



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

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



ELSEVIER

Contents lists available at ScienceDirect

Microbial Pathogenesis

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

Shell disorder analysis predicts greater resilience of the SARS-CoV-2 (COVID-19) outside the body and in body fluids

Gerard Kian-Meng Goh^{a,*}, A. Keith Dunker^b, James A. Foster^{c,d}, Vladimir N. Uversky^{e,f}

^a Goh's BioComputing, Singapore

^b Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, IN, USA

^c Department of Biological Sciences, University of Idaho, Moscow, ID, USA

^d Institute for Bioinformatics and Evolutionary Studies, University of Idaho, Moscow, ID, USA

^e Department of Molecular Medicine, Morsani College of Medicine, University of South Florida, Tampa, FL, USA

^f Institute for Biological Instrumentation, Russian Academy of Sciences, Pushchino, Moscow region, Russia

ARTICLE INFO

Keywords:

Intrinsically disordered protein
Nucleocapsid
CoV
Coronavirus
Spread
Membrane
SARS
MERS
COVID

ABSTRACT

The coronavirus (CoV) family consists of viruses that infects a variety of animals including humans with various levels of respiratory and fecal-oral transmission levels depending on the behavior of the viruses' natural hosts and optimal viral fitness. A model to classify and predict the levels of respective respiratory and fecal-oral transmission potentials of the various viruses was built before the outbreak of MERS-CoV using AI and empirically-based molecular tools to predict the disorder level of proteins. Using the percentages of intrinsic disorder (PID) of the nucleocapsid (N) and membrane (M) proteins of CoV, the model easily clustered the viruses into three groups with the SARS-CoV (M PID = 8%, N PID = 50%) falling into Category B, in which viruses have intermediate levels of both respiratory and fecal-oral transmission potentials. Later, MERS-CoV (M PID = 9%, N PID = 44%) was found to be in Category C, which consists of viruses with lower respiratory transmission potential but with higher fecal-oral transmission capabilities. Based on the peculiarities of disorder distribution, the SARS-CoV-2 (M PID = 6%, N PID = 48%) has to be placed in Category B. Our data show however, that the SARS-CoV-2 is very strange with one of the hardest protective outer shell, (M PID = 6%) among coronaviruses. This means that it might be expected to be highly resilient in saliva or other body fluids and outside the body. An infected body is likelier to shed greater numbers of viral particles since the latter is more resistant to antimicrobial enzymes in body fluids. These particles are also likelier to remain active longer. These factors could account for the greater contagiousness of the SARS-CoV-2 and have implications for efforts to prevent its spread.

1. Introduction

In December of 2019, a mysterious virus that causes pneumonia and produces pneumonia-like symptoms; i.e., dry cough and fever was noticed to be spreading in Wuhan, Hubei Province, China [1–4]. The culprit was quickly identified as a new respiratory coronavirus (CoV). This virus and its disease are currently known as Severe Acute Respiratory Syndrome coronavirus-2 (SARS-CoV-2 and coronavirus infectious disease 2019 (COVID-19) respectively [1–4]. SARS-CoV-2 is observed to be even more contagious (but fortunately less fatal) than SARS-CoV that was seen in 2003.

In 2011–2012, just before the outbreak of the Middle Eastern Respiratory Syndrome (MERS-CoV), we built an empirically-based model that measures the percentage of intrinsic disorder (PID) of the membrane (M) and nucleocapsid (N) proteins in viruses [5,6]. The main

tool uses AI technology to recognize intrinsic disorder, given the protein sequence. The model involves the listing of shell disorder by coronaviruses that is easily clustered into three groups, which incidentally correlate with the known levels of fecal-oral and respiratory transmission of the various coronaviruses. The model predicts that the 2003 SARS-CoV (M PID = 8% and N PID = 50%) is included into a category of CoVs with intermediate levels of both respiratory and fecal-oral transmission potentials.

An opportunity to test the validity of this model came when MERS-CoV first struck the Middle East in 2012. Upon the availability of the sequences of both MERS-CoV M and N proteins, the authors were able to confirm that MERS-CoV falls solidly into a category of viruses that have lesser respiratory transmission potential but have greater fecal-oral transmission capability [6,7]. The categorization of MERS-CoV within this category was supported by the observations that MERS-CoV

* Corresponding author.

E-mail address: gohsbiocomputing@yahoo.com (G.K.-M. Goh).

<https://doi.org/10.1016/j.micpath.2020.104177>

Received 17 February 2020; Received in revised form 18 March 2020; Accepted 27 March 2020

Available online 31 March 2020

0882-4010/ © 2020 Elsevier Ltd. All rights reserved.

