

Computational, Experimental, and Clinical Evidence of a Specific but Peculiar Evolutionary Nature of (COVID-19) SARS-CoV-2

Gerard Kian-Meng Goh,* A. Keith Dunker, James A. Foster, and Vladimir N. Uversky

Cite This: <https://doi.org/10.1021/acs.jproteome.2c00001>

Read Online

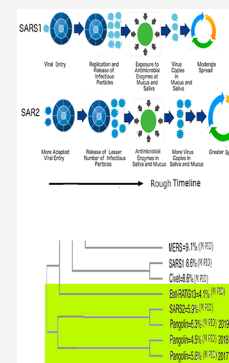
ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: The shell disorder models have predicted that SARS-CoV-2 is of a specific but peculiar evolutionary nature. All coronaviruses (CoVs) closely related to SARS-CoV-2 have been found to have the hardest outer shells (M protein) among CoVs. This hard shell (low M percentage of intrinsic disorder (PID)) is associated with burrowing animals, for example, pangolins, and is believed to be responsible for the high contagiousness of SARS-CoV-2 because it will be more resistant to antimicrobial enzymes found in saliva/mucus. Incoming clinical and experimental data do support this along with a prediction based on another aspect of the shell (N, inner shell) disorder models that SARS-CoV-1 is more virulent than SARS-CoV-2 because SARS-CoV-2 produces fewer virus copies in vital organs even if large amounts of infections particles are shed orally and nasally. A phylogenetic study using M reveals a closer relationship of SARS-CoV to pangolin-CoVs than the bat-RaTG13 found in Yunnan, China. Previous studies may have been confused by recombinations that were poorly handled. The shell disorder models suggest that a pangolin-CoV strain may have entered the human population in 2017 or before as an attenuated virus, which could explain why SARS-CoV is found to be highly adapted to humans.

KEYWORDS: pangolin, intrinsic, disorder, protein, nucleocapsid, virulence, shell, COVID, coronavirus, vaccine, immune, antibody, shell, nucleoprotein, membrane, matrix, attenuate, severe acute respiratory, omicron



INTRODUCTION: THE LINGERING MYSTERIES OF SARS-CoV-2

Some Characteristics of COVID-19 and SARS-CoV-2

The first observed coronavirus disease 2019 (COVID-19) outbreak occurred in December 2019 in Wuhan, China.¹ Because the initial cases were mainly associated with a Wuhan wet market, Huanan Seafood Wholesale Market, where live animals were sold, the possibility of a zoonotic transmission occurring there was suspected.² The virus, labeled severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), was sequenced and was found to be a betacoronavirus that has ~80% homology to the 2003 SARS-CoV (SARS-CoV-1). A search in the sample archives yielded a 2013 bat sample from Yunnan province, RaTG13, that had a 96% sequence similarity to SARS-CoV-2.^{3–5} Furthermore, CoV samples were retrieved from smuggled pangolins in Guangxi and Guangdong from 2017 to 2018 and in 2019, respectively, and they were found to have ~90% homology to SARS-CoV-2.^{6–9}

Natural or Unnatural Origin?

Right from the beginning, there were questions. Where did this virus come from? When and how did it enter humans? Because the Huanan Seafood Wholesale Market market is just a few miles from the Wuhan Institute of Virology, was the virus manufactured?¹⁰ To address some of the concerns, Andersen et al.¹¹ examined the S protein, which plays an important role in the viral entry into the cell by binding to the host ACE-2

(angiotensin converting enzyme 2) receptor. The group attempted to “design” the virus using computer software and found that the actual S protein binds using a different set of residues from the computer-suggested residues, which implies that the virus is simply too novel to be designed in the laboratory. The team mentioned that it is possible that the virus has been in humans for a long period of time. Other researchers have found evidence to suggest this as well. For example, computational studies have found that the SARS-CoV-2 S protein binds to ACE-2 with an affinity that is 20–30 times that of SARS-CoV-1 to ACE-2.^{11,12} Another study showed that the S protein has a binding affinity for the human ACE-2 that is greater than those of all of the other animal ACE-2 studied.¹³

These ACE-2/S protein studies^{11–13} suggest that SARS-CoV-2 could have been in humans for a long time. This seems to contradict what is popularly believed about the origin of COVID-19: SARS-CoV-2 entered humans in December 2019. Such an apparent contradiction is fodder for conspiracy theories that include the manufactured origin of the virus. This begs the question: If SARS-CoV-2 entered humans some time ago, why

Received: January 2, 2022

Table 1. Shell Disorder Models^a

shell disorder model	year published	description
parent viral shapeshifter model	2008 ²⁴	HIV is among the few viruses found with abnormally high outer shell disorder associated with the absence of effective vaccines and sexual transmission. The link between modes of transmission and shell disorder was first observed using this model.
CoV transmission shell disorder model	2012 ¹⁸	The level of shell disorder was found to be associated with levels of fecal–oral and respiratory transmission potentials. An abnormally hard SARS-CoV-2 outer shell (low M PID) was discovered during the COVID-19 pandemic. ²¹
virulence inner shell disorder model	2015 ²⁹	High disorder in the inner shell proteins was found to be associated with virulence of viruses such as Ebola virus, Nipah virus, and dengue virus. High outer shell disorder was found to be associated with higher morbidity, for example, Zika. ^{27–30}

^aThree closely related models were built using empirically based AI tools.

did the medical community not notice it then, instead only noticing it recently, in December 2019? Could a precursor have entered the human population as an attenuated strain before mutating to its current virulent form? In fact, our shell disorder models suggest that a pangolin-CoV strain did enter humans as an attenuated SARS-CoV-2 precursor in 2017 or before. Related to the virus' high adaptation to humans is the question of its contagiousness. Why is it so contagious that it exceeds SARS-CoV-1?² Some scientists believe that the S protein is responsible.¹⁴ While not necessarily contradicting this theme, we found peculiarities of SARS-CoV-2 that are likely to contribute to its infectivity.

Shell Disorder Models: Clinical and Experimental Reproducibility

When COVID-19 took the world by surprise, many scientists, including those in our group, had to resort to the use of computational tools and models to gain a quick understanding of the new enemy, given the limited information available then.^{15–22} Our group had previously published a series of papers based on closely related models (see Table 1) that have been able to provide insights into the evolution and nature of a variety of viruses, including CoVs, long before the onset of the COVID-19 pandemic.^{15–30} These closely related models were coined the “shell disorder models” because they involve the analysis and detection of disorder of the various shells of viruses to understand their evolutions using empirically based artificial intelligence (AI) tools.

Upon the application of our tools using the protein and genetic information pertaining to COVID-19, our shell disorder models came out with results that have a wide number of implications pertaining to the virus's transmissibility, virulence, and specific but peculiar evolutionary nature. The results were previously published during the pandemic. The only problem with our previously published COVID-19 research was that we had to rely solely on computational models, even if the main AI tool involved was empirically based.²³ The results needed to be reproduced experimentally and clinically. Just as laboratories throughout the world rushed to study the then mysterious enemy, much data have become available, and we now have sufficient data from independent sources to validate or contradict our models related to COVID-19. This Review will show that incoming experimental and clinical data do strongly reproduce the results of the shell disorder model, even to the finer details. A comparative study of the computational, experimental, and clinical research will be done to further demonstrate the consistencies of the evidence of the case.

It should also be noted that because there is no mention of the shell disorder models in the experimental and clinical papers involved, the scientists involved were apparently oblivious of our research.

More specifically, the shell disorder models have made a series of discoveries and predictions. These include the discovery that the hardest outer shell can only be found in SARS-CoV-2 and its closest relatives, not in SARS-CoV-1 and most of the other coronaviruses known, with the exception of those associated with burrowing animals, such as rabbits. We believe that this particular characteristic is responsible for the greater contagiousness, despite the lower virulence of SARS-CoV-2 when compared with SARS-CoV-1. Another prediction involves SARS-CoV-2 producing large amounts of particles orally and nasally but a smaller number of infectious particles in vital organs. These are just some examples of the many predictions made by the shell disorder models. Because all of these predictions were yet to be validated through experimental and clinical studies when our previous papers^{15–22} were published, recent research from various unrelated laboratories located throughout the world, for example, Germany, Australia, Holland, and the United States, will be reviewed later in this Review in a comparative manner as supporting evidence of reproducibility.

Parent Shell Disorder Model Foresaw the Successful Development of COVID-19 Vaccines

The parent shell disorder model that studied the intrinsic disorder of the various protein shells of viruses was initiated in 2005, when it was discovered that human immunodeficiency virus (HIV) is likely to have an abnormally disordered outer shell that is believed to be responsible for the failures in the search for an effective HIV vaccine.^{15,23,24} Conversely, the outer shell disorder of SARS-CoV-2 was found to not be similar to that of HIV but rather resembled that of classical viruses, for which vaccines have been developed.^{17,18} This led us to correctly foresee the successful development of effective vaccines for COVID-19 months before their approval by the U.S. Food and Drug Administration (FDA).^{16,17} The model involves the use of AI to detect and calculate the levels of intrinsic disorder in query proteins. In 2014, there was a spinoff of this model that detected positive correlations between inner shell disorder and virulence among a variety of viruses, such as Ebola (EBOV) and Nipah (NiV) viruses.^{11,12} This allows the shell disorder model to detect attenuated strains.

The parent shell disorder model^{23–25} is mentioned in this Review for many reasons. First, it was the parent model from which the other shell disorder models came from. It is important for the readers to understand the background pertaining to the origin of the shell disorder models. Second, even though the parent shell disorder model has been comparatively applied to a wide variety of viruses, especially HIV, viruses have proteins that are similar in structure, function, and composition,^{15–30} even if they are not related or closely related. Furthermore, they often share the same evolutionary background.^{15–30} Techniques used to study one or several viruses are often applicable to a different

virus. As already mentioned, this is the case when the parent model was used to predict the feasibility of COVID-19 vaccine development.^{16,17}

Shell Disorder Models: Contagiousness, Virulence, and Immunity

In 2011, before the MERS-CoV outbreak, there was an initiation of yet another spinoff model, the CoV shell disorder model,¹⁸ to predict the levels of fecal–oral and respiratory transmission of the various CoVs depending on the levels of disorder at mainly the inner shell (N). The model predicted SARS-CoV-1 to be of intermediate fecal–oral and respiratory transmission potential. Upon the various outbreaks, the model was further validated with MERS-CoV being placed in a category of CoVs with lower respiratory but higher fecal–oral transmission potential,¹⁹ whereas SARS-CoV-2, like SARS-CoV-1, has both intermediate fecal–oral and respiratory transmission potential.^{20,21} However, something else has been observed in SARS-CoV-2 but not SARS-CoV-1: SARS-CoV-2 has one of the hardest outer shells (low M disorder) among CoVs.^{20,21} This provides the initial evidence of the resilience of the virus, especially in body fluids, where it is exposed to harmful enzymes.^{1–32} As a result, heavy viral shedding occurs.

Computational, Experimental, and Clinical Evidence Points to a Specific but Peculiar Evolution of SARS-CoV-2 That Has Important Clinical and Epidemiological Implications

The shell disorder models do not stop there. A later search reveals that the hardness of the viral shell is associated with burrowing animals, such as rabbits and pangolins that often come into contact with buried feces.²² A more careful study of pangolin-CoV using the shell disorder models, including the virulence inner shell model, provided evidence that the 2017 strain is attenuated, as we will see later. Therefore, the shell disorder models not only support the natural evolution of SARS-CoV-2 but also point to a peculiar but specific evolution that has important clinical and epidemiological implications including immunity, virulence, and infectivity. The evidence put forth thus far was computational and empirically based and involved the use of AI. In this Review, however, we will also point to published experimental and clinical results that support important aspects of the shell disorder models pertaining to SARS-CoV-2, extending even to the finer details of the models' predictions. The results point to a specific but peculiar natural evolution of SARS-CoV-2. The WHO mission report¹⁰ underscores the importance of studies such as this in the attempt to uncover the origin of the virus.

