

ORIGINAL ARTICLE

Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations

Quan Li^{1,2} | Zanxia Cao³ | Proton Rahman¹ 

¹Department of Medicine, Faculty of Medicine, Memorial University, St. John's, NL, Canada

²Princess Margaret Cancer Centre, University Health Network, University of Toronto, Toronto, ON, Canada

³Shandong Provincial Key Laboratory of Biophysics, Institute of Biophysics, Dezhou University, Dezhou, China

Correspondence

Proton Rahman, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL A1C 5B8, Canada.
Email: prahman@mun.ca

Funding information

Atlantic Canada Opportunities Agency, Grant/Award Number: 781-19095-209843

Abstract

Background: There appears to be large regional variation for susceptibility, severity, and mortality for COVID-19 infections. Numerous potential factors could explain the wide variability in the number of infections and death among the countries. We examined genetic differences in the human angiotensin-converting enzyme 2 (hACE2) gene, as its receptor serves as a cellular entry for SARS-CoV-2. At present, there is a paucity of data regarding the differences for ACE2 polymorphisms and expression levels between ethnicities.

Methods: We compared the allele frequency of mutations between European and East Asians. Molecular dynamic simulation were performed to investigate the influences of significant mutant on protein structure. The binding free energies were calculated between S protein and hACE2. We also examined hACE2 gene expression in eight global populations from HapMap3.

Results: Four missense mutations showed significant minor allele frequency difference between Asians and Caucasians. Molecular dynamic demonstrated that two of these variants (K26R and I468V) may affect binding characteristics between S protein of the virus and hACE2 receptor. We also noted marginal differences in gene expression for some populations in HapMap3 as compared to the Chinese population.

Conclusion: Our studies reveal subtle changes in the genetics of hACE2 between human populations, but the magnitude of the difference was small and the significance is not clear in the absence of further in vitro and functional studies.

KEYWORDS

angiotensin-converting enzyme 2, COVID-19, ethnic variation, gene expression, molecular dynamics simulation, SARS-CoV-2

1 | INTRODUCTION

As of March 24, 2020, the worldwide tracking for COVID-19 infections reported 1,603,433 cases including 95,716 deaths

(Worldometers, 2020). Among the 18 countries that have reported more than 20,000 COVID-19 cases, the total number of cases per population ranged from 3,277 per million people in Spain to 57 per million cases in China and the

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *Molecular Genetics & Genomic Medicine* published by Wiley Periodicals LLC.

total death rate among these countries ranged from 330 per million in Spain to 2 per million in China (Worldometers, 2020). There are numerous potential factors to explain the wide variability in the number of infections and death among the countries including stringency of social isolation and contact tracing, the importation of the SARS-CoV-2, intensity of testing for the virus, as well as the preparedness and capacity of health care system to cope with the pandemic.

Differences among the host, particularly the sex of the patient and presence of selected comorbid diseases and immunosuppressive state have been identified as risk factors from severe COVID-19 (China-CCDC, 2020; World Health Organization, 2020). Preexisting medical conditions have been associated with increased prevalence of death including: cardiovascular disease (13.2%); diabetes (9.2%); chronic respiratory disease (8.0%); hypertension (8.4%); and recent diagnosis of cancer (7.6%) (Zhou, Yu, et al., 2020). Specifically, hypertensive patients had a hazard ratio of 1.70 for death (Wu et al., 2020) and 3.05 for in-hospital mortality (Zhou, Yu, et al., 2020). An increased association with the use of angiotensin-converting enzyme inhibitors (ACEIs) and angiotensin II receptor blockers (ARBs) has been reported for severe COVID-19 cases (Zhou, Yang, et al., 2020), however, the clinical relevance of this finding is uncertain and possibly influenced by the presence of confounders (Gracia-Ramos, 2020). A biologically plausible link between SARS-CoV-2 infections and ACE inhibition has been proposed (Patel & Verma, 2020).