Table 1

Categorization of coronaviruses by shell disorder to predict levels of Respiratory and Fecal-oral transmission (Statistical Analyses:Two-Way ANOVA, $p < 0.00$), (Regression analysis of category as dependent variable with M and N PIDs as independent variables: $p < 0.001$, $r^2 = 0.83$, Regression analysis with only N PID as an independent variable: $p < 0.001$, $r^2 = 0.77$).

Shell Disorder Category	Coronavirus	M PID % (UniProt/Genbank Accession Code) ^a	N PID % (UniProt/Genbank Accession Code) ^b	Remarks
A	HCoV-229E	23(P15422)	56(P15130)	Higher Levels of Respiratory Transmission Lower Levels of Fecal-oral Transmission
	IBV(Avian)	10(P69606)	56(Q8JMI6)	
B	Bovine	7.4(P69704)	53(Q8V432)	Intermediate Levels of Respiratory and Fecal-oral Transmission
	PEDV(Porcine)	8(P59771)	51(Q07499)	
	Canine(Resp.)	6.5(A3EXD6)	51(A3E2F7)	
	HCoV-OC43	7(Q4VID2)	51(P33469)	
	SARS-CoV	8(P59596)	50(P59595)	
	HCoV-NL63	11(Q6Q1R9)	49(Q6Q1R8)	
	Bat-HKU4	16(A3EXA0)	48(A3EXA1)	
	SARS-Cov-2 ^c	6(QHD43419.1) ^d	48(QHD43423.2) ^d	
	Bats	11.5(A3EXD6)	47(Q3LZX4)	
Bat-HKU5	11(A3EXD6)	47(A3EXD7) ^b		
C	MHV(Murine)	8(Q9JEB4)	46(Po3416)	Lower Levels of Respiratory Transmission Higher Levels of Fecal-oral Transmission
	MERS-CoV	9(K0BU37)	44(K0BVN3)	
	TGEV(Porcine)	14(P09175)	43(P04134)	
	Canine(Ent.)	8(B8RIR2)	40(Q04700)	
	HCoV-HKU1	4(Q14EA7)	37(Q0ZME3)	

^a M PID refers to the percentage of intrinsic disorder (PID) of the membrane protein (M). PID is measured by the number of residues predicted to be disordered divided by the total number of disorder. M PID predicts how long the virus can remain outside the body. M is considered to be one of the outer shell.

^b N protein refers to the nucleocapsid protein, which is an inner shell protein.

^c SARS-CoV-2 (COVID).

^d Genbank accession code, <https://www.ncbi.nlm.nih.gov/nuccore/MN908947>.

is not efficiently spreading among humans, unlike SARS-CoV, and requires close contact [8]. Furthermore, it is now known that camels are the natural reservoir for MERS-CoV, and viruses that are associated with farm animals, such as camels, tend to have greater fecal-oral transmission potentials [9].

In this study, we will have the opportunity to once again test the model. This time, it will involve the SARS-CoV-2. We shall see that not only the model is reliable and consistent, but is able to detect something very odd about this virus that could account for its quick spread.

2. Methods

2.1. Protein intrinsic disorder predisposition and PONDR® VLXT

Protein intrinsic disorder is found when portions of or an entire protein has no structure. It is also called by other names such as unstructured, natively unfolded [10–12]. The main tool used to develop the CoVs shell disorder model involves the use of neural networks that had been trained to recognize ordered and disordered regions of a protein given its sequence. A suite of such artificial intelligence tools (AI) have been named PONDR® (<http://www.pondr.com>).

Of a particular interest. PONDR® VLXT has been found to be accurate for structural proteins with protein-protein and protein-RNA/DNA interactions [13–17]. In fact, PONDR® VLXT has shown consistency and reproducibility when used for the study of structural viral shells that protect the virion from environmental damage. A wide range of viruses have been studied using PONDR® VLXT including equine infectious anemia virus (EIAV), human immunodeficiency virus (HIV), herpes simplex virus (HSV), smallpox virus, polio virus, Nipah and Ebola viruses, SARS-CoV, MERS-CoV, influenza virus, hepatitis C virus (HCV), hepatitis A virus (HAV), yellow fever virus (YFV), and Zika virus [2,5–7,18–27].

2.2. M and N proteins

The coronavirus virion contains four major structural proteins: the spike (S) protein, nucleocapsid (N) protein, membrane (M) protein, and

the envelope (E) protein [5–7,28–30]. These proteins are needed for production of a structurally complete viral particle [30–33]. The main focus of our intrinsic disorder-based model is at the viral M and N proteins. The M protein of CoV was chosen over the E protein because the M protein is a major transmembrane protein that is found in large numbers in the virion, whereas E protein is a minor protein of the envelope. Also, in previous studies, membrane or matrix proteins of various viruses were analysed with respect to their intrinsic disorder content and roles in protecting the respective virions. The N protein is also important for our model, as it has been shown that greater disorder in the inner shell is associated with the mode of infection and virulence in other viruses [23–26].

