Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Infection, Genetics and Evolution 85 (2020) 104507

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research Paper

A survey of genetic variants in SARS-CoV-2 interacting domains of ACE2, TMPRSS2 and TLR3/7/8 across populations

In-Hee Lee[a,1], Ji-Won Lee[b,c,1], Sek Won Kong[a,d,*]

[a] Computational Health Informatics Program, Boston Children's Hospital, Boston, MA 02115, USA
[b] Department of Nephrology, Boston Children's Hospital, Boston, MA 02115, USA
[c] Department of Pharmacology, Graduate School of Dental Medicine, Hokkaido University, Sapporo 060-8549, Japan
[d] Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA

ARTICLE INFO

Keywords:
COVID-19
SARS-CoV-2
ACE2
TMPRSS2
TLRs
Genetic variant

ABSTRACT

The COVID-19 pandemic highlighted healthcare disparities in multiple countries. As such morbidity and mortality vary significantly around the globe between populations and ethnic groups. Underlying medical conditions and environmental factors contribute higher incidence in some populations and a genetic predisposition may play a role for severe cases with respiratory failure. Here we investigated whether genetic variation in the key genes for viral entry to host cells—*ACE2* and *TMPRSS2*—and sensing of viral genomic RNAs (i.e., *TLR3/7/8*) could explain the variation in incidence across diverse ethnic groups. Overall, these genes are under strong selection pressure and have very few nonsynonymous variants in all populations. Genetic determinant for the binding affinity between SARS-CoV-2 and ACE2 does not show significant difference between populations. Non-genetic factors are likely to contribute differential population characteristics affected by COVID-19. Nonetheless, a systematic mutagenesis study on the receptor binding domain of ACE2 is required to understand the difference in host-viral interaction across populations.

1. Introduction

Coronavirus disease 2019 (COVID-19) caused by SARS-CoV-2 is a pandemic as of Mar. 2020. Initial reports from China revealed diverse risk factors, clinical courses and outcome for a relatively homogenous population (Zhou et al., 2020a). Morbidity and mortality vary between populations (Yancy, 2020). African Americans and Latinos are disproportionately affected by COVID-19 and show significantly higher mortality compared to the other race and ethnic groups in the US (Wadhera et al., 2020) and in the UK (Kirby, 2020). A "healthcare disparity" must be responsible for the high incidence among minorities although socioeconomic factors, underlying medical conditions, and the difference in genetic susceptibility to SARS-CoV-2 infection may contribute (Chen et al., 2020). Of note, a 3p21.31 gene cluster—*SLC6A20, LZTFL1, CCR9, FYCO1, CXCR6* and *XCR1*—is associated with genetic susceptibility for severe COVID-19 cases with respiratory failure (Ellinghaus et al., 2020). To find allelic variation across populations in the genes that are known be involved in viral entry to the host cells and sensing of viral RNA in host immune cells, we surveyed publicly available databases of genomic variants.

SARS-CoV-2 is an enveloped and positive single-stranded RNA (ssRNA) virus and initiates human cell entry by binding of spike (S) protein present on the viral envelope to angiotensin converting enzyme 2 (ACE2) receptor on the host cells (Zhou et al., 2020b). The SARS-CoV S protein/ACE2 interface has been elucidated at the atomic level, and the ACE2 was found to be a key factor of SARS-CoV transmission (Li et al., 2005b). The binding mode of SARS-CoV-2 receptor binding domain (RBD) to ACE2 is nearly identical to SARS-CoV (Lan et al., 2020). The S protein is cleaved into S1 and S2 by the type 2 transmembrane serine protease (TMPRSS2) and endosomal cysteine proteases cathepsin B and L (CatB/L) (Du et al., 2009). TMPRSS2 is believed to be of utmost importance for SARS-CoV-2 entry into host cells. Recent studies demonstrated that an inhibitor of the protease activity of TMPRSS2—camostat mesylate—attenuated SARS-CoV-2 entry into lung epithelial cells suggesting a promising candidate for potential intervention against COVID-19 (Hoffmann et al., 2020). The C-terminal domain of S1 subunit is responsible for binding of SARS-CoV-2 to ACE2 and the S2 subunit undergoes a conformational change that result in virus-membrane fusion and entry into the target cell (Du et al., 2009). Viral genomic RNA is then released and translated into viral polymerase

* Corresponding author at: 401 Park Drive, LM5528.4, Boston, MA 02115, USA.
E-mail address: sekwon.kong@childrens.harvard.edu (S.W. Kong).
[1] Contributed equally.

https://doi.org/10.1016/j.meegid.2020.104507
Received 30 May 2020; Received in revised form 28 July 2020; Accepted 18 August 2020
Available online 26 August 2020

proteins for viral replication. Innate immune response is the first line of host defense mechanism for SARS-CoV-2 infection. Toll-like receptors recognize the viral RNA – double-stranded RNA (dsRNA) by TLR3 and ssRNA by TLR7 and TLR8 – and trigger innate immune responses such as the expression of inflammatory genes for type I interferons and pro-inflammatory cytokines (Iwasaki and Pillai, 2014; Iwasaki and Yang, 2020).

