Review

On the Origin Debate and Plausible Future Endeavours of COVID-19 (Commentary)

Debdatta Nandy , Bijit Chatterjee , Sankar Basu *

Posted Date: 26 October 2023

doi: 10.20944/preprints202310.1717.v1

Keywords: SARS-CoV-2; COVID-19; murky origins; furin like cleavage sites (FLCS); spike; combination multi-epitope vaccines

Review

On the Origin Debate and Plausible Future Endeavours of COVID-19

Debdatta Nandy [1], Bijit Chatterjee [2], Sankar Basu [3,4,*]

[1] Department of Biochemistry, Asutosh College, 92, Shyama Prasad Mukherjee Rd, Jatin Das Park, Patuapara, Bhowanipore, Kolkata, West Bengal 700026, (affiliated to University of Calcutta), India; debdattanandy@gmail.com

[2] Department of Environmental Sciences, Asutosh College, 92, Shyama Prasad Mukherjee Rd, Jatin Das Park, Patuapara, Bhowanipore, Kolkata, West Bengal 700026, (affiliated to University of Calcutta), India; bijitchatterjee02@gmail.com

[3] Department of Microbiology, Asutosh College, 92, Shyama Prasad Mukherjee Rd, Jatin Das Park, Patuapara, Bhowanipore, Kolkata, West Bengal 700026, (affiliated to University of Calcutta), India

[4] 3BIO group, University de libre Brussels, Belgium (External Scientif Collaborator)

* Correspondence: sankarchandra.basu@asutoshcollege.in

Abstract: Globally, we are undergoing compelling changes in virtually every branch of the society and every aspect of our daily lives over the last couple of years due to the COVID-19 pandemic. In the words of Melinda Gates, the famous American Philanthropist, the pandemic has literally "magnified every existing inequality". There has been much talk and an ongoing debate on the origin of SARS-CoV-2 (or, the novel coronavirus) as to whether have evolved purely naturally or has the contaminating element of human intervention somewhere in its trajectory. Proponents of the pure 'natural evolution theory', some of whom are bigshots in the field of virology have seem to take up great effort to stamp their authority right from the very early days of the pandemic on the slow natural evolution of the virus from its ancestrors by several publications in high-profile journals. These well-cited papers however provide rationals that are somewhat contradictory and are backed up with data which maybe argued as far from being complete and/or all-inclusive. As time progressed, alternative theories to that of the pure natural evolution of the virus emerged with strong counter-logics and gradually compiling genomic and epidemiological data along with detailed structural studies of relevant molecular interactions. All said and done, till date, compelling evidences lack on either end of the spectrum making the COVID origin debate still very mucg wide open. Regardless of whether purely natural or human-intervented the damage caused by the virus is unmistakable and has been an alarming concern to medical science lately. To that end, a section of the scientific community is devoting committed efforts using immuno-informatics to design and develop possible multi-epitope combination vaccines to cover a plethora of viral antigens and their plausible future mutations, alongside with the vaccines that are currently available.

Keywords: SARS-CoV-2; COVID-19; murky origins; furin like cleavage sites (FLCS); spike; combination multi-epitope vaccines

---

"Mother Nature is a self-taught engineer."

— Stewart Stafford, Actor, Hollywood

"Covid has magnified every existing inequality"

— Melinda Gates, American Philanthropist

"We long to return to normal, but normal led to this. To avert the future pandemics we know are coming, we MUST grapple with all the ways normal failed us. We have to build something better. I hope this piece, in showing what went wrong, helps."

— Ed Yong, Science Journalist, The Atlantic

 doi:10.20944/preprints202310.1717.v1

It is very much mysterious to understand how a small cluster of coronaviruses ended up sweeping the whole world. On December 31, 2019, the World Health Organization (WHO) announced and made us aware of a 'viral pneumonia' in Wuhan (Tabibzadeh et al., 2021), subsequently thereafter characterized as a novel member of the severe acute respiratory syndrome coronavirus clade, namely, SARS-CoV-2 (Andersen et al., 2020, 2). The consequent diseased state was named COVID-19 which exhibited an exponentially amplified rate of human transmission compared to all earlier onsets of evolutionary ancestral variants of the virus, leading to a pandemic and causing a death toll of over 3 million people worldwide. The small invisible virus has immensely challenged public health, the economy, and food chains. The pandemic has raised alarming concerns about the whole basis of modern medical science and that of socioeconomic balance worldwide. The dynamics of the pandemic are multi-wave in nature which usually provokes medical research to tame the upcoming waves with gradual amplification of available data and knowledge (Cacciapaglia et al., 2021). Yet, the planet seemed to be already very much exhausted in its (finite) resources to confront the upcoming waves, having passed merely three waves. Though herd immunity (Randolph and Barreiro, 2020, 19) has been revealed to play a key barrier to the spread of the viral infection and severity (especially evident in the third world countries), individuals with lung and/or heart disorders have been severely affected by the virus, often causing death. There are also cases of previous coronavirus outbreaks in 2003 as well as in 2007 by SARS-CoV (or, SARS-CoV-1) and MERS respectively. Coronaviruses are enveloped viruses with a large, single-stranded, positive-sense RNA genome. The coronaviruses involved in the earlier outbreaks were seemingly much less hazardous and had been sequestered from individuals suffering from relatively mild respiratory infections, cold flue, etc. that includes alpha-coronaviruses: NL63, 229E and beta coronaviruses: HKU1, OC43 (Jiang et al., 2020).