2.3. Disorder and statistical analyses

The protein sequences of the selected CoV M and N proteins were downloaded from UniProt (<http://www.uniprot.org>) The sequences were fed into a MYSQL database using JAVA [19]. The sequences were fed into the PONDR® VLXT, and the corresponding results were archived in the database. PONDR® VLXT provides intrinsic disorder predisposition scores for each residue. Residues with the scores of 0.5 and above are considered disordered, and, conversely, residues with the scores below 0.5 are taken as ordered. We represented the outputs of these analyses in the form of the percentage of intrinsic disorder (PID), which is defined as the number of residues predicted to be disordered divided by the total number of residues in a query protein. Statistical analyses using regression analysis and Analysis of Variance (ANOVA) were done using R statistical package [34]. Regression analyses were done using category as the dependent variable with M and N PIDs as independent variables or N PID as a sole independent variable. Categories were assigned the values of 1, 2 and 3 to the three groups as according to the level of respiratory transmission potential. The amino acid sequences of the SARS-CoV-2 M and N proteins are available at NCBI (<https://www.ncbi.nlm.nih.gov/nuccore/MN908947>).

3. Results

3.1. Statistical analysis indicates clustering of CoVs into three statistically identifiable categories

As aforementioned, the model predicting the mode of viral transmission was built before the 2012 MERS-CoV outbreak. This model was based on the analysis of mainly the PIDs found for N proteins, the viruses were clustered into three groups according to their transmission mode, while the M PIDs provide an indicator of resilience of the virus in the harsh environment. The three groups are represented in Table 1. A Two-Way ANOVA ($p < 0.001$) indicates that the groups are statistically identifiable. The Table shows that category A consists of viruses that have greater respiratory transmission capabilities but lower fecal-oral transmission potentials. Category B involves coronaviruses that have intermediate levels of both respiratory and fecal-oral transmission potentials, whereas category C comprises of viruses with lower respiratory transmission. Furthermore, a higher inner shell disorder is often associated with the greater infectivity, especially with regard to respiratory transmission potential since disorder allows greater promiscuity of binding [6,7].

3.2. SARS-CoV-2 is in the category of CoVs with intermediate fecal-oral and respiratory transmission potentials

Table 1 shows that SARS-CoV is grouped into category B along with other viruses that have intermediate respiratory and fecal-oral transmission potentials. MERS-CoV falls into category C (lower respiratory potential but higher fecal-oral transmission capabilities). These are consistent with what we know about transmission behaviour of SARS-CoV and MERS-CoV [5–8,35]. It should be noted that category A are viruses with N PID above 53%, whereas category C comprises of those with N PID below 47%. Since SARS-CoV-2 has an N PID of 48%, which is above 46% but below 54%, it must be considered as being in category B.

3.3. SARS-CoV-2 has one of the hardest outer shell among CoVs (M PID = 6%)

As already mentioned, based on the similar analysis of the SARS-CoV-2 (M PID = 6%, N PID = 48%), it is clear that this virus has to be placed in Category B, along with SARS-CoV (Table 1). However, our data also pointed to something odd about the SARS-CoV-2. In fact, Table 1 and Fig. 1 show that data for the SARS-CoV-2, the M PID is 6%, which is second lowest in our sample of a fairly diverse selection of

CoVs. This means that the SARS CoV-2 has one of the hardest outer shells among its counterparts.

3.4. Disorder and statistical analyses

The model is consistent with the theme that viruses that remain in harsh environments require harder; i.e., less disordered, shells to survive. Furthermore, the outer shell is likely to play an even greater role in protecting the virion [20,22,27]. This has important implications pertaining to the characteristics and behaviors of the virus as we shall see below.