Here we surveyed the genetic variants in functional residues of *ACE2, TMPRSS2, CTSB/L* (CatB/L), and *TLR3/7/8* to investigate the difference in the genetic predisposition to the susceptibly of SARS-CoV-2 infection and the initiation of innate immune response. For ACE2, we investigated genetic variants in the residues on the interface to SARS-CoV-2 RBD from recent structural analyses (Hussain et al., 2020; Lan et al., 2020; Shang et al., 2020; Wrapp et al., 2020; Yan et al., 2020). Given the high sequence similarity between S proteins of SARS-CoV-2 and SARS-CoV, we also investigated the residues shown to inhibit interactions from in vitro mutagenesis analysis (Li et al., 2005b). We checked two residues reported to cause loss of cleavage activity of TMPRSS2 (Afar et al., 2001) and the enzymatically active sites for CatB/L. A total of 16 residues of TLR7 that are necessary for ssRNA-induced activation (Zhang et al., 2016) and the residues affecting reaction to ssRNAs from in vitro mutagenesis studies for TLR3 (Bell et al., 2006; de Bouteiller et al., 2005; Sarkar et al., 2007) and for TLR8 (Tanji et al., 2015) were checked for sequence variation. Additionally, we searched for nonsynonymous variants that would cause loss of gene function (i.e., frameshift, in-frame insertion/deletion, stop-gain, splice-disrupting, start-lost and stop-lost). The list of reported genetic variants in the genes and their allele frequencies (AFs) were compiled from three population-scale genomic variants databases— gnomAD (Karczewski et al., 2020), Korean Reference Genome Database (Jung et al., 2020), and TogoVar (a Japanese genetic variation database available at https://togovar.biosciencedbc.jp/) —and three whole-genome sequencing datasets (i.e., 1000 Genomes Project (Clarke et al., 2017), Gene-Tissue Expression (Consortium et al., 2017), and Simons Genome Diversity Project (Mallick et al., 2016)).

ACE2 is highly conserved with few nonsynonymous variants in the interacting domain with the SARS-CoV-2 RBM (Lan et al., 2020). Of 370 coding variants in *ACE2*, 248 were nonsynonymous variants with the highest AF of 1.6% (rs41303171). Within 33 residues interfacing the SARS-CoV-2 RBM, 19 variants (including 4 synonymous variants) were found with average AF of 0.03% (ranges 0–0.39%) (Table 1). Only one of the 19 variants (rs4646116; K26R) had global AF greater than 0.1% (AF = 0.39%). Rs4646116 (NC_000023.10:g.15618958 T > C) had the largest AF difference across populations: the lowest AF (0.007%) in East Asian and the highest (0.59%) in Non-Finnish European. The impact of this variant is not yet investigated with structural analysis but was not classified as deleterious (of possible impact on the structure and function of the protein) by in silico prediction algorithms such as SIFT and Polyphen2. The other variants were either very rare (i.e., population AF < 0.1%) or unique to a population or two. For the five known residues—K31, E35, D38, M82 and K353—that were reported to significantly change binding affinity to viral S protein (Li et al., 2005a), we found three variants: rs758278442 (K31K), rs1348114695 (E35K), and rs766996587 (M82I). However, all three were either synonymous or predicted to have little impact on protein. Rs758278442 showed significant AF difference across populations, especially among east Asian populations. It is found only among east Asian individuals in gnomAD – consists of 1909 Korean, 76 Japanese, and 7212 other east Asian individuals – with AF of 0.022%. The variant is also found at Korean Reference Genome Database (N = 1722) with AF of 0.029%, similar value to gnomAD. However, it was found with higher AF of 0.23% at Japanese genetic variation database (N = 3552). Rs1348114695 at residue 35 was found only in European and east Asian populations with very low frequencies: 0.001% and 0.014%, respectively. Lastly, rs766996587 at residue 82 was found only in African population (AF = 0.026%). Nonetheless, protein modeling predicts little topological difference between all ACE2 variants and wild-type ACE in their binding to S protein (Hussain et al., 2020). Therefore, we expect minimal genetic variance across populations critically affecting interaction between ACE2 and SARS-CoV-2. Fig. 1A illustrates the 19 variants over known functional protein domains of ACE2.

The proteolysis activity of TMPRSS2 is crucial for viral entry to host cells (Hoffmann et al., 2020). Two residues, V292 and M478, are reported to impact the catalytic activity of TMPRSS2 (Afar et al., 2001) but we found no variants at these residues (Supplementary Table 1). Reported variants for *TMPRSS2* contain 417 nonsynonymous variants including 40 loss-of-function variants. All of loss-of-function variants were very rare (AF < 0.01%). The rest of nonsynonymous variants were also of low frequencies (AF < 0.1%) mostly. Of the only 5 nonsynonymous variants with AF > 0.1%, rs12329760 (V192M, global AF = 24.88%) predicted deleterious and its AF ranged from 15.33% (Latino) to 38.38% (East Asian). Further studies are required to test whether rs12329760 could exert functional impact on TMPRSS2 activity. Thus, differences in TMPRSS2 activity caused either by variants at critical loci or by loss-of-function variants are unlikely. SARS-CoV-2 uses both TMPRSS2 and the endosomal cysteine proteases cathepsin B and L (CTSB and CTSL) for priming S protein (Hoffmann et al., 2020). UniProt entries for human CTSB and CTSL report 3 active sites. We found 3 variants in the active sites for CTSB (two missense variants and one synonymous variant), and one missense variant for CTSL (Table 1 and Fig. 1B). Although all missense variants on active sites of CTSB/L are predicted deleterious, they were of very low allele frequencies (AF < 0.01%). *CTSB* has 429 nonsynonymous variants including 51 loss-of-function variants (all with AF < 0.01%). *CTSL* has 211 nonsynonymous variants including 17 loss-of-function variants. Of note, one of 17 variants in *CTSL* (rs2378757, NC_000009.11:g.90343780A > C) is a common allele (global AF of 70.32%, population AF ranges from 62.66% to 98.48%). The variant changes stop codon to serine for one *CTSL* transcript isoform (ENST00000342020.5) but falls in intron for the other transcript isoforms.