Origin of SARS-CoV-2 – arguments in favor of the Natural evolution theory

There has been an ongoing debate on the evolutionary origin of SARS-CoV-2 – whether naturally evolved or engineered in a virology laboratory. After prolonged analysis, it was evident that the RBD_{Spike} sequence, favorable for host receptor binding in SARS-CoV-2 is different from those in earlier variants of the coronavirus lineage. Six critical amino acid positions have been identified in the RBD_{Spike} of SARS-CoV-2 that effectively (i.e., physically) bind to the human host receptors (hACE-2: human Angiotensin Converting Enzyme 2) out of which five were mutated with respect to SARS-CoV (Y442 → L455, L472 → F486, N479 → Q493, D480 → S494, T487 → N501). The proponents of the natural evolution theory argue on a not fully optimal (or, non-ideal) binding of the viral Spike to the host receptor is a signature of natural evolution (Andersen et al., 2020) – which has later been somewhat rationalized by implementation of the concept of complementarity in probing the RBD_{Spike}-ACE-2 binding across coronavirus variants (Basu et al., 2021). Proponents of the natural evolution theory also argue for the fact that a deadlier variant from one of the other reverse genetic systems of beta-coronavirus (Dudas et al., 2018) could have been used instead, if one were to engineer the virus in a laboratory to cause 'maximum harm'. For the progenitor of this deadly virus, the bats serve as the reservoir host. To that end, the RaTG13 virus isolated from *Rhinolophus affinis* bat shows 96% similarities with SARS-CoV-2 (Zhou et al., 2020). Although its RBD_{Spike} diverges from the RBD_{Spike} of SARS-CoV-2 resulting in a loose binding to hACE-2 (Wan et al., 2020). It has been recorded that the Malayan Pangolins (*Monis javanica*) imported into Guangdong Provinces harbors coronavirus analogous to SARS-CoV-2 (to which it shares the highest similarity) but not as much as to that of the RaTG13 virus. The polybasic cleavage site(s) and predicted O-linked glycans procured across different coronavirus variants have been characterized to undergo a series of mutations: non-synonymous substitutions, insertions and deletions at their S1/S2 junctions – which also emphasizes the natural evolution theory of SARS-CoV-2 wherein the mutations occur during the human passage of the virus.

Another viewpoint on how the coronavirus has emerged is said to be grasped by connecting the events at the places of seafood market and the virology lab. People suspect that the virus may be originated from a virus that intruded into an intermediate host and finally transmitted to human. The

 doi:10.20944/preprints202310.1717.v1

Wuhan seafood market is the epicenter of the virus where wild animals like bamboo rats, hedge wags, and hog badgers are sold as food items. The wild lives are raised on farms in an internally congested area where it's easy for animals to grow viruses and spread from pillar to post among each other. The cross-species barrier is feeble in viruses which makes them enable to undergo horizontal gene transfers and genome fusions oftento mutate fast. Furthermore, the ss-RNA genome uses the virus-coded RNA-dependent RNA polymerase (Gong, 2021) as a machinary to replicate which, unlike DNA polymerases, lacks the proof-reading activity and therefore facilitates the mutational influx.

Some wildlife that is pliable to coronavirus infections is devoured in China either for food or used in traditional medicine. The very first cases reported in China infected by SARS-CoV-2 were either the vendors who worked at the wet market, the shopkeepers, or the customers. China's south-westernmost province Yunnan is the habitat of many coronavirus-carrying bats that resides in an ancient cave (Karunarathna et al.). It is the dwelling place of the coronaviruses which closely resembles the pathogen SARS-CoV, the virus that caused the epidemic in 2003. The cave is very much congested and within it resides a large population of bats. The environment provides ample space for the viruses to blend and supply the niche for effective coronavirus hybridization. One of such hybrid viral offsprings was naturally selected to have suddenly intruded on us, the *homo sapiens*. Moreover, some wildlife farms are situated nearby that deliver animals to Wuhan which probably has made the virus move from a bat cave into the market transmitted by an intermediate host.