4. Discussion

4.1. Evidence of consistency and reproducibility

A superficial glance at the model that places SARS-CoV and SARS-CoV-2 in the same category, with intermediate levels of respiratory transmission potentials may mislead the reader to believe that the model is not reliable, given the fact that we have now seen that SARS-CoV-2 is much more contagious than SARS-CoV. We believe that this is not the case. If SARS-CoV-2 would be placed in category C, the same category as MERS-CoV, then the model would be definitely wrong in terms of reproducibility. On the other hand, the placing of SARS-CoV-2 in category A would be an inconsistency that would be difficult to explain. One would further notice that very few CoVs fall into category A, with the exception for HCoV-229E and infectious bronchitis virus (IBV), which make the chance of any CoVs including SARS-CoV-2, to be in category A slim. More importantly, because SARS-CoV-2 is zoonotic and the only non-human CoV in category A is the avian IBV, placing SARS-CoV-2 in category A could suggest that this virus is likely to be of avian origin, which is inconsistent with what we know about SARS-CoV-2. Phylogenetic study has confirmed that the closest relatives of SARS-CoV-2 are bat CoVs [4,36]. We also noticed that at least three of closest neighbors with similar N PIDs of 47–48% are those of bat origin (See Table 1). This suggests that bat CoVs are likely to have an optimal mix of fecal-oral and respiratory transmission potentials that reflect their N PID range of 47–48%, which provides for greater fitness for the spread of the viruses among bats and that SARS-CoV-2 is likely to be of bat origin too given its N PID is 48%. There are highly debatable suggestions that snakes or pangolins [3,37,38] may be an intermediary host of the virus. Even if any of these is later deemed as true, our data imply that the virus had not the chance to evolved much with any intermediary host except bats, unlike the cases of SARS-CoV and MERS-CoV, in which the model does suggest that they had evolved with hosts such

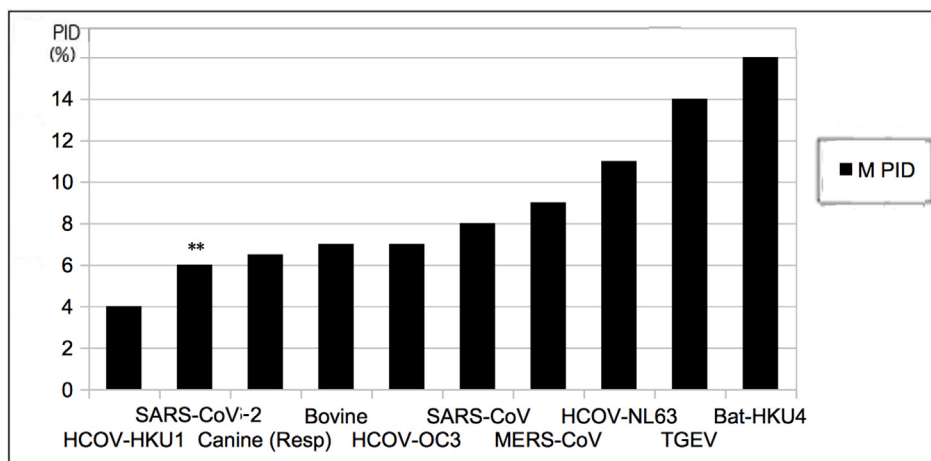


Fig. 1. Level of disorder (PID ie Percentage of Intrinsic Disorder) for the M protein. M PID for SARS-CoV-2 (**) is among the lowest in its family, which is indicative of its hard outer shell.

as civet cats and camels respectively. This suggestion is, however, based on N PID alone.

4.2. Evidence of correlations between modes of transmission and shell disorder

The CoV shell disorder model was originally conceived as part of a spin-off from a parent research project that studied relationships among the mode of transmission, shell disorder and the absence of vaccines for certain sexually transmitted viruses such as HIV, HSV and HCV [19,20,27]. At that time even after 2003 SARS-CoV, the wealth of CoV knowledge lies in the field of veterinary medicine [5,6]. Using what was then known about behaviors of porcine CoVs, a successful attempt to correlate M and N PIDs to modes of transmission was made [5,6]. The results were similar to those found in Table 1 and Fig. 1 without the data pertaining to MERS-CoV and SARS-CoV-2, of course, since the MERS-CoV and COVID-19 outbreaks had yet to occur then.

Evidence of a strong correlation of M and N PIDs to the mode of transmission ($r^2 = 0.83$, $p < 0.001$) can be seen in the regression analysis. Furthermore, a strong but oddly weaker correlation ($r^2 = 0.77$, $p < 0.001$) can also be found when N PID is the sole independent variable even if poor correlation ($r^2 = 0.21$) is found when M PID is the sole independent variable. This basically means that the model should include both M and N PIDs as the statistical model seems to suggest that M PID does contribute, even if slightly, to the mode of transmission in conjunction to the major contribution by N PID. This is consistent with the idea that both M and N play roles in protecting the virion from environmental damage and the need to protect the virion is dependent on mode of transmission. A qualitative inspection of the viruses in the three groups also reveals consistency. For example, as in Table 1, canine CoV-respiratory falls into category B, while canine CoV-enteritis is in category C. The porcine CoVs are also consistent with observation of their behaviors as we will later see.