■ VIRAL SHELL DISORDER MODELS

Protein Intrinsic Disorder and AI Tools

An important concept that was used in the mentioned shell disorder models is protein intrinsic disorder, which can be defined as entire proteins or portions of a protein that have no unique 3D structure.^{33–36} It has long been known that disordered proteins have important functions.^{33–41} Consequently, many tools used to predict disorder have been built. One of these is PONDR VLXT (<http://www.pondr.com>), which is known to be among the most sensitive in the detection of protein–protein/RNA/DNA/sugar/lipid interactions^{42–46} and has been successful in the study of the structural proteins of a large number of viruses, including influenza A virus, rabies virus, HIV, smallpox virus, herpes simplex virus (HSV), hepatitis C virus (HCV), and yellow fever virus (YFV).^{15–30,47,48}

PONDR VLXT is a neural network that is trained using sequences of known ordered and disordered proteins.^{49–51} The predictor is fed with the sequence of a protein, and the output is the prediction of order or disorder predisposition for each residue. An important measure that is frequently used in disorder status studies is the PID (percentage of intrinsic disorder), which is defined as the number of residues predicted to be disordered divided by the total number of residues and multiplied by 100. The PID provides a gauge for the level of disorder in a protein chain.

The sequences are available at UniProt (<https://www.uniprot.org>) and GenBank-NCBI (<https://www.ncbi.nlm.nih.gov/protein>). The sequences were entered into PONDR VLXT (<http://www.pondr.com>). Both sequences and PONDR VLXT results were automatically added to a MySQL server using a JAVA program.¹⁶ Measurements of sequence homology were obtained using BLASTP (<https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins>). Phylogenetic trees were constructed using EMBI-EBI Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>). They can also be constructed using a similar algorithm but with a bootstrapping option at T-REX (<http://www.trex.uqam.ca/index.php?action=newick&project=trex>). Figures were drawn using GIMP (<https://www.gimp.org/>) and OpenOffice Draw (<https://www.openoffice.org/download/>). Calculations necessary for the correlation, regression, and multivariate analyses were conducted using the R statistical package.⁵²

Viral Shell Disorder Models: Three Closely Related Projects

There are three closely related and highly interlinked shell-disorder-based models, as seen in Table 1. One involves the parent project, that is, HIV vaccine mystery/viral shapeshifter. The others are spinoffs (the CoV transmission shell disorder model and the virulence inner shell disorder model). What is common among them is the study of viral protein shell disorder. Each model has related but different implications.

Parent Project: The Viral Shapeshifters and the HIV Vaccine Mystery

In 2005, during the compilation of a database of viral shell proteins, something unusual was observed. The outer shell protein matrix of many HIV-1 strains was found to be abnormally disordered, which had not been seen in any of the other viruses before then. The results were first published in *Virology Journal* in 2008,^{23,24} and this unusual characteristic can be traced to the absence of an effective HIV vaccine despite a search that had spanned ~40 years.^{15,25,26} As more data became available, two other viruses, HSV and HCV, were added to the list of viruses with highly disordered outer shells, whereas virtually all other viruses (and especially those with effective vaccines available) have low or very low disorder levels of their outer shells.

Figure 1 illustrates the basic virion physiology with HIV and CoV as examples. We are able to see the differences and similarities between them in terms of the shell proteins. HIV has three shells, the matrix, capsid, and the nucleocapsid, whereas CoV has two shells, the membrane (M) and the nucleocapsid (N). Similar virion physiology can be found in other viruses, even if they have a single shell layer or multiple shell layers, as in the cases of poliovirus and smallpox (variola) virus, respectively.^{15,53}

As a shell disorder database was being built, a common characteristic of viruses that emerged is the tendency of their outer shell to be more highly ordered, that is, to have a low

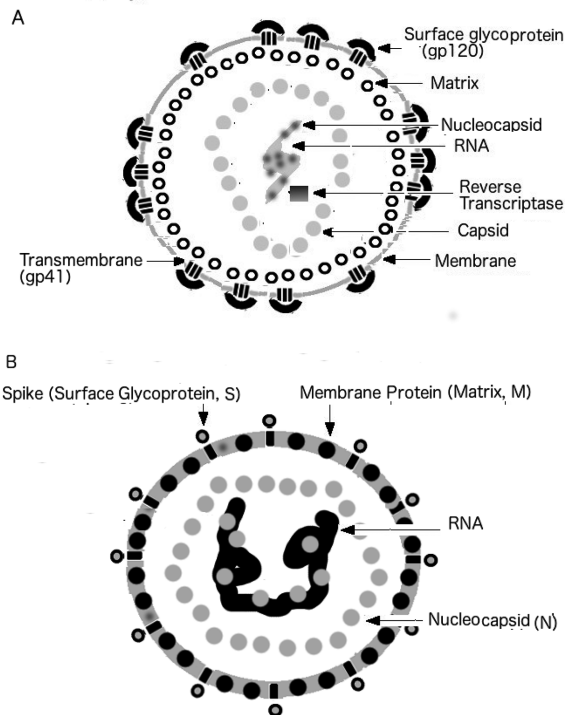


Figure 1. Comparative virion physiology. (A) HIV virion has three main shells: matrix (outer), capsid (intermediate), and nucleocapsid (inner shell). (B) Coronavirus (CoV). Reprinted with permission from ref 15. Copyright 2017 Gerard K. M. Goh.

PID.^{23,24} This can be attributed to the need of the virus to protect itself from environmental damage, as the outermost shell plays the greatest role in protecting the virion (as seen in Figure 2C). When the shells of HIV-1 were inspected, something abnormal was observed. Many of the HIV-1 strains have high disorder in the matrix (outer shell), which is not seen in virtually any virus.^{15,23,24} As the database of viral shell disorder grew to encompass a much larger variety of viruses, two other viruses were seen as having similar characteristics as HIV,^{15–17,25,26} namely, HCV and HSV. More interestingly, HCV and HSV have a strong association with sexual transmission, and effective vaccines have never been found for these viruses either. These are also characteristics of HIV. On the contrary, there are drastically lower disorder levels in the outer shells of classical viruses such as YFV, poliovirus, and smallpox virus, for which effective vaccines have been found (as seen in Figure 2B).

The mechanism of the “viral shapeshifting” immune evasion can be seen in Figure 2A. The highly disordered matrix (outer shell) allows for motions that increase the motions of the viral surface glycoprotein. The high level of motions at the surface therefore prevents strong binding of antibodies to the virus. The association of a soft disordered outer shell with sexual transmission is related to the fact that sexual transmission does not require the virus to remain in a harsh nonphysiological environment for a long time; therefore, a more protective outer shell is usually not necessary. As seen in Figure 2, the outer shell (M) PID looks nothing like those of HIV, HSV, or HCV, for which no effective vaccine has ever been found. We first made the accurate prediction of the feasibility of a COVID-19 vaccine in a paper published in October 2020, months before the first COVID-19 vaccine was approved by the U.S. FDA.^{16,17,34}

CoV Transmission Shell Disorder Model

A spinoff from the above-mentioned parent project came in 2011,¹⁸ which took place before the 2012 MERS-CoV outbreak. The N and M CoV PIDs were collected as an extension of our shell disorder database. N and M proteins are the inner and outer proteins of CoVs, as seen in Figure 1B. Because human coronaviruses (HCoVs) were previously only associated with mild colds, prior to the 2003 SARS outbreak, medical research of CoVs was sorely neglected, but knowledge of animal CoVs was plentiful due to their constant threat to the farming industry.^{15,19,20,55} We applied knowledge of animal CoVs, in particular, those of porcine CoVs, to our database of CoV shell disorder and discovered that CoVs clustered easily into three groups based mainly on the N PID.^{18,19} As shown in Table 2, Group A is made up of CoVs with lower fecal–oral but higher respiratory transmission, Group B consists of CoVs with intermediate fecal–oral and respiratory transmission potentials, and Group C includes viruses with lower levels of respiratory transmission and higher levels of fecal–oral transmission.

When the model was first built, it was determined that on the basis of its N PID (50%), 2003 SARS-CoV-1 fell into Group B, which includes viruses that have intermediate fecal–oral and respiratory transmission potentials. The model was first published in 2012.¹⁸ When the MERS-CoV outbreak came shortly after its publication, it became an opportunity to validate the model. MERS-CoV was assigned to Group C given an N PID of 43%.¹⁹ The fact that MERS-CoV is categorized as having higher fecal but lower respiratory transmission potential is consistent with what we now know about the virus.¹⁰ The virus is not easily spread among humans but spreads easily among camels, which are farmed in the Middle East.⁵⁶

Yet another chance to validate the model came with the COVID-19 outbreak. This time SARS-CoV-2 had to be placed in Group B, alongside SARS-CoV-1. There is, however, something irregular about this virus. When the shell disorder analysis was first performed on SARS-CoV-2, the virus was predicted to have one of the hardest outer shells (i.e., lowest M PID) within its CoV family. Because the outer shells of viruses play a vital role in protecting the virion from damage, suspicion was immediately cast on this factor as the culprit for the greater contagiousness of SARS-CoV-2. A harder shell basically means that the virus will be more resistant to the antimicrobial enzymes found in body fluids.^{20,31,32} A search for CoVs with similarly hard outer shells revealed that this hard shell is associated with CoVs of burrowing animals, such as pangolins and rabbits, even if the CoVs are not closely related to each other or to SARS-CoV-2, as in the case of rabbit-CoV.²² We will present both clinical and experimental evidence that SARS-CoV-2 has a hard shell and is shed in large quantities through the nose, nasal cavity, throat, and mouth.

Virulence Inner Shell Disorder Model

In 2014, yet another spinoff of the model was initiated. In this case, correlations between inner shell disorder and virulence were detected in a variety of viruses including flaviviruses, filoviruses, and the Nipah virus.^{27–30} Figure 3 shows some example of the correlations. Figure 3A represents the correlation between EBOV virulence and the inner shell (NP, nucleoprotein), whereas Figure 3B illustrates the same correlation for DENV CFRs and inner shell (C, capsid) disorder. A strong correlation has been found between EBOV virulence (CFR) and inner shell (NP) disorder with a correlation coefficient (r) of 0.92 ($p < 0.001$).²⁹ The DENV yielded a strong correlation of

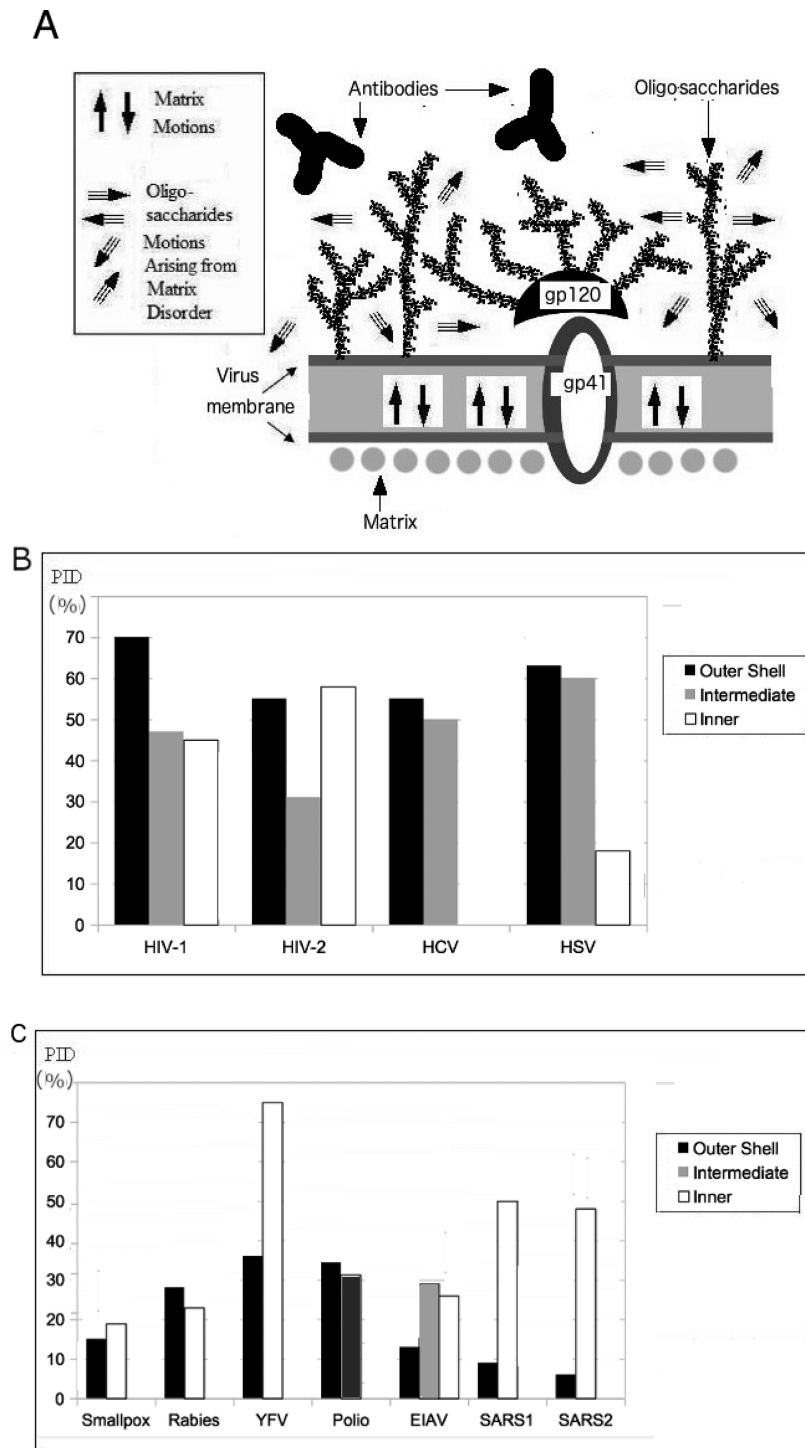


Figure 2. Parent shell disorder project: Viral shapeshifter. (A) Schematic diagram that illustrates the mechanism viral “shapeshifting” immune evasion in HIV. Reprinted with permission from ref 15. Copyright 2017 Gerard K. M. Goh. (B) Shell disorder (maximal PIDs) of viruses with no vaccines (HIV, HSV, and HCV). (C) Shell disorder of the classical viruses that have effective vaccines available. The mechanism involves the movement of the surface glycoprotein arising from motions allowed within the disordered matrix (outer shell). The outer shell PIDs of HIV, HSV, and HCV are noticeably higher than those of the classical viruses with known vaccines. SARS-CoV-2 is placed in panel B for comparative purposes. The M PID of SARS-CoV-2 resembles the outer shell PIDs of the classical viruses compared with those of HSV, HCV, and HIV.

