Journal of proteome research

pubs.acs.org/jpr

Perspective

COVID-19: The Effect of Host Genetic Variations on Host–Virus Interactions

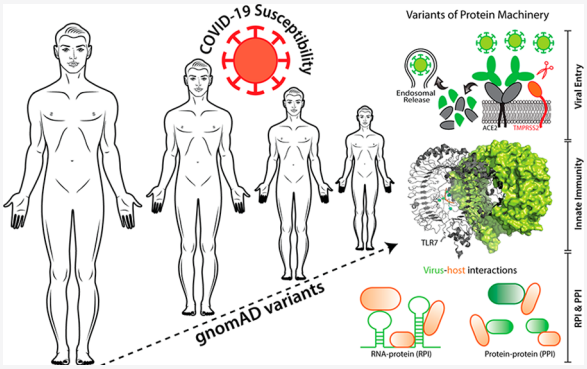
Suvobrata Chakravarty*

Cite This: https://dx.doi.org/10.1021/acs.jproteome.0c00637

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Spurred into action by the COVID-19 pandemic, the global scientific community has, in a short of period of time, made astonishing progress in understanding and combating COVID-19. Given the known human protein machinery for (a) SARS-CoV-2 entry, (b) the host innate immune response, and (c) virus–host interactions (protein–protein and RNA–protein), the potential effects of human genetic variation in this machinery, which may contribute to clinical differences in SARS-CoV-2 pathogenesis and help determine individual risk for COVID-19 infection, are explored. The Genome Aggregation Database (gnomAD) was used to show that several *rare* germline exome variants of proteins in these pathways occur in the human population, suggesting that carriers of these *rare* variants (especially for proteins of innate immunity pathways) are at risk for severe symptoms (like the severe symptoms in patients who are known to be *rare* variant carriers), whereas carriers of other variants could have a protective advantage against infection. The occurrence of genetic variation is thus expected to motivate the experimental probing of natural variants to understand the mechanistic differences in SARS-CoV-2 pathogenesis from one individual to another.



KEYWORDS: *exome variants, host–virus interactions, disease susceptibility, protein–protein and RNA–interactions, viral entry and innate immunity pathway proteins, loss-of-function tolerance*

INTRODUCTION

The worldwide spread of the coronavirus disease 2019 (COVID-19),[1] caused by the novel virus severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2),[2] has resulted in 1.33 million deaths globally, having advanced at a staggering rate of ∼22 000 new infections and ∼385 deaths every hour (WHO, November 18 situation report). During this health crisis, a knowledge of the risk factors of COVID-19 infection is valuable for determining the most appropriate measures, especially given our limited resources, to mitigate the threat. That is, in addition to efforts toward the discovery of a cure (e.g., vaccines or therapeutics), the identification of individual risk for separating low-risk from high-risk individuals is also important (e.g., in maintaining a minimal workforce to prevent the global economy from plunging into a depression). Determining the relationship between the genetic background of an individual and the susceptibility to and progression of the disease remains an important task in understanding and determining an individual's risk for COVID-19. Understanding the molecular basis of susceptibility is of vital interest and has been the subject of a number of important studies.[3−10] The emerging picture on susceptibility from these studies suggests that multiple genetic factors could contribute to the risk of SARS-CoV-2 infection and the severity of COVID-19 disease.[11,12] In general, proteins engaged in pathways related to the viral life cycle and host defense are important components of the genetic factors.[11,12] Polymorphic variants of human proteins engaged in these pathways are thus expected to make important contributions toward COVID-19 susceptibility.[11,12] Therefore, a discussion is presented here on polymorphic variants (carried by an individual) of pathway-specific proteins that have the potential to link the susceptibility to the disease to the genetic background of an individual.

This discussion has only been possible due to the availability of results from a number of molecular studies (Table 1) carried out in the last 8 months at a very rapid pace in several different disciplines, providing critical advances in understanding the molecular mechanisms of COVID-19 pathogenesis. Despite the health crisis, the efforts of scientists in unraveling the molecular

Received: August 17, 2020

Table 1. Breakthrough Reports on Molecular Mechanisms of SARS-CoV-2 Pathogenesis

study topic	examples of specific reports
SARS-CoV-2 spike (S protein) function, structure, and stabilization	(a) Discovery of the proteolysis of the S protein prior to cell fusion[35−38]
	(b) Structure determination of the S protein[39] and in situ structural analysis[40] of profusion spikes of intact[41] or inactivated virus[42]
	(c) Design of profusion stabilized,[43] thermostabilized,[44] conformationally locked,[45] or labeled[46] S protein
	(d) Deep mutational scanning[47] and analysis of S-protein variants for infectivity and antigenicity[48−51]
human ACE2 function and structure	(a) Identification of the ACE2 as the cellular receptor for the S protein[36,52,53] for membrane fusion
	(b) Determination of the tissue-level expression of the receptor and fusion-associated human proteins[54−56] and population level expression of proteins[10,11]
	(c) Structure elucidation of the interactions of the S protein with ACE2 (also chaperon for B^0AT1) in complex with[57] or without[58−60] the client protein B^0AT1
convalescent patient anti-S protein antibodies	(a) Identification of neutralizing antibodies (using single-cell RNA and VDJ sequencing of antigen-binding B cells[61]) from patients[61−64] and development of novel virus neutralizations tests (VNTs)[65,66]
	(b) Structure determination of these antibody−S-protein complexes[61,62,67−71]
high-affinity proteins targeting S-protein RBD	(a) Design of decoy ACE2 receptor[72] and *de novo* design of small proteins with picomolar affinity[73]
	(b) Design of nanobodies,[74] bivalent V_H domains,[75] antibody,[76] and IgA monoclonal antibody[77]
structure and function of other SARS-CoV-2 proteins	(a) Structure determination of RNA polymerase[78−80] and 2′-*O*-methyltransferase[81]
	(b) Structure determination of Nsp1[82,83] and M protease[84,85]
drug discovery	(a) Drug discovery through compound repurposing[86,87]
	(b) Small-molecule targeting of SARS-CoV-2 proteins[84,85,88,89]
proteomics	(a) Development of system-level virion proteomics[90] and phospho-proteomics survey of infected cells[91]
	(b) High-throughput experimental analysis of virus−host-protein−protein[92] and RNA−protein[93,94] interactions
animal models	(a) Creation of transgenic mice expressing human ACE2[95,96] or mouse-adapted redesigned RBD[97]
	(b) Use of golden hamster[98] and other model animals[99] to study transmission and pathogenesis
analyses of patients	(a) Clinical analyses of focused patients[100−103] for serological signatures[104]
	(b) Analysis of adaptive immune response,[105,106] host factors,[107] and innate immunity[3−8]
vaccines	(a) Lipid nanoparticle-encapsulated mRNA vaccine encoding RBD[108,109] or stabilized S protein[110]
	(b) Adenovirus vector-based vaccine expressing S protein[111,112] and recombinant S protein as vaccine[113]

pathology to combat the crisis through breakthroughs (Table 1) in a short time have been unparalleled. These results make possible an exploration of the possible relationship between the genetic background of an individual and the susceptibility to and the progression of the disease, which would be valuable to better understand COVID-19 pathogenesis. Genome-wide association study (GWAS) in the human population[13−16] has been the method of choice for identifying genetic risk factors associated with infectious diseases (e.g., mapping of HLA peptide binding cleft amino acid variants for disease susceptibility[14]). Although powerful, GWAS typically maps genetic loci (i.e., not genes necessarily), and probing the molecular mechanism may not be straightforward for a list of genetic loci. GWAS generally also fails to identify *rare* alleles. In general, GWAS of infectious disease susceptibility has been used for pathogens that have coexisted with human civilization since ancient times.[17] On an evolutionary time scale, the fixation of alleles providing selective advantage[18,19] by natural selection has influenced host susceptibility. COVID-19, in comparison, is a very recent threat, and the possibility of *rare* alleles providing selective advantage demands exploration. Consistent with the well-known fact that most human genes have several natural variants, the genes identified as interacting with SARS-CoV-2 also have several variants. Thus the question arises of whether the genotype of an individual could have a bearing on COVID-19 infection and prognosis. This question stems from previous observations of natural variants found to afford the host protection against specific pathogens. For example: (a) *HBB* missense variant against *Plasmodium falciparum*,[20] (b) *CCR5* deletion variant against HIV-1,[21,22] (c) *FUT2* variants against the Norwalk virus,[23] (d) SLC4A1 variants against *Plasmodium falciparum*,[24] and so on. CCR5Δ32 (i.e., CCR5 coreceptor deletion variant) impairs the T-cell entry of HIV-1, just as FUT2 and SLC4A1 variants impair the host-cell entry by the respective pathogens. In other words, certain natural variants of a host protein have the potential to influence the clinical outcome of a specific host−pathogen interaction if the host protein is essential to the pathogen life cycle. A recent *in silico* study on *ACE2* variants found in Iranian ethnic groups[25] suggests that ACE2 variants could possibly alter the receptor's interaction affinity, implying that subpopulations could have differences in intrinsic susceptibility to COVID-19.[25] Whereas the above example variants are known to provide protection against the pathogens, examples of variants responsible for severe infectious disease symptoms have also been reported[12] (e.g., herpes simplex virus 1 encephalitis,[26] viral respiratory infection[27] and influenza pneumonitis[28,29] in children, etc.). Many of the variants responsible for severe infectious disease symptoms are *rare*,[12,26−29] and the catalogues of *rare* variants, now available from next-generation sequencing studies of large populations (e.g., gnomAD[30]), therefore demand careful discussion in the context of disease susceptibility. In addition, recent pharmacogenomic studies[31,32] of receptor−drug interactions also emphasized the importance of *rare* variants, and like receptor−drug interactions, host-protein *rare* variants demand attention in the context of virus−host-protein−protein interactions as well. With several human proteins identified in the reports in Table 1 as being likely to play roles in the SARS-CoV-2 life cycle, the analysis of the variants of these proteins thus holds promise for relating the genetic backgrounds of patients to the clinical outcomes, complementing GWAS.[9] On the basis of the observations of some of the reports in Table 1, the discussion on variants is presented for host-protein machineries that are engaged in (a) viral entry, (b) innate immunity, and (c) virus−host-protein−protein and RNA−protein interactions to com-

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

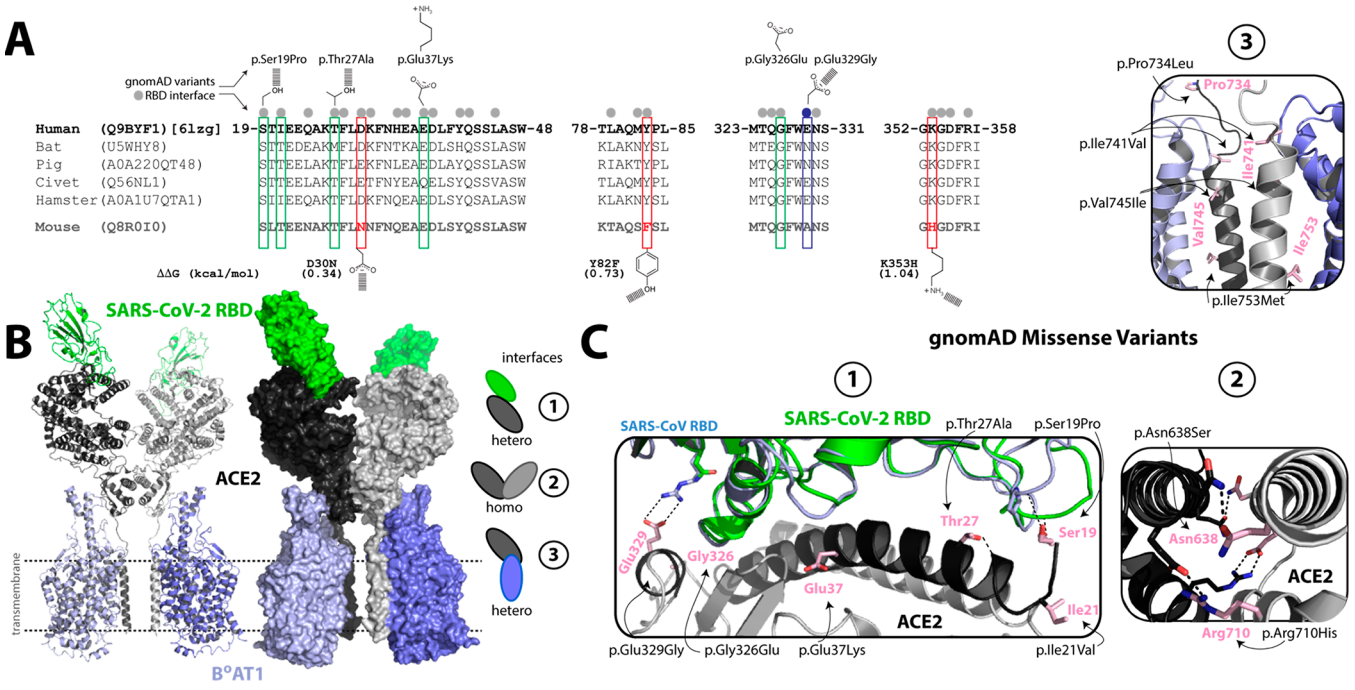


Figure 1. Missense variants at ACE2 interfaces: (A) Sequence comparison of the ACE2 peptidase domain (PD) for various mammalian species in the interfacial region with the SARS-CoV-2 S-protein receptor-binding domain (RBD). Human ACE2 PD positions in contact with the SARS-CoV-2 RBD (gray filled circle) showing that the mouse PD sequence differs at three hydrogen-bonding positions (red). Positions of gnomAD missense variants in contact with SARS-CoV-2 (green) and SARS-CoV (blue) are shown. gnomAD variants (e.g., pSer19Pro) are above the green/blue bars. (B) Structure (PDB 6M17) of the human ACE2 dimer (black and gray) in complex with the SARS-CoV-2 receptor-binding domain (RBD, green) and B^0AT1 protein (blue), with the membrane represented by dotted lines. The distinct interfaces (numbered 1−3, right) are for PD−RBD (heterodimer), ACE2−ACE2 (homodimer), and ACE2−B^0AT1 (heterodimer), respectively. (C) Positions (pink) of variants (e.g., p.Glu37Lys) at the three interfaces in panel A are shown.

plement other discussions[11,12,33,34] on genetic variants in relation to SARS-CoV-2 infections.

Human Protein Variants with Potential to Influence Viral Entry

Just as several of the above examples concern pathogen entry, variants related to SARS-CoV-2 entry are of great interest in understanding differences in pathogenesis. Human angiotensin-converting enzyme 2 (ACE2) is the cellular receptor of the SARS-CoV-2 Spike (S) protein,[36,52,53] and the ACE2 protein serves as the critical site for viral attachment. ACE2, a membrane-bound, counter-regulatory carboxypeptidase (responsible for the proteolysis of angiotensin I/II, neurotensin, kinetensin, and apelins),[114,115] is an essential component of the renin−angiotensin hormone system, playing a critical role in cardiovascular homeostasis.[116] Understandably, the probability of a loss-of-function-intolerant (pLI[117]) score for *ACE2* is 0.99 (Figure 2A), indicating that *ACE2* truncation variants are less likely to be tolerated and that the virus utilizes an essential human protein for entry. In contrast, the pLI scores for *CCR5* and *FUT2* are both 0.0, suggesting higher tolerance for deletion/truncation, which is consistent with the presence of deletion/truncation variants of CCR5 and FUT2, even as homozygotes, in the human population being known to provide protective advantages against the respective pathogens. For ACE2, we therefore predominantly focus on its missense variants, in particular, those distinct from the eight variants discussed in the Iranian population study.[25] *ACE2* is located on the X chromosome, and thus men are hemizygous for *ACE2* variants. The hemizygous nature of ACE2 variants demands attention in the context of SARS-CoV-2 interactions. *ACE2* is also included among genes that escape X chromosome inactivation (XCI)[118] (i.e., expressed from both the active and inactive X chromosomes in females), and the contributions of ACE2 variants toward SARS-CoV-2 interactions in females are therefore less likely to depend on XCI. Figure 2A summarizes the gnomAD variants of ACE2 (along with those of other proteins likely engaged in viral entry). The variants of the proteins are discussed in this section.