We need to keep in mind, however, that the outer shell is likely to play a greater and more immediate role in protecting the virion from environmental damage as it encases the entire virion, whereas the N protein is likely to protect only the RNA. M PID is therefore a more crucial indicator of the “hardness” of the virion. However, because of the higher correlation of N to the mode of transmission, our model is apparently suggesting that some other mechanisms are involved. In fact, studies of inner shell of other viruses such as Nipah virus, Ebola virus and flaviviruses [23–26] have indicated strong correlations between inner shell disorder and virulence. In the case of Nipah virus, the virulence and inner shell disorder were linked to the modes of transmission [26].

4.3. Greater viral loads in body fluids may be necessary for respiratory transmission: Roles of N disorder and protein binding promiscuity

The mechanism in which the virus acquires greater virulence via inner shell disorder arises from the ability of the viral protein to bind promiscuously to the host protein [6,23–26]. This ability provides for rapid replication of the viral proteins and particles. This quick replication is also an efficient way to evade the host immune system by reproducing a large number of viral particles before the immune system even recognizes the virus, which then often goes on to overwhelm the body leading to the death of the host [6,23–26]. One highly plausible explanation for the link between rapid replication and the modes of transmission is that viral load in the body fluids of the infected host needs to be sufficiently high before the virus can be infectious via the respiratory mode. This therefore explains the fact that rapid viral replication and inner shell disorder were found to be the common link between virulence and respiratory transmission in the case of Nipah virus [26].

We don't, however, know if there is correlation between CoV virulence and inner shell disorder, even though the authors believe it exists,

just like the other viruses. If so, N will make an excellent target for CoV vaccine development. The difficulty in the search for such links is that CoVs infect a large variety of animals and each CoV often uses a different receptor [4,6,39]. For example, MERS-CoV has a case-fatality rate of above 30% [8] for humans but is generally harmless to camels [27]. Furthermore, CoVs were generally not known to be virulent before the 2003 SARS-CoV outbreak and therefore lack adequate data for us to work with.

The reason that we can make the necessary above-mentioned extrapolations pertaining to inner shell disorder among unrelated viruses with somewhat greater confidence is that inner shell proteins of different species of viruses often share similar functions and structures [30,40,41]. For example, the CoV N proteins, like the nucleocapsids in other viruses, are responsible for assembling viral particles by forming complexes with other viral proteins and RNA [24,25,41]. The N protein also helps with the packaging and budding of viral particles in the host ER (Endoplasmic Reticulum) [23,30,40]. Such tasks require binding to host proteins, and the promiscuous protein binding capabilities of a more disordered N will certainly help towards rapid replication of the virus.

4.4. Hard outer shell might contribute to the high contagiousness of SARS-CoV-2

Since both SARS-CoV and SARS-CoV-2 have intermediate respiratory transmission potentials, how do we account for the greater contagiousness of the latter? We need to understand that the respiratory transmission potential of a single viral particle is just one of several factors that might contribute to the contagiousness of the virus. Other factors include how many particles are released by the host and how long does the particle remain in the environment. We mentioned that our model has detected something odd about SARS-CoV-2: this virus has one of the hardest outer shell within the family as seen in Fig. 1 and Table 1 (M PID = 6%). Since the outer shell plays the greater role in protecting the virion in comparison to the inner shell, the harder outer shell provides virus greater resilience to outside the body environment and to the presence of digestive enzymes found in the saliva, mucus, and other bodily fluids [20,22,42–46]. As a result, the virus with the harder outer shell is able to remain active for a longer time and, therefore a lesser number of viral particles is required for a chance to infect someone. Furthermore, because the virus is more resistant to the digestive enzymes in the bodily fluids, an infected body is likely to discharged more infectious particles.

Evidence of the protective role of the outer shells be seen in a wide variety of viruses. Viruses associated with saliva (e.g., YFV, ZIKV, EIAV, rabies) or fecal-oral transmission (e.g., poliovirus), have hard outer shells with low PID, whereas sexually transmitted viruses (e.g., HIV, HSV-2, HCV) have higher outer shell PIDs [6,20–27]. Also, viruses that are notorious for lasting in the environment for a long time such, as the smallpox virus, have low outer shell PIDs.

4.5. Lessons from the porcine cousins: PEDV vs. TGEV

Before the SARS-CoV outbreak, coronaviruses were not considered medically important, as, in the past, they had been primarily associated with cold viruses that cause minor sniffles. This is, however, not true in the field of veterinary medicine, where coronaviruses have posed a menace to livestock. This is especially so in the case of TGEV (transmissible gastroenteritis virus) and PEDV (porcine epidemic diarrhea virus). TGEV could move rapidly among farm pigs to quickly devastate the farming community if not controlled early [5–7]. Our analysis reveals that TGEV is of category C (higher fecal-oral, lower respiratory transmission potentials) with M and N PIDs at 14% and 43% respectively. The reason that it is able to move rapidly is that its high fecal-oral transmission potential represents a more efficient mode of transmission among the farm animals.