0.9 ($p < 0.001$).²⁷ It is interesting to note that only a modest correlation ($r = 0.6$, $p < 0.05$) can be obtained when the data encompass a large variety of flaviviruses (e.g., DENV, YFV, ZIKV) but the correlation jumps to 0.9 ($p < 0.001$) when the outer shell (M, membrane) disorder is included.²⁸

It would, however, be a mistake to assume that the outer membrane always has a correlation with virulence, as no correlation could be detected between DENV virulence and M disorder.

Depending on the virus, complex relationships are often found. In the case of EBOV, matrix and nucleocapsid disorder

Table 2. Categorization of Coronaviruses by Mainly N PID to Predict Levels of Respiratory and Fecal–Oral Transmission Potential ($p < 0.001$, $r = 0.9$)^a

Coronavirus	M PID	UniProt (U)/Genbank (G) Accession Code (M proteins) ^b	N PID	UniProt (U)/Genbank (G) Accession Code (N proteins) ^b	group/remarks
HCoV-229E	23	P15422 (U)	56	P15130 (U)	Group A; higher levels of respiratory transmission; lower levels of fecal–oral transmission
IBV (avian) ^d	10	P69606 (U)	56	Q8JMI6 (U)	
bovine	7.8	P69704 (U)	53.1	Q8V432 (U)	Group B; intermediate levels of respiratory and fecal–oral transmission
rabbit	5.7	H9AA37 (U)	52.2	H9AA59 (U)	
PEDV (porcine) ^d	8	P59771 (U)	51.7	Q07499 (U)	Group C; lower levels of respiratory transmission; higher levels of fecal–oral transmission
canine (resp.)	7	A3E2F6 (U)	50.5	A3E2F7 (U)	
HCoV-OC43	7	Q4VID2 (U)	51	P33469 (U)	
SARS-CoV-1	8.6	P59596 (U)	50.2	P59595 (U)	
HCoV-NL63	11	Q6Q1R9 (U)	49	Q6Q1R8 (U)	
SARS-CoV-2	5.9	P0DTC5 (U)	48.2	P0DTC9 (U)	
bats ^c	11.2 ± 5.3	A3EXD6 (U)	47.7 ± 0.9	Q3LZX4 (U)	
MHV (murine) ^d	8	Q9JEB4 (U)	46.8	P03416 (U)	
pangolin ^e	5.6 ± 0.9	QIA428617 (G)	46.6 ± 1.6	QIA48630 (G)	
MERS-CoV	9.1	K0BU37 (U)	44.3	K0BVN3 (U)	
TGEV (porcine) ^d	14	P09175 (U)	42.41	P04134 (U)	
canine (ent.)	8	B8RIR2 (U)	40	Q04700 (U)	
HCoV-HKU1	4.5	Q14EA7 (U)	37.4	Q0ZME3 (U)	

^aA slightly stronger correlation is obtained when M is also included as an independent variable. ^bUniProt (U) (<https://www.uniprot.org>). GenBank-NCBI (G) (<https://www.ncbi.nlm.nih.gov/protein>). ^cSummary figures on bats. Further details of the bat samples can be found in Table 3. Four out of 5 bat-CoVs are in Group B. High standard deviations are seen for N/M PIDs, as denoted by “±”. ^dMHV (murine hepatitis virus), IBV (infectious bronchitis virus), PEDV (porcine epidemic diarrhea virus), TGEV (transmissible, gastroenteritis virus). MHV is placed in Group C for convenience, but it is at the borderline, which means that it could also fall into Group B or be a hybrid of both Groups B and C. ^eDetails of the pangolin samples can be found in Table 3. Three out of 4 pangolin-CoVs are in Group C. Standard deviation is denoted by “±”.

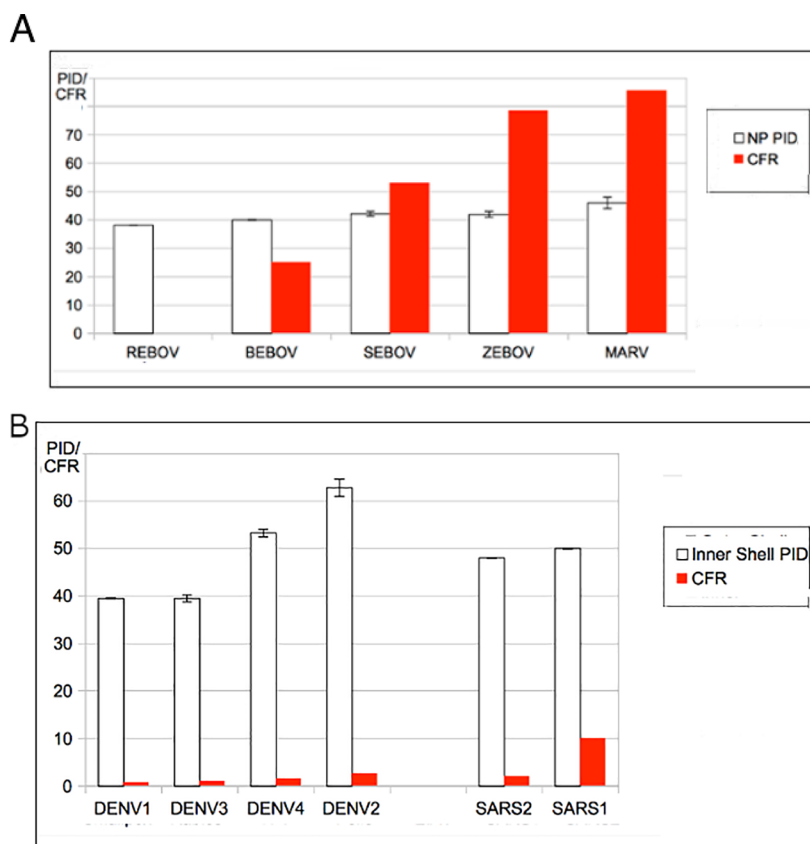


Figure 3. Virulence inner shell disorder model: “Trojan horse” immune evasion. (A) Correlation between filovirus/EBOV virulence (CFR) and NP PID (multivariate analysis: $p < 0.0001$, $r = 0.8$). (B) Relationships between virulence and inner shell disorder (DENV, SARS-CoV-1/2) (DENV: $r = 0.95$, $p < 0.001$).

have an inverse relationship ($r = -0.9$, $p < 0.05$) because the shells play compensating roles in protecting the virion.^{15,29} In the case of flaviviruses, the outer shell disorder often assists in the virulence by providing for better penetration into vital organs via greater efficiency in protein–protein/DNA/RNA/lipid interaction, as in the case of the highly pathogenic YFV. In the case of ZIKV, which has relatively high M disorder but low inner shell (C) disorder, the M disorder is manifested as greater fetal morbidity because this allows for greater penetration of the placenta.²⁸ SARS-CoV-1/2 does not penetrate the placenta easily, as revealed thus far by clinical studies,^{57,58} which is consistent with the fact that SARS-CoV-1/2 has a relatively hard outer shell (low M PIDs).

“Trojan Horse” Immune Evasion

The reason for the correlations between inner shell disorder and virulence has to do with an immune evasion strategy that is described as “Trojan horse”.^{15,16,27–30} This involves a strategy where the virus attempts to replicate rapidly before the host immune system notices the presence of the virus. In the process of doing so, however, the resulting high viral loads in vital organs often overwhelm the host, thereby killing it. It therefore often backfires on the host.¹⁵ Because the inner shell proteins play important roles in viral replication, intrinsic disorder enhances the efficiency of these proteins by allowing protein–protein/RNA/DNA/lipid bindings.^{33,34,36,42–46,59}

The inner shell proteins of viruses have similar, although not identical, functions in their roles in viral replication.^{53,60–63} They are commonly involved in the packaging of viral particles before their release.^{53,60} The coronavirus N (nucleocapsid) recruits RNA and other proteins to regions near the endoplasmic reticulum (ER) and Golgi apparatus, where they are assembled and packaged.⁶³ Similarly, the precursor of the inner shell, the C (capsid) protein, of the flavivirus migrates to the ER.⁵³ Upon doing so, it becomes embedded into the ER membrane, where it binds to viral mRNA and other viral proteins toward assembly. In the case of Ebola virus, its inner shell protein (NP) assists in building a tube-like structure, which is involved in the transportation of proteins that are assembled and budded as viral particles. As for the Nipah virus, measles virus, and other related viruses, the N binds to the P and L proteins to become the RNA polymerase, which is crucial for the replication of the RNA.⁶⁴ All of these require protein–protein/RNA/lipid interactions. Greater protein intrinsic disorder enhances the efficiency of such interactions by providing better molecular fitting.

DISORDER MODELS: A SPECIFIC EVOLUTIONARY NATURE OF SARS-CoV-2

Hard Outer Shell (Low M PID) and Burrowing Animals

We have already mentioned that SARS-CoV-2 differs from most of the other coronaviruses by having the hardest shell. Figure 4 shows that SARS-CoV-2 has one of the hardest outer shells (i.e., a lowest M PID) among a representative selection of CoVs. CoVs with similarly hard shells can only be found among CoVs of burrowing animals, such as pangolins and rabbits. Interestingly, rabbit-CoV is not closely related to SARS-CoV-2. It is therefore likely that the hard outer shells arose evolutionarily from the fact that burrowing animals are likelier to be in contact with fecal material that has been buried for months or even years. Interestingly, pangolins also have feeding habits that may enhance the chance of ingesting buried fecal material. Pangolins have strong arms that allow them to dig for

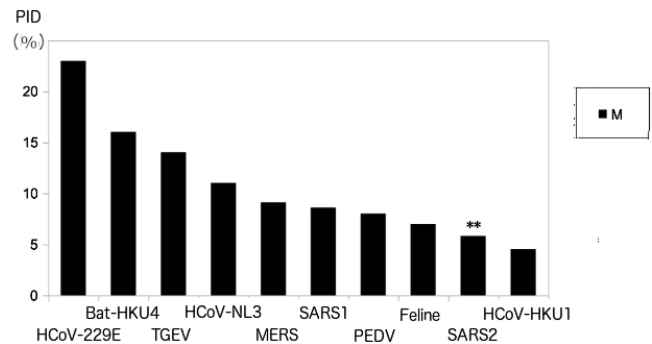


Figure 4. SARS-CoV-2 has one of the hardest outer shells (lowest M PID) among CoVs.

subterranean ants and termites. Whereas their sticky tongues allow termites and ants to be trapped as food before swallowing, the accidental ingestion of feces and soil is inevitable. The fecal–oral route is not only supported by the M PID data, as in Figure 4 and Table 2, but also supported by the CoV transmission model, as seen in Figure 5.