Not the early source but still worth to be mentioned – the Lab-leak hypothesis

Doubts, however, have been cast on the natural evolution theory of SARS-CoV-2 over the last couple of years. The most promising alternative hypothesis (as of now) is that of a lab escape (or, lab-leak) originating from the Wuhan Institute of Virology – a pioneer institute in virology worldwide with high-security laboratories (biosafety level 4) which has been engaged in collaborative research with the most proficient and reputed scientists internationally. A search for the intermediate host was going on as it was believed that according to the theory of zoonotic transfer, animals are used as the reservoirs of pathogenic viruses. The most victimized creatures are the bats which seemed to be rupturing the species barrier after intruding into an intermediate host and passing on the viral pathogen. The same phenomenon had been observed during the previous eruptions, 2003 (SARS-CoV) and 2007 (MERS) – wherein civets and camels played the role of the intermediate hosts respectively.

Scientists targeted the Wuhan wet market in China and conducted a survey there, mainly pointing out a small number of alive animals, suspected to serve as the intermediate host. But no trace of the pathogen was found in the animals so far. They concluded that this is a direct transmission from bat to human, but it remains only speculation without any authentication (or, documentary proof) in its support. The assumption that the virus reared through a poisonous food chain is extremely unlikely. These failures piled up indirect support in favor of the lab escape hypothesis (Balaram, 2021). Scientists from the Wuhan biosafety laboratories working on the 'gain of functions' mutation of SARS-CoV (2003) frequently visited the bat caves and collected samples for their ongoing experiment (Dance, 2021). An obvious argument in favor of a possible lab-leak hypothesis at the Wuhan Institute hence is that these people might have got pre-exposed to the virus during such visits and that the viral emission this way, allowed the human passage through inter-individual transmission. People, much fond of conspiracy theories hence called the virus a 'Chinese virus'. 'Gain of functions' studies of virus relies on repetitive passage of the virus through altering animal cultures and the evolutionary selection pressure to eventually hit the desired target. To that end, the American scientific establishment and eminent scientists (proponents of the natural evolution theory) postulated that the virus acquired its detrimental arms through adaptation, undergoing slow mutations due to the raised selection pressure, made to work in favor of an evolved pernicious human pathogen from its bat origin. Scientists allegedly claimed that the sequence of the SARS-CoV-2 glycoprotein is the possible cause behind its high pathogenicity. The wall of resistance collapsed when a deeply scrutinized article was published in the Bulletin of Atomic scientists by Nicholas Wade (The origin of COVID: Did people or nature open Pandora's box at Wuhan?, 2021), a long serving-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 October 2023 doi:10.20944/preprints202310.1717.v1

man of science journalism. Examining Furin cleavage site, the legendary biochemist David Baltimore also said that it was the smoking gun containing the arginine codons (Balaram, 2021).