Angiotensin-converting enzyme 2 (ACE2, OMIM: 3000335) cleaves peptides within the renin-angiotensin system (Battle, Wysocki, & Satchell, 2020). ACE2 is mostly expressed in kidneys and the GI tract, and smaller amounts are found in type 2 pneumocytes in the lung and peripheral blood. The ACE2 receptor serves as the host cell entry for SARS-CoV and more recently it has been shown to be receptor for cellular entry for SARS-CoV-2 (Du et al., 2009; Walls et al., 2020). As recently reported by Hoffman et al., 2020, the cellular entry of SARS-CoV-2 can be blocked by an inhibitor of the cellular serine protease TMPRSS2, which is employed by SARS-CoV-2 for S protein priming. Further attention has been drawn to the ACE2 receptor as antimalarials, which can interfere with ACE2 expression, have resulted in increased clearing of the SARS-CoV-2 clearance among COVID-19 patients (Gautret et al., 2020).

To explore the variability in genetic polymorphisms and expression in human ACE2 (hACE2), we set out to determine if there were any differences between the Asian and Caucasian populations for ACE2 polymorphisms and compare the variability of hACE2 expression in peripheral blood among eight different populations.

2 | METHODS

2.1 | Comparison of genetic polymorphisms of hACE2 and its possible functional alteration between Asians and Caucasians

In order to investigate whether differences in genetic variations exist between Caucasians and Asians and if these variants can influence the efficiency of cell entry of SARS-CoV-2, we retrieved the variants in the hACE2 from gnomAD v2.1 exomes (Lek et al., 2016). We used 56,885 Non-Finnish European (NFE) and 9,197 East Asians (EAS, including 1,909 Koreans) for our analysis.

Molecular dynamics (MD) simulation which is a canonical method for investigating the influences of mutant on protein structure, was then applied. As mutations N638S and N720 are far away from the RBD, as a result we mainly focused on the K26R as MD simulations demonstrated that this variant mutated more frequently in Non-Finnish European while I468V mutated frequently more in East Asians. The initial wild-type structure of the human ACE2 (hACE2) was selected from the X-ray crystal complex structure between hACE2 with SARS-CoV-2 spike (S) protein from Protein Data Bank (Berman, Henrick, Nakamura, & Markley, 2007) (PDB ID: 6LZG). The initial structures of mutants were constructed using FoldX (<http://foldxsuite.crg.eu/command/PositionScan>). The structural minimizations and molecular dynamics simulation was performed using the GROMACS 2018.4 (Kutzner et al., 2015, 2019; Pronk et al., 2013) software package. OPLS-AA force field was adopted for the protein complex and the SPCE water model for solvate the system. The periodic boundary conditions (PBC) in a cubic box was applied, and the distance to the edge of the box was set as 1.0 nm. After the protein complex system was solvated, the appropriate numbers of sodium and chloride ions were added to neutralize the system. Equilibration simulations were done under the NVT ensemble for 2 ns, and then, the NPT ensemble for 1 ns after an extensive energy minimization by using the steepest descent algorithm. The simulation time step was 2 fs. The temperature was maintained at 300K using the V-rescale algorithm with a tau-t of 0.1 ps. The pressure was maintained at 1 atm by the semi-isotropic Parrinello–Rahman method with a tau-p of 2 ps and a compressibility of 4.5×10^{-5} bar⁻¹. The Particle Mesh Ewald (PME) method was used for calculation the long-range electrostatic interactions. The LINCS algorithm was used for constraining all covalent bonds. Finally, a 100 ns long MD simulation was carried out for each system. The stability of the simulation was checked by computing root mean square deviations (RMSDs). RMSDs of C α were calculated for the mutant K26R/I468V and compared to the wild type shown in Figure 1. RMSDs for the complexes with K26R changed little during the 100 ns MD simulations.