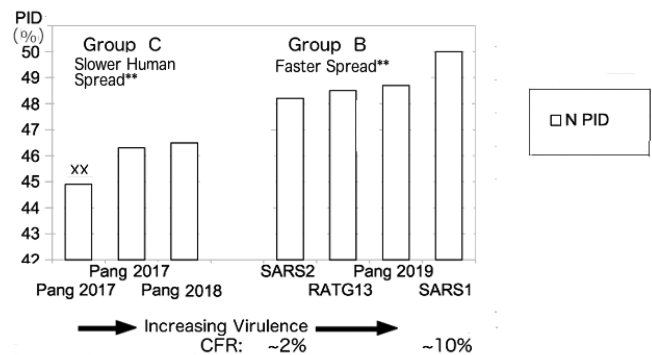


Figure 5. N PID of pangolin-CoV, bat-RaTG13, and SARS-CoV-1/2. Stepwise increment of pangolin-CoV N PIDs by the year. According to the virulence shell disorder model, the 2017 pangolin-CoV (marked “XX”) is the most attenuated of the pangolin samples found. It should also be noted that because pangolin-CoVs 2018 and 2019 fall close to the border of Groups B and C, they could be hybrids of the two groups. Nevertheless, this shows the high level of fecal–oral transmission potential of pangolin-CoVs. (***) The predicted faster or slower spread is extrapolated from N PID while assuming that the pangolin-CoV M PIDs are constant and not lower than the M PID of SARS-CoV-2.

An apparent contradiction can be found in the fact that rabbit-CoV is categorized as being in Group B (Table 2), whereas most pangolin-CoVs are in Group C (Table 3), even if both CoVs have among the hardest outer shells, which suggests contact with buried feces by burrowing animals, for example, rabbits and pangolins. The model does, however, detect differences in the N proteins of rabbits and pangolins. The higher N PID (52.2%) of rabbit-CoV is indicative of higher respiratory transmission potential, which points to the higher presence of fecal–respiratory transmission potential. This difference can also be observed in their feeding behaviors. Rabbits eat leaves of plants, whereas the ant meals of pangolins are often contaminated with feces and soil, as described above. Not only is there no inconsistency in the model as far as this is concerned but also it highlights the reproducibility of the CoV transmission shell disorder model.

Table 3. Grouping of Pangolin-CoVs and Bat-CoVs by Mainly N PID with SARS-CoV and SARS-CoV-2 as References

Coronavirus	sequence similarity M (%)	M PID (%)	accession: UniProt (U)/GenBank (G)	sequence similarity N (%)	N PID (%)	accession: UniProt (U)/GenBank (G)	group
SARS-CoV-2	100	5.9	P0DTC5 (U)	100	48.2	P0DTC9 (U)	B
SARS-CoV-1	90.5	8.6	P59596 (U)	90.5	50.2	P59595 (U)	B
civet-SARS-CoV	90.1	8.6	Q3ZTE9 (U)	90.01	49.1	Q3ZTE4 (U)	B
pangolin-CoV		5.6 ± 0.9 ^a			46.6 ± 1.6 ^a		
2019	98.2	6.3	QIG55948 (G)	98	48.7	QIG55953 (G)	B
2018	97.7	4.5	QIQ54051 (G)	93.8	46.3	QIQ54056 (G)	C
2017 ^b	98.2	5.9	QIA48617 (G)	94	44.9	QIA48630 (G)	C
				93.32	46.5	QIA48656 (G)	C
Bat-CoV		11.2 ± 15 ^a			47.7 ± 0.9 ^a		
RATG13	99.6	4.1	QHR63303 (G)	99.1	48.5	QHR63308 (G)	B
bat 512	35.5	15.3	Q0Q463 (U)	29.4	46.5	Q0Q462 (U)	C
HKU3	91	7.7	Q3LZX9 (U)	89.6	48	Q3LZX4 (U)	B
HKU4	42.7	16.4	A3EXA0 (U)	51.1	48.5	A3EXA1 (U)	B
HKU5	44.7	11.8	A3EXD6 (U)	47.9	47.1	A3EXD7 (U)	B

^aStandard deviation is denoted by “±”. ^bPossible vaccine strain for SARS-CoV-2 detected

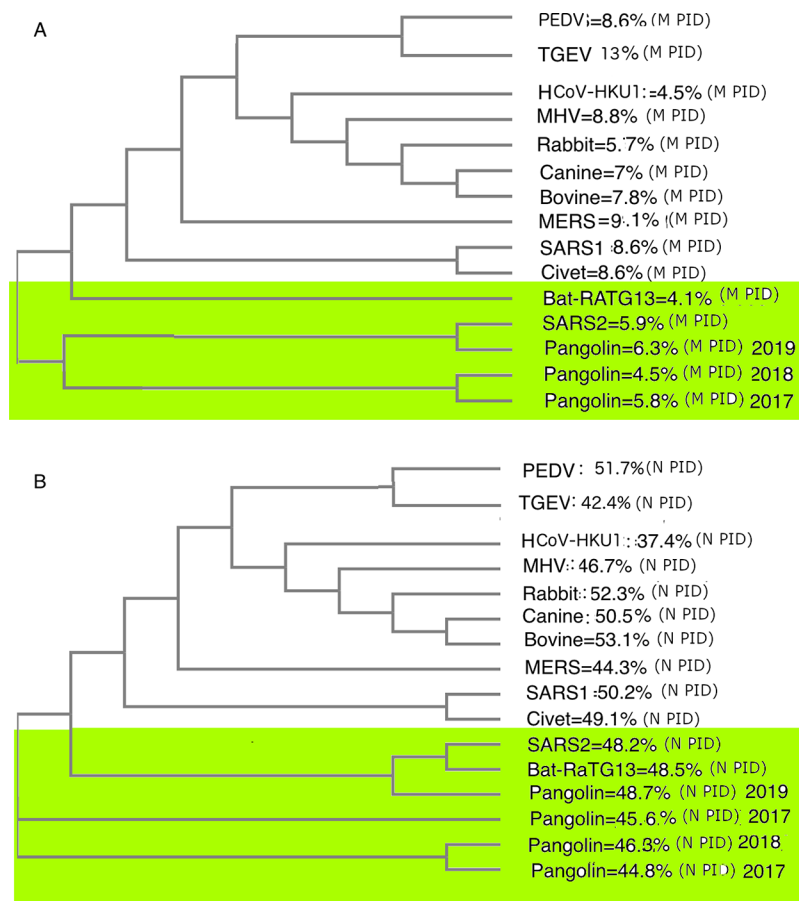


Figure 6. Phylogenetic tree based on M and N proteins: (A) M proteins and (B) N proteins. The regions shaded in green highlight important difference in the M phylogenetic tree not seen in other phylogenetic trees. Pangolin-CoVs are seen to be closer to SARS-CoV-2 than bat-RaTG13 in panel A.

Computational Evidence Suggests that SARS-CoV-2 Entered Humans via Pangolins in 2017 or Earlier as an Attenuated Strain

Table 3 and Figure 5 show that there are not just differences in the N PIDs of the various pangolin-CoVs but stepwise differences as the collection dates become farther from 2019, which suggests an interesting but peculiar evolution. The N PID of one of the 2017 pangolin-CoVs (marked “XX” in Figure 5) is

the lowest when compared with that of other pangolin-CoV samples. When we apply disorder analysis using CoV transmission models, we discover that most of the pangolin-CoV samples fall into C, the group that contains CoVs with higher fecal–oral transmission potential but lower respiratory transmission potential. When we apply the virulence inner shell disorder model, however, it reveals that most pangolin-CoVs, especially the 2017 pangolin-CoV, are likely attenuated versions

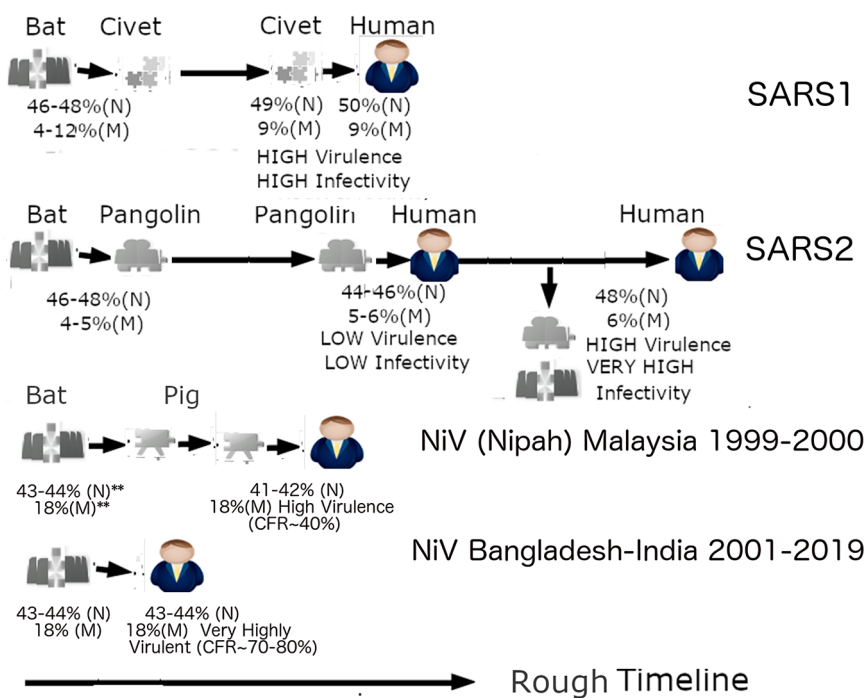


Figure 7. Zoonotic pathways toward human infections of SARS-CoV-1/2 and NiV. Qualitative timeline indicates that SARS-CoV-2 ancestral strains took a longer time to evolve in pangolins and humans before mutating to the current virulent form.

of SARS-CoV-2. An exception is the 2019 pangolin-CoV, which has similar N PID to SARS-CoV-2. Apparently, the virus could have been attenuated by the natural fecal–oral behaviors of pangolins.

Figure 5 shows that there is a relationship between virulence and N PID in SARS-CoV-1/2. The broad SARS-CoV-2 CFR range given is merely a reflection of estimates commonly found in the literature and is based on the current number of cases and deaths.^{65,66} More importantly, there is a consensus within the scientific community that SARS-CoV-1 was generally more lethal than SARS-CoV-2.^{66,67}

Phylogenetic Tree of the M Protein Provides Evidence of a Closer Relationship between SARS-CoV-2 and Pangolin-CoV

Phylogenetic studies have been done to examine the relationship between pangolin-CoV and SARS-CoV-2. Most of the studies available are based on the virus' genome-wide analysis. In fact, at least one study discounted the likelihood of pangolin-CoV, in particular, the 2019 pangolin-CoV, as being a direct ancestor of the SARS-CoV-2.⁷ Because it is genome-wide, the chance of recombinations occurring in at least one of the regions increases, which could lead to a mistake in the phylogenetic analysis.^{68,69} Most phylogenetic tools and algorithms currently available do not handle recombination well and are likely to give the wrong results in cases where recombination did occur. One solution is to constrain the phylogenetic analysis to a highly conserved proteins, such as M. The M protein is an ideal target, as seen in Table 3, which shows that pangolin-CoV M has a homology of 97 to 98% to SARS-CoV-2 compared with ~91% for genome-wide sequence similarity.⁹ Indeed, we will see that the result for M phylogenetic analysis is different from what we have seen in previous studies.

Because the M and N proteins are crucial to our research, further investigation revealed a specific pattern of evolution of pangolin-CoVs. Figure 6 shows the phylogenetic trees of N and

M proteins. Whereas much of the two phylogenetic trees resemble each other, important differences should be noticed, as shown in the regions shaded in green. Previous phylogenetic studies, especially those using the entire CoV genome, paint similar relationships among SARS-CoV-2, pangolin-CoV, and bat-RaTG13, as seen in Figure 6A, which uses the N protein. The phylogenetic tree based on M shows, however, that RaTG13 is not closer to SARS-CoV-2 than pangolin-CoVs are to SARS-CoV-2, despite the greater sequence similarity between RaTG13 and SARS-CoV-2. This is also in sharp contrast with the phylogenetic studies based on the entire CoV genome or proteins other than M, where RaTG13 is depicted as having the closest relationship to SARS-CoV-2.