The genetic material of SARS-CoV-2 is RNA which is injected into the host cells upon viral invasion (Masters, 2006). It is stereotypical for such invading viral pathogens to disrupt all the vital processes occurring in the living host to replicate. It's potential to cause the disease is influenced by the Spike protein interacting specifically with the hACE-2 receptor of the lung cells through its receptor-binding domain (RBD_{Spike}). The interaction between the two molecules is compulsory (or, necessary) but not sufficient for viral entry (Millet and Whittaker, 2018). A kinetically driven 'down-to-earth' conformational change of RBD_{Spike} (transformation from 'lying down' to 'standing up' conformation) activates the receptor-binding domain, upon its proximity to hACE-2 which then rises to the occasion and binds to the host receptor. This is one of the key steps in the viral host cell entry which is intricately and temporally coupled to a priming event of (Furin-like) host protease mediated proteolytic cleavage at the S1/S2 junction setting free the already 'up' RBD_{Spike} for the host receptor binding. Furin and other host proteases cleave the arginine-rich polybasic $FLCS_{Spike}$ in SARS-CoV-2 which remains a disordered loop in the unbound Spike and gets transiently ordered (exhibiting meta-stability) (Roy et al., 2022) by forming a dynamically interchangeable ionic bond network at the Spike – Furin interface that facilitates the cleavage immensely. Such polybasic motifs are largely absent in SARS-CoV making them only rarely cleaved by non-specific host proteases (Bertram et al., 2011; Yuan et al., 2017) and thus make a big difference in the transmissibility of the two events of viral invasion. The RBD_{Spike} has a β-barrel as its predominant secondary structure. Upon being cleaved and released free from the rest of the Spike, the RBD_{Spike} approaches the lung cell receptors and binds to them transiently by forming quasi-stable complexes due to poor electrostatic match at the RBD_{Spike}-hACE-2 interface. The interaction, however, is characterized by high mutual affinity due to optimal shape complementarity between the molecular partners (Basu et al., 2021). The transient nature of the interaction enables the RBD_{Spike}-hACE-2 binding to act as molecular switches for the subsequent membrane fusion and viral host cell entry. The standing up (or 'up') conformation of the RBD_{Spike} is preferable for binding with the befitting surface, docked to the solvent-exposed Spike binding site of hACE-2. Comparative cryo-EM studies have revealed that in the case of SARS-CoV (2003), the 'up' state is a prevailing state while in its younger analog, SARS-CoV-2, the 'up' state is only stimulated near hACE-2 (Cai et al., 2020) which otherwise helps the virus to escape the host immune surveillance (Walls et al., 2020). It is very much prominent that over time the younger homolog developed a piece of machinery facilitating the binding of RBD_{Spike} to hACE-2 in a more transient way (by modulating and lowering the electrostatic matching at their interface making the binding less stable) in SARS-CoV-2 in comparison with SARS-CoV. This phenomenon helps other adjoining cells to become more vulnerable to viral entries than in the earlier event (SARS-CoV). The other noticeable feature is that the virus secluded from the host's body shows a unique Furin cleavage site (consisting of the 'PRRAR' motif) like never before. Thus, we come to a central question of whether these pivotal differences between SARS-CoV-2 and its ancestor leave footprints of slow natural evolution or an inevitable human intervention somewhere in the pipeline. One of the key differences between the bat coronavirus (Bat RmYN01, Bat RaTG13), responsible for the disease caused in the early 21st century and SARS-CoV-2, is indeed the 'PRRAR' motif, the polybasic arginine-rich lucrative bait found on the $FLCS_{Spike}$ that attracts Furin to cleave effectively at the S1/S2 junction of the $FLCS_{Spike}$ (Matyášek and Kovařík, 2020). One other key question is how the viruses become so well adapted to a hostile environment of human host cells from the very beginning! Usually, viruses take a long time to get accustomed by gaining mutations one by one until it gets stronger to manifest the full-bloomed infection, though, in the case of SARS-CoV-2, such a lag-phase seems almost redundant. The virus seems to be familiar to the host environment from day-1 and very much well balanced in that environment to spread infections and bring about the demolition of the biochemical machinery of the host cell. It looks as if it is already very familiar with the host cell environment, this point merits attention. Such familiarity probably hints toward pre-exposure of the virus to the human cellular environment and leaves much scope to the 'lab escape hypothesis' to discuss and explain the phenomena. As a matter of fact, Chinese virologist Dr. Zheng Lee used to grow her viruses in

 doi:10.20944/preprints202310.1717.v1

humanized mice that carry functioning human genes, and tissues. Consequently, an engineered hACE-2 receptor in the streamline of these humanized mice would result in the production of threatening viruses to cause infection in human cells as it got acquired with the host cell before (Fujiwara, 2018). Such engineered hACE-2 receptors work as a switch to check the infectivity of the pathogen by introducing a new cleavage site to enhance human infection. In earlier coronaviruses, HKU1, and OCH3, it was observed that they contain highly basic cleavage sites, RRKRR and RRSRG, respectively. There are many more other proteases instead of Furin to accelerate the virus host cell entry but Furin does this more competently. The other (non-specific) proteases, for example, TMPRSS-2, lysosomal cathepsins, NE, PR3, Cat-G, NSPH etc. secreted by activated neutrophils moving around the viral pathogen also intrudes into the host cells (Shang et al., 2020, 2; Mustafa et al., 2021) while the new virus (SARS-CoV-2) had already gained capabilities to avoid the host immune surveillance by its 'down-to-earth' conformation of the RBD_{Spike}. Furin attacks the polybasic arginine-rich region of the $FLCS_{Spike}$, mainly the 'PRRAR' segment that protrudes out as a lucrative bait to attract Furin. Extending from comparative genomic and epidemiological studies (Walls et al., 2020; Chan and Zhan, 2021), doubts have been casted on possible human intervention (rather than raw nature) to select for the 'PRRAR' segment, arguing that the insertion of the sequence is only to humanize the bat. Such intervention can easily be executed by alluring nucleotides at the gene level or by genetic engineering. The 'PRRAR' segment is encoded by a 15 nucleotide long ORF (5'-CTCGGCGGGCACGT-3') wherein the two consecutive arginine residues have been encoded by CGG triplet codons which happen to be the least frequent codon used to code for arginine. But yet it is recruited to code for two adjacent arginine residues giving rise to $_{681}$PRRAR$_{685}$ in the SARS-CoV-2 $FLCS_{Spike}$ elevating its pathogenicity (Balaram, 2021). The arginine residues are therefore satirically referred to as Baltimore's smoking gun. A new variant found in India, titled B.1.617 further bears a Furin cleavage site that has a single 'critical' mutation (P681 → R681) making the motif more basic and disordered (Mlcochova et al., 2021). A single turnover of a letter in the coding triplet is enough for adding this extra disorder in the Spike so that Furin can cleave at the S1/S2 function even more smoothly. People have argued that this perhaps adds a more authentic signature of human intervention in designing the virus (Balaram, 2021). It is however still not clear as to which of the two events is more critical to elevate the viral pathogenicity - the transient binding of RBD_{Spike} to hACE-2 or the Furin mediated cleavage at the $FLCS_{Spike}$.