ACE2 Variants. Topologically, a single transmembrane helix of ACE2 separates the extracellular peptidase domain (PD) from the cytoplasmic domain.[57] The PD is generally referred to as the receptor for coronaviruses (CoVs), and the C-terminal domain of S protein, referred to as the receptor-binding domain (RBD), physically interacts with the PD.[57−60] Recent structures of the PD−RBD complex[57−60] highlight the similarities and differences in receptor recognition between SARS-CoV-2 and other CoVs.[58,59] To appreciate the possible consequences of human PD sequence variants, it is important to look at how changes in mammalian ACE2 sequence influence SARS-CoV-2 RBD recognition.[119,120] Genetic variants are discussed in the same context as the PD sequence.[119,120] SARS-CoV-2 infects humans, bats, pigs, civets,[52] and golden hamsters[98] but not mice.[52] This observation was the key reason for creating a transgenic mouse (a cost-effective animal model[95]) expressing human ACE2[95,96] to study COVID-19 pathogenesis. A comparison of the mouse PD sequence with those of SARS-CoV-2 hosts (Figure 1A) sheds light on how the infection may be mediated by PD−RBD complexation. Although sequences of human, bat, pig, hamster, and civet PDs do differ from one another, they still interact with the SARS-CoV-2 RBD, whereas subtle changes in the mouse PD sequence likely compromise the

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

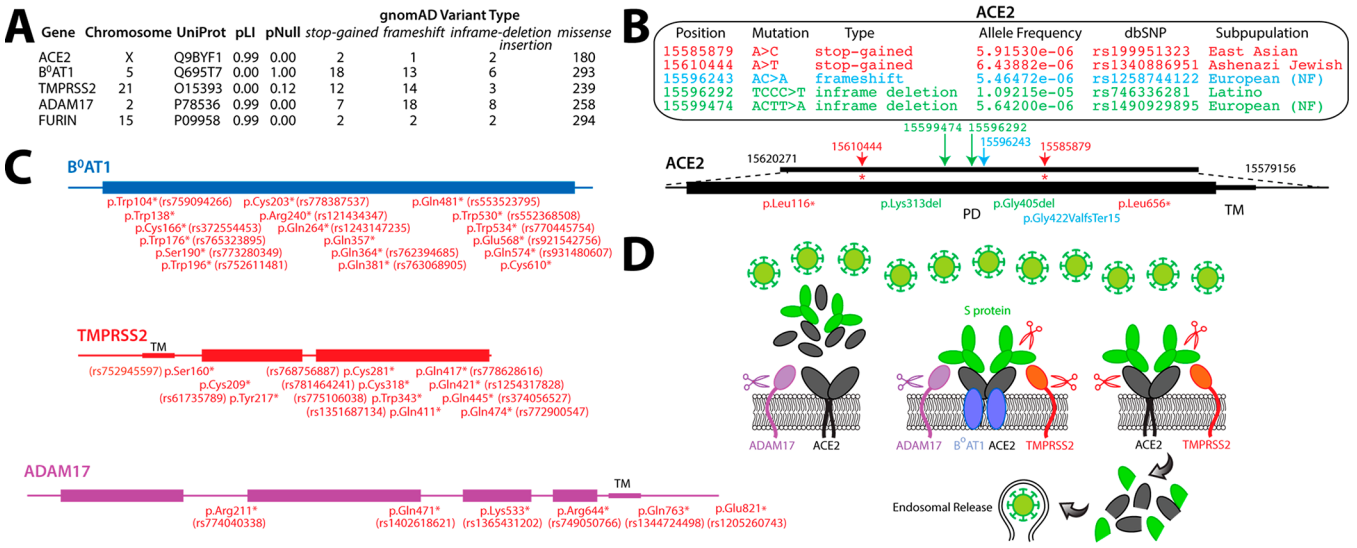


Figure 2. Variants of protein machinery engaged in viral entry: (A) List of genes, their chromosomal location, the UniProt accession number of the corresponding protein, pLI and pNull probabilities, and the number of different gnomAD variants observed for each gene. Because the sum (pLI + pNull + pRec = 1) is one, only pLI and pNull probabilities are provided. (B) *ACE2* nonmissense variants, stop-gained (red), frameshift (blue), and inframe deletion/insertion, are listed (top) and mapped onto the protein (bottom) showing the peptidase domain (PD) and the transmembrane (TM) regions. The listed variants (top) includes nucleotide position and substitution, mutation type, allele frequency, dbSNP IDs, and the subpopulation in which the variant is identified (NF, non-Finnish). (C) Stop-gained variants (red) of other proteins, B^0AT1 (blue), TMPRSS2 (red), ADAM17 (purple). The dbSNP IDs are in parentheses. (D) Cartoon representing how B^0AT1 could compete/interfere with TMPRSS2 and ADAM17 proteolysis of ACE2. The cartoon was inspired from the study of Brest et al.,[11] and for convenience, the proteins are colored in the same way as in panels B and C.

interaction (Figure 1A). On the basis of the structure of the human ACE2 PD−RBD complex,[57−60] the mouse PD has conservative (with respect to the human) substitutions (D30N, Y83F, and K353H) at three hydrogen-bonding interfacial sites (Figure 1), which together are likely sufficient to negatively perturb the interaction. This observation suggests that subtle changes in hydrogen bonding and charged residues at interfacial sites could impact the receptor−S-protein interaction. This is also consistent with adapting mouse ACE2 for the SARS-CoV-2 RBD interaction by a few point mutations.[97] The predicted loss of binding free energy ($\Delta\Delta G$) for the three mouse PD substitutions is 0.34, 0.73, and 1.04 kcal/mol (using BeAtMuSiC[121,122]). Given the behavior of PD interfacial residues, Genome Aggregation Database (gnomAD,[30] an unbiased large population database) missense variants that map to the PD−RBD interface are of great interest. Six human ACE2 missense variants, including p.Ser19Pro (PolyPhen score,[123] 0.767), p.Ile20Val (0.000), p.Thr27Ala (0.000), p.Glu37Lys (0.712), p.Gly326Glu (0.090), and p.Glu329Gly (0.027), map to the interface (Figure 2B). The low PolyPhen scores (<0.9) indicate that these variants are unlikely to be deleterious (i.e., unlikely to perturb the essential peptidase function, being distal from the catalytic site). Ser19 is engaged in hydrogen bonding with the SARS-CoV-2 RBD backbone atoms, whereas the hydrogen bonding of Glu329 is prominent in the SARS-CoV RBD (Figure 2B). In addition, p.Glu37Lys and p.Glu329Gly result in interfacial charge alteration (Figure 2B). Ser19, Thr27, and Glu37 also engage in intramolecular hydrogen bonds, and variants at these positions could perturb PD−RBD interactions. The predicted $\Delta\Delta G$ for these variants (0.04, 0.18, 0.92, 0.34, 0.62, and 0.33 kcal/mol) is also within the range of the mouse PD substitutions. Given the known role of hydrogen bonds and charged residues in the mouse PD, these variants could contribute to altering the PD−RBD affinity.

Position-specific amino acid enrichment/depletion in the ACE2 sequence observed by Chan et al.[72] in the designed decoy ACE2 receptor by deep mutagenesis[72] is also an important reference for predicting the expected behavior of ACE2 missense variants. In the deep mutagenesis study, the substitutions corresponding to three gnomAD missense variants (p.Ser19Pro, p.Thr27Ala, and p.Thr92Ile) were observed to be highly enriched[72] (i.e., Pro for Ser19, Ile for Thr92), suggesting that these variants are anticipated to enhance PD−RBD affinity. Asn90 and Thr92 constitute the N-glycosylation motif of the ACE2 protein, and substitutions at both positions were highly enriched, suggesting that N-glycans at these positions likely interfere with RBD binding,[72] whereas the proline substitution at Ser19 (the first turn of the interfacial helix) likely entropically stabilizes the helix.[72] Substitutions associated with a number of missense variants showing medium-range depletion (p.Glu35Lys, p.Glu37Lys, p.Asn51Ser, p.Lys68Glu, p.Phe72Val, p.Gly326Glu, p.Glu329Gly, p.Gly352Val, and p.Gln388Leu) or enrichment (e.g., p.Lys26Arg, p.Asn64Lys, p.Gln102Pro and p.His378Arg) were also observed,[72] suggesting that these variants could contribute to weakening or enhancing the PD−RBD affinity, respectively. The dose of a viral inoculum[124] is an important factor for infection or severity of infection; a weakened/enhanced affinity could influence the dose required for infection in individuals carrying variants, and these observations encourage the investigation of the consequences of host variants in SARS-CoV-2 pathogenesis.

These variants, however, are *rare* (gnomAD[30] allele frequencies, 10^{-3} to 10^{-5}) and are unlikely to occur in ACE2 at the same time. The occurrence frequency of these variants also varies between human subpopulations. For example, in gnomAD,[30] the p.Thr27Ala variant is seen only in the Latino population (allele frequency, 7.3×10^{-5}), p.Glu37Lys is observed in the Finnish European (3.2×10^{-4}) and African

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

(1.0×10^{-4}) populations, and p.Ser19Pro is observed only in the African (3.3×10^{-3}) population. With a population of ~1.3 billion in Africa (and assuming 50% men), the p.Ser19Pro variant, with a frequency of 3.3×10^{-3}, is likely to be present in a significant number of people. In women, these variants are likely to occur predominantly as heterozygotes (ACE2 escapes XCI). In heterozygotes, oligomeric protein complexation can be affected (e.g., the sickle-cell trait of heterozygotes[20] or oligomeric haptoglobin of heterozygotes and homozygotes[125]). The recent structural observation of ACE2 chaperoning B^0AT1 (also known as SLC6A19) suggests that ACE2 likely dimerizes within the membrane. The positions of the PD−RBD interface variants are distal from the homodimer interface, suggesting that the variants are also less likely to affect homodimerization. The ACE dimers in heterozygous women could also be affected by PD−RBD interface variants. The ACE2−B^0AT1 interface is also important in this respect, especially in the kidney and the intestine, where B^0AT1 is predominantly expressed. Proteolytic cleavage of ACE2 by TMPRSS2 for the endosomal release of SARS-CoV-2 is an important step, and B^0AT1 could possibly compete with TMPRSS2 for ACE2 access (Figure 2D). ACE2−B^0AT1 interfacial variants could therefore also indirectly influence pathogenesis. The ACE2 transmembrane helix is the primary site of B^0AT1 interaction[57] (Figure 1B), and four ACE2 missense variants (p.Pro734Leu, pIle741Val, p.Val745Ile, and p.Ile753Met) can be mapped onto the ACE2−B^0AT1 interface (Figure 1C), with predicted $\Delta\Delta G$ values in a similar range as above. Among the four variants, only p.Pro734Leu has a PolyPhen score of 1.0 (i.e., likely deleterious), and this variant is seen in the South Asian population with an allele frequency 1.1×10^{-4} (also observed in 1 out of ~5000 South Asian men in gnomAD). That is, despite being predicted to be deleterious, men possessing the variant are known to have survived. It is, however, unclear whether these variants could interfere with B^0AT1 interactions to influence TMPRSS2 access. Missense variants (p.Asn638Ser and p.Arg710His) at the ACE2 homodimer interface (Figure 1C) could also influence viral entry (i.e., homodimer interface disruption could enhance TMPRSS2 accessibility). Asn638 and Arg710 are engaged in ACE2 intersubunit hydrogen bonding and salt bridge formation, respectively. In particular, p.Arg710His has a PolyPhen score of 1.0 (i.e., deleterious). p.Arg710His is observed in non-Finnish Europeans (also observed in 1 out of ~40 000 Europeans) and the East Asian population, with allele frequencies of 6.1×10^{-5} and 1.5×10^{-4}, respectively. The predicted $\Delta\Delta G$ of these two naturally occurring mutants is 1.95 and 0.24 kcal/mol, respectively, and they would be useful in revealing mechanistic details of the endosomal release of the trimeric S protein docked onto the dimeric ACE2 protein, similar in flavor to a previous *in vitro* mechanistic study[126] of HIV-1 entry disruption by CCR5 nonsense variant C100X (a rare allele, 6.6×10^{-4}). Similarly, one of the ACE2 disulfide (Cys133−Cys141) disrupting variants (p.Cys141Tyr, 5.8×10^{-6}) would be useful. In addition to the ACE2 missense variants discussed here in the context of ACE2 structure and complex, a discussion of ACE2 nonmissense variants would also be very important. gnomAD reports ACE2 nonmissense variants (Figure 2A, B), although only a few due to high pLI score. There, respectively, are two, one, and two ACE2 rare nonsense (stop-gained, 5.9×10^{-6} and 6.4×10^{-6}), frameshift (5.4×10^{-6}), and inframe deletion (1.1×10^{-5} and 5.6×10^{-6}) variants reported in gnomAD (Figure 2B). All nonmissense variants map to PD, suggesting that PD truncation and structural perturbations from the nonmissense variants are likely to negatively impact PD−RBD interactions and viral entry in carrier individuals.

Auxiliary Protein Variants. In addition to ACE2 (which directly interacts with SARS-CoV-2 proteins), other human proteins could also influence ACE2−S-protein interactions (Figure 2A,D). For example, interactions of B^0AT1 with ACE2 could possibly compete with TMPRSS2 and ADAM17 to influence ACE2 proteolysis (Figure 2D). Proteins (e.g., B^0AT1) with the potential to exert indirect influence on host−viral protein interactions (e.g., ACE2−S-protein) are referred to as *auxiliary* proteins here for convenience. The list of *auxiliary* proteins (e.g., B^0AT1, TMPRSS2, ADAM17, etc.) that likely influence viral entry is provided in Figure 2A, highlighting different categories of variants associated with these *auxiliary* proteins. In this regard, B^0AT1 missense variants at the ACE2−B^0AT1 interface would be important in influencing the proteolytic outcome of ACE2. Six out of the 14 B^0AT1 missense variants at the ACE2−B^0AT1 interface have 30 $Å^2$ or more of their surface area in contact with ACE2, with the predicted $\Delta\Delta G$ of four of these being quite high (1.20 to 2.02 kcal/mol). Of these four, p.Arg214Gly is engaged in intermolecular hydrogen bonding and is carried by non-Finnish Europeans (allele frequency, 6.1×10^{-5}) and African (3.5×10^{-5}) populations. Unlike ACE2, pLI for B^0AT1 is 0.0 (and pNull = 1.0, see later), and several B^0AT1 stop-gained and other nonmissense variants are observed in gnomAD (Figure 2A, C). Truncations/structural perturbations of B^0AT1 due to nonmissense variants are likely to favor ACE2 proteolysis.

Analyzing expression quantitative trait loci (eQTL) polymorphic variants for examining ACE2 expression in the genotype-tissue expression (GTEx) database, Chen et al.[10] in a recent report showed that at the population level, ACE2 expression is negatively correlated with COVID-19 severity.[10] To substantiate the counterintuitive observation, follow-up *in silico* expression analyses by Brest et al.[11] of variants of not only ACE2 but also the transmembrane proteases TMPRSS2 and ADAM17 suggest that protease expressions are likely to also contribute to disease severity.[11] The logical support for the argument[11] is that ACE2 is shed from membranes upon proteolysis[127] (Figure 2D). Proteolysis by ADAM17 results in ACE2 shedding into extracellular space, as ADAM17 operates on ACE2 and not the S protein.[11,127] TMPRSS2, on the contrary, functions on ACE2 as well as the S protein, leading to endosomal release (i.e., virus uptake) upon membrane fusion.[11] The authors suggest that for low viral loads, ACE2 shedding could serve as a barrier for infection, whereas for high viral loads, proteolysis would lead to infection[11] (Figure 2D). In short, these proteases and their variants are likely to play an important role in the clinical outcome. The variants probed by human population level GTEx expression analyses are genome variants (i.e., not exome variants necessarily) typically occurring within ±10 kb of the gene of interest with minor allele frequencies between 0.05 and 0.5. The focus here is on *rare* (allele frequencies, 10^{-3} to 10^{-5}) exome variants, as these variants have the potential (see later) to influence clinical outcomes of an individual (with the prospect of understanding mechanisms from protein analysis), complementing the population level study. The list of gnomAD *rare* nonmissense variants for TMPRSS2 and ADAM17 shows several variants (Figure 2A,C), especially for TMPRSS2 (pLI = 0.0), that could influence viral entry.