There are important reasons for this discrepancy. The fact that CoVs closely related to SARS-CoV-2 all have the hardest outer shell (lowest M PID) in the family adds to the reason that the M is likely to be very conserved among the related CoVs. Because the M proteins are highly structured and conserved, only mutations, not recombinations, are likely to have occurred, and this provides for a more accurate phylogenetic snapshot of the ancestral tree of CoVs closely related to SARS-CoV-2. Further support of this can be found in Table 3. The sequence similarity of M proteins of pangolin-CoVs is in the range of 97 to 98.6%, whereas RaTG13 has a sequence homology of 99%, which is in contrast with approximately 90 and 96% homologies for pangolin-CoV and RaTG13 genome to SARS-CoV-2, respectively. Because higher sequence similarities of M proteins of COVID-19-related viruses, including pangolin-CoVs, have been obtained, especially when compared with the sequence similarities of other proteins genome-wide (Table 2), this is further evidence of the more conserved nature of the M proteins among COVID-19-related viruses, and the phylogenetic study using M is therefore likely to be providing a more accurate snapshot, as already mentioned. In summary, the chance of pangolin-CoV being the intermediary increases dramatically when using M for our phylogenetic analysis. It should also be

noted that recombinations are notorious for causing confusion in previous studies, as most phylogenetic algorithms are not designed to handle them. This was therefore likely the case in all previous pangolin-CoV studies, and the choice of M in our phylogenetic study sidesteps this hurdle.

Specific Evolution of SARS-CoV-2 via Pangolins

Figure 7 illustrates the specific evolution of SARS-CoV-2. More importantly, it points to an evolution with implications that have thus far been seen as unique. Because of the pangolin association with fecal–oral transmission, the ancestral strains of SARS-CoV-2 were likely attenuated and had greater fecal–oral transmission potential. There is evidence that other viruses could have been attenuated by animals that have fecal–oral transmission behaviors.

One example is the case of NiV. (See Figure 7.) It was first discovered in Malaysia when pigs ate virus-laden fruits that had fallen on the ground after being consumed by bats. Farm workers were then infected by the pigs (which are farm animals that are bred in close contact with each other, and thus fecal–oral transmission is an important factor).⁷⁰ The case fatality rate (CFR) was ~38%, which is in sharp contrast with the outbreaks in Bangladesh–India after 2001 that had CFRs of 70–80%.⁷¹ Disorder analysis revealed differences in inner shell disorder (N PID) of 41 to 42% versus 43 to 44%.

No such attenuation could be seen in the case of the 2003 SARS-CoV-2 and civet cat, however. A search for signs of attenuation among civet-CoV came out empty. This necessarily lead to the scenario seen in Figure 7, as denoted by “SARS1”. It is therefore likely that the virus entered the human population as a virulent strain that alerted the medical community almost immediately, unlike “SARS2”, that is, SARS-CoV-2.

A curious but baffling trend that is worth mentioning is the stepwise increase in pangolin-CoV N PID for each year (Figure 5). It is indeed very difficult to account for this intriguing pattern. The only plausible explanation is the possibility of cross-species transmission from humans back to pangolins for each subsequent year. The first introduction of pangolin-CoV to humans is illustrated in Figure 6, where arrows point to the possibilities of viruses moving back to pangolins and bats. It is, however, not difficult to imagine the common situation where humans dispose of feces together with sweet and fruity foodstuff that attracts ants and other insects and that would eventually be scavenged by pangolins and bats.

It is also plausible that the stepwise N PIDs are actually capturing a snapshot of the evolutionary gain of function for the human versions. There was an obvious evolutionary pressure toward greater respiratory transmission potentials among its human variants that forced the virus toward becoming more virulent, as respiratory transmission requires an adequate viral load in the saliva and mucus for it to be viable, and, likewise, an overwhelming viral load in vital organs can cause death.^{15,16,27–30} The involved mechanisms of action have already been discussed. It also should be noted that the possibility of an attenuated virus mutating and evolving into a more virulent strain should not be surprising, and it is actually quite common. The most notorious example is seen in the poliovirus of the 1950s.¹⁵

EXPERIMENTAL AND CLINICAL EVIDENCE

Strong Experimental Evidence of SARS-CoV-2 Resilience Is Consistent with the Shell Disorder Model

We have seen computational evidence of the hard outer shell (low M DID) of SARS-CoV-2 arising from being evolutionarily intertwined with a burrowing animal, which raises the question: Is there experimental evidence for this? As it turns out, the answer is “yes”. Both SARS-CoV-1 and SARS-CoV-2 were tested on hard surfaces like plastic and steel, but the investigators, van Doremalen et al.,⁷² were unable to find any significant differences in the stability of the two viruses, as infectious particles could be collected up to 72 h. They were, however, able to detect hints of stronger resilience for SARS-CoV-2 on cardboard, as this virus remained viable for 24 h, in contrast with SARS-CoV-1, which lasted only 8 h. van Doremalen et al. were unfortunately unable to reaffirm their result, owing to high variability.⁷²

SARS-CoV-2 Resilience Is Many Times That of Other CoVs: Experimental Data

Riddell et al.⁷³ conducted a similar experiment but, this time, in the dark, devoid of harmful UV light. They also took note of the temperature and humidity. The group revealed that at 20 °C and 50% RH (relative humidity), SARS-CoV-2 remained viable even after 28 days. This is in sharp contrast with MHV (murine hepatitis virus) and TGEV (transmissible gastroenteritis virus), which were found by another group, Casanova et al.,⁷⁴ to last for only 3 and 5 days under the same conditions, respectively.

The stark results are astonishingly consistent with the CoV shell disorder model with respect to the hardness of the outer shell (M PID). The M PIDs of SARS-CoV-2, MHV, and TGEV were 6, 8, and 14% just as the viruses lasted >28, 5, and 3 days, respectively, at 20 °C and 50% RH. The correlation concurs with the basic shell disorder tenet that stipulates that shells, especially the outer one, protect the virion from damage. Whereas SARS-CoV-1 was not studied by Riddell et al. or Casanova et al., the M PID of SARS-CoV-1 (8.6%) is very close to that of MHV (8%),^{18–21} and the necessary interpolation and inference can be made.

Hard Outer Shell Protects the Virus from the Environment and Antimicrobial Enzymes Found in Body Fluid

The fact that greater evidence of resilience arose when the experiments were conducted away from any light is likely a reflection of the nature of the environmental protection that the hard outer shell confers. This further supports the theory that ancestral coronaviruses were likely to have been buried in the soil covered by fecal materials away from UV light for a long time while awaiting contact with the next pangolin. This hypothesis was constructed from the fact that the few coronaviruses found with such hard outer shells are generally associated with burrowing animals. (See Table 2.) Furthermore, it also highlights the environment that the virus is exposed to in the nasal regions, throat, and mouth, physiologically devoid of UV light but exposed to antimicrobial enzymes in body fluids. The results therefore illuminate the ability of the virus to resist the antimicrobial enzymes, thus giving rise to heavy viral shedding that provides for greater contagiousness.

Comparative Clinical Study: SARS-CoV-1 versus SARS-CoV-1

An important study that concerns us is the clinical research done by Wolfel et al.⁷⁵ The study involves nine COVID-19 patients warded in a Munich hospital. The levels of viral shedding were

carefully measured using RT-PCR (reverse transcription–polymerase chain reaction) and antigenic tests in each stage of the illness, recorded, and compared with past clinical studies of patients from the 2003 SARS outbreak. Signs of shedding were seen in all nine COVID-19 patients on the first onset of symptoms compared with the only 38 2003 SARS patients that tested positive on the first day out of a total of 98. Heavy shedding was found in first week for COVID-19, which peaked before the fifth day with an RNA load of 7×10^8 copies per nasal swab. Shedding could last as long as 28 days given the sample size of nine patients. As for the 2003 SARS, shedding began 7–10 days after the first onset of symptoms, with a peak viral load of 5×10^5 RNA copies per swab. SARS-CoV-2, unlike SARS-CoV-1, is easily detectable in throat swabs. We believe that this could be the result of the higher resistance of SARS-CoV-2 to antimicrobial enzymes in the saliva. Also, infectious particles were found in the throat, nasal, lung, and stool samples but not in urine or blood, which could indicate the presence of high fecal–oral and fecal–respiratory transmission potentials of the virus. In summary, much larger amounts of viral particles are shed for SARS-CoV-2 patients for a much longer period of time compared with those of SARS-CoV-1. This could also be indicative of the differences in the lifecycle of the two viruses, as we will see later.

Clinical Evidence of Heavy Shedding Provides Clues to Transmissibility and Viral Resilience in Body Fluids

There is actually clinical evidence of SARS-CoV-2 relationships among outer shell hardness, transmissibility, heavier shedding, and viral resilience in body fluid. Wolfel et al. performed a study of patients and found that COVID-19 patients typically shed viral particles in volumes that are 1000 times those of SARS patients.⁷⁵ The shedding begins when the patients start showing the mildest symptoms and continues even after the symptoms are over. Whereas the shell disorder analysis has uncovered the unusual hardness of the outer shell and the clinical study of Wolfel et al. showed a large amount of viral shedding that surpasses that of SARS-CoV-1, these two facts do not, on the contrary, necessitate a causal relationship and could just be coincidental. For additional but more compelling evidence, we need to turn to the laboratory experiment that shows that SARS-CoV-1 conversely produces more infectious particles in cells, which will be discussed in the next paragraph.

Strong Experimental Evidence Supporting the Disorder Models of Infection and Virulence: SARS-CoV-1 versus SARS-CoV-2

A group lead by Eric J. Snijder in The Netherlands⁷⁶ infected separate sets of Vero E6 cells with SARS-CoV-1 and SARS-CoV-2. The presence of infectious particles was detected and quantified using antibodies and electron microscopy. The number of infectious particles of SARS-CoV-1 was found to be greater than that of SARS-CoV-2 by a factor of ~ 50 . The levels of intracellular RNA were also quantified using polymerase chain reaction (PCR). Conversely, a higher level of RNA was observed for SARS-CoV-2.

Experiment Reproduced the Predictions of the Shell Disorder Models

The results are astonishing, as Ogando et al.⁷⁶ have basically reproduced the predictions made by the shell disorder models. The models stipulate that SARS-CoV-1 (CFR = 2%) is more virulent than SARS-CoV-2 (CFR = 10%) because SARS-CoV-1 (N PID = 50%) has a higher N disorder than SARS-CoV-2 (N PID = 48%). The reason for the greater virulence is the higher

viral load at vital organs and the greater ability of the virus to replicate more rapidly via a disordered inner shell before the host immune system is able to detect its presence. This begs the question: If SARS-CoV-1 replicates faster than SARS-CoV-2, why is the virus not shedding in greater quantities, as in the case of SARS-CoV-1, as we have seen above? The only possible explanation is that it is not as resistant to the antimicrobial enzymes found in saliva and mucus, as highlighted by the shell disorder model.

Clinical and Experimental Validation of the Shell Disorder Models: Viral Load in the Vital Organ (Virulence) versus Body Fluids (Respiratory Transmission)

Whereas the experiment⁷⁶ shows that SARS-CoV-1 is more aggressive than SARS-CoV-2 due to its producing a greater number of virus copies, Wolfel et al.⁷⁵ have shown that SARS-CoV-2 is more contagious than SARS-CoV-1 because it sheds much more infectious particles. A superficial analysis of this matter would lead us to believe that the experimental and clinical evidence is contradictory. A closer look, however, reveals that the only plausible way that the different sets of data are actually consistent with each other is the scenario described by the CoV transmission shell disorder model, which tells us that the reason that SARS-CoV-2 is more contagious is attributed to its hard outer shell that resists the antimicrobial actions of the enzymes found in the saliva and mucus.^{31,32}

There are also hints in the results of Wolfel et al. that this is the case. The viral loads in the saliva and mucus of COVID-19 patients could reach 1000 times those of SARS patients, whereas the viral loads from lung samples are similar between the different sets of patients.⁷⁵ Whereas according to the shell disorder models, SARS-CoV-1 patients should have higher viral loads in the lungs than those of SARS-CoV-2, the clinical results pertaining to the lung samples could reflect the differences in the life cycles of SARS and COVID-19, which could affect the timing of the release of viral particles. Nevertheless, the difference in the viral loads in body fluids of the two types of patients is startling and very telling.

A further support of this fact can be found in the observation made by Wolfel et al. that the viral loads of samples found in the lower respiratory region tend to be lower than those obtained in the upper respiratory region in the case of SARS-CoV-2. This fits perfectly with our hypothesis, as it implies that large quantities of infectious particles are able to accumulate in the mucus and saliva because of the virus's resistance to the antimicrobial enzymes found there. It could also provide a hint of support that SARS-CoV-2 is less virulent than SARS-CoV-1 because less infectious particles are being produced in vital organs such as the lungs, which is consistent with the lower disorder of N and the aspect of the shell disorder models that involves “Trojan Horse” immune evasion, as mentioned above.