Paired comparison analysis: Reasons behind the emergence of SARS-CoV-2

There also exist alternative perspectives on the origin of SARS-CoV-2 further adding to its complexity. Some extreemists even speculated that the virus might have been made to evolve to be utilized as biological weaponry. Again, some talk in favor of zoonotic spillover models (Plowright et al., 2017; Frutos et al., 2021) which is a multivariate complex interplay of ecological, epidemiological, behavioral and other possible determinants of pathogen exposure, as well as the human internal factors leading to the varying degree of susceptibility of individuals to infection. The latter of the two events must also involve overcoming a hierarchical series of barriers to cause spillover infections in humans (Plowright et al., 2017). Extending on the first of the two hypotheses, the obvious question was whether the virus was deliberately released to cause harm to mankind globally. Some also speculated if the leading corporate multinationals worldwide together wanted to depopulate the world this way simultaneously feeding the pharmaceutical industries for years and decades. Some rumors even fueled speculation that the viruses might have originated in Canada while a troop of Chinese agents smuggled the virus from Canada to China. The Wuhan level 4 biosafety laboratory, well known worldwide for conducting ('gain of functions') research on lethal viruses, is the center of all these charges as it is merely 32 km away from the Wuhan seafood market which has been tagged as the epicenter of the coronavirus outbreak. WHO along with the 'disease control and prevention laboratory' after a careful inspection reported that it's not a biological weaponry. Notably, they paid more attention to the seafood and animal market in Wuhan rather than blaming a world-famous pioneer laboratory in virology. However, an allegation could not be undermined about the shipping of some deadly viruses from the Canadian lab to China and these are powerful bioweapons. A piece

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 26 October 2023 doi:10.20944/preprints202310.1717.v1

of news also popped up in the television media that the Chinese agents working in the lab were expelled and later found involved with smuggling the virus again (probably as an act of revenge). But the scientists threw out the claim as a trumped-up story.

Yet, another alternative hypothesis postulates that SARS-CoV-2 is a recombinant between Sarbecovirus from Malayan Pangolins and RaTG13 from *R. affinis*. Although there are similarities between their RBD_{Spikes}, there is a dispute regarding their habitation as *R. affinis* and Pangolins did not share the same habitat. Moreover, further research revealed that no traces of Sarbecoviruses was detected in Chinese Pangolins, *Manis pentadactyla,* although, the virus was found in the anal swab sample of the *R. affinis* bat. The Pangolins *M. javanica* illegally imported from the Indo-Malayan region harbors Sarbecovirus also ruling out for the chances of recombination occurring between the *R. affinis* bat and Chinese Pangolins.

There is also suspicion that SARS-CoV-2 might have originated from an unintended laboratory accident as the same thing happened in the case of SARS-CoV. But investigation says that there is no clear evidence regarding such an accident as the low mortality rate (in comparison to SARS-CoV) characteristic of asymptomatic patients could have easily buffered such an undue event (Frutos et al., 2021). Also, none of the staff members of Wuhan Institute of Virology tested positive (Fact Sheet: Activity at the Wuhan Institute of Virology). Moreover, it's not clear whether the template for possible engineering (the RaTG13 virus) exists in reality or merely a predicted sequence, *in-silico,* which is unable to make a lab escape.