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

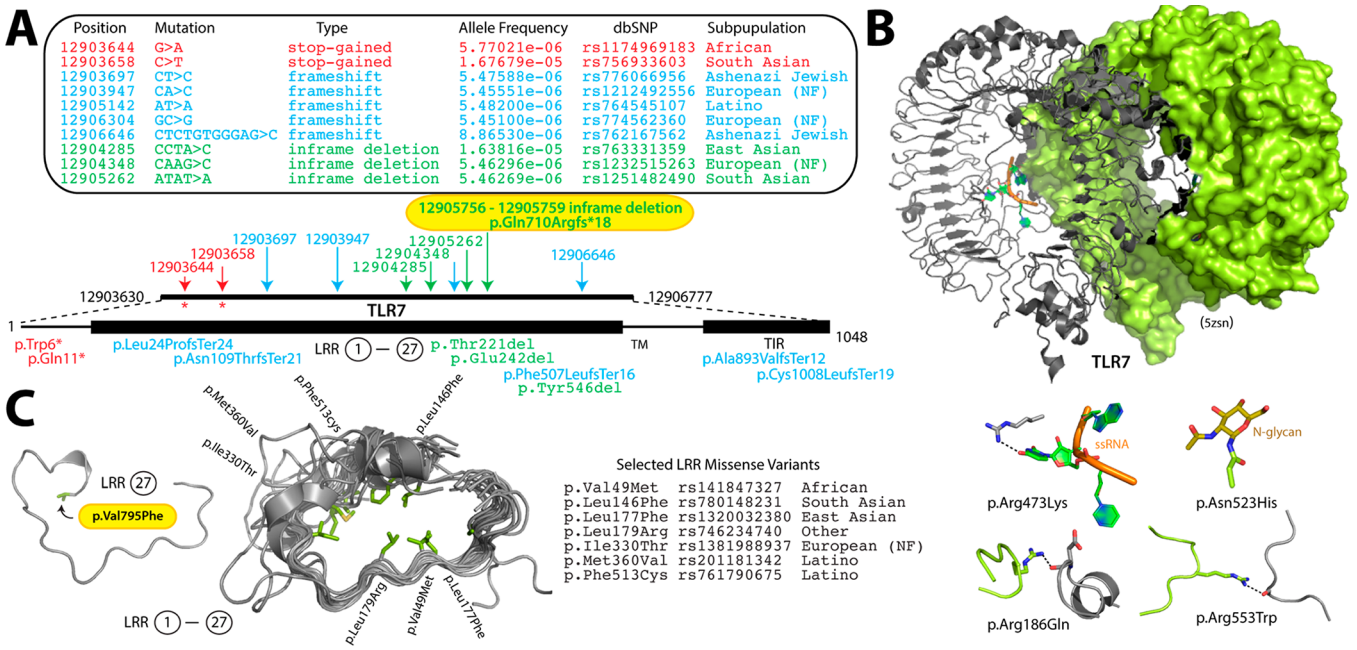


Figure 3. TLR7 variants: (A) *TLR7* nonmissense variants (as described in Figure 2B) are listed (top) and mapped onto the protein (bottom) showing the leucine rich region (LRR), the transmembrane (TM) and Toll/IL-1 receptor (TIR) regions. The mutants shown with a yellow background are reported in the van der Made et al.[3] study of brothers with severe symptoms. (B) *Rhesus macaque* TLR7 dimer structure with bound ssRNA (top) showing human TLR7 missense variants at interfacial positions and glycosylation site (bottom). (C) Select missense variants within TLR7 LRR (right) are mapped onto the 22−29 residue LRR structural motifs (middle) with the p.Val795Phe variant shown within the structural motif (left). TLR7 LRR structural motifs (1−27) were superposed using MUSTANG[128] structural alignment.

Human Protein Variants with Potential to Influence Immune Response

Post-viral-entry protein machinery engaged in the host immune response is discussed next, as a number of recent studies[3−8] point to the contributions of innate immunity toward disease severity. For example, an important recent study from the Casanova laboratory[5] reports that ∼10% of patients with life-threatening COVID-19 pneumonia have neutralizing autoantibodies (auto-Abs) against interferons,[5] whereas the anti-interferon auto-Abs are absent in asymptomatic patients.[5] Interferons (cytokine subgroup) are key signaling proteins of innate immunity, especially in viral infections. It had previously been shown that phenocopies of auto-Abs against cytokines are similar to clinical phenotypes of germline variants of cytokines and cytokine-receptors.[5] In other words, the absence of functional innate immunity signaling proteins due to either germline mutations or auto-Abs against them could compromise the innate immune response, leading to a severe infectious. Here variants of proteins associated with innate immunity signaling pathway for the production of interferons and other cytokines (e.g., interleukins) are discussed in regard to COVID-19 disease severity.

Toll-like Receptor 7 (TLR7) Variants. van der Made et al.[3] reported an important case study highlighting the role of monogenic *rare* genetic variants in severe COVID-19 in four young men (i.e., two pairs of brothers) from unrelated families with no prior history of major chronic diseases.[3] Rapid whole genome sequencing of the four (including one deceased) paired brother patients showed putative loss-of-function mutations in *TLR7*.[3] TLR7 functions as a pattern recognition receptor (e.g., recognizing ssRNA) and plays an important role in the viral (e.g., RNA viruses) immune response. TLR7 is encoded from the X chromosome. The two *TLR7* variants found in the brothers are frameshift p.Gln710Argfs*18 (in one family) and missense p.Val795Phe (in the other) variants that map to the leucine-rich repeat (LRR) structural region of TLR7 (Figure 3A,B). These variants are not observed in gnomAD, suggesting that the two variants are *rare* (i.e., occur with very low allele frequency to be observed in ∼140 000 individuals in gnomAD). The functional assay of primary immune cells of the patients (i.e., carrying the TLR7 variants) by probing with TLR7 agonist showed the defective regulation of type-I IFN-related genes in comparison with normal cells.[3] This suggested that the functional (and likely structural) consequences of the *rare* variants compromise TLR7 (and downstream innate immunity signaling protein) functions, further underscoring the importance of the study of *rare* variants. The p.Gln710Argfs*18 frameshift variant is expected to alter the spacing of nonpolar residue patterns at distinctive intervals within the 22−29 residue repeating LRR motif, and p.Gln710Argfs*18 is thus expected to perturb the LRR fold. Given the functional consequence of the LRR frameshift (p.Gln710Argfs*18) variant in the above patient brothers, the identification of TLR7 nonmissense variants is of great importance. pLI for TLR7 is 0.98; therefore, only a few nonmissense (two stop-gained, three frameshift, and three inframe deletion) TLR7 variants are observed in gnomAD (Figure 3A), and all eight TLR7 nonmissense variants map to LRRs (Figure 3A). These nonmissense variants are likely to similarly compromise TLR7 function, and men carrying these nonmissense variants could be at risk for severe COVID-19.

The structure of human TLR7 is unavailable, but estimates of structural perturbations due to missense variants of human TLR7 can be made by utilizing the structure of *Rhesus macaque* TLR7[129] (98% sequence identity). On the basis of the *Rhesus macaque* TLR7 structure, gnomAD missense variants that are expected to be similar to p.Val795Phe (in the patient brothers) in their behavior are also important for discussion. The

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

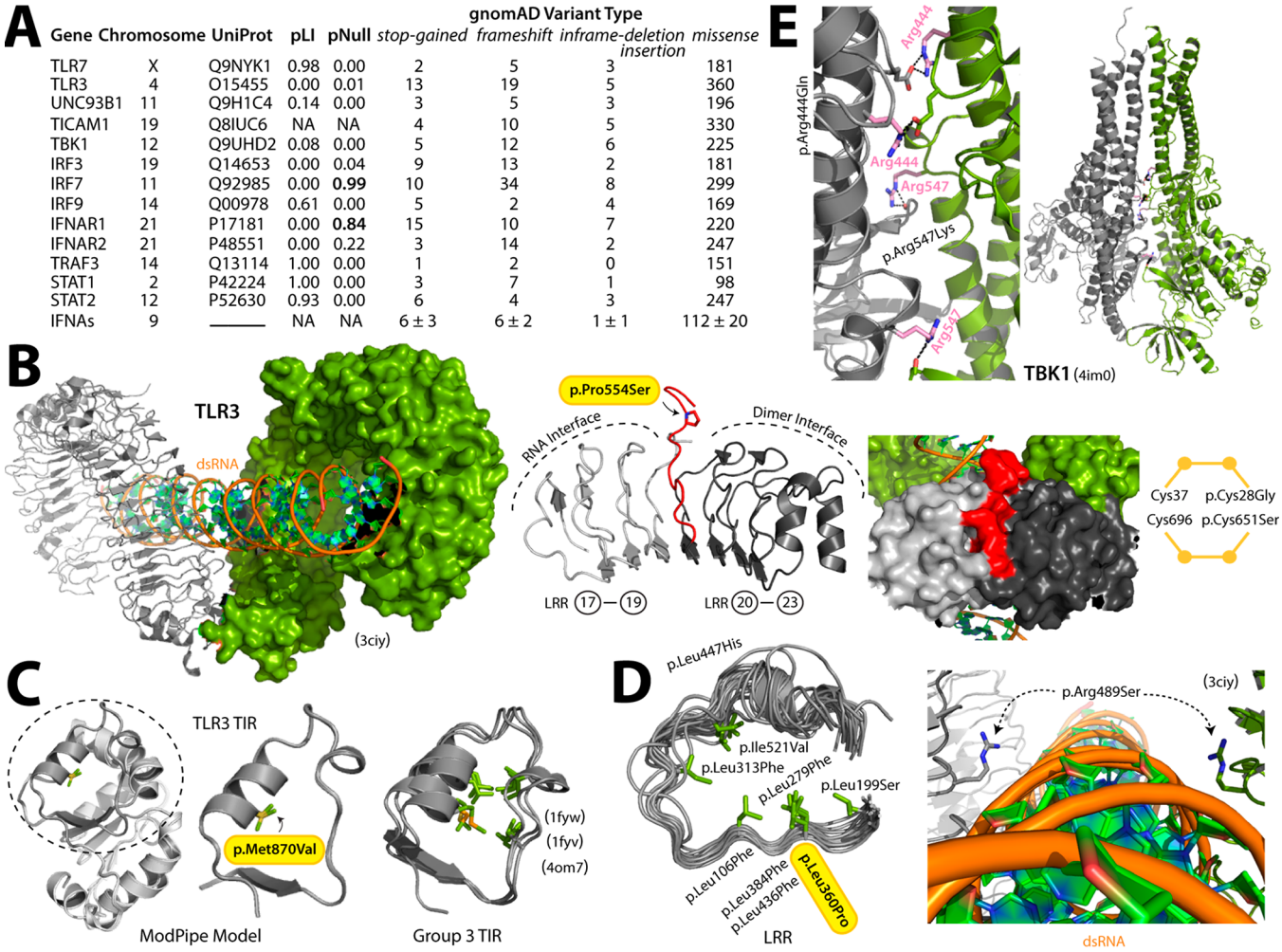


Figure 4. Variants of innate immunity proteins: (A) List of genes with their chromosomal location, UniProt accession number of the corresponding protein, pLI and pNull probabilities, and the number of different gnomAD variants. For the interferon genes, only the average values of the 14 genes are presented. (B) Mouse TLR3 dimer structure with bound dsRNA (left), showing p.Pro554Ser located between two sets of interfacial LRRs (middle) and the pair of disulfide disrupting variants (right). (C) TLR3 TIR domain modeled structure (from ModPipe[132]) showing p.Met870Val (middle) located within the nonpolar residues cluster specific for Group 3 TIR[133] (right). (D) Select missense variants within TLR3 LRR structural motifs (left) and the TLR3−dsRNA interfacial variant (right). (E) TBK1 homodimer structure showing interfacial missense variants (left). The mutants shown with a yellow background in panels B, C, and D are autosomal dominant variants reported in some patients with severe symptoms[4] or in herpes simplex virus-1 encephalitis (HSE).[28]

substituted valine in p.Val795Phe is at an *inward* facing position within the LRR structural motif (Figure 3C), and the estimated energetic contribution of Val795 (FoldX[130] alascan) is ∼2 kcal/mol. Several missense variants that substitute nonpolar residues (i.e., Leu/Ile) at *inward* facing positions within the LRRs (similar to that of Val795 and with estimated energetic contributions of 2−4 kcal/mol) are also present in gnomAD (Figure 3C), suggesting that the observed *rare* variants would be important for determining the risk of COVID-19 severity. The structure of *Rhesus macaque* TLR7[129] in complex with ssRNA in the functionally competent homodimer state (Figure 3B) also shows the presence of interfacial missense variants such as p.Arg473Lys (TLR7−ssRNA interface), p.Arg186Gln, and p.Arg553Trp (TLR7−TLR7 interface), along with the variant p.Asn523His at a glycosylation site. All of these variants, through structural or functional perturbation of TLR7, have the potential to contribute to the malfunctioning of TLR7-mediated signaling. For X-linked *TLR7*, men are expected to be affected more by TLR7 variants. The impact of TLR7 variants on females is less clear, as some studies report *TLR7* to *escape* XCI,[131] whereas other studies have not included *TLR7* among the list of genes *escaping* XCI.[118] However, the autosomal dominant behavior of TLR3 variants (see later) suggests that females could also be affected by TLR7 variants, likely due to the dimerization during RNA recognition.

Variants of Other Innate Immunity Proteins. Another important recent study[4] from the Casanova laboratory shows that *rare* loss-of-function variants in proteins associated with TLR3 and the interferon regulatory factor 7 (IRF7)-dependent type-I interferon (IFN) signaling pathway are enriched in patients with life-threatening COVID-19 pneumonia relative to asymptomatic infections.[4] In the study, 13 genetic loci whose variants had been known to be associated with immunity to the influenza virus were probed experimentally to identify autosomal-dominant/-recessive loss-of-function variants in these loci in ∼4% of patients with life-threatening COVID-19 pneumonia.[4] Information on the loss-of-function variants of these 13 genes present in the human population would therefore be valuable for predicting susceptibility to the severe disease. In addition, because auto-Abs against interferons (i.e., resulting in

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

A

	Gene	Chromosome	UniProt	pLI	pNull	gnomAD Variant Type: *stop-gained*	*frameshift*	*inframe-deletion insertion*	*missense*
RPI	MCM7	7	P33993	0.00	1.00	23	35	9	394
	CHD1L	1	Q86WJ1	0.00	1.00	22	36	15	397
	APRT	16	P07741	0.00	1.00	7	16	5	123
	ATIC	2	P31939	0.00	0.99	15	21	4	306
	ANXA5	4	P08758	0.00	0.99	4	11	2	121
PPI	CNTRL	9	Q7Z7A1	0.00	1.00	59	79	21	906
	NINL	20	Q9Y2I6	0.00	1.00	38	43	11	714
	ANO6	12	Q4KMQ2	0.00	1.00	28	25	7	379
	UGGT2	13	Q9NYU1	0.00	1.00	31	55	12	625
	HS6ST2	X	Q96MM7	0.00	0.05	5	4	3	166
	PDZD11	X	Q5EBL8	0.20	0.04	2	0	0	17
MHCI-assembly	TAPBP	6	O15533	0.00	0.56	15	16	1	208
	PDIA3	15	P30101	1.00	0.00	2	6	2	201
	CALR	19	P27797	0.89	0.00	1	5	14	187
	TAP1	6	Q03518	0.00	0.00	10	10	5	262
	TAP2	6	Q03519	0.00	0.01	16	11	3	239
	B2M	15	P61769	0.56	0.01	1	0	0	35