There is an antigenically related cousin, PEDV, which has similar characteristics as TGEV [29]. Like TGEV, diarrhea and vomiting are main symptoms of PEDV infection. While TGEV is highly infectious, PEDV is less contagious among farm pigs. A peculiar characteristic has been observed by farmers and veterinarians: PEDV can reappear out of nowhere from within the same pen that was previously occupied by infected pigs, even after the disease has passed on several months ago. An inspection of the M and N PIDs reveals the reason showing that PEDV belongs to the category B, along with SARS-CoV and SARS-CoV-2, unlike TGEV and MERS-CoV. This means that PEDV has a higher respiratory transmission potential than TGEV and that implies greater chances of fecal-respiratory transmission for the former. Obviously, viral particles in fecal materials that was somehow inadvertently left behind go on to infect a new population of pigs several months later [5–7].

4.6. Airborne virus from feces, mucus, vomit and saliva: A hidden foe

Infectious droplets from mucus, vomit, or feces have actually been seen during the SARS-CoV outbreak. A stark example happened in a 2003 outbreak at Amoy Gardens, a housing complex in Hong Kong, where a huge cluster of infected patients was found and the inefficient sewerage and toilet ventilation systems were responsible for the spread [5,35,47]. We need to be reminded that PEDV, SARS-CoV, and SARS-CoV-2 are all under the same category B, and that SARS-CoV-2 has the hardest outer shell within the entire CoV family. SARS-CoV-2 is thus likely to be able remain infectious in the environment for the longest time regardless of being in feces, mucus, vomit, sweat, or saliva.

4.7. Evidence of advantage of greater resilience in the environment: PEDV

A puzzling thing can be observed is when we notice that PEDV (M PID = 8%) has a harder outer shell than TGEV (M PID = 14%), given that TGEV is predicted as having higher fecal-oral transmission potential. We would think that having greater fecal-oral potential requires harder outer shell among CoVs. Apparently, the result is implying that this is not necessarily true, as we have seen, and having higher fecal-oral transmission potential can actually provide greater efficiency in the spread among farm animals, as we have seen in TGEV, such that the virus does not have to stay long outside the body. The ability to stay longer in the environment could, however, provide advantages to a virus with greater respiratory transmission potential, especially with respect to spread via airborne virus from feces or body fluids. This is likely the case for SARS-CoV-2.

5. Conclusion

We have presented here a set of data that not only place the SARS-CoV-2 within the category of CoVs that have intermediate levels of both respiratory and fecal-oral transmission potentials, alongside SARS-CoV and PEDV, but have also predicted SARS-CoV-2 to have one of the hardest outer shells among most CoVs. This peculiarity is likely responsible for its high level of contagiousness, since hardness of its outer shell could provide the virus with the greater resilience to the conditions outside the body and in body fluid, as the harder shell will better protect the virion from damage as a result of the hostile environment and action of the digestive enzymes found in bodily fluids. As a result, it is likely that the infected body can shed more infectious particles that have a greater chance of infecting a person over its lifetime. Chances of infections via indirect contact and airborne virus from feces and bodily fluids are therefore greater. The results of our research could have important implications for epidemiologists and public health officials.

Author contributions

GKMG. conceived the idea, collected and analyzed the data and

wrote the first draft. VNU. helped with the collection and analysis of literature data, reviewed and revised the draft. AKD and JAF reviewed the manuscript and provided the resources necessary for the research.

Declaration of competing interest

GKMG. is an independent researcher and the owner of Goh's BioComputing, Singapore.

GKMG. has also written a book ("Viral Shapeshifters: Strange Behaviors of HIV and Other Viruses") on a related subject. The authors have no other potential conflict of interests.