2.2 | Variability in gene expression of hACE2 gene among various ethnic populations

The second goal was to assess the expression variability of hACE2 in eight global populations from the HapMap3 project (Montgomery et al., 2010; Stranger et al., 2012). The following eight populations were compared: Utah residents with Northern and Western European ancestry from the CEPH collection (CEU); Han Chinese in Beijing China (CHB); Gujarati Indians in Houston, Texas (GIH), Japanese in Tokyo, Japan (JPT); Luhya in Webuye, Kenya (LWK); Mexican ancestry in Los Angeles (MEX), California, Maasai in Kinyawa, Kenya (MKK); and Yoruba in Ibadan, Nigeria (YRI). Raw expression data from Sentrix Human-6 Expression BeadChip version 2 were extracted, background-corrected, log₂ scaled, and quantile normalized. Expression data from the populations with admixture (GIH, LWK, MEX, and MKK) were also normalized for population admixture using EIGENSTRAT (Price et al., 2006). The normalized expression data were used for the assessing the expression level for different sets of analysis.

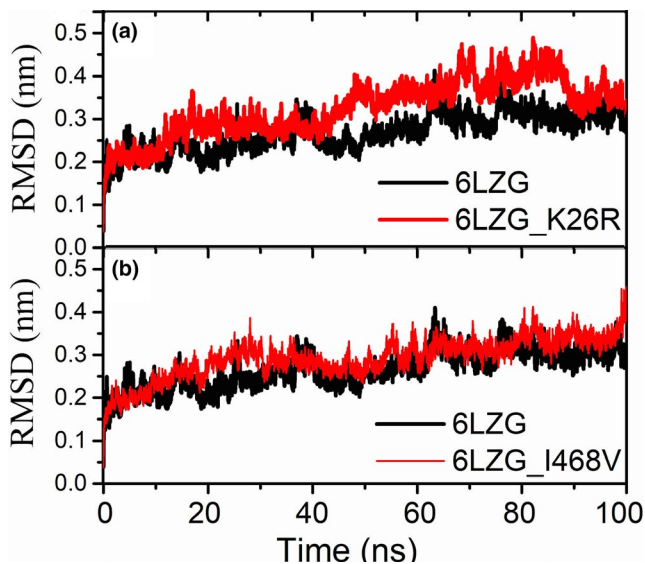


FIGURE 1 RMSDs of hACE2 wild type (black line) and mutated hACE2 protein (red line) during Molecular dynamics simulation

As only publicly available data used and the information and the data linkage did not generate any identifiable information, ethics approval was not required. The data sets analyzed during the current study are available from the lead author, Dr. Quan Li.

3 | RESULTS

3.1 | Comparison of genetic polymorphisms of hACE2 and its possible functional alteration between Asians and Caucasians

In gnomAD v2.1 exomes, there were 5,693 mutations of which 177 were missense mutations. In comparing the allele frequency of Caucasian and Asian cohorts, Fisher's exact *t* tests were performed and adjusted for multiple testing. Four genetic variants reached statistical significance (adjusted *P*-value less than 0.05), although their effect size was modest (Details in Table 1). One variant was located at the $\alpha 1$ helix of ACE2 gene (K26R), two in the middle of SARS-CoV-2 (N638S and I468V), and one in the tail (N720D). The K26R variant was of immediate interest, since the spike glycoprotein (S protein) of virus contains the receptor binding domain (RBD), which directly binds to the peptidase domain (PD) of hACE2. The PD of hACE2 mainly engages the $\alpha 1$ helix in the recognition of the RBD (Yan et al., 2020). The hACE2 X-ray structure for these four mutations is illustrated in Figure 2.

From the molecular dynamic simulation, mutant I468V showed RMSD increase during time 20–30 ns, but no obvious RMSD change comparing with wild type after 40 ns. The RMSD values for the K26R mutant are slightly higher and structure change compared with wild type.