Next we checked genetic variants in TLRs that sense viral RNAs and initiate innate immune responses. There were 7 variants—4 synonymous and 3 nonsynonymous—in the 16 residues of ssRNA interacting domain of TLR7 (Table 1 and Fig. 1C). Most variants were of extremely low frequencies (AF < 0.01%) except for one synonymous variant, rs769401373 (D135D), found only in east Asian population (AF = 0.46%). *TLR7* harbors 232 nonsynonymous variants including 8 loss-of-function variants. As in *TMPRSS2*, AFs of loss-of-function variants were also very low (AF < 0.01%). The UniProt entries for TLR3 and TLR8 list 10 sites (6 for TLR3 (Bell et al., 2006; de Bouteiller et al., 2005; Sarkar et al., 2007) and 4 for TLR8 (Tanji et al., 2015)) from in vitro mutagenesis study that impact their response to viral infection (sensing of dsRNA or ssRNA, respectively). For these loci, two missense variants on *TLR3* and one missense variant with one synonymous variant on *TLR8* were found (Table 1 and Fig. 1C). All of these variants in *TLRs* were very rare (AF < 0.01%) across all populations.

To summarize, the critical loci for host-viral interaction and sensing viral genomic RNA are highly conserved in all populations with few very rare variants. Especially, *ACE2* and *TLR7* seem to be under strong selection pressure as reflected in their relatively lower number of loss-of-function variants than expected in large variant databases such as gnomAD (Karczewski et al., 2020): three observed variants out of 31 expected ones for *ACE2* and two observed variants out of 20.7 expected ones for *TLR7*. Moreover, nonsynonymous variants in these genes were mostly of very low frequencies which suggests the chance of gene function altered by these variants would be unlikely, compared to the incidence of COVID-19 around the globe. Other factors such as existing medical conditions and environmental risk factors could contribute the regulation of expression of these key genes in susceptible individuals; however, further studies are required to elucidate potential associations.

The majority of infected individuals experience no or mild symptoms of upper respiratory tract infection; however, for some

Table 1
Genetic variants in the genes related to host-viral interaction and sensing of viral RNAs.

Genes	Residues	AA changes from mutagenesis studies	Residue loci (b37)	Reported variants within the residues				Variant allele frequencies	
				Variants	RS ID	Impact	AA Change	gnomAD[1]	
								Global	African
ACE2	S19[7,10]		X:15618978–15,618,980	NC_000023.10:-g.15618980A > G	rs73635825	Missense	S > P	0.031%	0.332%
	A25[11]	24–26, QAK-KAE	X:15618957–15,618,965	NC_000023.10:-g.15618960G > A	rs761614932	Synonymous	=	0.001%	
	K26[11]			NC_000023.10:-g.15618958 T > C	rs4646116	Missense	K > R	0.388%	0.095%
				NC_000023.10:-g.15618959 T > C	rs1299103394	Missense	K > E	0.001%	
	T27[7,9]		X:15618954–15,618,956	NC_000023.10:-g.15618956 T > C	rs781255386	Missense	T > A	0.001%	
	K31[7,9–11]	K31D	X:15618942–15,618,944	NC_000023.10:-g.15618942C > T	rs758278442	Synonymous	=	0.002%	
	H34[7–10]		X:15618933–15,618,935	NC_000023.10:-g.15618933G > A	rs368655410	Synonymous	=	0.063%	
	E35[7,9,10]		X:15618930–15,618,932	NC_000023.10:-g.15618932C > T	rs1348114695	Missense	E > K	0.002%	
	E37[7,9,10]		X:15618924–15,618,926	NC_000023.10:-g.15618926C > T	rs146676783	Missense	E > K	0.004%	0.011%
	K68[11]	K68D	X:15613109–15,613,111	NC_000023.10:-g.15613111 T > C	rs755691167	Missense	K > E	0.001%	
	M82[7–11]	82–84, MYP-NFS	X:15613061–15,613,069	NC_000023.10:-g.15613067C > T	rs766996587	Missense	M > I	0.002%	0.026%
	P84[11]			NC_000023.10:-g.15613063G > T	rs759134032	Missense	P > T	0.001%	
	E329[7]		X:15599427–15,599,429	NC_000023.10:-g.15599428 T > C	rs143936283	Missense	E > G	0.003%	
	D355[7,9,11]	D355A	X:15599349–15,599,351	NC_000023.10:-g.15599351C > T	rs961360700	Missense	D > N	0.001%	
	P389[11]	P389A	X:15596342–15,596,344	NC_000023.10:-g.15596343G > T	rs762890235	Missense	P > H	0.004%	
	P426[11]	425–427, SPD-PSN	X:15596228–15,596,236	NC_000023.10:-g.15596233G > C	rs1238146879	Missense	P > A	0.001%	
				NC_000023.10:-g.15596231G > A	rs1335386721	Synonymous	=	0.001%	
	D427[11]			NC_000023.10:-g.15596230C > A	rs1316056737	Missense	D > Y	0.001%	0.015%
	R559[11]	R559S	X:15589907–15,589,909	NC_000023.10:-g.15589907C > G	rs1016777825	Missense	R > S	0.001%	
TLR7[12]	F351	F351A	X:12904678–12,904,680	NC_000023.10:-g.12904680 T > C	rs200549906	Synonymous	=	0.002%	
	L557	L557A	X:12905296–12,905,298	NC_000023.10:-g.12905296C > T	rs1419393304	Missense	L > F	0.002%	
	T586	T586A	X:12905383–12,905,385	NC_000023.10:-g.12905385 T > A	rs185622718	Synonymous	=	0.001%	
	L105	L105A	X:12903940–12,903,942	NC_000023.10:-g.12903940C > T	rs773554481	Synonymous	=	0.001%	
	D135	D135A	X:12904030–12,904,032	NC_000023.10:-g.12904032 T > C	rs769401373	Synonymous	=	0.033%	
	R186	R186A	X:12904183–12,904,185	NC_000023.10:-g.12904184G > A	rs868177091	Missense	R > Q	0.001%	