B

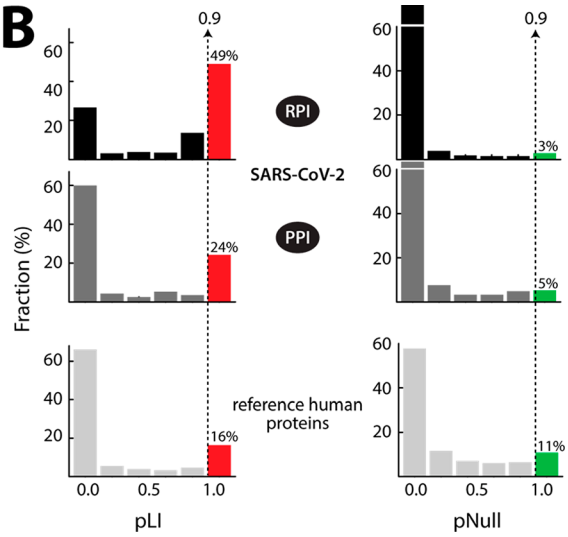


Figure 5. SARS-CoV-2 RPI and PPI protein: (A) List of genes (as described in Figures 2A and 4A) of RPI (top), PPI (middle), and MHCI-assembly proteins (bottom). For convenience, only a select few among the 13 RPI and 24 PPI with high pNull are shown. (B) pLI (left) and pNull (right) probability distributions of SARS-CoV-2 RPI (top) and PPI (middle) proteins compared with reference human proteins (bottom). The dotted line (pLI or pNull = 0.9) represents the cutoff. For the RPI proteins, only those showing at least two-fold enrichment in both postinfection time points (24 and 48 h)[94] were considered.

the depletion of interferons) also results in severe COVID-19 disease,[5] information about loss-of-function variants of interferon genes present in the human population will also be valuable. The loss-of-function variants of the 13 and 14 interferon (*IFNA1*−*8*, *IFNA10*, *IFNA14*, *IFNA16*, *IFNA17*, *IFNA21*, *IFNB1*, and *IFNW1*) genes are listed in Figure 4A. It is important to note that the pLI score of many of these genes (e.g., *TLR3*, *IRF7*, *IFNAR1*, etc.) is ∼0.0, suggesting that loss-of-function variants of these genes are tolerated, especially *TLR3*, *IRF7*, and *IFNAR1*, which have 15, 10, and 17 truncation variants, respectively, in gnomAD along with having several other nonmissense variants (Figure 4A). It is interesting to note that for some of the loss-of-function variants (e.g., in *TLR3*, *TBK1*, *IRF3*, etc.) identified in the study[4] (or those known from previous studies), a single copy of the variant (i.e., heterozygous) could result in severe disease.[4] That is, with respect to disease severity, some of the variant alleles are dominant. In other words, a heterozygous individual can otherwise be healthy (i.e., in absence of the infection, unlikely to be affected by the loss-of-function) but could develop severe disease when challenged by the infection. For the autosomal dominant nature (i.e., with respect to disease severity) of several variants of the genes identified in the study,[4] other variants of these genes observed in gnomAD also have the potential to result in malfunction of the corresponding proteins. For example, like the TLR3 dominant p.Ser339 fs (frameshift) and p.Trp769* (stop-gained) loss-of-function variants observed in patients with severe pneumonia,[4] the list of other TLR3 stop-gained (e.g., p.Trp11*, p.Ser88*, p.Arg394*, etc.), frameshift (e.g., p.Leu243MetfsTer16, p.Leu255GlufsTer2, p.Lys345ArgfsTer10, etc.), and deletion/insertion (e.g., p.Asn285del, p.Asp348del, p.Leu482del, etc.) variants, IRF7 stop-gained (e.g., p.Trp30*, p.Trp214*, p.Trp391*, etc.) and frameshift (e.g., p.Asp207AlafsTer206, p.Leu185SerfsTer26, p.Glu95ArgfsTer3, etc.) variants, and others observed in gnomAD are of great interest for understanding the potential for risks of severe symptoms. In gnomAD, *rare* variants are typically observed in heterozygotes, and the *rare* variants of innate immunity proteins have the potential to function as autosomal dominant with respect to disease severity.

Whereas there are several variants of proteins (Figure 4A) with the potential to interfere with the type I interferon (IFN) pathway (and therefore demand attention), only a few variants (e.g., missense) of TLR3 and TBK1 (Figure 4B−E) are discussed here in the context of protein structure for convenience. The autosomal dominant TLR3 missense variants p.Pro554Ser and p.Met870Val probed in the study[4] above are, respectively, within the LRR and TIR (Toll/IL-1 receptor) regions of the molecule. In the mouse TLR3 LRR structure (∼80% sequence identity with human), p.Pro554Ser is placed in the loop connecting two sets of LRRs engaged in the dsRNA and the dimer interface, respectively (Figure 4B). With an estimated energetic contribution of ∼2 kcal/mol (FoldX[130]), Pro55 likely plays an important role in the assembly for the TLR3 signaling complex. The known p.Leu360Pro dominant loss-of-function TLR3 missense variant, impacting herpes simplex virus-1 encephalitis (HSE),[28] resides in an *inward* facing position within the LRR motif. Similar to p.Leu360Pro, several missense variants that substitute nonpolar residues (i.e., Leu/Ile) at *inward* facing positions within the LRRs (with estimated energetic contributions of 2−4 kcal/mol) are also observed in gnomAD (Figure 4D, left). In addition, TLR3−dsRNA interface variant (p.Arg489Ser) (Figure 4D, right) and TLR3 disulfide disrupting variants (p.Cys28Gly and p.Cys651Ser) are also observed (Figure 4B, right). TLR3 TIR belongs to Group 3 TIR.[133] Group 3 TIR possesses a distinct cluster of bulky nonpolar residues (Figure 4C), and the modeled human TLR3 TIR structure suggests that p.Met870Val is likely to interfere with the nonpolar residue cluster (Figure 4C). The TBK1 homodimer structure indicates that gnomAD missense variants p.Arg444Gln and p.Arg547Lys in each protomer are likely to perturb several interfacial hydrogen-bonded interactions and that these variants have the potential to perturb TBK1 function (Figure 4E).

In addition to variants of the innate immunity pathway proteins, the genetic variability across the human leukocyte

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

antigen (HLA) A, B, and C genes could also contribute to SARS-CoV-2 susceptibility and disease severity. In this regard, *in silico*[134−136] assessment of the binding of SARS-CoV-2 peptides onto major histocompatibility complex (MHC) class I binding pocket across HLA-A, -B, and -C genotypes is also useful for the analysis of susceptibility.[134−136] Because the presentation of high-affinity peptide by MHC class I requires the peptide-loading complex[137,138] protein machinery, information about the variants of genes (TAPBP, PDIA3, CALR, TAP1, and TAP2, referred to here as MHCI-assembly genes) of the peptide-loading complex will also be useful in this regard. The variants of MHCI-assembly genes are therefore listed in Figure 5A (bottom), and it is interesting to note that pLI probabilities for TAP1, TAP2, and TAPBP are 0.0.

Human Protein Variants with Potential to Influence Virus−Host-Protein−Protein and RNA−Protein Interactions

The focus of the discussion above on variants was on a small set of genes that are engaged in a specific pathway (e.g., viral entry and innate immunity). Variants of proteins of other pathways could also contribute to susceptibility. In general, for a system-level view of variants in the context of host−virus interactions, proteins identified through proteomic-scale approaches provide a useful reference. In this regard, a discussion on variants associated with human proteins recently identified by Gordon et al.[92] and Flynn et al.[94] that interact with SARS-CoV-2 proteins and RNA, respectively, would be valuable. Gordon et al.[92] and Flynn et al.,[94] respectively, report 330 and 309 human proteins engaged in high-confidence SARS-CoV-2−human protein−protein interactions[92] (referred to as SARS-CoV-2 PPI proteins) and RNA−protein interactions[94] (referred to as SARS-CoV-2 RPI proteins). For convenience, the discussion on variants of SARS-CoV-2 PPI and RPI proteins is in the context of tolerance to loss-of-function variants. Genes can be categorized as (a) tolerant to loss-of-function variants as homozygous (pNull probability high), (b) tolerant to loss-of-function variants as heterozygous (pRec probability high), and (c) intolerant to loss-of-function variants as heterozygous (pLI probability high).[117] As discussed above, *TLR3*, *IRF7*, and *IFNAR1* all have low pLI (i.e., 0.0, Figure 4A). *IRF7* and *IFNAR1* have high pNull (0.99 and 0.84, respectively), and *TLR3* additionally has small pNull (i.e., 0.01, Figure 4A). That is, *IRF7* and *IFNAR1* are expected to be tolerant to loss-of-function variants as homozygous, whereas *TLR3* is expected to be tolerant to loss-of-function as heterozygous. It is interesting to note that autosomal recessive variants (i.e., homozygous) for disease severity were observed for *IRF7* (e.g., p.Pro364 fs/p.Pro364 fs)[4] and *IFNAR1* (e.g., p.Trp73Cys/Trp73Cys),[4] which is consistent with their high pNull probability for an individual to carry both loss-of-function alleles. Autosomal dominant (i.e., heterozygous) variants were observed for *TLR3* (e.g., p.Trp769*/wildtype),[4] which is consistent with the low pNull probability for *TLR3*. In fact, autosomal dominant variants identified in the above study[4] were predominantly in genes with low pNull probability (e.g., *TLR3*, *UNC93B1*, *TBK1*, *IRF3*, etc., Figure 4A). In short, the set of pLI, pNull, and pRec probabilities provides important information for expecting a variant to be observed as homozygous/heterozygous. Genes with low pLI (equal to 0) and high pNull (>0.8) probability would have a higher chance of having homozygous variants in a population. For example, *FUT2* and *CCR5* (pLI = 0 and 0; pNull = 1.0 and 0.98) genes have homozygous truncation variants present in the human population. Because these proteins directly participate in the entry of their respective virus, homozygous truncation variants of FUT2 and CCR5 provide protective advantage. Therefore, for SARS-CoV-2 PPI and RPI proteins that directly interact with viral machinery (and likely participate in the essential steps of the viral life cycle), a discussion on the behavior of pLI and pNull probabilities of these genes is very useful (Figure 5). There are 24 and 13 genes of SARS-CoV-2 PPI and RPI proteins, respectively, with pLI = 0 and pNull >0.8 (Figure 5A), and individuals carrying homozygous truncation variants of these genes are likely to be found in the human population. For example, homozygous individuals carrying nonsense or frame-shift variants of these genes (e.g., CNTRL, NINL, ANO6, CHD1L, ANXA5, NLRX1, etc.) are observed in gnomAD. The absence of functional forms of these proteins in homozygous individuals could provide a protective advantage. In general, with respect to the disease susceptibility, homozygous variants of genes with high pNull probability are expected to (a) provide protection if the protein is essential for the viral life cycle and (b) result in severe symptoms if the protein is a vital player in innate immunity. The distribution of the pLI score (Figure 5B) shows that ∼24 and ∼49% of human SARS-CoV-2 PPI and RPI proteins, respectively, have pLI > 0.9, whereas only ∼16% of human proteins, in comparison, have pLI > 0.9 (Figure 5B). This suggests that the virus tends to interact with machinery (especially for viral RNA maintenance) that is less prone to tolerating truncation variants than that expected from the reference background. This could be an adaptation to minimize host variability for viral survival. The distribution of pNull probability suggests that ∼10% of human genes are tolerant to loss-of-function genes as homozygous (i.e., pNull ≥0.9), whereas that for SARS-CoV-2 PPI and RPI protein genes is ∼5 and ∼3%, respectively, which is lower than the reference background (Figure 5B) and is consistent with an adaptation to minimize host variability. Among the SARS-CoV-2 PPI and RPI protein genes, 19 are X-linked (e.g., POLA1, PDZD11, DDX3X, MID1IP1, etc.), with none having high pNull probability, but three X-linked genes (PDZD11, HS6ST2, and MID1IP1) have low pLI, and their nonsense variants may provide protection in men. Finally, other variants (e.g., 5′ UTR, synonymous, splice-acceptor/donor, etc.) affecting tissue-level protein expression or stability-compromising missense variants that indirectly affect protein expression levels of the above-discussed pathway specific proteins could also be valuable for assessing the role of genotype in pathogenesis and susceptibility.

CONCLUSIONS

In recognition of the important role that variants could play in disease susceptibility, the human protein variants involved in SARS-CoV-2−host interactions have recently been reported, analyzed, and discussed.[33,34] The emphasis here has been to provide a comprehensive view of the roles that human protein variants could play not only in these interactions but also in all major aspects of viral biology, such as viral entry, innate immunity, and host−virus interactions (protein−protein and protein−RNA). Because *rare* variants have been observed to contribute to disease severity in otherwise healthy individuals,[4] a careful look at those involving pathway-specific proteins in a large reference database is needed, given that they typically occur in heterozygotes. The findings extracted from the gnomAD database suggest that heterozygous carriers of several rare variants are present in the human population and that disease risks may be associated with these individuals. In

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

addition, the observed influence of *rare* homozygous recessive variants on disease severity[4] encouraged an assessment of tolerance for loss-of-function (involving pLI, pNull, and pRec probabilities[117]) of the genes discussed above. The degree of this tolerance helps estimate the likelihood of observing *rare* variant carriers (both homozygous and heterozygous) in a population. Although the discussion was centered around gnomAD, the *rare* variants not observed in this database (e.g., the variants of TLR7 and TLR3 discussed above) could play equally important roles. Because the profiling of an individual's variants can now be performed with relative ease (e.g., using quantitative PCR[139] or MassARRAY[140−142]), this is especially true for specific loci, which would avoid whole-genome sequencing and thereby facilitate the assessment of COVID-19 disease risks. Such profiling would be valuable for optimizing national responses to this health crisis while we await an effective therapeutic or vaccine. In addition, knowledge about the protein structural consequences of these variants (Figures 1, 3, and 4) enables probing of the mechanistic basis of COVID-19 susceptibility. The significance of the discussion is thus to advocate: (a) rapid sequencing of specific loci to facilitate the assessment of COVID-19 disease risk and (b) experimental probing of the structural consequences of known variants to unravel the mechanistic basis of susceptibility.

■ ASSOCIATED CONTENT

Special Issue Paper

This paper was intended for the Proteomics in Pandemic Disease Special Issue, published as the November 6, 2020 issue of *J. Proteome Res.* (Vol. 19, No. 11).

■ AUTHOR INFORMATION

Corresponding Author

Suvobrata Chakravarty − *Chemistry & Biochemistry, South Dakota State University, Brookings, South Dakota 57007, United States; BioSNTR, Brookings, South Dakota 57007, United States;* orcid.org/0000-0002-0044-3931; Phone: 1-605-688-5694; Email: suvobrata.chakravarty@sdstate.edu; Fax: 1-605-688-6364

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jproteome.0c00637

Notes

The author declares no competing financial interest.

■ ACKNOWLEDGMENTS

I thank all of the investigators who contributed to and created the Exome Aggregation Consortium for publicly sharing the genetic variation datasets. I also thank all reviewers for valuable suggestions. I thank Logan A. Smith and Avori K. Bestemeyer for reading the text carefully and Dr. Moul Day for suggestions and encouragement. Research in the S.C. laboratory is supported by the U.S. National Institutes of Health through grants from the National Institutes of General Medical Sciences (1R15-GM134502-01, 1R15GM116040-01A1), National Science Foundation/EPSCoR award no. AII-1355423, and the state of South Dakota through BioSNTR (Biochemical Spatiotemporal NeTwork Resource). Any opinions, findings, and conclusions or recommendations expressed in this material are those of authors and do not necessarily reflect the views of the National Science Foundation and National Institutes of Health.