The effect of missense on protein-ligand complex can be assessed by various experimental technologies (Kastritis & Bonvin, 2013), but are very time-consuming. The in silico binding free energies calculation can be performed to determine the effects of mutations. Here, the binding free energies between S protein and hACE2 (wild type or mutant) were calculated using the *g_mmpbsa* (Table 2). These two mutations K26R/I468V were both predicted to slightly increase the binding free energy and may slightly decrease the binding

TABLE 1 Four missense mutations in hACE2 show significant allele frequency difference between EAS and NFE

Chromosome	Position (hg38)	Reference Allele	Alternative Allele	MAF ^a (EAS)	MAF (NFE)	<i>p</i> -value	Exon	cDNA change	Protein change
chrX	15564175	T	C	0	0.0258	1.05E-23	18	A2158G	N720D
chrX	15600835	T	C	7.23E-05	0.0059	8.68E-07	2	A77G	K26R
chrX	15567810	T	C	0.0036	0	2.01E-14	16	A1913G	N638S
chrX	15575706	T	C	0.0112	2.50E-05	1.42E-33	11	A1402G	I468V

^aMAF stands for minor allele frequency; EAS stands for East Asians; NFE stands for Non-Finnish European.

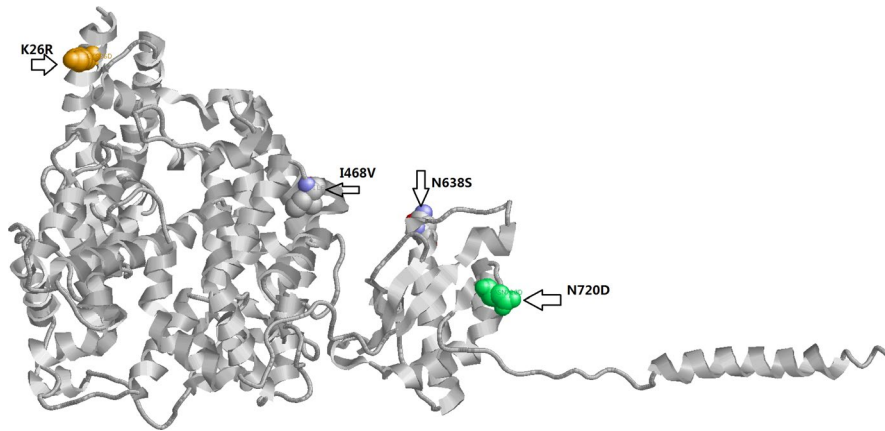


FIGURE 2 Human ACE2 structure and four missense mutations that differ between Caucasians and Asians

TABLE 2 Binding free energy (in units of kJ/mol) for S protein and hACE2 (wild type and mutated) complex

	VDW energy	Electrostatic energy	Polar solvation energy	SASA energy	Binding energy
6LZG_wild type	-343.6 ± 27.4	-1515.1 ± 66.5	716.7 ± 48.4	-45.5 ± 3.1	-1187.5 ± 69.9
6LZG_K26R	-348.2 ± 31.1	-1405.2 ± 63.1	701.4 ± 53.6	-47.5 ± 4.5	-1099.4 ± 54.1
6LZG_I468V	-353.3 ± 23.7	-1498.2 ± 48.9	804.5 ± 71.1	-48.8 ± 4.1	-1095.9 ± 83.1

affinity. While, K26R mutated more frequently in Caucasian, and I468V mutated more frequently among Asians.

3.2 | Variability in gene expression of hACE2 gene among various ethnic populations

Only marginal differences were noted in the hACE2 expression levels of these eight populations (CEU; CHB; GIH; JPT; LWK; MEX, MKK; and YRI), Figure 3. The hACE2 expression level of CHB population (42 females and 38 males) was then compared to other cohorts using ANOVA analysis, after stratification for sex (Table 3). Marginal differences were noted among the populations. There was a statistically higher hACE2 expression in men for LWK and GIH and JPT populations, LWK and YRI populations for females. However, the effect sizes were small and its clinical relevance is not clear.