References

- [1] WHO, Novel coronavirus (2019-nCoV), <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
- [2] G.K. Goh, A.K. Dunker, J.A. Foster, V.N. Uversky, Rigidity of outer shell predicted by protein disorder model sheds light on COVID-19(Wuhan-2019-nCoV) infectivity, *Biomolecules* 199 (2020) e221.
- [3] X. Li, J. Zai, Q. Zhao, Q. Nie, Y. Li, B.T. Foley, A. Chaillon, Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2, *J. Med. Virol.* 27 (2020), <https://doi.org/10.1002/jmv.25731>.
- [4] P. Zhou, X.L. Yang, X.G. Wang, B. Hu, L. Zhang, W. Zhang, et al., A pneumonia outbreak associated with a new coronavirus of probable bat origin, *Nature* 579 (2020) 270–273, <https://doi.org/10.1038/s41586-020-2012-7>.
- [5] G.K. Goh, A.K. Dunker, V.N. Uversky, Understanding viral transmission behavior via protein intrinsic disorder prediction: Coronaviruses, *J. Pathog.* 2012 (2012) 738590.
- [6] G.K. Goh, *Viral Shapeshifters: Strange Behaviors of Hiv and Other Viruses*, Simplicity Research Institute, Singapore, 2017.
- [7] G.K. Goh, A.K. Dunker, V.N. Uversky, Prediction of intrinsic disorder in MERS-CoV/HCoV- EMC supports a high oral-fecal transmission, *PLoS Curr.* 5 (2013).
- [8] 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)).
- [9] M. Ferguson, M.D. Van Kerkhov, Identification of MERS-CoV in dromedary camels, *Lancet Infect. Dis.* 14 (2014) 93–94.
- [10] P. Tompa, Intrinsically unstructured proteins, *Trends Biochem. Sci.* 27 (2002) 527–533.
- [11] V.N. Uversky, J.R. Gillespie, A.L. Fink, Why are "natively unfolded" proteins unstructured under the physiological conditions? *Proteins Struct. Funct. Genet.* 41 (2000) 415–427.
- [12] P.E. Wright, H.J. Dyson, Intrinsically unstructured proteins: Re-assessing the protein structure-paradigm, *J. Mol. Biol.* 293 (1999) 321–331.
- [13] X. Li, P. Romero, M. Rani, A.K. Dunker, Z. Obradovic, Predicting protein disorder for N-, C-, and internal regions, *Genome Inform. Ser. Workshop Genome Inform.* 10 (1999) 30–40.
- [14] E. Garner, P. Romero, A.K. Dunker, C. Brown, Z. Obradovic, Predicting binding regions within disordered proteins, *Genome Inform.* 10 (1999) 41–50.
- [15] P. Romero, Z. Obradovic, X. Li, E.C. Garner, C.J. Brown, A.K. Dunker, Sequence complexity of disordered protein, *Proteins* 42 (2001) 38–48.
- [16] Y. Cheng, C.J. Oldfield, J. Meng, P. Romero, V.N. Uversky, A.K. Dunker, Mining alpha-helix-forming molecular recognition features with cross species sequence alignments, *Biochemistry* 46 (2007) 13468–13477.
- [17] C.J. Oldfield, Y. Cheng, M.S. Cortese, P. Romero, V.N. Uversky, A.K. Dunker, Coupled folding and binding with alpha-helix-forming molecular recognition elements, *Biochemistry* 44 (2005) 12454–12470.
- [18] B. Xue, R.W. Williams, C.J. Oldfield, G.K. Goh, A.K. Dunker, V.N. Uversky, Viral disorder or disordered viruses: do viral proteins possess unique features? *Protein Pept. Lett.* 17 (2010) 932–951.
- [19] G.K. Goh, A.K. Dunker, V.N. Uversky, Protein intrinsic disorder toolbox for comparative analysis of viral proteins, *BMC Genom.* 9 (Suppl. 2) (2008) S4.
- [20] G.K. Goh, A.K. Dunker, V.N. Uversky, A comparative analysis of viral matrix proteins using disorder predictors, *Virol. J.* 5 (2008) 126.
- [21] G.K. Goh, A.K. Dunker, V.N. Uversky, Protein intrinsic disorder and influenza virulence: the 1918 H1N1 and H5N1 viruses, *Virol. J.* 6 (2009) 69.
- [22] G.K. Goh, A.K. Dunker, V.N. Uversky, Shell disorder, immune evasion and transmission behaviors among human and animal retroviruses, *Mol. Biosyst.* 11 (2015) 2312–2323.
- [23] G.K. Goh, A.K. Dunker, V.N. Uversky, Detection of links between Ebola nucleocapsid and virulence using disorder analysis, *Mol. Biosyst.* 11 (2015) 2337–2344.
- [24] G.K. Goh, A.K. Dunker, V.N. Uversky, Correlating flavivirus virulence and levels of intrinsic disorder in shell proteins: protective roles vs. Immune evasion, *Mol. Biosyst.* 12 (2016) 1881–1891.
- [25] G.K. Goh, A.K. Dunker, J.A. Foster, V.N. Uversky, Zika and flavivirus disorder: virulence and fetal morbidity, *Biomolecules* 9 (2019) e710.
- [26] G.K. Goh, A.K. Dunker, J.A. Foster, V.N. Uversky, Nipah shell disorder, mode of transmission and virulence, *Microb. Pathog.* 141 (2020) 103976.
- [27] G.K. Goh, A.K. Dunker, J.A. Foster, V.N. Uversky, HIV vaccine mystery and viral shell disorder, *Biomolecules* 9 (2019) 178.
- [28] N. Acheson, *Fundamentals of Molecular Virology*, Wiley, 2007.