■ REFERENCES

(1) Wiersinga, W. J.; Prescott, H. C. What Is COVID-19? *JAMA* **2020**, *324*, 816.
(2) Callaway, E.; Cyranoski, D.; Mallapaty, S.; Stoye, E.; Tollefson, J. Coronavirus by the Numbers. *Nature* **2020**, *579*, 482.
(3) van der Made, C. I.; Simons, A.; Schuurs-Hoeijmakers, J.; van den Heuvel, G.; Mantere, T.; Kersten, S.; van Deuren, R. C.; Steehouwer, M.; van Reijmersdal, S. V.; Jaeger, M.; et al. Presence of Genetic Variants Among Young Men With Severe COVID-19. *JAMA* **2020**, *324*, 663−673.
(4) Zhang, Q.; Bastard, P.; Liu, Z.; Le Pen, J.; Moncada-Velez, M.; Chen, J.; Ogishi, M.; Sabli, I. K. D.; Hodeib, S.; Korol, C.; et al. Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* **2020**, *370*, eabd4570.
(5) Bastard, P.; Rosen, L. B.; Zhang, Q.; Michailidis, E.; Hoffmann, H. H.; Zhang, Y.; Dorgham, K.; Philippot, Q.; Rosain, J.; Beziat, V.; et al. Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* **2020**, *370*, eabd4585.
(6) Hadjadj, J.; Yatim, N.; Barnabei, L.; Corneau, A.; Boussier, J.; Smith, N.; Péré, H.; Charbit, B.; Bondet, V.; Chenevier-Gobeaux, C.; Breillat, P.; Carlier, N.; Gauzit, R.; Morbieu, C.; Pène, F.; Marin, N.; Roche, N.; Szwebel, T.-A.; Merkling, S. H.; Treluyer, J.-M.; Veyer, D.; Mouthon, L.; Blanc, C.; Tharaux, P.-L.; Rozenberg, F.; Fischer, A.; Duffy, D.; Rieux-Laucat, F.; Kernéis, S.; Terrier, B. Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **2020**, *369*, 718−724.
(7) Lucas, C.; Wong, P.; Klein, J.; Castro, T. B. R.; Silva, J.; Sundaram, M.; Ellingson, M. K.; Mao, T.; Oh, J. E.; Israelow, B.; et al. Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **2020**, *584*, 463−469.
(8) Arunachalam, P. S.; Wimmers, F.; Mok, C. K. P.; Perera, R. A. P. M.; Scott, M.; Hagan, T.; Sigal, N.; Feng, Y.; Bristow, L.; Tak-Yin Tsang, O.; Wagh, D.; Coller, J.; Pellegrini, K. L.; Kazmin, D.; Alaaeddine, G.; Leung, W. S.; Chan, J. M. C.; Chik, T. S. H.; Choi, C. Y. C.; Huerta, C.; Paine McCullough, M.; Lv, H.; Anderson, E.; Edupuganti, S.; Upadhyay, A. A.; Bosinger, S. E.; Maecker, H. T.; Khatri, P.; Rouphael, N.; Peiris, M.; Pulendran, B. Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **2020**, *369*, 1210−1220.
(9) The Severe Covid-19 GWAS Group. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. *N. Engl. J. Med.* **2020**, *383*, 1522−1534.
(10) Chen, J.; Jiang, Q.; Xia, X.; Liu, K.; Yu, Z.; Tao, W.; Gong, W.; Han, J. J. Individual variation of the SARS-CoV-2 receptor ACE2 gene expression and regulation. *Aging Cell* **2020**, *19*, No. e13168.
(11) Brest, P.; Refae, S.; Mograbi, B.; Hofman, P.; Milano, G. Host polymorphisms may impact SARS-CoV-2 infectivity. *Trends Genet.* **2020**, *36*, 813−815.
(12) Casanova, J.-L.; Su, H. C.; Abel, L.; Aiuti, A.; Almuhsen, S.; Arias, A. A.; Bastard, P.; Biggs, C.; Bogunovic, D.; Boisson, B.; et al. A Global Effort to Define the Human Genetics of Protective Immunity to SARS-CoV-2 Infection. *Cell* **2020**, *181*, 1194−1199.
(13) Chapman, S. J.; Hill, A. V. Human genetic susceptibility to infectious disease. *Nat. Rev. Genet.* **2012**, *13*, 175−188.
(14) Tian, C.; Hromatka, B. S.; Kiefer, A. K.; Eriksson, N.; Noble, S. M.; Tung, J. Y.; Hinds, D. A. Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nat. Commun.* **2017**, *8*, 599.
(15) Newport, M. J.; Finan, C. Genome-wide association studies and susceptibility to infectious diseases. *Briefings Funct. Genomics* **2011**, *10*, 98−107.
(16) Wang, L.; Pittman, K. J.; Barker, J. R.; Salinas, R. E.; Stanaway, I. B.; Williams, G. D.; Carroll, R. J.; Balmat, T.; Ingham, A.; Gopalakrishnan, A. M.; et al. An Atlas of Genetic Variation Linking Pathogen-Induced Cellular Traits to Human Disease. *Cell Host Microbe* **2018**, *24*, 308−323.
(17) Klebanov, N. Genetic Predisposition to Infectious Disease. *Cureus* **2018**, *10*, e3210.

(18) López, C.; Saravia, C.; Gomez, A.; Hoebeke, J.; Patarroyo, M. A. Mechanisms of genetically-based resistance to malaria. *Gene* **2010**, *467*, 1−12.
(19) Hedrick, P. W. Population genetics of malaria resistance in humans. *Heredity* **2011**, *107*, 283−304.
(20) Allison, A. C. Protection afforded by sickle-cell trait against subtertian malarial infection. *British medical journal* **1954**, *1*, 290.
(21) Dean, M.; Carrington, M.; Winkler, C.; Huttley, G. A.; Smith, M. W.; Allikmets, R.; Goedert, J. J.; Buchbinder, S. P.; Vittinghoff, E.; Gomperts, E.; et al. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science* **1996**, *273*, 1856−1862.
(22) Michael, N. L.; Chang, G.; Loum, L. G.; Mascola, J. R.; Dondero, D.; Birx, D. L.; Sheppard, H. W. The role of viral phenotype and CCR-5 gene defects in HIV-1 transmission and disease progression. *Nat. Med.* **1997**, *3*, 338−340.
(23) Lindesmith, L.; Moe, C.; Marionneau, S.; Ruvoen, N.; Jiang, X.; Lindblad, L.; Stewart, P.; LePendu, J.; Baric, R. Human susceptibility and resistance to Norwalk virus infection. *Nat. Med.* **2003**, *9*, 548−553.
(24) Liu, S.-C.; Palek, J.; Yi, S.; Nichols, P.; Derick, L.; Chiou, S.-S.; Amato, D.; Corbett, J.; Cho, M.; Golan, D. Molecular basis of altered red blood cell membrane properties in Southeast Asian ovalocytosis: role of the mutant band 3 protein in band 3 oligomerization and retention by the membrane skeleton. *Blood* **1995**, *86*, 349−358.
(25) Hadi-Alijanvand, H.; Rouhani, M. Studying the Effects of ACE2Mutations on Stability, Dynamics and Dissociation Process of SARS-CoV-2 S1/hACE2 Complexes. *J. Proteome Res.* **2020**, *19*, 4609−4623.
(26) Zhang, S.-Y.; Jouanguy, E.; Ugolini, S.; Smahi, A.; Elain, G.; Romero, P.; Segal, D.; Sancho-Shimizu, V.; Lorenzo, L.; Puel, A.; Picard, C.; Chapgier, A.; Plancoulaine, S.; Titeux, M.; Cognet, C.; von Bernuth, H.; Ku, C.-L.; Casrouge, A.; Zhang, X.-X.; Barreiro, L.; Leonard, J.; Hamilton, C.; Lebon, P.; Héron, B.; Vallée, L.; Quintana-Murci, L.; Hovnanian, A.; Rozenberg, F.; Vivier, E.; Geissmann, F.; Tardieu, M.; Abel, L.; Casanova, J.-L. TLR3 Deficiency in Patients with Herpes Simplex Encephalitis. *Science* **2007**, *317*, 1522−1527.
(27) Asgari, S.; Schlapbach, L. J.; Anchisi, S.; Hammer, C.; Bartha, I.; Junier, T.; Mottet-Osman, G.; Posfay-Barbe, K. M.; Longchamp, D.; Stocker, M.; et al. Severe viral respiratory infections in children with *IFIH1* loss-of-function mutations. *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 8342−8347.
(28) Lim, H. K.; Huang, S. X. L.; Chen, J.; Kerner, G.; Gilliaux, O.; Bastard, P.; Dobbs, K.; Hernandez, N.; Goudin, N.; Hasek, M. L.; Garcia Reino, E. J.; Lafaille, F. G.; Lorenzo, L.; Luthra, P.; Kochetkov, T.; Bigio, B.; Boucherit, S.; Rozenberg, F.; Vedrinne, C.; Keller, M. D.; Itan, Y.; Garcia-Sastre, A.; Celard, M.; Orange, J. S.; Ciancanelli, M. J.; Meyts, I.; Zhang, Q.; Abel, L.; Notarangelo, L. D.; Snoeck, H. W.; Casanova, J. L.; Zhang, S. Y. Severe influenza pneumonitis in children with inherited TLR3 deficiency. *J. Exp. Med.* **2019**, *216*, 2038−2056.
(29) Hernandez, N.; Melki, I.; Jing, H.; Habib, T.; Huang, S. S. Y.; Danielson, J.; Kula, T.; Drutman, S.; Belkaya, S.; Rattina, V.; Lorenzo-Diaz, L.; Boulai, A.; Rose, Y.; Kitabayashi, N.; Rodero, M. P.; Dumaine, C.; Blanche, S.; Lebras, M. N.; Leung, M. C.; Mathew, L. S.; Boisson, B.; Zhang, S. Y.; Boisson-Dupuis, S.; Giliani, S.; Chaussabel, D.; Notarangelo, L. D.; Elledge, S. J.; Ciancanelli, M. J.; Abel, L.; Zhang, Q.; Marr, N.; Crow, Y. J.; Su, H. C.; Casanova, J. L. Life-threatening influenza pneumonitis in a child with inherited IRF9 deficiency. *J. Exp. Med.* **2018**, *215*, 2567−2585.
(30) 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.; et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **2020**, *581*, 434−443.
(31) Hauser, A. S.; Chavali, S.; Masuho, I.; Jahn, L. J.; Martemyanov, K. A.; Gloriam, D. E.; Babu, M. M. Pharmacogenomics of GPCR Drug Targets. *Cell* **2018**, *172*, 41−54.
(32) Wang, B.; Yan, C.; Lou, S.; Emani, P.; Li, B.; Xu, M.; Kong, X.; Meyerson, W.; Yang, Y. T.; Lee, D.; Gerstein, M. Building a Hybrid Physical-Statistical Classifier for Predicting the Effect of Variants Related to Protein-Drug Interactions. *Structure* **2019**, *27*, 1469−1481.
(33) Gupta, R.; Charron, J.; Stenger, C. L.; Painter, J.; Steward, H.; Cook, T. W.; Faber, W.; Frisch, A.; Lind, E.; Bauss, J.; Li, X.; Sirpilla, O.; Soehnlen, X.; Underwood, A.; Hinds, D.; Morris, M.; Lamb, N.; Carcillo, J. A.; Bupp, C.; Uhal, B. D.; Rajasekaran, S.; Prokop, J. W. SARS-CoV-2 (COVID-19) structural and evolutionary dynamicome: Insights into functional evolution and human genomics. *J. Biol. Chem.* **2020**, *295*, 11742−11753.
(34) Sirpilla, O.; Bauss, J.; Gupta, R.; Underwood, A.; Qutob, D.; Freeland, T.; Bupp, C.; Carcillo, J.; Hartog, N.; Rajasekaran, S.; et al. SARS-CoV-2 encoded proteome and human genetics: from interaction-based to ribosomal biology impact on disease and risk processes. *J. Proteome Res.* **2020**, *19*, 4275−4290.
(35) Hoffmann, M.; Kleine-Weber, H.; Pohlmann, S. A Multibasic Cleavage Site in the Spike Protein of SARS-CoV-2 Is Essential for Infection of Human Lung Cells. *Mol. Cell* **2020**, *78*, 779−784.
(36) Hoffmann, M.; Kleine-Weber, H.; Schroeder, S.; Krüger, N.; Herrler, T.; Erichsen, S.; Schiergens, T. S.; Herrler, G.; Wu, N.-H.; Nitsche, A.; et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **2020**, *181*, 271−280.
(37) Walls, A. C.; Park, Y.-J.; Tortorici, M. A.; Wall, A.; McGuire, A. T.; Veesler, D. Structure, function, and antigenicity of the SARS-CoV-2 spike glycoprotein. *Cell* **2020**, *181*, 281−292.
(38) Shang, J.; Wan, Y.; Luo, C.; Ye, G.; Geng, Q.; Auerbach, A.; Li, F. Cell entry mechanisms of SARS-CoV-2. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 11727−11734.
(39) Wrapp, D.; Wang, N.; Corbett, K. S.; Goldsmith, J. A.; Hsieh, C.-L.; Abiona, O.; Graham, B. S.; McLellan, J. S. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **2020**, *367*, 1260−1263.
(40) Turoňová, B.; Sikora, M.; Schürmann, C.; Hagen, W. J.; Welsch, S.; Blanc, F. E.; von Bülow, S.; Gecht, M.; Bagola, K.; Hörner, C.; et al. In situ structural analysis of SARS-CoV-2 spike reveals flexibility mediated by three hinges. *Science* **2020**, *370*, 203−208.
(41) Yao, H.; Song, Y.; Chen, Y.; Wu, N.; Xu, J.; Sun, C.; Zhang, J.; Weng, T.; Zhang, Z.; Wu, Z.; Cheng, L.; Shi, D.; Lu, X.; Lei, J.; Crispin, M.; Shi, Y.; Li, L.; Li, S. Molecular Architecture of the SARS-CoV-2 Virus. *Cell* **2020**, *183*, 730−738.
(42) Liu, C.; Mendonça, L.; Yang, Y.; Gao, Y.; Shen, C.; Liu, J.; Ni, T.; Ju, B.; Liu, C.; Tang, X.; Wei, J.; Ma, X.; Zhu, Y.; Liu, W.; Xu, S.; Liu, Y.; Yuan, J.; Wu, J.; Liu, Z.; Zhang, Z.; Liu, L.; Wang, P.; Zhang, P. The Architecture of Inactivated SARS-CoV-2 with Postfusion Spikes Revealed by Cryo-EM and Cryo-ET. *Structure* **2020**, *28*, 1218−1224.
(43) Hsieh, C.-L.; Goldsmith, J. A.; Schaub, J. M.; DiVenere, A. M.; Kuo, H.-C.; Javanmardi, K.; Le, K. C.; Wrapp, D.; Lee, A. G.; Liu, Y.; Chou, C.-W.; Byrne, P. O.; Hjorth, C. K.; Johnson, N. V.; Ludes-Meyers, J.; Nguyen, A. W.; Park, J.; Wang, N.; Amengor, D.; Lavinder, J. J.; Ippolito, G. C.; Maynard, J. A.; Finkelstein, I. J.; McLellan, J. S. Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science* **2020**, *369*, 1501−1505.
(44) Xiong, X.; Qu, K.; Ciazynska, K. A.; Hosmillo, M.; Carter, A. P.; Ebrahimi, S.; Ke, Z.; Scheres, S. H. W.; Bergamaschi, L.; Grice, G. L.; et al. A thermostable, closed SARS-CoV-2 spike protein trimer. *Nat. Struct. Mol. Biol.* **2020**, *27*, 934−941.
(45) Henderson, R.; Edwards, R. J.; Mansouri, K.; Janowska, K.; Stalls, V.; Gobeil, S. M. C.; Kopp, M.; Li, D.; Parks, R.; Hsu, A. L.; Borgnia, M. J.; Haynes, B. F.; Acharya, P. Controlling the SARS-CoV-2 spike glycoprotein conformation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 925−933.
(46) Zhou, T.; Teng, I. T.; Olia, A. S.; Cerutti, G.; Gorman, J.; Nazzari, A.; Shi, W.; Tsybovsky, Y.; Wang, L.; Wang, S.; Zhang, B.; Zhang, Y.; Katsamba, P. S.; Petrova, Y.; Banach, B. B.; Fahad, A. S.; Liu, L.; Lopez Acevedo, S. N.; Madan, B.; Oliveira de Souza, M.; Pan, X.; Wang, P.; Wolfe, J. R.; Yin, M.; Ho, D. D.; Phung, E.; DiPiazza, A.; Chang, L. A.; Abiona, O. M.; Corbett, K. S.; DeKosky, B. J.; Graham, B. S.; Mascola, J. R.; Misasi, J.; Ruckwardt, T.; Sullivan, N. J.; Shapiro, L.; Kwong, P. D. Structure-Based Design with Tag-Based Purification and In-Process Biotinylation Enable Streamlined Development of SARS-CoV-2 Spike Molecular Probes. *Cell Rep.* **2020**, *33*, 108322.