4 | DISCUSSION

This study that examined the genetic polymorphisms in the hACE2 gene between Caucasians and Asians. We noted that there were four missense mutations, of which K26R and I468V were of most interest based on their location. The propensity for these variants to mutate were different among these two polymorphisms. Among Caucasians,

K26R mutated more, while among Asians I468V mutated more frequently. In silico studies have noted that these two

mutations may affecting the binding characteristics between S protein and hACE2, however, in the absence of further functional studies, the significance of these alternations is not clear.

We also found that there was some variability of genetic expression of hACE2 among various populations, but the magnitude of the differences was small and so it is unclear if this has any impact based on this subtle differences in susceptibility or severity of COVID-19 among various ethnicities. We also did not notice a differential expression of hACE2 between the sexes, even though it is widely reported that males have a worse prognosis than females for complications of COVID-19 (Zhou, Yu, et al., 2020). Our findings are generally consistent with a recently published study by Chen, Shan, & Qian, 2020, that reported the ACE2 expression in Asians was similar to that of other races. Our expression study was conducted in peripheral blood and the results for site specific expression (alveolar, GI, or renal tissue) may yield different results.

In conclusion, our results do reveal some differences in genetic polymorphisms between Asians and Caucasians which may potentially alter the binding of the virus to the hACE2 receptor and some subtle variation in genetic expression of hACE2 among different populations. However, our findings need to be replicated and further in vitro studies should be performed, before the significance of these findings can be determined. At present the reasons behind the variability in COVID-19 presentation among the various countries needs to be further explored, as we were unable

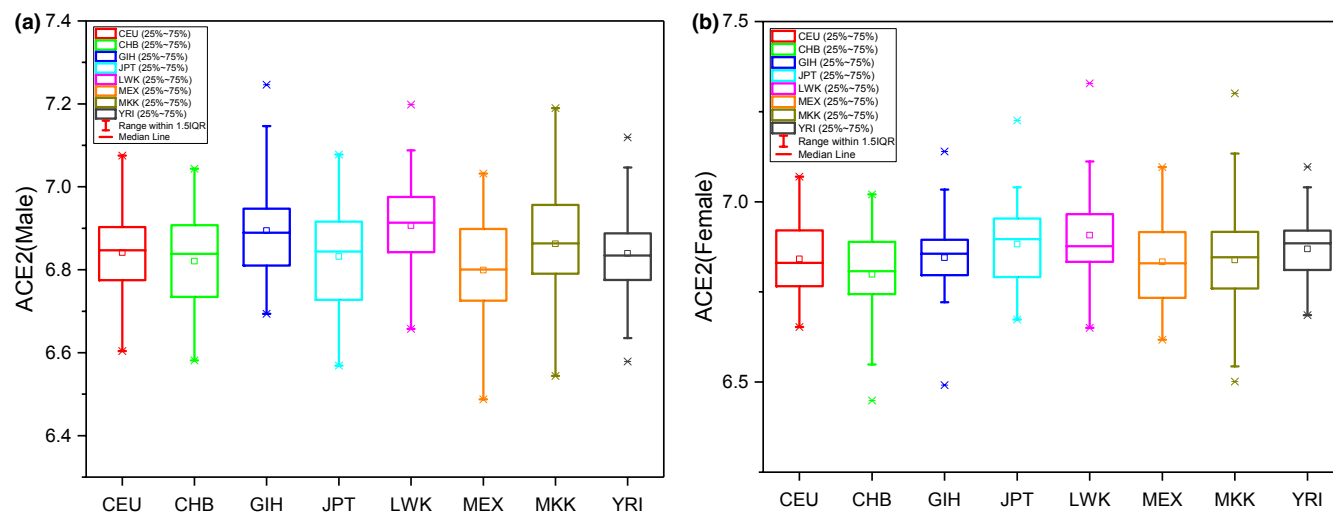


FIGURE 3 hACE2 expression among eight HapMap3 populations. A) for Males and B) for Females