(*continued on next page*)

Table 1 (*continued*)

Genes	Residues	AA changes from mutagenesis studies	Residue loci (b37)	Reported variants within the residues				Variant allele frequencies	
				Variants	RS ID	Impact	AA Change	gnomAD[1]	
								Global	African
	R473	R473A	X:12905044–12,905,046	NC_000023.10:-g.12905045G > A	rs754381606	Missense	R > K	0.001%	
CTSB (CatB)[13]	C108		8:11708378–11,708,380	NC_000008.10:-g.11708378G > A	rs759843078	Synonymous	=	0.002%	
	H278		8:11703258–11,703,260	NC_000008.10:-g.11703259 T > C	rs1373655221	Missense	H > R	0.0004%	Only found in Finnish population (0.005%)
				NC_000008.10:-g.11703260G > A	rs1225109229	Missense	H > Y	0.0004%	
CTSL (CatL)[14]	C138		9:90343515–90,343,517	NC_000009.11:-g.90343515 T > C	rs757571238	Missense	C > R	0.001%	
TLR3[15]	H539	H539E	4:187004455–187,004,457	NC_000004.11:-g.187004456A > G	rs776387492	Missense	H > R	0.001%	
	Y759	Y759F	4:187005115–187,005,117	NC_000004.11:-g.187005115 T > C	rs768605211	Missense	Y > H	0.001%	
TLR8[16]	Y348	Y348A	X:12938201–12,938,203	NC_000023.10:-g.12938202A > G	rs1175381548	Missense	Y > C	0.001%	Only found in Finnish population (0.006%)
				NC_000023.10:-g.12938203 T > C	rs768875789	Synonymous	=	0.001%	

Genes	Variant allele frequencies								
	gnomAD[1]				1KGP[2]	SGDP[3]	GTEx[4]	KRGDB[5]	TogoVar[6]
	Latino	European	East Asian	South Asian					
ACE2			0.007%		0.030%				
	0.325%	0.587%	0.007%	0.131%	0.210%	0.333%	0.477%		
		0.001%							
	0.007%								
			0.022%			0.333%		0.029%	0.230%
		0.033%	0.027%	0.026%					0.040%
		0.001%	0.014%						
						0.333%			
				0.011%					
				0.005%					
		0.007%							
		0.003%							
	0.018%	0.002%							
		0.001%							
		0.001%							
	0.004%								
TLR7[12]		0.004%							
		0.001%							
			0.007%		0.030%				

(*continued on next page*)

Table 1 (*continued*)

Genes	Variant allele frequencies								
	gnomAD[1]				1KGP[2]	SGDP[3]	GTEx[4]	KRGDB[5]	TogoVar[6]
	Latino	European	East Asian	South Asian					
		0.001%							
			0.458%		0.050%				
				0.005%					
				0.005%					
CTSB (CatB)[13]				0.013%					
	Only found in Finnish population (0.005%)								
		0.001%							
CTSL (CatL)[14]		0.003%							
TLR3[15]			0.005%	0.003%				0.029%	
				0.007%					
TLR8[16]	Only found in Finnish population (0.006%)								
			0.007%						