It should also be noted that the clinical work of Wolfel et al. in the mentioned paper mainly involved measurements of viral loads using RT-PCR and antigenic quantifications, whereas the experimental work of Ogando et al. involved measurements of viral loads (infectious particles and viral RNA) using gel electrophoresis and virus titration in plaque assays. Any comparative study of the results from the two laboratories should not be hindered by the differences in their approaches because interpolation is possible. Slightly more copies of RNA are found in SARS-CoV-2 compared with SARS-CoV-1 in various stages of the life cycle when grown in Vero E6 cells.⁷⁶ This is negligible in comparison with the amount of RNA found

in the mucus of SARS-CoV-2, which exceeds 1000-fold in comparison with SARS-CoV-1.⁷⁵

COVID-19 Vaccine Success: Yet Another Validation of the Shell Disorder Models Pertaining to SARS-CoV-2

While the world sighs with relief at the arrival of effective COVID-19 vaccines, we must not forget that the search for effective HIV, HCV, and HSV vaccines has taken approximately 40, 30, and 100 years and has thus far been met with complete failure.^{15,25,26,77} The world was gripped by the nightmare scenario in which this could happen in the COVID-19 vaccine search. From the beginning, the shell disorder models have had plenty to say about this, and if any were proven wrong, then their main tenet would have been totally invalidated. Figure 2B shows that the outer shell of SARS-CoV-2 does not resemble those of HIV, HSV, or HCV in terms of outer shell disorder. Instead, it resembles the classical viruses, for which effective vaccines have been found. In fact, given the samples of coronaviruses we have examined, none of their outer shell disorders resemble that of HIV. This is consistent with what we know about CoVs. CoVs have strong fecal–oral transmission potential, in contrast with HIV, HCV, and HSV's association with sexual transmission.^{15,22,25,26} The rapidity of the discovery of COVID-19 should also not be surprising given its hard (low M PID) shell among CoVs and other viruses, as can be seen in Figures 2, 3C, and 4A. We actually foresaw the successful development of COVID-19 vaccines months before their approvals, as seen in our previous publications.

■ ADDRESSING SARS-CoV-2's MYSTERIOUS ADAPTATION TO HUMANS

S Protein and Furin Cleavage Sites: Do the S-Protein Enigmas Contradict the Shell Disorder Theories? No!

Whereas shell disorder models are mainly based on the N and M proteins that offer a set of evolutionary implications, most other studies have been centered on the S (spike) protein. The SARS-CoV-2 S protein facilitates viral entry by attaching to the ACE-2 receptor. As mentioned above, the SARS-CoV-2 S protein binds to ACE-2 with an affinity that is 20–30 times that of SARS-CoV-1 in a computational study.^{11,12} SARS-CoV-2 has also been found to have the greatest affinity to human ACE-2 by another computational study.¹³

Furthermore, SARS-CoV-2 S protein has a unique furin cleavage site that is not seen among its close relatives, SARS-CoV-1, pangolin-CoVs, and bat-RaTG13.^{2,11,14} The furin site plays a role in the cleavage of the S protein into S1 and S2, which is an important process prior to viral entry. The SARS-CoV-2 furin cleavage site is somewhat unique by being polybasic, which allows for more efficient cleavage.^{11,14} Because of all of these established characteristics of the S protein, it has become a somewhat common belief that the S protein is responsible for the greater transmissibility of SARS-CoV-2.¹⁴ If so, the question is then, is the S protein or the N/M protein responsible for the greater contagiousness of SARS-CoV-2? Or is it both?

Whereas the S protein may still contribute to the transmissibility, we will see in the next paragraph that published experimental and clinical data support the greater contagiousness arising from the order and disorder of M and N proteins, respectively. Furthermore, we will see that the shell disorder theories provide a pathway by which SARS-CoV-2 could evolve with humans for a long period of time, as suggested by many studies on the S protein.

Revisiting the Experiment: Definitive Answers and Some Hints

The experiments⁷⁶ provided both definitive answers and some hints for the S-protein enigma. Because it has been shown that the SARS-CoV-2 has an S protein that is more adapted to the human ACE-2 receptors in many ways when compared with SARS-CoV-1, we could reasonably expect SARS-CoV-2 to produce more infectious particles if we claim that the S protein is responsible for its greater contagiousness. The experiment, however, clearly shows that this is not the case. The number of copies of infectious particles of SARS-CoV-1 exceeds that of SARS-CoV-2 by about 50-fold. Therefore, the greater efficiency of the S protein in cleavage and binding to the ACE-2 does not lead to greater contagiousness with respect to the increase in viral loads, even in body fluids.

Nevertheless, there are important signs of more efficient viral entry for SARS-CoV-2 seen in the results of the experiment reported by Ogando et al.⁷⁶ Even though there were fewer SARS-CoV-2 infectious particles, there were, conversely, also greater quantities of viral RNA copies in the case of SARS-CoV-2. The greater presence of viral RNA in certain stages of the cell cycle is likely evidence of greater ease of viral entry. Such evidence not only is consistent with the shell disorder models but also suggests how the models could explain the peculiarities of the data. The greater presence of RNA but lower presence of viral particles suggests that plenty of RNA has been replicated but is awaiting packaging and release. The virus is, however, unable to release more particles, unlike SARS-CoV-1, because its N protein is not sufficiently disordered. Apparently, the data point to the greater efficiency in the replication, packaging, assembly, and budding of SARS-CoV-1. This is consistent with what is understood about the role of protein disorder in replication, assembly, packaging, and viral budding, as already described above. There is, however, a hint on how the excessive RNA could affect the transmissibility. Excessive RNA may mean that there are more RNA waiting in line within the cell in the queue to be assembled as part of the infectious particle. This could mean that there is a longer infectivity duration at the cellular level. There is actually clinical evidence for such in the case of SARS-CoV-2, in which some patients still shed infectious particles even months after recovery. Nevertheless, further research needs to be done.

S Protein: Evidence of SARS-CoV-2 Evolution with Humans

We have seen compelling evidence that points to the high adaptation of SARS-CoV-2 to human ACE-2. There is yet further evidence of this. For example, computational studies have found that the SARS-CoV-2 S protein is more adapted to the human ACE-2 than those of a variety of animals. We have also seen that the polybasic furin cleavage site is found only in SARS-CoV-2 and not even among its closest relatives, including pangolin-CoVs, RaTG13, and SARS-CoV-1. If that is the case, how did SARS-CoV-2 acquire its polybasic characteristic? An Indian group found a matching sequence in an unrelated human enzyme, eNac.⁷⁸ This sequence has the closest match to the human eNac compared with those of many other animals.

If SARS-CoV-2 Has Been Present in Humans for a Long Time, How Did It Escape the Notice of the Medical Community for All of These Years?

Assuming that all of this plausible evidence that the virus has been in humans for a long period of time is correct, how did SARS-CoV-2 remain in humans for so long without the notice of the medical community? Whereas it is possible that the virus had

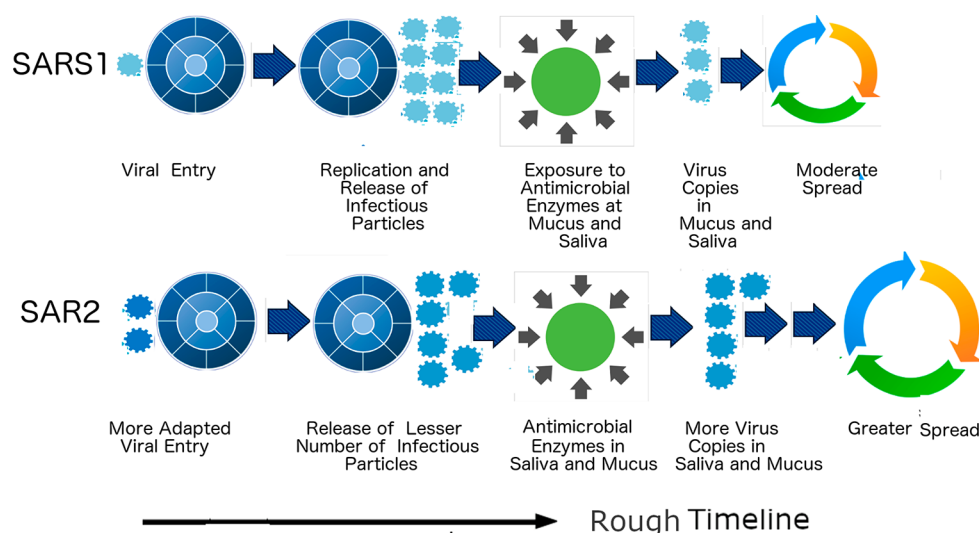


Figure 8. Schematic differences in the replication cycles of SARS-CoV-1 (SARS1) and SARS-CoV-2 (SARS2), as seen the shell disorder models and published experimental and clinical data. One and two virus copies are shown next to the cells for SARS1 and SARS2, respectively, to illustrate the more efficient cell entry of SARS2, which could induce the ease of entry of more than one viral particle. The additional number of infectious particles produced by SARS1 compared with SARS2 is just a qualitative illustration that more particles are produced by SARS1. The same goes for the number of particles left after exposure to antimicrobial enzymes in body fluids.

been in its virulent form within humans for a long time without being noticed by doctors, it is logically difficult to imagine this to be the case. This becomes a paradox for the S-protein enigmas. This is where the shell disorder models come in as a complementary theory, not a competing one. The data presented by the shell disorder models imply that SARS-CoV-2 entered the human population via pangolins as an attenuated strain a few years ago, perhaps in 2017, or even earlier. This provides ample time for the virus to adapt to humans. In the past, attenuated viruses and vaccines have mutated into virulent strains. An example is the Sabin polio vaccine.^{15,22}

SARS-CoV-1 versus SARS-CoV-2: Replication Cycle

Figure 8 summarizes the differences in the replication cycles of SARS-CoV-1 (SARS1) and SARS-CoV-2 (SARS2). Upon entry, both viruses undergo replication. However, SARS-CoV-1 is more efficient in replicating and releasing the infectious particles, as shown in the experiment by Ogando et al.⁷⁶ and predicted by the shell disorder models. Because SARS-CoV-1 produces more infectious particles, especially in vital organs, it is more virulent than SARS-CoV-2. Whereas the experiment reaffirms the presence of a greater number of infectious particles upon the initial infection of the cell, clinical evidence suggests that large quantities of SARS-CoV-2 are shed. This can only be possible given the information from Ogando et al.⁷⁶ if SARS-CoV-2 is more resistant to the antimicrobial enzymes found in the saliva and mucus because of its harder outer shell, as detected by the shell disorder models. The large quantities of the virus found in the saliva and mucus contribute to the contagiousness of COVID-19.

There is also the S-protein enigma that hypothesizes that the S protein contributes to the transmissibility of the virus. The results from the experiments by Ogando et al.⁷⁶ suggest that the S proteins contribute to the transmissibility in other ways. The presence of large amounts of viral RNA detected in SARS-CoV-2 implies the possibility that huge amounts of incomplete viral particles have been found. This could account for the clinical evidence that some people are still infectious months after recovery,⁷⁹ and it is represented by a longer timeline for SARS2

in Figure 8. This phenomenon points to the likelihood that greater protein intrinsic disorder in SARS-CoV-1 leads to greater efficiency in the packaging and release of the particles and not necessarily in RNA replication. This is consistent with the tenet that protein intrinsic disorder provides for more efficient protein–protein/RNA/lipid interactions and that the N protein is certainly involved in the packing, assembly, and release of the virus.

China Conundrum and Pangolin Factor

A curious statistic that has not escaped the scientific community has to do with the low rate of COVID-19 infection in China, especially when we compare the number of cases in the USA and Europe. As of today, the total number of cases in China is close to 91 000, whereas the number of cases in the USA is above 32 million.⁶⁵ The numbers are simply astonishing because the populations of China and USA are over 1 billion and 300 million, respectively. Nobody has been able to provide a satisfactory answer to this puzzle. A highly plausible answer yet to be suggested is that an attenuated strain had entered the population in China many years ago as a mild cold that provided immunity. If this is the case, then it is likely that a pangolin-CoV entered the human population many years before 2017. However, it will be exceptionally difficult, if not impossible, to prove that an attenuated strain had passed through China's population or otherwise because studies have shown that many people who were infected had COVID-19 antibodies that began to drop to undetectable levels after 6 months.^{80,81} Nevertheless, an infected patient may still be protected regardless of the presence of antibodies, as immune system cells such as T cells and memory B cells may provide long-term protection.^{53,82,83} We need to keep in mind, however, that the idea that there is already mass immunity among people in China is currently a speculative but compelling observation that warrants further investigation. We must also realize that this possibility has not crossed the minds of the scientific community until now.