TABLE 3 Comparison of ACE2 expression between Asians (CHB) and other seven populations in HapMap3

Populations	ACE2(Male)			ACE2(Female)		
	Samples	log2FC	P-value	Samples	log2FC	p-value
CEU	56	0.02	0.43	56	0.042	.08
GIH	41	0.073	0.007	41	0.046	.08
JPT	42	0.011	0.69	40	0.083	.002
LWK	43	0.085	0.0017	40	0.11	6.12E-05
MEX	22	-0.022	0.50	23	0.034	.27
MKK	70	0.042	0.09	68	0.040	.09
YRI	55	0.018	0.47	53	0.070	.005

to provide convincing evidence to suggest there are differences in allele frequency or expression of hACE2 among Caucasians and Asians.

ACKNOWLEDGMENT

This work was supported by Atlantic Innovation Fund (781-19095-209843).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

AUTHOR CONTRIBUTIONS

Q.L. and P.R. conceived research, P.R. supervised the research. Q.L., Z.C., and P.R. devised the data analysis strategies. Q.L., Z.C., and P.R. contributed to the data interpretation and writing of the manuscript. All of the authors have read and approved the final manuscript.

ORCID

Proton Rahman  <https://orcid.org/0000-0002-4521-2029>

REFERENCES

- Battle, D., Wsocki, J., & Satchell, K. (2020). Soluble angiotensin-converting enzyme 2: A potential approach for coronavirus infection therapy? *Clinical Science (Lond)*, *134*(5), 543–545. <https://doi.org/10.1042/CS20200163>
- Berman, H., Henrick, K., Nakamura, H., & Markley, J. L. (2007). The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data. *Nucleic Acids Research*, *35*(Database issue), D301–D303. <https://doi.org/10.1093/nar/gkl971>
- Chen, Y., Shan, K., & Qian, W. (2020). Asians and Other Races Express Similar Levels of and Share the Same Genetic Polymorphisms of the SARS-CoV-2 Cell-Entry Receptor. Preprints, 2020020258.
- China-CCDC. (2020). The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19). Retrieved 23 March 2020, from <http://weekly.chinacdc.cn/en/article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51>
- Du, L., He, Y., Zhou, Y., Liu, S., Zheng, B. J., & Jiang, S. (2009). The spike protein of SARS-CoV—a target for vaccine and therapeutic development. *Nature Reviews Microbiology*, *7*(3), 226–236. <https://doi.org/10.1038/nrmicro2090>
- Gautret, P., Lagier, J.-C., Parola, P., Hoang, V. T., Meddeb, L., Mailhe, M., ... Raoult, D. (2020). Hydroxychloroquine and azithromycin as

