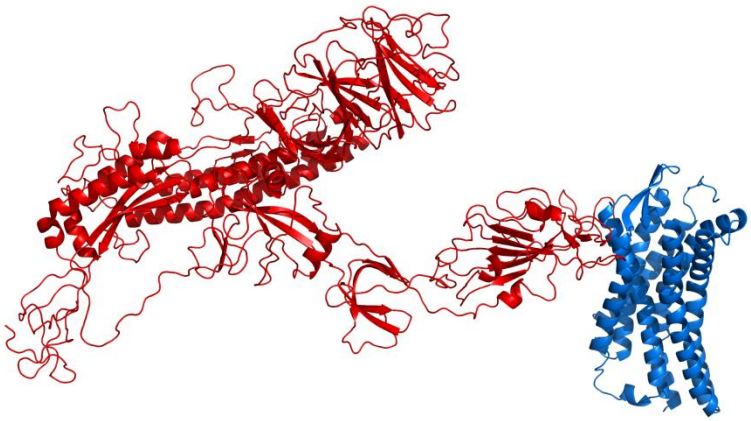

Figure 3. Predicted 3D structure of the complex formed by 2019-nCoV Spike protein (red) and human AGTR2 protein.

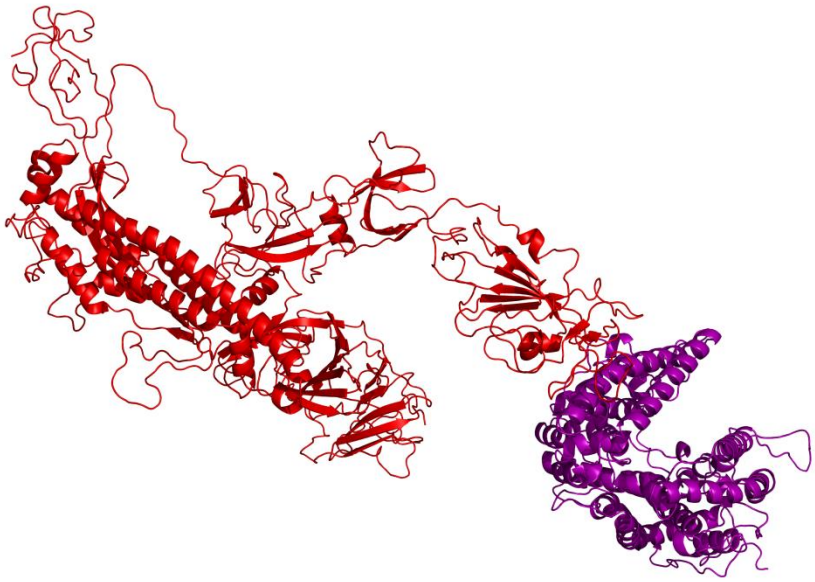

Figure 4. Predicted 3D structure of the complex formed by 2019-nCoV Spike protein (red) and human ACE2 protein.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 14 February 2020 doi:10.20944/preprints202002.0194.v1

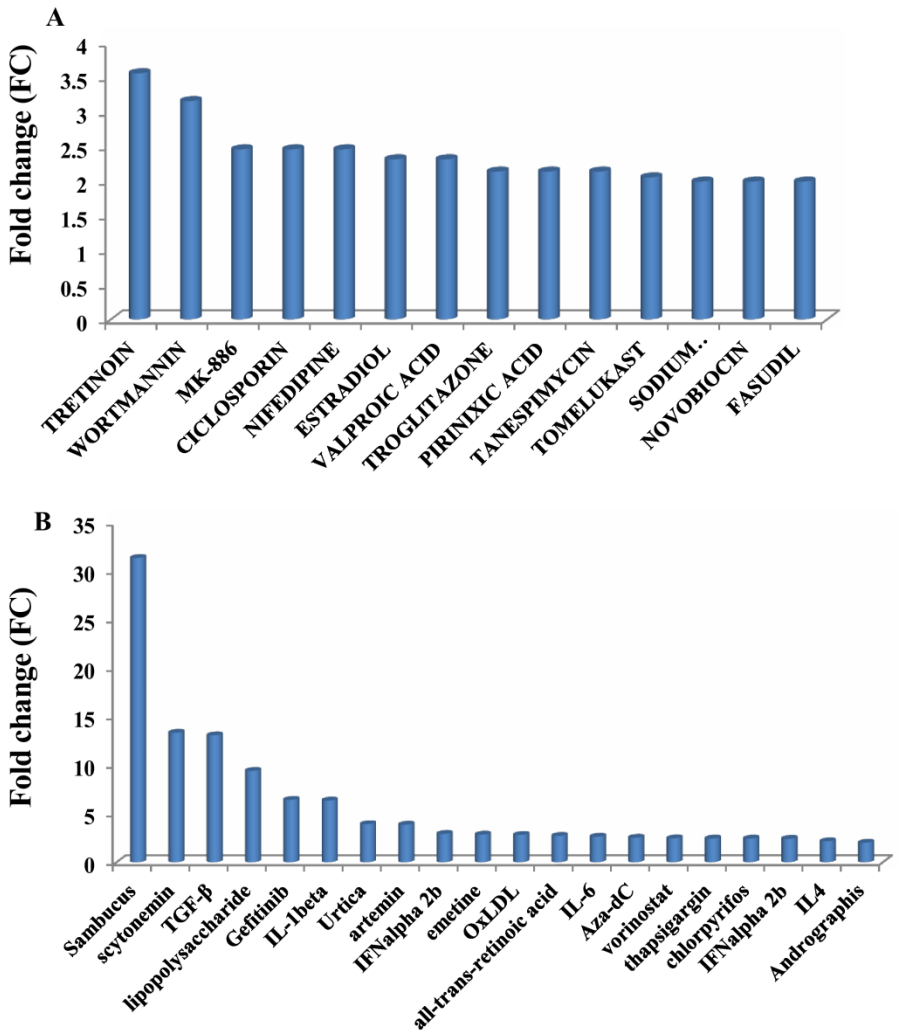

Figure 5. Possible agents screened from CMap (A) and JMap (B). The y axis means the decreased fold change of AGTR2 expression induced by the corresponding agents.