Current therapeutic measures and future directions

A covid vaccine is a vaccine that is manufactured to build up a protective shield around us and make us safe from getting affected by SARS-CoV-2. The vaccines are attributed worldwide for their role in diminishing the strength and lethality caused by COVID-19. The viral host cell entry can be prevented by blocking the RBD_{Spike}-hACE-2 interaction either by structurally mimicking the host-receptor (Cao et al., 2020) to entrap the viral Spike well before reaching the lung cells or by pre-occupying the receptor with the abundant use of non-reactive bio-friendly structural mimics of RBD_{Spike} (Basu, 2021). The RBD_{Spike} is very much sturdy towards acquiring mutations on its own, making it less difficult to engineer its suitable non-virulent mimics with a better electrostatic match at the RBD_{Spike}-hACE-2 interface, offering high-affinity binding. These mimics, being much smaller compared to the full-length Spikes carrying the viruses will hence approach faster than the native RBD_{Spike} towards the hACE-2 receptor, thereby, effectively inhibiting the native Spike protein binding to the host cells (Basu et al., 2021). Apart from that, David Baker and co-workers have designed S mini peptides (Cao et al., 2020) mimicking the RBD_{Spike}-binding 'ACE-2 peptidase domain α1 helix' in hACE-2 which can effectively entrap the viral Spike (via binding to RBD_{Spike}) much before the virus can reach out to the host (lung) cells. Also, the PRRAR motif of the $FLCS_{Spike}$ uniquely present in SARS-CoV-2 and yet absent in its ancestors helps priming the virus by a host-protease mediated proteolytic cleavage that sets free the RBD_{Spike} – which stands out to be a major contributory factor in the enhancement of viral transmissibility in COVID-19. For these reasons, it is natural to target the RBD_{Spike}-hACE-2 interaction (and eventually block it) as a therapeutic measure to prevent the viral host cell entry. In parallel, virus replication and spreading are inhibited by targeting RdRP (RNA dependent RNA-polymerase) (Gong, 2021) wherein replacing one of the RNA-nucleotides with a suitable nucleotide analog that can easily fit in the same enzyme-substrate binding site, can effectively terminate the replication of the viral RNA. As alternatives to the mainstream approaches, parallel therapeutic directions have also been explored using nanobodies (Huo et al., 2020, 2), and influenza virus-like particle (VLP) vaccines (Chu et al., 2021) etc. Plenty of natural products (Antonio et al., 2020; Huang et al., 2020; Omrani et al., 2021) have also been surveyed and have shown much promise to be aided as plausible 'herbal medicines' for the future. Keeping in mind the multi-wave dynamics of the COVID-19 pandemic with the growing number of emerging variants, a section of the scientific community is aiming to procure apt interpretable supervised learning / artificial intelligence (AI) – based predictors which would need to be trained continually with more and more available data to predict plausible future mutations of the virus. Though, these methods are currently only in their early days of development. Another section

 doi:10.20944/preprints202310.1717.v1

of the community is devoting committed efforts using immuno-informatics to design and develop possible multi-epitope combination vaccines (Naz et al., 2020) which are aimed to cover a plethora of viral antigens (epitopes) and their plausible future mutations (much like the decades-long challenge in developing anti-influenza and/or anti-HIV vaccines (Enuh et al., 2018)).

References

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., and Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine* 26, 450–452. https://doi.org/10.1038/s41591-020-0820-9.

Antonio, A. da S., Wiedemann, L. S. M., and Veiga-Junior, V. F. (2020). Natural products' role against COVID-19. *RSC Adv.* 10, 23379–23393. https://doi.org/10.1039/D0RA03774E.

Balaram, P. (2021). The murky origins of the Coronavirus SARS-CoV-2, the causative agent of the COVID-19 pandemic. *Current Science*, 1663–1666.

Basu, S. (2021). Special Issue on 'Coronavirus: Vaccines and Other Therapeutics' (2020–2021). *Vaccines* 9, 1083. https://doi.org/10.3390/vaccines9101083.

Basu, S., Chakravarty, D., Bhattacharyya, D., Saha, P., and Patra, H. K. (2021). Plausible blockers of Spike RBD in SARS-CoV2—molecular design and underlying interaction dynamics from high-level structural descriptors. *J Mol Model* 27, 191. https://doi.org/10.1007/s00894-021-04779-0.

Bertram, S., Glowacka, I., Müller, M. A., Lavender, H., Gnirss, K., Nehlmeier, I., et al. (2011). Cleavage and activation of the severe acute respiratory syndrome coronavirus spike protein by human airway trypsin-like protease. *J Virol* 85, 13363–13372. https://doi.org/10.1128/JVI.05300-11.