- a treatment of COVID-19: Results of an open-label non-randomized clinical trial. *International Journal of Antimicrobial Agents*, <https://doi.org/10.1016/j.ijantimicag.2020.105949>
- Gracia-Ramos, A. E. (2020). Is the ACE2 Overexpression a Risk Factor for COVID-19 Infection? *Archives of Medical Research*, <https://doi.org/10.1016/j.arcmed.2020.03.011>
- Hoffmann, M., Kleine-Weber, H., Schroeder, S., Krüger, N., Herrler, T., Erichsen, S., ... Pöhlmann, S. (2020). SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and Is blocked by a clinically proven protease inhibitor. *Cell*, <https://doi.org/10.1016/j.cell.2020.02.052>
- Kastritis, P. L., & Bonvin, A. M. (2013). On the binding affinity of macromolecular interactions: Daring to ask why proteins interact. *Journal of the Royal Society, Interface*, *10*(79), 20120835. <https://doi.org/10.1098/rsif.2012.0835>
- Kutzner, C., Pall, S., Fechner, M., Esztermann, A., de Groot, B. L., & Grubmuller, H. (2015). Best bang for your buck: GPU nodes for GROMACS biomolecular simulations. *Journal of Computational Chemistry*, *36*(26), 1990–2008. <https://doi.org/10.1002/jcc.24030>
- Kutzner, C., Pall, S., Fechner, M., Esztermann, A., de Groot, B. L., & Grubmuller, H. (2019). More bang for your buck: Improved use of GPU nodes for GROMACS 2018. *Journal of Computational Chemistry*, *40*(27), 2418–2431. <https://doi.org/10.1002/jcc.26011>
- Lek, M., Karczewski, K. J., Minikel, E. V., Samocha, K. E., Banks, E., Fennell, T., ... MacArthur, D. G. (2016). Analysis of protein-coding genetic variation in 60,706 humans. *Nature*, *536*(7616), 285–291. <https://doi.org/10.1038/nature19057>
- Montgomery, S. B., Sammeth, M., Gutierrez-Arcelus, M., Lach, R. P., Ingle, C., Nisbett, J., ... Dermitzakis, E. T. (2010). Transcriptome genetics using second generation sequencing in a Caucasian population. *Nature*, *464*(7289), 773–777. <https://doi.org/10.1038/nature08903>
- Patel, A. B., & Verma, A. (2020). COVID-19 and angiotensin-converting enzyme inhibitors and angiotensin receptor blockers: What is the evidence? *JAMA*, <https://doi.org/10.1001/jama.2020.4812>
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, *38*(8), 904–909. <https://doi.org/10.1038/ng1847>
- Pronk, S., Páll, S., Schulz, R., Larsson, P., Bjelkmar, P., Apostolov, R., ... Lindahl, E. (2013). GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, *29*(7), 845–854. <https://doi.org/10.1093/bioinformatics/btt055>
- Stranger, B. E., Montgomery, S. B., Dimas, A. S., Parts, L., Stegle, O., Ingle, C. E., ... Dermitzakis, E. T. (2012). Patterns of cis regulatory variation in diverse human populations. *PLoS Genetics*, *8*(4), e1002639. <https://doi.org/10.1371/journal.pgen.1002639>
- Walls, A. C., Park, Y. J., Tortorici, M. A., Wall, A., McGuire, A. T., & Velesler, D. (2020). Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell*, <https://doi.org/10.1016/j.cell.2020.02.058>
- World Health Organization. (2020). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. Retrieved 23 March 2020, from <https://www.who.int/docs/default-source/coronavirus/who-china-joint-mission-on-covid-19-final-report.pdf>
- Worldometers. (2020). COVID-19 Coronavirus Pandemic. Retrieved 23 March 2020, from <https://www.worldometers.info/coronavirus/>
- Wu, C., Chen, X., Cai, Y., Xia, J., Zhou, X., Xu, S., ... Song, Y. (2020). Risk factors associated with acute respiratory distress syndrome and death in patients with Coronavirus disease 2019 pneumonia in Wuhan, China. *JAMA Internal Medicine*, <https://doi.org/10.1001/jamainternmed.2020.0994>
- Yan, R., Zhang, Y., Li, Y., Xia, L., Guo, Y., & Zhou, Q. (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, *367*(6485), 1444–1448. <https://doi.org/10.1126/science.abb2762>
- Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., ... Cao, B. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *Lancet*, *395*(10229), 1054–1062. [https://doi.org/10.1016/S0140-6736\(20\)30566-3](https://doi.org/10.1016/S0140-6736(20)30566-3)
- Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., ... Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. <https://doi.org/10.1038/s41586-020-2012-7>

How to cite this article: Li Q, Cao Z, Rahman P. Genetic variability of human angiotensin-converting enzyme 2 (hACE2) among various ethnic populations. *Mol Genet Genomic Med*. 2020;8:e1344. <https://doi.org/10.1002/mgg3.1344>