[1] The genome aggregate database (gnomAD), v2.1.1. https://gnomad.broadinstitute.org. Allele frequencies for European are from Non-Finnish European population.
[2] 1000 Genomes Project (1KGP), phase 3. https://www.internationalgenome.org
[3] Simons Genome Diversity Project (SGDP). https://www.simonsfoundation.org/simons-genome-diversity-project/
[4] Gene-Tissue Expression project (GTEx), v8 whole genomes. https://gtexportal.org/home/
[5] Korean Reference Genome Database (KRGDB). http://coda.nih.go.kr/coda/KRGDB/index.jsp
[6] NBDC's integrated database of Japanese genomic variation (TogoVar). https://togovar.biosciencedbc.jp
[7] Shang et al., Nature, 2020
[8] Yan et al., Science, 2020
[9] Lan et al., Nature, 2020
[10] Hussain et al., J Med Vir, 2020
[11] Based on mutagenesis studies from UniProt protein information for Q9BYF1 (ACE2_HUMAN). https://www.uniprot.org/uniprot/Q9BYF1
[12] The ligand-binding sites for small ligands and ssRNA from Zhang et al., Immunity, 2016
[13] Based on active sites from UniProt protein information for P07858 (CATB_HUMAN). https://www.uniprot.org/uniprot/P07858
[14] Based on active sites from UniProt protein information for P07711 (CATL1_HUMAN). https://www.uniprot.org/uniprot/P07711
[15] Based on mutagenesis studies from UniProt protein information for O15455 (TLR3_HUMAN). https://www.uniprot.org/uniprot/O15455
[16] Based on mutagenesis studies from UniProt protein information for Q9NR97 (TLR8_HUMAN). https://www.uniprot.org/uniprot/Q9NR97