Whereas we believe that much of the computational, experimental, and clinical evidence of a specific but peculiar evolution of SARS-CoV-2 is compelling and has been shown to

be reproducible thus far, further reproducibility may be needed to convince the hardened skeptics. The possible directions for further investigation are simply too innumerable to be listed here. Immunologists could, for instance, attempt to search for blood samples stored in laboratory and hospital refrigerators since 2017 or before to detect the presence of COVID-19-related viruses or antibodies. Epidemiologists could retrieve data pertaining to the beginning of the pandemic to study the patterns of spread to see if there is any hint of immunity in the population, especially among those in China and Southeast Asia, where pangolins are common. The relative slowness of the original Wuhan virus, in comparison with later variants such as Delta, could provide possible hints of any existence of a precursor strain. Even within the expertise of our group, there is still much to be done. Much of this Review was written before the identification of the Omicron variant. Incoming data pertaining to variants such as Omicron offer further insights into the reproducibility of the shell disorder models. Our group has been carefully studying the data, and whereas there is still work to be done, the data indicate that the inner shell disorder of Omicron resembles more that of 2017 pangolin-CoV than SARS-CoV-2 (unpublished data), which could account for the attenuation of the Omicron variant. The outer shell of Omicron is harder than those of SARS-CoV-2 and most pangolin-CoVs. These exciting developments lend greater credibility to the idea of a silent spread of an attenuated strain before the COVID-19 pandemic. For further details regarding shell disorder and Omicron, we will, unfortunately, need to await future publications by our group.

■ SUMMARY AND CONCLUSIONS

Computational, Experimental, and Clinical Evidence of the Specific Evolution of SARS-CoV-2

We have seen that the shell disorder models elucidated a specific but peculiar evolutionary pathway. The hint of this pathway was first noticed when the shell disorder models found that the SARS-CoV-2 has among the hardest outer shell disorder among CoVs, and it was later discovered that this hard outer shell is associated with burrowing animals, in particular, pangolins. This characteristic is validated by at least one experimental study. The hard outer shell nature of SARS-CoV-2 has wide implications, including those that are of clinical, epidemiological, and immunological importance. The volume of virus shed by COVID-19 patients is attributed by computational, clinical, and experimental studies to the hard protective outer shell that provides the virus' resistance to antimicrobial enzymes found in the saliva and mucus.

The shell disorder models accurately predicted not only the feasibility of COVID-19 vaccine development based on its M PID but also the existence of an attenuated strain that had entered the human population from pangolins several years ago. The latter was done using N–M shell disorder analysis and phylogenetic analysis based on the M protein, which reveals a much closer relationship between SARS-CoV-2 and pangolin-CoV than previously thought. The same shell disorder analysis also predicts that SARS-CoV-1 is more virulent because of its higher N disorder that allows for the more efficient replication of infectious particles, especially in vital organs, but has a lower viral load. No only is this supported by the experimental study of Ogando et al.⁷⁶ that showed larger quantities of infectious particles found in SARS-CoV-1, but also the experimental results hint at the possibility that the shell disorder models are actually

complementary to studies showing that the higher adaptation of the S protein to human ACE-2 contributes to transmissibility. More specifically, the results showing a larger amount of viral RNA, not infectious particles in cells infected by SARS-CoV-2, suggest that the S protein may be prolonging the infectious cycle. This is supported by the observation that many patients still shed infectious particles months after recovery.

Specific but Natural Evolution

Conspiracy theories will persist as the precise virus that first entered humans from an animal intermediary is likely to be extinct, especially if the virus had entered humans at least a few years ago. Given that it is likely to be extinct, we need to rely on genetic and proteomic analyses such as the shell disorder models. As mentioned, the shell disorder model has plenty to say about the evolutionary nature of SARS-CoV-2. The models explain that the likely reason that the S protein is so adapted to the human ACE-2 is that it first entered humans as an attenuated strain a few years ago, not that it was engineered in a laboratory.²² The silent spread, helped by its attenuated nature, escaped the notice of the medical communities because it could easily have been mistaken for a common cold. Furthermore, if the virus was engineered in a laboratory, then how was anyone able to acquire the knowledge, which was previously unheard of, pertaining to the virus that we have just mentioned, including the contribution of its hard outer shell to its contagiousness?

There are also questions as to whether the pangolin-CoV samples were actually from pangolin-CoVs because the samples were obtained from caged animals that were confiscated from smugglers, and the pangolins could have been infected by other caged animals.⁸⁴ For this, our data support the likelihood that the samples are those of pangolin-CoV because evidence based on our CoV search and phylogenetic tree points to the fact that the CoVs with the hardest outer shell are associated with unrelated burrowing animals such as rabbits and pangolins. Whereas one of the main goals of this Review is to highlight incoming clinical and experimental data so as to ascertain the reproducibility of the shell disorder models relevant to COVID-19, much could still be done.

Implications Are Far-Reaching

The scientific implications are aplenty, and there are simply too many to list in this Review. As already mentioned, the immunological significance includes the ease of vaccine development and the existence of attenuated strains in nature. We have also seen how the outer shell affects the contagiousness of the virus. This should provide further hints regarding the control that would be of interest to epidemiologists. The envisaged way, if proven true, that the S protein affects the transmissibility via prolonged infectious periods could provide options for treatments. The intricate evolutionary pathway that ancestral strains of SARS-CoV-2 had to undergo provides us with insights into and possible patterns on what to look out for, especially when we monitor for new potential pandemics and possible zoonotic or nonzoonotic COVID-19 reemergence after the current mass vaccination effort.

■ AUTHOR INFORMATION

Corresponding Author

Gerard Kian-Meng Goh – *Goh's BioComputing, Singapore 548957, Republic of Singapore*;  orcid.org/0000-0003-3635-0806; Email: gohsbiocomputing@yahoo.com

Authors

A. Keith Dunker – Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana 46202, United States

James A. Foster – Department of Biological Sciences, University of Idaho, Moscow, Idaho 83844, United States; Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, Idaho 83844, United States

Vladimir N. Uversky – Department of Molecular Medicine, USF Health Byrd Alzheimer's Research Institute, Morsani College of Medicine, University of South Florida, Tampa, Florida 33620, United States; Laboratory of New Methods in Biology, Institute for Biological Instrumentation of the Russian Academy of Sciences, Federal Research Center "Pushchino Scientific Center for Biological Research of the Russian Academy of Sciences", Pushchino 142290, Russia;

orcid.org/0000-0002-4037-5857

Complete contact information is available at:

<https://pubs.acs.org/10.1021/acs.jproteome.2c00001>

Author Contributions

G.K.-M.G. conceived the idea, collected and analyzed the data, and wrote the first draft. V.N.U. helped with the collection and analysis of literature data and reviewed and revised the draft. A.K.D. and J.A.F. reviewed the manuscript and provided the resources necessary for the research.

Notes

The authors declare the following competing financial interest(s): G.K.-M.G. is an independent researcher and the owner of Goh's BioComputing, Singapore. G.K.-M.G. has also written a book (*Viral Shapeshifters: Strange Behaviors of HIV and Other Viruses*) on a related subject. The authors have no other potential conflicts of interest.

REFERENCES

- (1) WHO. Middle Eastern respiratory syndrome coronavirus (MERS-CoV). [https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-\(mers-cov\)](https://www.who.int/news-room/fact-sheets/detail/middle-east-respiratory-syndrome-coronavirus-(mers-cov)) (accessed February 9, 2022).
- (2) Hu, B.; et al. Characteristics of SARS-CoV-2 and COVID-19. *Nat. Rev. Microbiol.* **2021**, *19*, 141–154.
- (3) Li, X.; Zai, J.; Zhao, Q.; Nie, Q.; Li, Y.; Foley, B. T.; Chaillon, A. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J. Med. Virol.* **2020**, *92*, 602.
- (4) Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; et al. Adendum: A henumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *588*, E6.
- (5) Zhou, P.; Yang, X. L.; Wang, X. G.; Hu, B.; Zhang, L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270–273.
- (6) Xiao, K.; et al. Isolation of SARS-CoV-2 related coronavirus from Malayan pangolins. *Nature* **2020**, *583*, 286–289.
- (7) Liu, P.; Jiang, J. Z.; Wan, X. F.; et al. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog.* **2020**, *16*, e1008421.
- (8) Fan, H. H.; Wang, L. Q.; Liu, W. L.; An, X. P.; Liu, Z. D.; He, X. Q.; Song, L. H.; Tong, Y. G. Repurposing of clinically approved drugs for treatment of coronavirus 2019 in a 2019-novel coronavirus (2019 nCoV) related virus model. *Chin. Med. J. (Engl.)* **2020**, *133*, 1051–1056.
- (9) Lam, T. T.; et al. Identifying SARS-CoV-2 related coronaviruses in Malayan pangolins. *Nature* **2020**, *583*, 282–285.
- (10) WHO. Coronavirus diseases (COVID-19) pandemic. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019> (accessed February 9, 2022).
- (11) Andersen, K. G.; Rambaut, A.; Lipkin, W. I.; Holmes, E. C.; Garry, R. F. The proximal origin of SARS-CoV-2. *Nat. Med.* **2020**, *26*, 450–452.
- (12) Shang, J.; Ye, G.; Shi, K.; Wan, Y.; Luo, C.; Aihara, H.; Geng, Q.; Auerbach, A.; Li, F.; et al. Structural basis of receptor recognition by SARS-CoV-2. *Nature* **2020**, *581*, 221–224.
- (13) Piplani, S.; Singh, P. K.; Winkler, D. A.; Petrovsky, N.; et al. In silico comparison of SARS-CoV-2 spike protein-ACE2 binding affinities across species and implications for virus origin. *Sci. Rep.* **2021**, *11*, 13063.
- (14) Hofmann, H.; Pyrc, K.; van der Hoek, L.; Geier, M.; Berkhout, B.; Pohlmann, S. Human coronavirus NL63 employs the severe acute respiratory syndrome coronavirus receptor for cellular entry. *Proc. Natl. Acad. Sci.* **2005**, *102*, 7988–7993.
- (15) Goh, G. K. *Viral Shapeshifters: Strange Behaviors of HIV and Other Viruses*, 1st ed.; Simplicity Research Institute: Singapore, 2017; pp 1–120.
- (16) Goh, G.; Dunker, A. K.; Foster, J.; Uversky, V. A Novel Strategy for the Development of Vaccines for SARS-CoV-2 (COVID-19) and Other Viruses Using AI and Viral Shell Disorder. *J. Proteome Res.* **2020**, *19*, 4355–4363.
- (17) Goh, G.; Dunker, A. K.; Foster, J.; Uversky, V. Feasibility of vaccine development for SARS-CoV-2 and other viruses using the shell disorder analysis. *Pac. Symp. Biocomput.* **2021**, *26*, 143–153.
- (18) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Understanding viral transmission behavior via protein intrinsic disorder prediction: Coronaviruses. *J. Pathog.* **2012**, *2012*, 1–13.
- (19) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Prediction of intrinsic disorder in MERS CoV/HCoV-EMC supports a high oral-fecal transmission. *PLoS Curr.* **2013**, *5*.
- (20) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. Rigidity of the outer shell predicted by a protein disorder model sheds light on COVID-19 (Wuhan-2019-nCoV) infectivity. *Biomolecules* **2020**, *10*, 331.
- (21) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. Shell disorder analysis predicts greater resilience of SARS-CoV-2 (COVID-19) outside the body and in body fluids. *Microb. Pathog.* **2020**, *144*, 104177.
- (22) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. Shell disorder analysis suggests that pangolins offered a window for a silent spread of an attenuated SARS-CoV-2 precursor among humans. *J. Proteome Res.* **2020**, *19*, 4543–52.
- (23) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Protein intrinsic disorder toolbox for comparative analysis of viral proteins. *BMC Genom.* **2008**, *9* (Suppl.2), S4.
- (24) Goh, G. K.; Dunker, A. K.; Uversky, V. N. A comparative analysis of viral matrix proteins using disorder predictors. *Virol. J.* **2008**, *5*, 126.
- (25) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. HIV Vaccine Mystery and Viral Shell Disorder. *Biomolecules* **2019**, *9*, 178.
- (26) Goh, G.; Uversky, V. Shell disorder and the HIV vaccine mystery: Lessons from the legendary Oswald Avery. *J. Biomol. Struct. Dyn.* **2021**, *1*–10.
- (27) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Correlating Flavivirus virulence and levels of intrinsic disorder in shell proteins: Protective roles vs. immune evasion. *Mol. Biosyst.* **2016**, *12*, 1881–1891.
- (28) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. Zika and flavivirus shell disorder: Virulence and fetal morbidity. *Biomolecules* **2019**, *9*, 710.
- (29) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Detection of links between Ebola nucleocapsid and virulence using disorder analysis. *Mol. Biosyst.* **2015**, *11*, 2337–2344.
- (30) Goh, G.K.-M.; Dunker, A. K.; Foster, J. A.; Uversky, V. N. Nipah shell disorder: Mode of transmission and virulence. *Microb. Pathog.* **2020**, *141*, 103976.