- [29] C.M. Lin, X. Gao, O. Tomochir, A.N. Vlasova, M.A. Oeselli, L. Saif, Antigenic relationships among porcine epidemic diarrhea virus and transmissible gastroenteritis virus strains, *J. Virol.* 89 (2015) 3332–3342.
- [30] R. McBride, M. van Zyl, Fielding BC the coronavirus nucleocapsid is a multi-functional protein, *Viruses* 6 (2014) 2991–3018, <https://doi.org/10.3390/v6082991>.
- [31] P.S. Masters, The molecular biology of coronaviruses, *Adv. Virus Res.* 66 (2006) 193–292.
- [32] E. Mortola, P. Roy, Efficient assembly and release of SARS coronavirus-like particles by a heterologous expression system, *FEBS Lett.* 576 (2004) 174–178.
- [33] C. Wang, X. Zheng, W. Gai, Y. Zhao, H. Wang, H. Wang, et al., MERS-CoV virus-like particles produced in insect cells induce specific humoral and cellular immunity in rhesus macaques, *Oncotarget* 8 (2017) 12686–12694.
- [34] R Core Team, R: A Language and Environment for Statistical Computing, (2019) Vienna, Austria.
- [35] T. Abraham, Twenty-first Century Plague: the Story of SARS. Hong Kong, Hong Kong University Press, 2004.
- [36] R. Lu, X. Zhao, J. Li, et al., Genomic characterisation and epidemiology of 2019 novel coronavirus, *Lancet* (20) (2020) 30251–30258, [https://doi.org/10.1016/S0140-6736\(20\)30251-8](https://doi.org/10.1016/S0140-6736(20)30251-8) pii: S0140-6736.
- [37] H.H. Fan, L.Q. Wang, W.L. Liu, X.P. An, Z.D. Liu, X.Q. He, L.H. Song, Y.G. Tong, Repurposing of clinically approved drugs for treatment of coronavirus disease 2019 in a 2019-novel coronavirus (2019-nCoV) related coronavirus model, *Chin. Med. J.* 6 (2020), <https://doi.org/10.1097/CM9.0000000000000797>.
- [38] W. Ji, W. Wang, X. Zhao, J. Zai, X. Li, Cross-species transmission of the newly identified coronavirus 2019-nCoV, *J. Med. Virol.* 92 (2020) 433–440, <https://doi.org/10.1002/jmv.25682>.
- [39] D.R. Adney, M. Letko, I.K. Ragan, D. Scott, N. van Doremalen, R.A. Bowen, V.J. Munster, Bactrian camels shed large quantities of Middle East respiratory syndrome coronavirus (MERS-CoV) after experimental infection, *Emerg. Microb. Infect.* 8 (2019) 717–723, <https://doi.org/10.1080/22221751.2019.1618687>.
- [40] Y. Liu, S. Stone, R.N. Harty, Characterization of filovirus protein-protein interactions in mammalian cells using bimolecular complementation, *J. Infect. Dis.* 204 (Suppl 3) (2011) S817–S824, <https://doi.org/10.1093/inf>.
- [41] R. Bhuvanankantham, M.K. Chong, M.L. Ng, Specific interaction of capsid protein and importin alpha/beta influences West Nile virus production, *Biochem. Biophys. Res. Commun.* 389 (2009) 63–69, <https://doi.org/10.1016/j.bbrc.2009.08.108>.
- [42] Malamud, W. Abrams, C. Barber, D. Weissman, M. Rehtanz, E. Golub, Antiviral activities in saliva, *Adv. Dent. Res.* 23 (2011) 34–37.
- [43] J. Hankiewicz, C. Swierczek, Lysozyme in body fluids, *Clin. Chim. Acta* 57 (1974) 205–209.
- [44] S.A. Ragland, A.K. Cris, From bacterial killing to immune modulation: recent insights into functions of lysozyme, *PLoS Pathog.* 13 (2019) e1006512.
- [45] J. Schaubert, R.L. Gallo, Antimicrobial peptides and skin defense immune system, *J. Allergy Clin. Immunol.* 122 (2009) 261266.
- [46] A.M. Cole, P. Dewan, T. Ganz, Innate antimicrobial activity of nasa secretions, *Infect. Immun.* 67 (1999) 3267–3275.
- [47] I.T. Yu, Y. Li, T.W. Wong, W. Tam, T. Chan, J.H. Lee, B.Y. Leung, T. Ho, Evidence of airborne transmission of the Severe Acute respiratory syndrome virus, *NEJM* 350 (2004) 1731–1739.