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

(47) Starr, T. N.; Greaney, A. J.; Hilton, S. K.; Ellis, D.; Crawford, K. H. D.; Dingens, A. S.; Navarro, M. J.; Bowen, J. E.; Tortorici, M. A.; Walls, A. C.; King, N. P.; Veesler, D.; Bloom, J. D. Deep Mutational Scanning of SARS-CoV-2 Receptor Binding Domain Reveals Constraints on Folding and ACE2 Binding. *Cell* **2020**, *182*, 1295−1310.
(48) Li, Q.; Wu, J.; Nie, J.; Zhang, L.; Hao, H.; Liu, S.; Zhao, C.; Zhang, Q.; Liu, H.; Nie, L.; Qin, H.; Wang, M.; Lu, Q.; Li, X.; Sun, Q.; Liu, J.; Zhang, L.; Li, X.; Huang, W.; Wang, Y. The Impact of Mutations in SARS-CoV-2 Spike on Viral Infectivity and Antigenicity. *Cell* **2020**, *182*, 1284−1294.
(49) Chen, J.; Wang, R.; Wang, M.; Wei, G. W. Mutations Strengthened SARS-CoV-2 Infectivity. *J. Mol. Biol.* **2020**, *432*, 5212−5226.
(50) Yurkovetskiy, L.; Wang, X.; Pascal, K. E.; Tomkins-Tinch, C.; Nyalile, T. P.; Wang, Y.; Baum, A.; Diehl, W. E.; Dauphin, A.; Carbone, C.; Veinotte, K.; Egri, S. B.; Schaffner, S. F.; Lemieux, J. E.; Munro, J. B.; Rafique, A.; Barve, A.; Sabeti, P. C.; Kyratsous, C. A.; Dudkina, N. V.; Shen, K.; Luban, J. Structural and Functional Analysis of the D614G SARS-CoV-2 Spike Protein Variant. *Cell* **2020**, *183*, 739−751.
(51) Plante, J. A.; Liu, Y.; Liu, J.; Xia, H.; Johnson, B. A.; Lokugamage, K. G.; Zhang, X.; Muruato, A. E.; Zou, J.; Fontes-Garfias, C. R.; Mirchandani, D.; Scharton, D.; Bilello, J. P.; Ku, Z.; An, Z.; Kalveram, B.; Freiberg, A. N.; Menachery, V. D.; Xie, X.; Plante, K. S.; Weaver, S. C.; Shi, P. Y. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **2020**, DOI: 10.1038/s41586-020-2895-3.
(52) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270−273.
(53) Letko, M.; Marzi, A.; Munster, V. Functional assessment of cell entry and receptor usage for SARS-CoV-2 and other lineage B betacoronaviruses. *Nature microbiology* **2020**, *5*, 562−569.
(54) Lukassen, S.; Chua, R. L.; Trefzer, T.; Kahn, N. C.; Schneider, M. A.; Muley, T.; Winter, H.; Meister, M.; Veith, C.; Boots, A. W.; Hennig, B. P.; 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.
(55) Hou, Y. J.; Okuda, K.; Edwards, C. E.; Martinez, D. R.; Asakura, T.; Dinnon, K. H., 3rd; Kato, T.; Lee, R. E.; Yount, B. L.; Mascenik, T. M.; Chen, G.; Olivier, K. N.; Ghio, A.; Tse, L. V.; Leist, S. R.; Gralinski, L. E.; Schafer, A.; Dang, H.; Gilmore, R.; Nakano, S.; Sun, L.; Fulcher, M. L.; Livraghi-Butrico, A.; Nicely, N. I.; Cameron, M.; Cameron, C.; Kelvin, D. J.; de Silva, A.; Margolis, D. M.; Markmann, A.; Bartelt, L.; Zumwalt, R.; Martinez, F. J.; Salvatore, S. P.; Borczuk, A.; Tata, P. R.; Sontake, V.; Kimple, A.; Jaspers, I.; O'Neal, W. K.; Randell, S. H.; Boucher, R. C.; Baric, R. S. SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* **2020**, *182*, 429−446.
(56) Hikmet, F.; Mear, L.; Edvinsson, A.; Micke, P.; Uhlen, M.; Lindskog, C. The protein expression profile of ACE2 in human tissues. *Mol. Syst. Biol.* **2020**, *16*, e9610.
(57) Yan, R.; Zhang, Y.; Li, Y.; Xia, L.; Guo, Y.; Zhou, Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science* **2020**, *367*, 1444−1448.
(58) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221−224.
(59) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K. Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* **2020**, *181*, 894−904.
(60) Lan, J.; Ge, J.; Yu, J.; Shan, S.; Zhou, H.; Fan, S.; Zhang, Q.; Shi, X.; Wang, Q.; Zhang, L.; et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* **2020**, *581*, 215−220.
(61) Cao, Y.; Su, B.; Guo, X.; Sun, W.; Deng, Y.; Bao, L.; Zhu, Q.; Zhang, X.; Zheng, Y.; Geng, C.; Chai, X.; He, R.; Li, X.; Lv, Q.; Zhu, H.; Deng, W.; Xu, Y.; Wang, Y.; Qiao, L.; Tan, Y.; Song, L.; Wang, G.; Du, X.; Gao, N.; Liu, J.; Xiao, J.; Su, X. D.; Du, Z.; Feng, Y.; Qin, C.; Qin, C.; Jin, R.; Xie, X. S. Potent Neutralizing Antibodies against SARS-CoV-2 Identified by High-Throughput Single-Cell Sequencing of Convalescent Patients' B Cells. *Cell* **2020**, *182*, 73−84.
(62) Zost, S. J.; Gilchuk, P.; Case, J. B.; Binshtein, E.; Chen, R. E.; Nkolola, J. P.; Schafer, A.; Reidy, J. X.; Trivette, A.; Nargi, R. S. Potently neutralizing and protective human antibodies against SARS-CoV-2. *Nature* **2020**, *584*, 443.
(63) Pinto, D.; Park, Y. J.; Beltramello, M.; Walls, A. C.; Tortorici, M. A.; Bianchi, S.; Jaconi, S.; Culap, K.; Zatta, F.; De Marco, A.; Peter, A.; Guarino, B.; Spreafico, R.; Cameroni, E.; Case, J. B.; Chen, R. E.; Havenar-Daughton, C.; Snell, G.; Telenti, A.; Virgin, H. W.; Lanzavecchia, A.; Diamond, M. S.; Fink, K.; Veesler, D.; Corti, D. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **2020**, *583*, 290−295.
(64) Zost, S. J.; Gilchuk, P.; Chen, R. E.; Case, J. B.; Reidy, J. X.; Trivette, A.; Nargi, R. S.; Sutton, R. E.; Suryadevara, N.; Chen, E. C.; Binshtein, E.; Shrihari, S.; Ostrowski, M.; Chu, H. Y.; Didier, J. E.; MacRenaris, K. W.; Jones, T.; Day, S.; Myers, L.; Eun-Hyung Lee, F.; Nguyen, D. C.; Sanz, I.; Martinez, D. R.; Rothlauf, P. W.; Bloyet, L. M.; Whelan, S. P. J.; Baric, R. S.; Thackray, L. B.; Diamond, M. S.; Carnahan, R. H.; Crowe, J. E., Jr. Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nat. Med.* **2020**, *26*, 1422−1427.
(65) Tan, C. W.; Chia, W. N.; Qin, X.; Liu, P.; Chen, M. I.; Tiu, C.; Hu, Z.; Chen, V. C.; Young, B. E.; Sia, W. R.; Tan, Y. J.; Foo, R.; Yi, Y.; Lye, D. C.; Anderson, D. E.; Wang, L. F. A SARS-CoV-2 surrogate virus neutralization test based on antibody-mediated blockage of ACE2-spike protein-protein interaction. *Nat. Biotechnol.* **2020**, *38*, 1073−1078.
(66) Nie, J.; Li, Q.; Wu, J.; Zhao, C.; Hao, H.; Liu, H.; Zhang, L.; Nie, L.; Qin, H.; Wang, M.; Lu, Q.; Li, X.; Sun, Q.; Liu, J.; Fan, C.; Huang, W.; Xu, M.; Wang, Y. Quantification of SARS-CoV-2 neutralizing antibody by a pseudotyped virus-based assay. *Nat. Protoc.* **2020**, *15*, 3699−3715.
(67) Yuan, M.; Wu, N. C.; Zhu, X.; Lee, C.-C. D.; So, R. T.; Lv, H.; Mok, C. K.; Wilson, I. A. A highly conserved cryptic epitope in the receptor binding domains of SARS-CoV-2 and SARS-CoV. *Science* **2020**, *368*, 630−633.
(68) Barnes, C. O.; West, A. P., Jr.; Huey-Tubman, K. E.; Hoffmann, M. A. G.; Sharaf, N. G.; Hoffman, P. R.; Koranda, N.; Gristick, H. B.; Gaebler, C.; Muecksch, F.; Lorenzi, J. C. C.; Finkin, S.; Hagglof, T.; Hurley, A.; Millard, K. G.; Weisblum, Y.; Schmidt, F.; Hatziioannou, T.; Bieniasz, P. D.; Caskey, M.; Robbiani, D. F.; Nussenzweig, M. C.; Bjorkman, P. J. Structures of Human Antibodies Bound to SARS-CoV-2 Spike Reveal Common Epitopes and Recurrent Features of Antibodies. *Cell* **2020**, *182*, 828−842.
(69) Barnes, C. O.; Jette, C. A.; Abernathy, M. E.; Dam, K. A.; Esswein, S. R.; Gristick, H. B.; Malyutin, A. G.; Sharaf, N. G.; Huey-Tubman, K. E.; Lee, Y. E.; Robbiani, D. F.; Nussenzweig, M. C.; West, A. P., Jr.; Bjorkman, P. J. SARS-CoV-2 neutralizing antibody structures inform therapeutic strategies. *Nature* **2020**, DOI: 10.1038/s41586-020-2852-1.
(70) Zhou, D.; Duyvesteyn, H. M. E.; Chen, C. P.; Huang, C. G.; Chen, T. H.; Shih, S. R.; Lin, Y. C.; Cheng, C. Y.; Cheng, S. H.; Huang, Y. C.; Lin, T. Y.; Ma, C.; Huo, J.; Carrique, L.; Malinauskas, T.; Ruza, R. R.; Shah, P. N. M.; Tan, T. K.; Rijal, P.; Donat, R. F.; Godwin, K.; Buttigieg, K. R.; Tree, J. A.; Radecke, J.; Paterson, N. G.; Supasa, P.; Mongkolsapaya, J.; Screaton, G. R.; Carroll, M. W.; Gilbert-Jaramillo, J.; Knight, M. L.; James, W.; Owens, R. J.; Naismith, J. H.; Townsend, A. R.; Fry, E. E.; Zhao, Y.; Ren, J.; Stuart, D. I.; Huang, K. A. Structural basis for the neutralization of SARS-CoV-2 by an antibody from a convalescent patient. *Nat. Struct. Mol. Biol.* **2020**, *27*, 950−958.
(71) Wu, N. C.; Yuan, M.; Liu, H.; Lee, C. D.; Zhu, X.; Bangaru, S.; Torres, J. L.; Caniels, T. G.; Brouwer, P. J. M.; van Gils, M. J.; Sanders, R. W.; Ward, A. B.; Wilson, I. A. An Alternative Binding Mode of IGHV3−53 Antibodies to the SARS-CoV-2 Receptor Binding Domain. *Cell Rep.* **2020**, *33*, 108274.