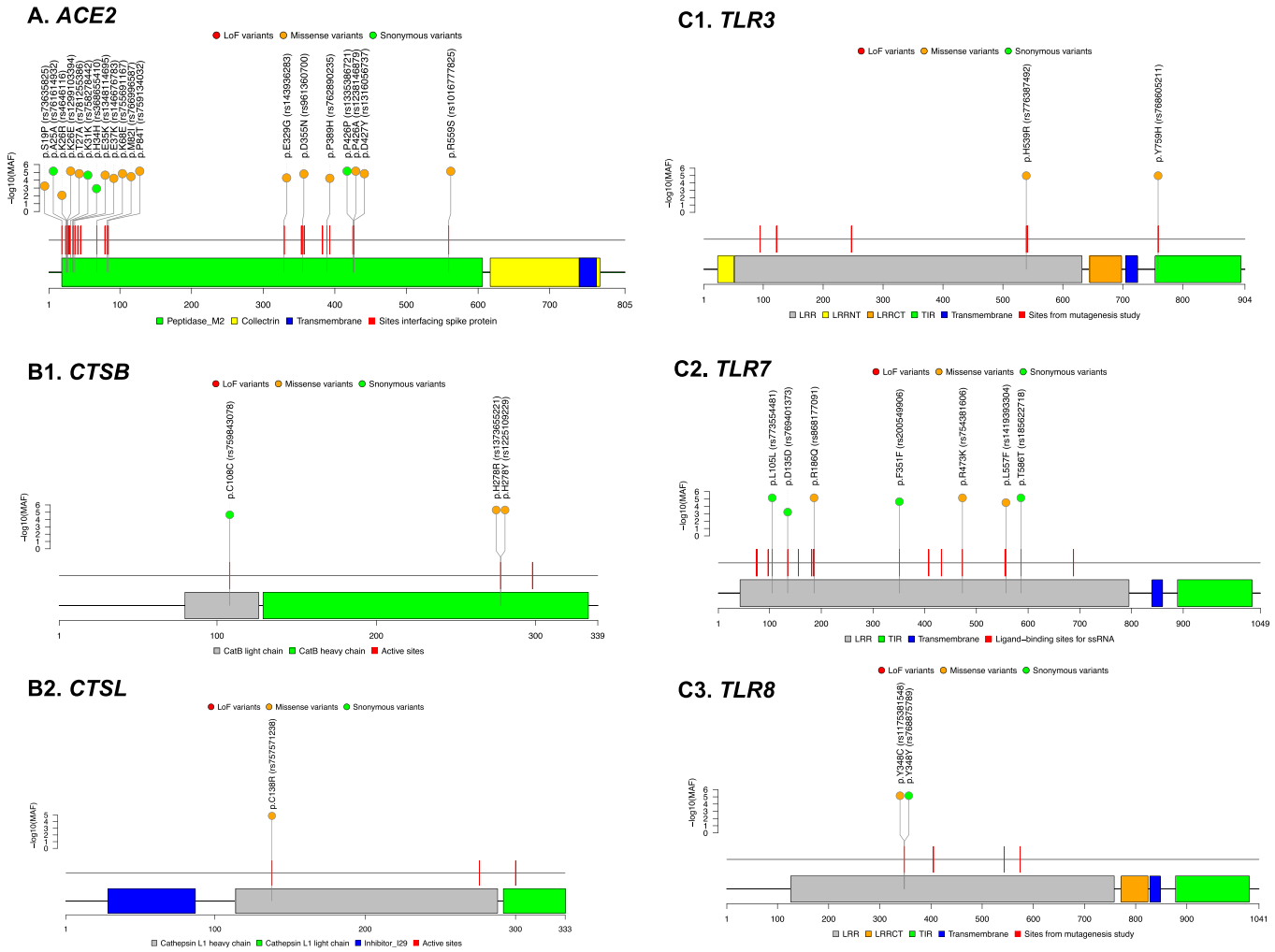

Fig. 1. Location of genetic variants relative to known functional domains of (A) ACE2, (B) CTLB/L and (C) TLR3/7/8. For each gene, x-axis represents positions in protein sequence. The block diagram directly above the x-axis depicts major protein domains in different colored boxes. The vertical red lines above domains correspond to the critical residues investigated in this study. Each of the circles with grey lines represents variant found on the critical loci. The circles are colored differently based on their calculated effect on protein: loss-of-function (LoF) variants (red), missense variants (orange), and synonymous variants (green). The height of each circle denotes variant allele frequency (in -log10 scale). The higher the circle, the lower the allele frequency. Of note, TMPRSS2 does not have any reported genetic variant in enzymatically active functional domain. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

individuals, the consequence of SARS-CoV-2 infection could be fatal. One of the contributing factors may be the viral load due to differential affinity of viral spike proteins to ACE2 and the efficiency of cleavage by TMPRSS2 that are essential for virus to enter and replicate inside of host cells. We did not find genetic variation between populations while there is a significant difference in incidence and mortality between race and ethnic groups in the U.S. Therefore, underlying medical conditions, age, environmental factors (e.g., air pollution, smoking, and humidity), and a healthcare disparity influence morbidity and mortality from COVID-19 considering the allelic spectrum for the key genes associated with viral entry. Nonetheless, genetic susceptibility may play a role for severe cases with respiratory failure (Ellinghaus et al., 2020).

The population-scale genotype databases and datasets used in this study have limitations from relatively small sample size and imbalanced and incomplete representation of various human populations. Thus, there could be unreported variants in *ACE2, TMPRSS2*, and *TLR3/7/8* that may be associated with change of susceptibility to COVID-19. With additional population-scale genomic databases for diverse populations, it will be possible to identify the individuals with rare genetic variants such as rs758278442 in the interacting domain of ACE2 and the genetic predisposition to cytokine storm that causes an acute progress of illness in young people. In parallel, a systematic mutagenesis analysis of the RBM of ACE2 is highly required to understand the difference in host-viral interaction across populations (Lan et al., 2020).

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2020.104507.

Declaration of Competing Interest

On behalf of all authors, the corresponding author states that there is no conflict of interest.

Acknowledgements

SWK was supported in part by the grants from the National Institution of Health (R01MH107205, R24OD024622 and U01TR002623). JWL was supported by Grant-in-Aids for Scientific Research from the Japan Society for the Promotion of Science (19K10044).

References

Afar, D.E., Vivanco, I., Hubert, R.S., Kuo, J., Chen, E., Saffran, D.C., Raitano, A.B., Jakobovits, A., 2001. Catalytic cleavage of the androgen-regulated TMPRSS2 protease results in its secretion by prostate and prostate cancer epithelia. Cancer Res. 61, 1686–1692.

Bell, J.K., Askins, J., Hall, P.R., Davies, D.R., Segal, D.M., 2006. The dsRNA binding site of human toll-like receptor 3. Proc. Natl. Acad. Sci. U. S. A. 103, 8792–8797.

de Bouteiller, O., Merck, E., Hasan, U.A., Hubac, S., Benguigui, B., Trinchieri, G., Bates, E.E., Caux, C., 2005. Recognition of double-stranded RNA by human toll-like receptor 3 and downstream receptor signaling requires multimerization and an acidic pH. J. Biol. Chem. 280, 38133–38145.

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., Zhang, L., 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet 395, 507–513.

Clarke, L., Fairley, S., Zheng-Bradley, X., Streeter, I., Perry, E., Lowy, E., Tasse, A.M., Flicek, P., 2017. The international genome sample resource (IGSR): a worldwide collection of genome variation incorporating the 1000 genomes project data. Nucleic Acids Res. 45, D854–D859.

Consortium, G.T., Laboratory, D.A., Coordinating Center -Analysis Working, G., Statistical Methods groups-Analysis Working, G., Enhancing, G.g., Fund, N.I.H.C., Nih/Nci, Nih/Nhgri, Nih/Nimh, Nih/Nida, Biospecimen Collection Source Site, N., Biospecimen Collection Source Site, R., Biospecimen Core Resource, V., Brain Bank Repository-University of Miami Brain Endowment, B., Leidos Biomedical-Project, M., Study, E., Genome Browser Data, I., Visualization, E.B.I., Genome Browser Data, I., Visualization-Ucsc Genomics Institute, U.o.C.S.C., Lead, a., Laboratory, D.A., Coordinating, C., management, N.I.H.p., Biospecimen, c., Pathology, e, Q.T.L.m.w.g., Battle, A., Brown, C.D., Engelhardt, B.E., Montgomery, S.B, 2017. Genetic effects on gene expression across human tissues. Nature 550, 204–213.

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. Nat Rev Microbiol 7, 226–236.