Cacciapaglia, G., Cot, C., and Sannino, F. (2021). Multiwave pandemic dynamics explained: how to tame the next wave of infectious diseases. *Sci Rep* 11, 6638. https://doi.org/10.1038/s41598-021-85875-2.

Cai, Y., Zhang, J., Xiao, T., Peng, H., Sterling, S. M., Walsh, R. M., et al. (2020). Distinct conformational states of SARS-CoV-2 spike protein. *Science*. https://doi.org/10.1126/science.abd4251.

Cao, L., Goreshnik, I., Coventry, B., Case, J. B., Miller, L., Kozodoy, L., et al. (2020). De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* 370, 426–431. https://doi.org/10.1126/science.abd9909.

Chan, Y. A., and Zhan, S. H. (2021). The Emergence of the Spike Furin Cleavage Site in SARS-CoV-2. *Molecular Biology and Evolution*, msab327. https://doi.org/10.1093/molbev/msab327.

Chu, K.-B., Kang, H.-J., Yoon, K.-W., Lee, H.-A., Moon, E.-K., Han, B.-K., et al. (2021). Influenza Virus-like Particle (VLP) Vaccines Expressing the SARS-CoV-2 S Glycoprotein, S1, or S2 Domains. *Vaccines* 9, 920. https://doi.org/10.3390/vaccines9080920.

Dance, A. (2021). The shifting sands of 'gain-of-function' research. *Nature* 598, 554–557. https://doi.org/10.1038/d41586-021-02903-x.

Dudas, G., Carvalho, L. M., Rambaut, A., and Bedford, T. (2018). MERS-CoV spillover at the camel-human interface. *eLife* 7. https://doi.org/10.7554/eLife.31257.

Enuh, H. A., Ezedunukwe, I. R., and Oragwu, C. I. (2018). Influenza vaccine for people with HIV/AIDS. *Cochrane Database Syst Rev* 2018, CD003509. https://doi.org/10.1002/14651858.CD003509.pub4.

Fact Sheet: Activity at the Wuhan Institute of Virology *United States Department of State*. Available at: https://2017-2021.state.gov/fact-sheet-activity-at-the-wuhan-institute-of-virology/ [Accessed May 28, 2022].

Frutos, R., Gavotte, L., and Devaux, C. A. (2021). Understanding the origin of COVID-19 requires to change the paradigm on zoonotic emergence from the spillover to the circulation model. *Infection, Genetics and Evolution* 95, 104812. https://doi.org/10.1016/j.meegid.2021.104812.

Fujiwara, S. (2018). Humanized mice: A brief overview on their diverse applications in biomedical research. *J Cell Physiol* 233, 2889–2901. https://doi.org/10.1002/jcp.26022.

Gong, P. (2021). "Chapter Eight - Structural basis of viral RNA-dependent RNA polymerase nucleotide addition cycle in picornaviruses," in *The Enzymes* Viral Replication Enzymes and their Inhibitors Part A., eds. C. E. Cameron, J. J. Arnold, and L. S. Kaguni (Academic Press), 215–233. https://doi.org/10.1016/bs.enz.2021.06.002.

Huang, J., Tao, G., Liu, J., Cai, J., Huang, Z., and Chen, J. (2020). Current Prevention of COVID-19: Natural Products and Herbal Medicine. *Frontiers in Pharmacology* 11, 1635. https://doi.org/10.3389/fphar.2020.588508.

Huo, J., Le Bas, A., Ruza, R. R., Duyvesteyn, H. M. E., Mikolajek, H., Malinauskas, T., et al. (2020). Neutralizing nanobodies bind SARS-CoV-2 spike RBD and block interaction with ACE2. *Nat Struct Mol Biol* 27, 846–854. https://doi.org/10.1038/s41594-020-0469-6.

Jiang, S., Shi, Z., Shu, Y., Song, J., Gao, G. F., Tan, W., et al. (2020). A distinct name is needed for the new coronavirus. *The Lancet* 395, 949. https://doi.org/10.1016/S0140-6736(20)30419-0.

Karunarathna, S. C., Dong, Y., Karasaki, S., Tibpromma, S., Hyde, K. D., Lumyong, S., et al. Discovery of novel fungal species and pathogens on bat carcasses in a cave in Yunnan Province, China. *Emerg Microbes Infect* 9, 1554–1566. https://doi.org/10.1080/22221751.2020.1785333.

Masters, P. S. (2006). The molecular biology of coronaviruses. *Adv Virus Res* 66, 193–292. https://doi.org/10.1016/S0065-3527(06)66005-3.