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

(72) Chan, K. K.; Dorosky, D.; Sharma, P.; Abbasi, S. A.; Dye, J. M.; Kranz, D. M.; Herbert, A. S.; Procko, E. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **2020**, *369*, 1261−1265.
(73) Cao, L.; Goreshnik, I.; Coventry, B.; Case, J. B.; Miller, L.; Kozodoy, L.; Chen, R. E.; Carter, L.; Walls, A. C.; Park, Y.-J.; Strauch, E.-M.; Stewart, L.; Diamond, M. S.; Veesler, D.; Baker, D. De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **2020**, *370*, 426−431.
(74) Huo, J.; Le Bas, A.; Ruza, R. R.; Duyvesteyn, H. M. E.; Mikolajek, H.; Malinauskas, T.; Tan, T. K.; Rijal, P.; Dumoux, M.; Ward, P. N.; Ren, J.; Zhou, D.; Harrison, P. J.; Weckener, M.; Clare, D. K.; Vogirala, V. K.; Radecke, J.; Moynie, L.; Zhao, Y.; Gilbert-Jaramillo, J.; Knight, M. L.; Tree, J. A.; Buttigieg, K. R.; Coombes, N.; Elmore, M. J.; Carroll, M. W.; Carrique, L.; Shah, P. N. M.; James, W.; Townsend, A. R.; Stuart, D. I.; Owens, R. J.; Naismith, J. H. Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat. Struct. Mol. Biol.* **2020**, *27*, 846−854.
(75) Li, W.; Schafer, A.; Kulkarni, S. S.; Liu, X.; Martinez, D. R.; Chen, C.; Sun, Z.; Leist, S. R.; Drelich, A.; Zhang, L.; Ura, M. L.; Berezuk, A.; Chittori, S.; Leopold, K.; Mannar, D.; Srivastava, S. S.; Zhu, X.; Peterson, E. C.; Tseng, C. T.; Mellors, J. W.; Falzarano, D.; Subramaniam, S.; Baric, R. S.; Dimitrov, D. S. High Potency of a Bivalent Human VH Domain in SARS-CoV-2 Animal Models. *Cell* **2020**, *183*, 429−441.
(76) Parray, H. A.; Chiranjivi, A. K.; Asthana, S.; Yadav, N.; Shrivastava, T.; Mani, S.; Sharma, C.; Vishwakarma, P.; Das, S.; Pindari, K.; Sinha, S.; Samal, S.; Ahmed, S.; Kumar, R. Identification of an anti-SARS-CoV-2 receptor-binding domain-directed human monoclonal antibody from a naive semisynthetic library. *J. Biol. Chem.* **2020**, *295*, 12814−12821.
(77) Ejemel, M.; Li, Q.; Hou, S.; Schiller, Z. A.; Tree, J. A.; Wallace, A.; Amcheslavsky, A.; Kurt Yilmaz, N.; Buttigieg, K. R.; Elmore, M. J.; et al. A cross-reactive human IgA monoclonal antibody blocks SARS-CoV-2 spike-ACE2 interaction. *Nat. Commun.* **2020**, *11*, 4198.
(78) Gao, Y.; Yan, L.; Huang, Y.; Liu, F.; Zhao, Y.; Cao, L.; Wang, T.; Sun, Q.; Ming, Z.; Zhang, L.; et al. Structure of the RNA-dependent RNA polymerase from COVID-19 virus. *Science* **2020**, *368*, 779−782.
(79) Yin, W.; Mao, C.; Luan, X.; Shen, D.-D.; Shen, Q.; Su, H.; Wang, X.; Zhou, F.; Zhao, W.; Gao, M.; et al. Structural basis for inhibition of the RNA-dependent RNA polymerase from SARS-CoV-2 by remdesivir. *Science* **2020**, *368*, 1499−1504.
(80) Wang, Q.; Wu, J.; Wang, H.; Gao, Y.; Liu, Q.; Mu, A.; Ji, W.; Yan, L.; Zhu, Y.; Zhu, C.; Fang, X.; Yang, X.; Huang, Y.; Gao, H.; Liu, F.; Ge, J.; Sun, Q.; Yang, X.; Xu, W.; Liu, Z.; Yang, H.; Lou, Z.; Jiang, B.; Guddat, L. W.; Gong, P.; Rao, Z. Structural Basis for RNA Replication by the SARS-CoV-2 Polymerase. *Cell* **2020**, *182*, 417−428.
(81) Rosas-Lemus, M.; Minasov, G.; Shuvalova, L.; Inniss, N. L.; Kiryukhina, O.; Brunzelle, J.; Satchell, K. J. High-resolution structures of the SARS-CoV-2 2′-O-methyltransferase reveal strategies for structure-based inhibitor design. *Sci. Signaling* **2020**, *13*, eabe1202.
(82) Schubert, K.; Karousis, E. D.; Jomaa, A.; Scaiola, A.; Echeverria, B.; Gurzeler, L. A.; Leibundgut, M.; Thiel, V.; Muhlemann, O.; Ban, N. SARS-CoV-2 Nsp1 binds the ribosomal mRNA channel to inhibit translation. *Nat. Struct. Mol. Biol.* **2020**, *27*, 959−966.
(83) Thoms, M.; Buschauer, R.; Ameismeier, M.; Koepke, L.; Denk, T.; Hirschenberger, M.; Kratzat, H.; Hayn, M.; Mackens-Kiani, T.; Cheng, J.; Straub, J. H.; Stürzel, C. M.; Fröhlich, T.; Berninghausen, O.; Becker, T.; Kirchhoff, F.; Sparrer, K. M. J.; Beckmann, R. Structural basis for translational shutdown and immune evasion by the Nsp1 protein of SARS-CoV-2. *Science* **2020**, *369*, 1249−1255.
(84) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; Duan, Y.; Yu, J.; Wang, L.; Yang, K.; Liu, F.; Jiang, R.; Yang, X.; You, T.; Liu, X.; Yang, X.; Bai, F.; Liu, H.; Liu, X.; Guddat, L. W.; Xu, W.; Xiao, G.; Qin, C.; Shi, Z.; Jiang, H.; Rao, Z.; Yang, H. Structure of M(pro) from SARS-CoV-2 and discovery of its inhibitors. *Nature* **2020**, *582*, 289−293.
(85) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **2020**, *368*, 409−412.
(86) Riva, L.; Yuan, S.; Yin, X.; Martin-Sancho, L.; Matsunaga, N.; Pache, L.; Burgstaller-Muehlbacher, S.; De Jesus, P. D.; Teriete, P.; Hull, M. V.; Chang, M. W.; Chan, J. F.; Cao, J.; Poon, V. K.; Herbert, K. M.; Cheng, K.; Nguyen, T. H.; Rubanov, A.; Pu, Y.; Nguyen, C.; Choi, A.; Rathnasinghe, R.; Schotsaert, M.; Miorin, L.; Dejosez, M.; Zwaka, T. P.; Sit, K. Y.; Martinez-Sobrido, L.; Liu, W. C.; White, K. M.; Chapman, M. E.; Lendy, E. K.; Glynne, R. J.; Albrecht, R.; Ruppin, E.; Mesecar, A. D.; Johnson, J. R.; Benner, C.; Sun, R.; Schultz, P. G.; Su, A. I.; Garcia-Sastre, A.; Chatterjee, A. K.; Yuen, K. Y.; Chanda, S. K. Discovery of SARS-CoV-2 antiviral drugs through large-scale compound repurposing. *Nature* **2020**, *586*, 113−119.
(87) Li, Z.; Li, X.; Huang, Y. Y.; Wu, Y.; Liu, R.; Zhou, L.; Lin, Y.; Wu, D.; Zhang, L.; Liu, H.; Xu, X.; Yu, K.; Zhang, Y.; Cui, J.; Zhan, C. G.; Wang, X.; Luo, H. B. Identify potent SARS-CoV-2 main protease inhibitors via accelerated free energy perturbation-based virtual screening of existing drugs. *Proc. Natl. Acad. Sci. U. S. A.* **2020**, *117*, 27381−27387.
(88) Rut, W.; Lv, Z.; Zmudzinski, M.; Patchett, S.; Nayak, D.; Snipas, S. J.; El Oualid, F.; Huang, T. T.; Bekes, M.; Drag, M.; Olsen, S. K. Activity profiling and crystal structures of inhibitor-bound SARS-CoV-2 papain-like protease: A framework for anti−COVID-19 drug design. *Science Advances* **2020**, *6*, eabd4596.
(89) Yang, Q.; Hughes, T. A.; Kelkar, A.; Yu, X.; Cheng, K.; Park, S.; Huang, W.-C.; Lovell, J. F.; Neelamegham, S. Inhibition of SARS-CoV-2 viral entry upon blocking N-and O-glycan elaboration. *eLife* **2020**, *9*, e61552.
(90) Bojkova, D.; Klann, K.; Koch, B.; Widera, M.; Krause, D.; Ciesek, S.; Cinatl, J.; Munch, C. Proteomics of SARS-CoV-2-infected host cells reveals therapy targets. *Nature* **2020**, *583*, 469−472.
(91) Bouhaddou, M.; Memon, D.; Meyer, B.; White, K. M.; Rezelj, V. V.; Correa Marrero, M.; Polacco, B. J.; Melnyk, J. E.; Ulferts, S.; Kaake, R. M.; Batra, J.; Richards, A. L.; Stevenson, E.; Gordon, D. E.; Rojc, A.; Obernier, K.; Fabius, J. M.; Soucheray, M.; Miorin, L.; Moreno, E.; Koh, C.; Tran, Q. D.; Hardy, A.; Robinot, R.; Vallet, T.; Nilsson-Payant, B. E.; Hernandez-Armenta, C.; Dunham, A.; Weigang, S.; Knerr, J.; Modak, M.; Quintero, D.; Zhou, Y.; Dugourd, A.; Valdeolivas, A.; Patil, T.; Li, Q.; Huttenhain, R.; Cakir, M.; Muralidharan, M.; Kim, M.; Jang, G.; Tutuncuoglu, B.; Hiatt, J.; Guo, J. Z.; Xu, J.; Bouhaddou, S.; Mathy, C. J. P.; Gaulton, A.; Manners, E. J.; Felix, E.; Shi, Y.; Goff, M.; Lim, J. K.; McBride, T.; O'Neal, M. C.; Cai, Y.; Chang, J. C. J.; Broadhurst, D. J.; Klippsten, S.; De Wit, E.; Leach, A. R.; Kortemme, T.; Shoichet, B.; Ott, M.; Saez-Rodriguez, J.; tenOever, B. R.; Mullins, R. D.; Fischer, E. R.; Kochs, G.; Grosse, R.; Garcia-Sastre, A.; Vignuzzi, M.; Johnson, J. R.; Shokat, K. M.; Swaney, D. L.; Beltrao, P.; Krogan, N. J. The Global Phosphorylation Landscape of SARS-CoV-2 Infection. *Cell* **2020**, *182*, 685−712.
(92) Gordon, D. E.; Jang, G. M.; Bouhaddou, M.; Xu, J.; Obernier, K.; White, K. M.; O'Meara, M. J.; Rezelj, V. V.; Guo, J. Z.; Swaney, D. L.; Tummino, T. A.; Huttenhain, R.; Kaake, R. M.; Richards, A. L.; Tutuncuoglu, B.; Foussard, H.; Batra, J.; Haas, K.; Modak, M.; Kim, M.; Haas, P.; Polacco, B. J.; Braberg, H.; Fabius, J. M.; Eckhardt, M.; Soucheray, M.; Bennett, M. J.; Cakir, M.; McGregor, M. J.; Li, Q.; Meyer, B.; Roesch, F.; Vallet, T.; Mac Kain, A.; Miorin, L.; Moreno, E.; Naing, Z. Z. C.; Zhou, Y.; Peng, S.; Shi, Y.; Zhang, Z.; Shen, W.; Kirby, I. T.; Melnyk, J. E.; Chorba, J. S.; Lou, K.; Dai, S. A.; Barrio-Hernandez, I.; Memon, D.; Hernandez-Armenta, C.; Lyu, J.; Mathy, C. J. P.; Perica, T.; Pilla, K. B.; Ganesan, S. J.; Saltzberg, D. J.; Rakesh, R.; Liu, X.; Rosenthal, S. B.; Calviello, L.; Venkataramanan, S.; Liboy-Lugo, J.; Lin, Y.; Huang, X. P.; Liu, Y.; Wankowicz, S. A.; Bohn, M.; Safari, M.; Ugur, F. S.; Koh, C.; Savar, N. S.; Tran, Q. D.; Shengjuler, D.; Fletcher, S. J.; O'Neal, M. C.; Cai, Y.; Chang, J. C. J.; Broadhurst, D. J.; Klippsten, S.; Sharp, P. P.; Wenzell, N. A.; Kuzuoglu-Ozturk, D.; Wang, H. Y.; Trenker, R.; Young, J. M.; Cavero, D. A.; Hiatt, J.; Roth, T. L.; Rathore, U.; Subramanian, A.; Noack, J.; Hubert, M.; Stroud, R. M.; Frankel, A. D.; Rosenberg, O. S.; Verba, K. A.; Agard, D. A.; Ott, M.; Emerman, M.; Jura, N.; von Zastrow, M.; Verdin, E.; Ashworth, A.; Schwartz, O.; d'Enfert, C.; Mukherjee, S.; Jacobson, M.; Malik, H. S.; Fujimori, D. G.;

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

Ideker, T.; Craik, C. S.; Floor, S. N.; Fraser, J. S.; Gross, J. D.; Sali, A.; Roth, B. L.; Ruggero, D.; Taunton, J.; Kortemme, T.; Beltrao, P.; Vignuzzi, M.; Garcia-Sastre, A.; Shokat, K. M.; Shoichet, B. K.; Krogan, N. J. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **2020**, *583*, 459−468.

(93) Schmidt, N.; Lareau, C. A.; Keshishian, H.; Melanson, R.; Zimmer, M.; Kirschner, L.; Ade, J.; Werner, S.; Caliskan, N.; Lander, E. S. A direct RNA-protein interaction atlas of the SARS-CoV-2 RNA in infected human cells. *bioRxiv* **2020**, 2020.07.15.204404 DOI: 10.1101/2020.07.15.204404.

(94) Flynn, R. A.; Belk, J. A.; Qi, Y.; Yasumoto, Y.; Schmitz, C. O.; Mumbach, M. R.; Limaye, A.; Wei, J.; Alfajaro, M. M.; Parker, K. R. Systematic discovery and functional interrogation of SARS-CoV-2 viral RNA-host protein interactions during infection. *bioRxiv* **2020**, No. 2020.10.06.327445, DOI: 10.1101/2020.10.06.327445.

(95) Jiang, R. D.; Liu, M. Q.; Chen, Y.; Shan, C.; Zhou, Y. W.; Shen, X. R.; Li, Q.; Zhang, L.; Zhu, Y.; Si, H. R.; Wang, Q.; Min, J.; Wang, X.; Zhang, W.; Li, B.; Zhang, H. J.; Baric, R. S.; Zhou, P.; Yang, X. L.; Shi, Z. L. Pathogenesis of SARS-CoV-2 in Transgenic Mice Expressing Human Angiotensin-Converting Enzyme 2. *Cell* **2020**, *182*, 50−58.

(96) Bao, L.; Deng, W.; Huang, B.; Gao, H.; Liu, J.; Ren, L.; Wei, Q.; Yu, P.; Xu, Y.; Qi, F.; Qu, Y.; Li, F.; Lv, Q.; Wang, W.; Xue, J.; Gong, S.; Liu, M.; Wang, G.; Wang, S.; Song, Z.; Zhao, L.; Liu, P.; Zhao, L.; Ye, F.; Wang, H.; Zhou, W.; Zhu, N.; Zhen, W.; Yu, H.; Zhang, X.; Guo, L.; Chen, L.; Wang, C.; Wang, Y.; Wang, X.; Xiao, Y.; Sun, Q.; Liu, H.; Zhu, F.; Ma, C.; Yan, L.; Yang, M.; Han, J.; Xu, W.; Tan, W.; Peng, X.; Jin, Q.; Wu, G.; Qin, C. The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. *Nature* **2020**, *583*, 830−833.

(97) Dinnon, K. H., 3rd; Leist, S. R.; Schafer, A.; Edwards, C. E.; Martinez, D. R.; Montgomery, S. A.; West, A.; Yount, B. L., Jr.; Hou, Y. J.; Adams, L. E.; Gully, K. L.; Brown, A. J.; Huang, E.; Bryant, M. D.; Choong, I. C.; Glenn, J. S.; Gralinski, L. E.; Sheahan, T. P.; Baric, R. S. A mouse-adapted model of SARS-CoV-2 to test COVID-19 countermeasures. *Nature* **2020**, *586*, 560−566.

(98) Sia, S. F.; Yan, L. M.; Chin, A. W. H.; Fung, K.; Choy, K. T.; Wong, A. Y. L.; Kaewpreedee, P.; Perera, R.; Poon, L. L. M.; Nicholls, J. M.; Peiris, M.; Yen, H. L. Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature* **2020**, *583*, 834−838.

(99) Munoz-Fontela, C.; Dowling, W. E.; Funnell, S. G. P.; Gsell, P. S.; Riveros-Balta, A. X.; Albrecht, R. A.; Andersen, H.; Baric, R. S.; Carroll, M. W.; Cavaleri, M.; Qin, C.; Crozier, I.; Dallmeier, K.; de Waal, L.; de Wit, E.; Delang, L.; Dohm, E.; Duprex, W. P.; Falzarano, D.; Finch, C. L.; Frieman, M. B.; Graham, B. S.; Gralinski, L. E.; Guilfoyle, K.; Haagmans, B. L.; Hamilton, G. A.; Hartman, A. L.; Herfst, S.; Kaptein, S. J. F.; Klimstra, W. B.; Knezevic, I.; Krause, P. R.; Kuhn, J. H.; Le Grand, R.; Lewis, M. G.; Liu, W. C.; Maisonnasse, P.; McElroy, A. K.; Munster, V.; Oreshkova, N.; Rasmussen, A. L.; Rocha-Pereira, J.; Rockx, B.; Rodriguez, E.; Rogers, T. F.; Salguero, F. J.; Schotsaert, M.; Stittelaar, K. J.; Thibaut, H. J.; Tseng, C. T.; Vergara-Alert, J.; Beer, M.; Brasel, T.; Chan, J. F. W.; Garcia-Sastre, A.; Neyts, J.; Perlman, S.; Reed, D. S.; Richt, J. A.; Roy, C. J.; Segales, J.; Vasan, S. S.; Henao-Restrepo, A. M.; Barouch, D. H. Animal models for COVID-19. *Nature* **2020**, *586*, 509−515.

(100) Guan, W.-j.; Ni, Z.-y.; Hu, Y.; Liang, W.-h.; Ou, C.-q.; He, J.-x.; Liu, L.; Shan, H.; Lei, C.-l.; Hui, D. S.; et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708−1720.

(101) Grasselli, G.; Zangrillo, A.; Zanella, A.; Antonelli, M.; Cabrini, L.; Castelli, A.; Cereda, D.; Coluccello, A.; Foti, G.; Fumagalli, R.; et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *JAMA* **2020**, *323*, 1574−1581.

(102) Price-Haywood, E. G.; Burton, J.; Fort, D.; Seoane, L. Hospitalization and mortality among black patients and white patients with Covid-19. *N. Engl. J. Med.* **2020**, *382*, 2534−2543.

(103) Richardson, S.; Hirsch, J. S.; Narasimhan, M.; Crawford, J. M.; McGinn, T.; Davidson, K. W.; Barnaby, D. P.; Becker, L. B.; Chelico, J. D.; Cohen, S. L.; et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. *JAMA* **2020**, *323*, 2052−2059.

(104) Atyeo, C.; Fischinger, S.; Zohar, T.; Slein, M. D.; Burke, J.; Loos, C.; McCulloch, D. J.; Newman, K. L.; Wolf, C.; Yu, J.; Shuey, K.; Feldman, J.; Hauser, B. M.; Caradonna, T.; Schmidt, A. G.; Suscovich, T. J.; Linde, C.; Cai, Y.; Barouch, D.; Ryan, E. T.; Charles, R. C.; Lauffenburger, D.; Chu, H.; Alter, G. Distinct Early Serological Signatures Track with SARS-CoV-2 Survival. *Immunity* **2020**, *53*, 524−532.

(105) Grifoni, A.; Weiskopf, D.; Ramirez, S. I.; Mateus, J.; Dan, J. M.; Moderbacher, C. R.; Rawlings, S. A.; Sutherland, A.; Premkumar, L.; Jadi, R. S.; Marrama, D.; de Silva, A. M.; Frazier, A.; Carlin, A. F.; Greenbaum, J. A.; Peters, B.; Krammer, F.; Smith, D. M.; Crotty, S.; Sette, A. Targets of T Cell Responses to SARS-CoV-2 Coronavirus in Humans with COVID-19 Disease and Unexposed Individuals. *Cell* **2020**, *181*, 1489−1501.

(106) Le Bert, N.; Tan, A. T.; Kunasegaran, K.; Tham, C. Y. L.; Hafezi, M.; Chia, A.; Chng, M. H. Y.; Lin, M.; Tan, N.; Linster, M. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **2020**, *584*, 457.