Ellinghaus, D., Degenhardt, F., Bujanda, L., Buti, M., Albillos, A., Invernizzi, P., Fernandez, J., Prati, D., Baselli, G., Asselta, R., Grimsrud, M.M., Milani, C., Aziz, F., Kassens, J., May, S., Wendorff, M., Wienbrandt, L., Uellendahl-Werth, F., Zheng, T., Yi, X., de Pablo, R., Chercoles, A.G., Palom, A., Garcia-Fernandez, A.E., Rodriguez-Frias, F., Zanella, A., Bandera, A., Protti, A., Aghemo, A., Lleo, A., Biondi, A., Caballero-Garralda, A., Gori, A., Tanck, A., Carreras Nolla, A., Latiano, A., Fracanzani, A.L., Peschuck, A., Julia, A., Pesenti, A., Voza, A., Jimenez, D., Mateos, B., Nafria Jimenez, B., Quereda, C., Paccapelo, C., Gassner, C., Angelini, C., Cea, C., Solier, A., Pestana, D., Muniz-Diaz, E., Sandoval, E., Paraboschi, E.M., Navas, E., Garcia Sanchez, F., Ceriotti, F., Martinelli-Boneschi, F., Peyvandi, F., Blasi, F., Tellez, L., Blanco-Grau, A., Hemmrich-Stanisak, G., Grasselli, G., Costantino, G., Cardamone, G., Foti, G., Aneli, S., Kurihara, H., ElAbd, H., My, I., Galvan-Femenia, I., Martin, J., Erdmann, J., Ferrusquia-Acosta, J., Garcia-Etxebarria, K., Izquierdo-Sanchez, L., Bettini, L.R., Sumoy, L., Terranova, L., Moreira, L., Santoro, L., Scudeller, L., Mesonero, F., Roade, L., Ruhlemann, M.C., Schaefer, M., Carrabba, M., Riveiro-Barciela, M., Figuera Basso, M.E., Valsecchi, M.G., Hernandez-Tejero, M., Acosta-Herrera, M., D'Angio, M., Baldini, M., Cazzaniga, M., Schulzky, M., Cecconi, M., Wittig, M., Ciccarelli, M., Rodriguez-Gandia, M., Bocciolone, M., Miozzo, M., Montano, N., Braun, N., Sacchi, N., Martinez, N., Ozer, O., Palmieri, O., Faverio, P., Preatoni, P., Bonfanti, P., Omodei, P., Tentorio, P., Castro, P., Rodrigues, P.M., Blandino Ortiz, A., de Cid, R., Ferrer, R., Gualtierotti, R., Nieto, R., Goerg, S., Badalamenti, S., Marsal, S., Matullo, G., Pelusi, S., Juzenas, S., Aliberti, S., Monzani, V., Moreno, V., Wesse, T., Lenz, T.L., Pumarola, T., Rimoldi, V., Bosari, S., Albrecht, W., Peter, W., Romero-Gomez, M., D'Amato, M., Duga, S., Banales, J.M., Hov, J.R., Folseraas, T., Valenti, L., Franke, A., Karlsen, T.H., Severe Covid, G.G., 2020. Genomewide association study of Severe Covid-19 with respiratory failure. N. Engl. J. Med. https://doi.org/10.1056/NEJMoa2020283.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., Muller, M.A., Drosten, C., Pohlmann, S., 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 181 (271–280), e278.

Hussain, M., Jabeen, N., Raza, F., Shabbir, S., Baig, A.A., Amanullah, A., Aziz, B., 2020. Structural variations in human ACE2 may influence its binding with SARS-CoV-2 spike protein. J. Med. Virol. 92, 1580–1586.

Iwasaki, A., Pillai, P.S., 2014. Innate immunity to influenza virus infection. Nat Rev Immunol 14, 315–328.

Iwasaki, A., Yang, Y., 2020. The potential danger of suboptimal antibody responses in COVID-19. Nat Rev Immunol 20, 339–341.

Jung, K.S., Hong, K.W., Jo, H.Y., Choi, J., Ban, H.J., Cho, S.B., Chung, M., 2020. KRGDB: the large-scale variant database of 1722 Koreans based on whole genome sequencing. Database (Oxford). https://doi.org/10.1093/database/baz146.

Karczewski, K.J., Francioli, L.C., Tiao, G., Cummings, B.B., Alfoldi, J., Wang, Q., Collins, R.L., Laricchia, K.M., Ganna, A., Birnbaum, D.P., Gauthier, L.D., Brand, H., Solomonson, M., Watts, N.A., Rhodes, D., Singer-Berk, M., England, E.M., Seaby, E.G., Kosmicki, J.A., Walters, R.K., Tashman, K., Farjoun, Y., Banks, E., Poterba, T., Wang, A., Seed, C., Whiffin, N., Chong, J.X., Samocha, K.E., Pierce-Hoffman, E., Zappala, Z., O'Donnell-Luria, A.H., Minikel, E.V., Weisburd, B., Lek, M., Ware, J.S., Vittal, C., Armean, I.M., Bergelson, L., Cibulskis, K., Connolly, K.M., Covarrubias, M., Donnelly, S., Ferriera, S., Gabriel, S., Gentry, J., Gupta, N., Jeandet, T., Kaplan, D., Llanwarne, C., Munshi, R., Novod, S., Petrillo, N., Roazen, D., Ruano-Rubio, V., Saltzman, A., Schleicher, M., Soto, J., Tibbetts, K., Tolonen, C., Wade, G., Talkowski, M.E., Genome Aggregation Database, C, Neale, B.M., Daly, M.J., MacArthur, D.G., 2020. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 581, 434–443.