Matyášek, R., and Kovařík, A. (2020). Mutation Patterns of Human SARS-CoV-2 and Bat RaTG13 Coronavirus Genomes Are Strongly Biased Towards C>U Transitions, Indicating Rapid Evolution in Their Hosts. *Genes (Basel)* 11, E761. https://doi.org/10.3390/genes11070761.

Millet, J. K., and Whittaker, G. R. (2018). Physiological and molecular triggers for SARS-CoV membrane fusion and entry into host cells. *Virology* 517, 3–8. https://doi.org/10.1016/j.virol.2017.12.015.

Mlcochova, P., Kemp, S. A., Dhar, M. S., Papa, G., Meng, B., Ferreira, I. A. T. M., et al. (2021). SARS-CoV-2 B.1.617.2 Delta variant replication and immune evasion. *Nature* 599, 114–119. https://doi.org/10.1038/s41586-021-03944-y.

Mustafa, Z., Zhanapiya, A., Kalbacher, H., and Burster, T. (2021). Neutrophil Elastase and Proteinase 3 Cleavage Sites Are Adjacent to the Polybasic Sequence within the Proteolytic Sensitive Activation Loop of the SARS-CoV-2 Spike Protein. *ACS Omega* 6, 7181–7185. https://doi.org/10.1021/acsomega.1c00363.

Naz, A., Shahid, F., Butt, T. T., Awan, F. M., Ali, A., and Malik, A. (2020). Designing Multi-Epitope Vaccines to Combat Emerging Coronavirus Disease 2019 (COVID-19) by Employing Immuno-Informatics Approach. *Frontiers in Immunology* 11. Available at: https://www.frontiersin.org/article/10.3389/fimmu.2020.01663 [Accessed May 29, 2022].

Omrani, M., Keshavarz, M., Nejad Ebrahimi, S., Mehrabi, M., McGaw, L. J., Ali Abdalla, M., et al. (2021). Potential Natural Products Against Respiratory Viruses: A Perspective to Develop Anti-COVID-19 Medicines. *Frontiers in Pharmacology* 11, 2115. https://doi.org/10.3389/fphar.2020.586993.

Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., et al. (2017). Pathways to zoonotic spillover. *Nat Rev Microbiol* 15, 502–510. https://doi.org/10.1038/nrmicro.2017.45.

Randolph, H. E., and Barreiro, L. B. (2020). Herd Immunity: Understanding COVID-19. *Immunity* 52, 737–741. https://doi.org/10.1016/j.immuni.2020.04.012.

Roy, S., Ghosh, P., Bandyopadhyay, A., and Basu, S. (2022). Capturing a Crucial 'Disorder-to-Order Transition' at the Heart of the Coronavirus Molecular Pathology—Triggered by Highly Persistent, Interchangeable Salt-Bridges. *Vaccines* 10, 301. https://doi.org/10.3390/vaccines10020301.

Shang, J., Wan, Y., Luo, C., Ye, G., Geng, Q., Auerbach, A., et al. (2020). Cell entry mechanisms of SARS-CoV-2. *PNAS* 117, 11727–11734. https://doi.org/10.1073/pnas.2003138117.

Tabibzadeh, A., Esghaei, M., Soltani, S., Yousefi, P., Taherizadeh, M., Tameshkel, F. S., et al. (2021). Evolutionary study of COVID-19, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) as an emerging coronavirus: Phylogenetic analysis and literature review. *Veterinary Medicine and Science* 7, 559. https://doi.org/10.1002/vms3.394.

The origin of COVID: Did people or nature open Pandora's box at Wuhan? (2021). *Bulletin of the Atomic Scientists*. Available at: https://thebulletin.org/2021/05/the-origin-of-covid-did-people-or-nature-open-pandoras-box-at-wuhan/ [Accessed September 30, 2021].

Walls, A. C., Park, Y.-J., Tortorici, M. A., Wall, A., McGuire, A. T., and Veesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281-292.e6. https://doi.org/10.1016/j.cell.2020.02.058.

Wan, Y., Shang, J., Graham, R., Baric, R. S., and Li, F. (2020). Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *J Virol* 94, e00127-20. https://doi.org/10.1128/JVI.00127-20.

Yuan, Y., Cao, D., Zhang, Y., Ma, J., Qi, J., Wang, Q., et al. (2017). Cryo-EM structures of MERS-CoV and SARS-CoV spike glycoproteins reveal the dynamic receptor binding domains. *Nat Commun* 8, 15092. https://doi.org/10.1038/ncomms15092.

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., et al. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* 579, 270–273. https://doi.org/10.1038/s41586-020-2012-7.