(107) Zhang, X.; Tan, Y.; Ling, Y.; Lu, G.; Liu, F.; Yi, Z.; Jia, X.; Wu, M.; Shi, B.; Xu, S.; Chen, J.; Wang, W.; Chen, B.; Jiang, L.; Yu, S.; Lu, J.; Wang, J.; Xu, M.; Yuan, Z.; Zhang, Q.; Zhang, X.; Zhao, G.; Wang, S.; Chen, S.; Lu, H. Viral and host factors related to the clinical outcome of COVID-19. *Nature* **2020**, *583*, 437−440.

(108) Sahin, U.; Muik, A.; Derhovanessian, E.; Vogler, I.; Kranz, L. M.; Vormehr, M.; Baum, A.; Pascal, K.; Quandt, J.; Maurus, D.; Brachtendorf, S.; Lorks, V.; Sikorski, J.; Hilker, R.; Becker, D.; Eller, A. K.; Grutzner, J.; Boesler, C.; Rosenbaum, C.; Kuhnle, M. C.; Luxemburger, U.; Kemmer-Bruck, A.; Langer, D.; Bexon, M.; Bolte, S.; Kariko, K.; Palanche, T.; Fischer, B.; Schultz, A.; Shi, P. Y.; Fontes-Garfias, C.; Perez, J. L.; Swanson, K. A.; Loschko, J.; Scully, I. L.; Cutler, M.; Kalina, W.; Kyratsous, C. A.; Cooper, D.; Dormitzer, P. R.; Jansen, K. U.; Tureci, O. COVID-19 vaccine BNT162b1 elicits human antibody and TH1 T cell responses. *Nature* **2020**, *586*, 594−599.

(109) Mulligan, M. J.; Lyke, K. E.; Kitchin, N.; Absalon, J.; Gurtman, A.; Lockhart, S.; Neuzil, K.; Raabe, V.; Bailey, R.; Swanson, K. A.; Li, P.; Koury, K.; Kalina, W.; Cooper, D.; Fontes-Garfias, C.; Shi, P. Y.; Tureci, O.; Tompkins, K. R.; Walsh, E. E.; Frenck, R.; Falsey, A. R.; Dormitzer, P. R.; Gruber, W. C.; Sahin, U.; Jansen, K. U. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. *Nature* **2020**, *586*, 589−593.

(110) Corbett, K. S.; Edwards, D. K.; Leist, S. R.; Abiona, O. M.; Boyoglu-Barnum, S.; Gillespie, R. A.; Himansu, S.; Schafer, A.; Ziwawo, C. T.; DiPiazza, A. T.; Dinnon, K. H.; Elbashir, S. M.; Shaw, C. A.; Woods, A.; Fritch, E. J.; Martinez, D. R.; Bock, K. W.; Minai, M.; Nagata, B. M.; Hutchinson, G. B.; Wu, K.; Henry, C.; Bahl, K.; Garcia-Dominguez, D.; Ma, L.; Renzi, I.; Kong, W. P.; Schmidt, S. D.; Wang, L.; Zhang, Y.; Phung, E.; Chang, L. A.; Loomis, R. J.; Altaras, N. E.; Narayanan, E.; Metkar, M.; Presnyak, V.; Liu, C.; Louder, M. K.; Shi, W.; Leung, K.; Yang, E. S.; West, A.; Gully, K. L.; Stevens, L. J.; Wang, N.; Wrapp, D.; Doria-Rose, N. A.; Stewart-Jones, G.; Bennett, H.; Alvarado, G. S.; Nason, M. C.; Ruckwardt, T. J.; McLellan, J. S.; Denison, M. R.; Chappell, J. D.; Moore, I. N.; Morabito, K. M.; Mascola, J. R.; Baric, R. S.; Carfi, A.; Graham, B. S. SARS-CoV-2 mRNA vaccine design enabled by prototype pathogen preparedness. *Nature* **2020**, *586*, 567−571.

(111) Mercado, N. B.; Zahn, R.; Wegmann, F.; Loos, C.; Chandrashekar, A.; Yu, J.; Liu, J.; Peter, L.; McMahan, K.; Tostanoski, L. H.; He, X.; Martinez, D. R.; Rutten, L.; Bos, R.; van Manen, D.; Vellinga, J.; Custers, J.; Langedijk, J. P.; Kwaks, T.; Bakkers, M. J. G.; Zuijdgeest, D.; Rosendahl Huber, S. K.; Atyeo, C.; Fischinger, S.; Burke, J. S.; Feldman, J.; Hauser, B. M.; Caradonna, T. M.; Bondzie, E. A.; Dagotto, G.; Gebre, M. S.; Hoffman, E.; Jacob-Dolan, C.; Kirilova, M.; Li, Z.; Lin, Z.; Mahrokhian, S. H.; Maxfield, L. F.; Nampanya, F.; Nityanandam, R.; Nkolola, J. P.; Patel, S.; Ventura, J. D.; Verrington, K.; Wan, H.; Pessaint, L.; Van Ry, A.; Blade, K.; Strasbaugh, A.; Cabus, M.; Brown, R.; Cook, A.; Zouantchangadou, S.; Teow, E.; Andersen, H.; Lewis, M. G.; Cai, Y.; Chen, B.; Schmidt, A. G.; Reeves,

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX

R. K.; Baric, R. S.; Lauffenburger, D. A.; Alter, G.; Stoffels, P.; Mammen, M.; Van Hoof, J.; Schuitemaker, H.; Barouch, D. H. Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. *Nature* **2020**, *586*, 583−588.

(112) van Doremalen, N.; Lambe, T.; Spencer, A.; Belij-Rammerstorfer, S.; Purushotham, J. N.; Port, J. R.; Avanzato, V. A.; Bushmaker, T.; Flaxman, A.; Ulaszewska, M.; Feldmann, F.; Allen, E. R.; Sharpe, H.; Schulz, J.; Holbrook, M.; Okumura, A.; Meade-White, K.; Perez-Perez, L.; Edwards, N. J.; Wright, D.; Bissett, C.; Gilbride, C.; Williamson, B. N.; Rosenke, R.; Long, D.; Ishwarbhai, A.; Kailath, R.; Rose, L.; Morris, S.; Powers, C.; Lovaglio, J.; Hanley, P. W.; Scott, D.; Saturday, G.; de Wit, E.; Gilbert, S. C.; Munster, V. J. ChAdOx1 nCoV-19 vaccine prevents SARS-CoV-2 pneumonia in rhesus macaques. *Nature* **2020**, *586*, 578−582.

(113) Yang, J.; Wang, W.; Chen, Z.; Lu, S.; Yang, F.; Bi, Z.; Bao, L.; Mo, F.; Li, X.; Huang, Y.; Hong, W.; Yang, Y.; Zhao, Y.; Ye, F.; Lin, S.; Deng, W.; Chen, H.; Lei, H.; Zhang, Z.; Luo, M.; Gao, H.; Zheng, Y.; Gong, Y.; Jiang, X.; Xu, Y.; Lv, Q.; Li, D.; Wang, M.; Li, F.; Wang, S.; Wang, G.; Yu, P.; Qu, Y.; Yang, L.; Deng, H.; Tong, A.; Li, J.; Wang, Z.; Yang, J.; Shen, G.; Zhao, Z.; Li, Y.; Luo, J.; Liu, H.; Yu, W.; Yang, M.; Xu, J.; Wang, J.; Li, H.; Wang, H.; Kuang, D.; Lin, P.; Hu, Z.; Guo, W.; Cheng, W.; He, Y.; Song, X.; Chen, C.; Xue, Z.; Yao, S.; Chen, L.; Ma, X.; Chen, S.; Gou, M.; Huang, W.; Wang, Y.; Fan, C.; Tian, Z.; Shi, M.; Wang, F. S.; Dai, L.; Wu, M.; Li, G.; Wang, G.; Peng, Y.; Qian, Z.; Huang, C.; Lau, J. Y.; Yang, Z.; Wei, Y.; Cen, X.; Peng, X.; Qin, C.; Zhang, K.; Lu, G.; Wei, X. A vaccine targeting the RBD of the S protein of SARS-CoV-2 induces protective immunity. *Nature* **2020**, *586*, 572−577.

(114) Donoghue, M.; Hsieh, F.; Baronas, E.; Godbout, K.; Gosselin, M.; Stagliano, N.; Donovan, M.; Woolf, B.; Robison, K.; Jeyaseelan, R.; et al. A novel angiotensin-converting enzyme−related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1−9. *Circ. Res.* **2000**, *87*, e1−e9.

(115) Vickers, C.; Hales, P.; Kaushik, V.; Dick, L.; Gavin, J.; Tang, J.; Godbout, K.; Parsons, T.; Baronas, E.; Hsieh, F.; et al. Hydrolysis of biological peptides by human angiotensin-converting enzyme-related carboxypeptidase. *J. Biol. Chem.* **2002**, *277*, 14838−14843.

(116) Wang, W.; McKinnie, S. M.; Farhan, M.; Paul, M.; McDonald, T.; McLean, B.; Llorens-Cortes, C.; Hazra, S.; Murray, A. G.; Vederas, J. C.; et al. Angiotensin-converting enzyme 2 metabolizes and partially inactivates pyr-apelin-13 and apelin-17: physiological effects in the cardiovascular system. *Hypertension* **2016**, *68*, 365−377.

(117) Lek, M.; Karczewski, K. J.; Minikel, E. V.; Samocha, K. E.; Banks, E.; Fennell, T.; O'Donnell-Luria, A. H.; Ware, J. S.; Hill, A. J.; Cummings, B. B.; et al. Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **2016**, *536*, 285−291.

(118) Tukiainen, T.; Villani, A. C.; Yen, A.; Rivas, M. A.; Marshall, J. L.; Satija, R.; Aguirre, M.; Gauthier, L.; Fleharty, M.; Kirby, A.; et al. Landscape of X chromosome inactivation across human tissues. *Nature* **2017**, *550*, 244−248.

(119) Li, R.; Qiao, S.; Zhang, G. Analysis of angiotensin-converting enzyme 2 (ACE2) from different species sheds some light on cross-species receptor usage of a novel coronavirus 2019-nCoV. *J. Infect.* **2020**, *80*, 469−496.

(120) Luan, J.; Lu, Y.; Jin, X.; Zhang, L. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem. Biophys. Res. Commun.* **2020**, *526*, 165−169.

(121) Dehouck, Y.; Kwasigroch, J. M.; Rooman, M.; Gilis, D. BeAtMuSiC: Prediction of changes in protein-protein binding affinity on mutations. *Nucleic Acids Res.* **2013**, *41*, W333−339.

(122) Pucci, F.; Bernaerts, K. V.; Kwasigroch, J. M.; Rooman, M. Quantification of biases in predictions of protein stability changes upon mutations. *Bioinformatics* **2018**, *34*, 3659−3665.

(123) Adzhubei, I. A.; Schmidt, S.; Peshkin, L.; Ramensky, V. E.; Gerasimova, A.; Bork, P.; Kondrashov, A. S.; Sunyaev, S. R. A method and server for predicting damaging missense mutations. *Nat. Methods* **2010**, *7*, 248−249.

(124) Reed, L. J.; Muench, H. A Simple Method of Estimating Fifty Per Cent Endpoints. *Am. J. Epidemiol.* **1938**, *27*, 493−497.

(125) Kristiansen, M.; Graversen, J. H.; Jacobsen, C.; Sonne, O.; Hoffman, H.-J.; Law, S. A.; Moestrup, S. K. Identification of the haemoglobin scavenger receptor. *Nature* **2001**, *409*, 198−201.

(126) Blanpain, C.; Lee, B.; Tackoen, M.; Puffer, B.; Boom, A.; Libert, F.; Sharron, M.; Wittamer, V.; Vassart, G.; Doms, R. W.; et al. Multiple nonfunctional alleles of CCR5 are frequent in various human populations. *Blood* **2000**, *96*, 1638−1645.

(127) Heurich, A.; Hofmann-Winkler, H.; Gierer, S.; Liepold, T.; Jahn, O.; Pöhlmann, S. TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *Journal of virology* **2014**, *88*, 1293−1307.

(128) Konagurthu, A. S.; Whisstock, J. C.; Stuckey, P. J.; Lesk, A. M. MUSTANG: a multiple structural alignment algorithm. *Proteins: Struct., Funct., Genet.* **2006**, *64*, 559−574.

(129) Zhang, Z.; Ohto, U.; Shibata, T.; Krayukhina, E.; Taoka, M.; Yamauchi, Y.; Tanji, H.; Isobe, T.; Uchiyama, S.; Miyake, K.; et al. Structural analysis reveals that Toll-like receptor 7 is a dual receptor for guanosine and single-stranded RNA. *Immunity* **2016**, *45*, 737−748.

(130) Guerois, R.; Nielsen, J. E.; Serrano, L. Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations. *J. Mol. Biol.* **2002**, *320*, 369−387.

(131) Souyris, M.; Cenac, C.; Azar, P.; Daviaud, D.; Canivet, A.; Grunenwald, S.; Pienkowski, C.; Chaumeil, J.; Mejía, J. E.; Guéry, J.-C. TLR7 escapes X chromosome inactivation in immune cells. *Science Immunology* **2018**, *3*, eaap8855.

(132) Pieper, U.; Webb, B. M.; Dong, G. Q.; Schneidman-Duhovny, D.; Fan, H.; Kim, S. J.; Khuri, N.; Spill, Y. G.; Weinkam, P.; Hammel, M.; et al. ModBase, a database of annotated comparative protein structure models and associated resources. *Nucleic Acids Res.* **2014**, *42*, D336−D346.

(133) Toshchakov, V. Y.; Neuwald, A. F. A survey of TIR domain sequence and structure divergence. *Immunogenetics* **2020**, *72*, 181−203.

(134) Nguyen, A.; David, J. K.; Maden, S. K.; Wood, M. A.; Weeder, B. R.; Nellore, A.; Thompson, R. F. Human Leukocyte Antigen Susceptibility Map for Severe Acute Respiratory Syndrome Coronavirus 2. *J. Virol.* **2020**, *94*, e00510-20.

(135) La Porta, C. A. M.; Zapperi, S. Estimating the Binding of Sars-CoV-2 Peptides to HLA Class I in Human Subpopulations Using Artificial Neural Networks. *Cell Syst* **2020**, *11*, 412−417.

(136) Zeng, H.; Gifford, D. K. Quantification of Uncertainty in Peptide-MHC Binding Prediction Improves High-Affinity Peptide Selection for Therapeutic Design. *Cell Syst* **2019**, *9*, 159−166.

(137) Hulpke, S.; Tampé, R. The MHC I loading complex: a multitasking machinery in adaptive immunity. *Trends Biochem. Sci.* **2013**, *38*, 412−420.

(138) Blees, A.; Januliene, D.; Hofmann, T.; Koller, N.; Schmidt, C.; Trowitzsch, S.; Moeller, A.; Tampé, R. Structure of the human MHC-I peptide-loading complex. *Nature* **2017**, *551*, 525−528.

(139) Katsanis, S. H.; Katsanis, N. Molecular genetic testing and the future of clinical genomics. *Nat. Rev. Genet.* **2013**, *14*, 415−426.

(140) Gabriel, S.; Ziaugra, L.; Tabbaa, D. SNP genotyping using the Sequenom MassARRAY iPLEX platform. *Current protocols in human genetics* **2009**, *60*, 2.12.11−12.12.18.

(141) Qiu, F.; Tang, R.; Zuo, X.; Shi, X.; Wei, Y.; Zheng, X.; Dai, Y.; Gong, Y.; Wang, L.; Xu, P.; et al. A genome-wide association study identifies six novel risk loci for primary biliary cholangitis. *Nat. Commun.* **2017**, *8*, 1−8.

(142) Refae, S.; Gal, J.; Ebran, N.; Otto, J.; Borchiellini, D.; Peyrade, F.; Chamorey, E.; Brest, P.; Milano, G.; Saada-Bouzid, E. Germinal Immunogenetics predict treatment outcome for PD-1/PD-L1 checkpoint inhibitors. *Invest. New Drugs* **2020**, *38*, 160−171.

https://dx.doi.org/10.1021/acs.jproteome.0c00637
J. Proteome Res. XXXX, XXX, XXX−XXX