Kirby, T., 2020. Evidence mounts on the disproportionate effect of COVID-19 on ethnic minorities. Lancet Respir. Med. 8, 547–548.

Lan, J., Ge, J., Yu, J., Shan, S., Zhou, H., Fan, S., Zhang, Q., Shi, X., Wang, Q., Zhang, L., Wang, X., 2020. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature 581, 215–220.

Li, F., Li, W., Farzan, M., Harrison, S.C., 2005a. Structure of SARS coronavirus spike receptor-binding domain complexed with receptor. Science 309, 1864–1868.

Li, W., Zhang, C., Sui, J., Kuhn, J.H., Moore, M.J., Luo, S., Wong, S.K., Huang, I.C., Xu, K., Vasilieva, N., Murakami, A., He, Y., Marasco, W.A., Guan, Y., Choe, H., Farzan, M., 2005b. Receptor and viral determinants of SARS-coronavirus adaptation to human ACE2. EMBO J. 24, 1634–1643.

Mallick, S., Li, H., Lipson, M., Mathieson, I., Gymrek, M., Racimo, F., Zhao, M., Chennagiri, N., Nordenfelt, S., Tandon, A., Skoglund, P., Lazaridis, I., Sankararaman, S., Fu, Q., Rohland, N., Renaud, G., Erlich, Y., Willems, T., Gallo, C., Spence, J.P., Song, Y.S., Poletti, G., Balloux, F., van Driem, G., de Knijff, P., Romero, I.G., Jha, A.R., Behar, D.M., Bravi, C.M., Capelli, C., Hervig, T., Moreno-Estrada, A., Posukh, O.L., Balanovska, E., Balanovsky, O., Karachanak-Yankova, S., Sahakyan, H., Toncheva, D., Yepiskoposyan, L., Tyler-Smith, C., Xue, Y., Abdullah, M.S., Ruiz-Linares, A., Beall, C.M., Di Rienzo, A., Jeong, C., Starikovskaya, E.B., Metspalu, E., Parik, J., Villems, R., Henn, B.M., Hodoglugil, U., Mahley, R., Sajantila, A., Stamatoyannopoulos, G., Wee, J.T., Khusainova, R., Khusnutdinova, E., Litvinov, S., Ayodo, G., Comas, D., Hammer, M.F., Kivisild, T., Klitz, W., Winkler, C.A., Labuda, D., Bamshad, M., Jorde, L.B., Tishkoff, S.A., Watkins, W.S., Metspalu, M., Dryomov, S., Sukernik, R., Singh, L., Thangaraj, K., Paabo, S., Kelso, J., Patterson, N., Reich, D., 2016. The Simons genome diversity project: 300 genomes from 142 diverse populations. Nature 538, 201–206.

Sarkar, S.N., Elco, C.P., Peters, K.L., Chattopadhyay, S., Sen, G.C., 2007. Two tyrosine residues of toll-like receptor 3 trigger different steps of NF-kappa B activation. J. Biol. Chem. 282, 3423–3427.

Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., Li, F., 2020. Structural basis of receptor recognition by SARS-CoV-2. Nature 581, 221–224.

Tanji, H., Ohto, U., Shibata, T., Taoka, M., Yamauchi, Y., Isobe, T., Miyake, K., Shimizu, T., 2015. Toll-like receptor 8 senses degradation products of single-stranded RNA. Nat. Struct. Mol. Biol. 22, 109–115.

Wadhera, R.K., Wadhera, P., Gaba, P., Figueroa, J.F., Joynt Maddox, K.E., Yeh, R.W., Shen, C., 2020. Variation in COVID-19 hospitalizations and deaths across New York City boroughs. JAMA. 323 (21), 2192–2195.

Wrapp, D., Wang, N., Corbett, K.S., Goldsmith, J.A., Hsieh, C.L., Abiona, O., Graham, B.S., McLellan, J.S., 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. Science 367, 1260–1263.

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, 1444–1448.

Yancy, C.W., 2020. COVID-19 and African Americans. JAMA. 323 (19), 1891–1892.

Zhang, Z., Ohto, U., Shibata, T., Krayukhina, E., Taoka, M., Yamauchi, Y., Tanji, H., Isobe, T., Uchiyama, S., Miyake, K., Shimizu, T., 2016. Structural analysis reveals that toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. Immunity 45, 737–748.

Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., Cao, B., 2020a. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 395, 1054–1062.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., Chen, H.D., Chen, J., Luo, Y., Guo, H., Jiang, R.D., Liu, M.Q., Chen, Y., Shen, X.R., Wang, X., Zheng, X.S., Zhao, K., Chen, Q.J., Deng, F., Liu, L.L., Yan, B., Zhan, F.X., Wang, Y.Y., Xiao, G.F., Shi, Z.L., 2020b. A pneumonia outbreak associated with a new coronavirus of probable bat origin. Nature 579, 270–273.