- (31) Malamud, D.; Abrams, W.; Barber, C.; Weissman, D.; Rehtanz, M.; Golub, E. Antiviral activities in saliva. *Adv. Dent. Res.* **2011**, *23*, 34–37.
- (32) Cole, A. M.; Dewan, P.; Ganz, T. Innate antimicrobial activity of nasa secretions. *Infect. Immun.* **1999**, *67*, 3267–3275.
- (33) Dunker, A.K.; Lawson, J.D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z.; et al. Intrinsically disordered proteins. *J. Mol. Graph. Model.* **2001**, *19* (1), 26–59.
- (34) Tompa, P. Intrinsically unstructured proteins. *Trends Biochem. Sci.* **2002**, *27*, 527–533.
- (35) Uversky, V. N.; Gillespie, J. R.; Fink, A. L. Why are “natively unfolded” proteins unstructured under physiologic conditions? *Proteins* **2000**, *41*, 415–427.
- (36) Wright, P. E.; Dyson, H. J. Intrinsically unstructured proteins: Re-assessing the protein structure-function paradigm. *J. Mol. Biol.* **1999**, *293*, 321–331.
- (37) van der Lee, R.; Buljan, M.; Lang, B.; Weatheritt, R. J.; Daughdrill, G. W.; Dunker, A. K.; Fuxreiter, M.; Gough, J.; Gsponer, J.; Jones, D. T.; Kim, P. M.; Kriwacki, R. W.; Oldfield, C. J.; Pappu, R. V.; Tompa, P.; Uversky, V. N.; Wright, P. E.; Babu, M. M.; et al. Classification of intrinsically disordered regions and proteins. *Chem. Rev.* **2014**, *114*, 6589–6631.
- (38) Uversky, V. N.; Dunker, A. K. Unstructured protein non-folding. *Biochim. Biophys. Acta.* **2010**, *1804*, 1231–64177.
- (39) Dunker, A. K.; et al. Function and structure of inherently disordered proteins. *Curr. Opin. Struct. Biol.* **2008**, *18*, 756–64.
- (40) Dunker, A. K.; et al. Flexible nets. The roles of intrinsic disorder in protein interaction network. *FEBS.* **2005**, *272*, 5129–48.
- (41) Uversky, V. N.; Oldfield, C. J.; Dunker, A. K. Showing your ID: Intrinsically disordered as an ID for recognition, regulation and cell signaling. *J. Mol. Recognit.* **2005**, *18*, 343–84.
- (42) Yan, J.; Dunker, A. K.; Uversky, V. N.; Kurgan, L.; et al. Molecular recognition features (MoRF) in three domains of life. *Mol. Biosyst.* **2016**, *12* (3), 697–710.
- (43) Cheng, Y.; Oldfield, C. J.; Meng, J.; Romero, P.; Uversky, V. N.; Dunker, A. K. Mining alpha-helix-forming molecular recognition features with cross species sequence alignments. *Biochemistry* **2007**, *46*, 13468–13477.
- (44) Vacic, V.; et al. Characterization of molecular recognition features, MORF and their binding partners. *J. Proteome Res.* **2007**, *6*, 2351–60.
- (45) Mohan, A.; et al. Analysis of molecular recognition features (MORFs). *J. Mol. Biol.* **2006**, *362*, 1043–59.
- (46) Oldfield, C. J.; Cheng, Y.; Cortese, M. S.; Romero, P.; Uversky, V. N.; Dunker, A. K. Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry.* **2005**, *44*, 12454–12470.
- (47) Xue, B.; Williams, R. W.; Oldfield, C. J.; Goh, G. K.-M.; Dunker, A.; Uversky, V. N. Viral disorder or disordered viruses: Do viral proteins possess unique features? *Protein Pept. Lett.* **2010**, *17* (8), 932–951.
- (48) Goh, G. K.; Dunker, A. K.; Uversky, V. N. Protein intrinsic disorder and influenza virulence: The 1918 H1N1 and H5N1 viruses. *Viral J.* **2009**, *6*, 69.
- (49) Romero, P.; Obradovic, Z.; Li, X.; Garner, E. C.; Brown, C. J.; Dunker, A. K. Sequence complexity of disordered protein. *Proteins* **2001**, *42*, 38–48.
- (50) Garner, E.; Romero, P.; Dunker, A. K.; Brown, C.; Obradovic, Z. Predicting Binding Regions within Disordered Proteins. *Genome Inform. Ser. Workshop Genome Inform.* **1999**, *10*, 41–50.
- (51) Li, X.; Romero, P.; Rani, M.; Dunker, A. K.; Obradovic, Z. Predicting protein disorder for N-, C-, and internal regions. *Genome Inform. Ser. Workshop Genome Inform.* **1999**, *10*, 30–40.
- (52) R Core Team. *R: A Language and Environment for Statistical Computing*; The R Foundation: Vienna, Austria, 2019.
- (53) Acheson, N. H. *Fundamentals of Molecular Virology*; Wiley: Hoboken, NJ, 2007; pp 1-403.
- (54) Skowronski, D. M.; De Serres, G. Safety and efficacy of BNT162b2 mRNA COVID-19 vaccine. *New. Engl. J. Med.* **2021**, *384*, 1576–1577.
- (55) Lin, C.-M.; Gao, X.; Oka, T.; Vlasova, A. N.; Esseili, M. A.; Wang, Q.; Saif, L. J. Antigenic relationships among porcine epidemic diarrhea virus and transmissible gastroenteritis virus strains. *J. Virol.* **2015**, *89*, 3332–3342.
- (56) Ferguson, N. M.; Van Kerkhove, M. D. Identification of MERS-CoV in dromedary camels. *Lancet Infect. Dis.* **2014**, *14*, 93–94.
- (57) Wadman, M. COVID-19 unlikely to cause birth defects but doctors await fall births. *Science* **2020**, *369*, 607.
- (58) Komine-Aizawa, S.; Takada, K.; Hayakawa, S. Placental barrier against COVID-19. *Placenta* **2020**, *99*, 45–49.
- (59) Macossay-Castillo, M.; Marvelli, G.; Guharoy, M.; Jain, A.; Kihara, D.; Tompa, P.; Wodak, S. J.; et al. The balancing act of intrinsically disordered proteins enabling functions while minimizing promiscuity. *J. Mol. Biol.* **2019**, *431*, 1650–1670.
- (60) McBride, R.; van Zyl, M.; Fielding, B. C. The coronavirus nucleocapsid is a multifunctional protein. *Viruses* **2014**, *6*, 2991–3018.
- (61) Masters, P. S. The molecular biology of coronaviruses. *Adv. Vir. Res.* **2006**, *66*, 193–292.
- (62) Watanabe, S.; Noda, T.; Kawaoka, Y. Functional mapping of the nucleoprotein of the Ebola virus. *J. Virol.* **2006**, *80*, 3743–3751.
- (63) Mortola, E.; Roy, P. Efficient assembly and release of SARS coronavirus-like particle by a heterologous expression system. *FEBS Lett.* **2004**, *576*, 174–178.
- (64) Longhi, S.; Oglesbee, M. Structural disorder within the measles virus nucleoprotein and phosphoprotein. *Protein Pept. Lett.* **2010**, *17*, 961–978.
- (65) Worldometer. COVID-19 Coronavirus Pandemic, 2021. <https://www.worldometers.info/coronavirus/> (accessed February 9, 2022).
- (66) Rajgor, D. D.; Lee, M. H.; Archuleta, S.; Bagdasarian, N.; Quek, S. C. The many estimates of COVID-19 case-fatality rate. *Lancet Infect. Dis.* **2020**, *20*, 776–777.
- (67) Abraham, T. *Twenty-First Century Plague: The Story of SARS*; The Johns Hopkins University Press: Baltimore, MD, 2004; pp 1–176.
- (68) Flores-Alanis, A.; Sandner-Miranda, L.; Delgado, G.; Cravioto, A.; Morales-Espinosa, R.; et al. The receptor binding domain of SARS-CoV-2 spike protein is the result of an ancestral recombination between bat = CoV RaTG13 and pangolin-CoV MP789. *BMC. Res. Notes* **2020**, *13*, 398.
- (69) Posada, D. How does recombination affect phylogeny estimation. *Trends Ecol. Evol.* **2000**, *15*, 489–490.
- (70) Lo Presti, A.; et al. Origin and evolution of Nipah virus. *J. Med. Virol.* **2016**, *88*, 380–8.
- (71) Mire, C. E.; Satterfield, B. A.; Geisbert, J. B.; Agans, K. N.; Borisevich, V.; Yan, L.; Chan, Y.-P.; Cross, R. W.; Fenton, K. A.; Broder, C. C.; Geisbert, T. W.; et al. Pathogenic differences between Nipah virus Bangladesh and Malaysia strains in primates: Implications for antibody therapy. *Sci. Rep.* **2016**, *6*, 30916.
- (72) van Doremalen, N.; et al. Aerosol and surface stability of SARS-CoV-2 as compared to SARS-CoV-1. *New Engl. J. Med.* **2020**, *382*, 1564–156.
- (73) Riddell, S.; Goldie, S.; Hill, A.; Eagles, D.; Drew, T. W.; et al. The effect of temperature on persistence of SARS-CoV-2 on common surfaces. *Viral J.* **2020**, *17*, 145.
- (74) Casanova, L. M.; et al. Effects of air temperature and relative humidity on coronavirus survival on surfaces. *Appl. Environ. Microbiol.* **2010**, *76*, 2712–7.
- (75) Wolfel, R.; Corman, V. M.; Guggemos, W.; et al. Virological assessment of hospitalized patients with COVID-19. *Nature* **2020**, *581*, 465–469.
- (76) Ogando, N. S.; et al. SARS coronavirus-2 replication on Vero E6 cells replication kinetics, rapid adaption and cytopathology. *J. Gen. Virol.* **2020**, *101*, 925–940.
- (77) Uversky, V. N. On the irrationality of rational design of an HIV vaccine in light of protein intrinsic disorder. *Arch. Virol.* **2021**, *166*, 1283–1296.

(78) Anand, P.; Puranik, A.; Aravamudan, M.; Venkatakrishnan, A. J.; Soundararajan, V. SARS- CoV-2 strategically mimics proteolytic activation of human EnaC. *Elife* **2020**, *9*, e58603.

(79) Mahase, E. COVID-19: Past infections provide 83% protection for five months but may not stop transmission, study finds. *BMJ*. **2021**, *372*, n124.

(80) Lumley, S. F.; et al. Antibody status and incidence of SARS-CoV-2 infections in health care workers. *New Engl. J. Med.* **2021**, *384*, 533–540.

(81) Seow, J.; et al. Longitudinal observation and decline of neutralizing antibody responses in three months following SARS-CoV-2 infections in humans. *Nat. Microbiol.* **2020**, *5*, 1598–1607.

(82) Karlsson, A. C.; Humbert, M.; Buggert, M. The known unknowns of T cell immunity to COVID-19. *Sci. Immunol.* **2020**, *5*, eabe8063.

(83) Schulien, I.; et al. Characterization of pre-existing and induced SARS-CoV-2 specific CD8(+) T cells. *Nat. Med.* **2021**, *27*, 78–85.

(84) Chan, Y. A.; Zhan, S. H. Single source of pangolin-CoV with a near identical Spike RBD of SARS-CoV-2. *BioRxiv* **2020**, 2020.07.07.184374.