

Structural Design on Virus and its Diversity

Nadia Naseer

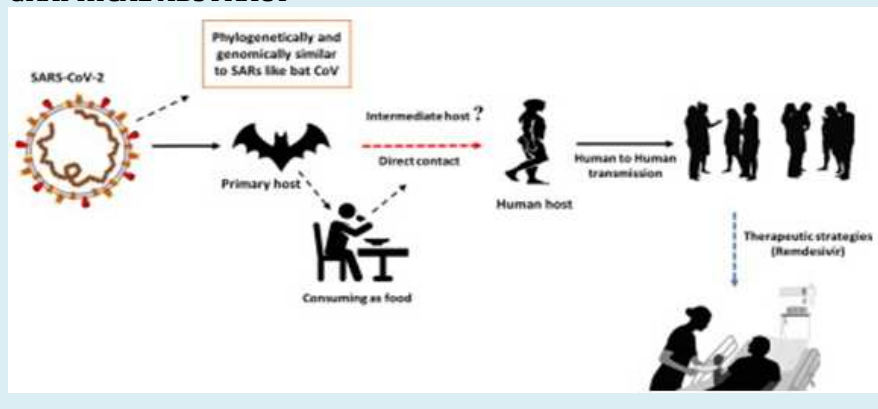
Student, St Ann's College for Women, Hyderabad, India

ABSTRACT

The coronavirus disease 19 (COVID-19) is a highly transmittable and pathogenic viral infection caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), which emerged in Wuhan, China and spread around the world. Genomic analysis revealed that SARS-CoV-2 is phylogenetically related to severe acute respiratory syndrome-like (SARS-like) bat viruses, therefore bats could be the possible primary reservoir. The intermediate source of origin and transfer to humans is not known, however, the rapid human to human transfer has been confirmed widely.

In this document we will analyze the structure and diversity of the pathogen and we will also discuss the previous emergence of human coronaviruses like Severe Acute Respiratory Syndrome Coronavirus (SARS-CoV) and Middle East Respiratory Syndrome (MERS-CoV).

GRAPHICAL ABSTRACT



KEYWORDS: Coronavirus, SARS-CoV-2/COVID-19, Origin, Outbreak, Spread, Natural Analysis: Protein homology modelling Molecular Docking and Molecular dynamics simulation.

INTRODUCTION:

The coronavirus disease (COVID-19) was identified as the cause of an outbreak of respiratory illness in Wuhan, Hubei province, China in the end December 2019 [1]. After an outbreak of coronavirus was experienced in China, the virus killed more than eighteen hundred and infected over seventy thousand individuals within first fifty days of the epidemic.

The World Health Organization (WHO) initially named this pathogen as (2019-nCoV) [2]. On January 10, the first whole-genome sequencing of (2019-nCoV) was released which helped the researchers to identify the type of virus and its diversity.

This virus was reported to be a member of the β group of coronaviruses and The International Committee on Taxonomy of Viruses (ICTV) named the virus as SARS-CoV-2 and the disease as COVID-19 [3],[4],[5].

Taxonomy of Coronaviruses

Coronaviruses is virus genera within the virus family Coronaviridae, under order Nidovirales.

Coronaviruses are a group of related RNA viruses that cause diseases in mammals and birds humans, these viruses cause

How to cite this paper: Nadia Naseer "Structural Design on Virus and its Diversity" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-4 | Issue-4, June 2020, pp.1534-1542, URL: www.ijtsrd.com/papers/ijtsrd31225.pdf



IJTSRD31225

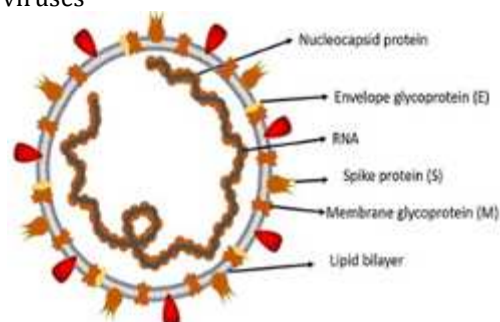
Copyright © 2020 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



respiratory tract infections that can range from mild to lethal. Mild illnesses include some cases of the common cold (which is caused also by certain other viruses, predominantly rhinoviruses), while more lethal varieties can cause SARS, MERS, which can also cause present disease COVID-19.

Structure of Coronaviruses

Corona represents crown-like spikes on the outer surface of the virus; thus, it was named as a coronavirus. Coronaviruses are minute in size (65–125 nm in diameter) and contain a single-stranded RNA as a nucleic material, size ranging from 26-32 kbp (Fig.1) in length and is among one of the largest among viruses



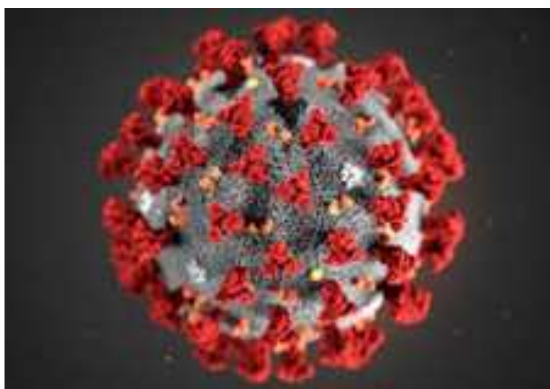


Figure1. Structure of respiratory syndrome causing human coronavirus.

They are enveloped viruses with a positive-sense single-stranded RNA genome and a Nucleocapsid of helical symmetry. This is wrapped in icosahedral protein shell, They have characteristic club-shaped spikes that project from their surface.

Etymology

The name "coronavirus" is derived from Latin corona, meaning "crown" or "wreath", The name was coined by **June Almeida and David Tyrrell** who first observed and studied human coronaviruses [6],[7]. The name refers to the characteristic appearance of Virions (the infective form of the virus) by electron microscopy, which have a fringe of large, bulbous surface projections, this morphology is created by the viral spike Peplomers, which are proteins on the surface of the virus.

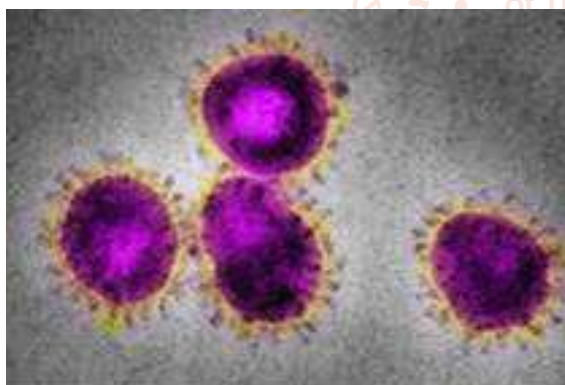


Figure2. Images of virions under Electron microscopy.

Classification of Coronaviruses

The classification of Coronaviruses has been based on genomic organizations, similarities in genomic sequence, antigenic properties viral proteins, replication strategies and structural characteristics virions, pathogenic, cytopathogenic, and physicochemical properties.

Coronaviridae family is the largest one of the four families, *Coronaviridae* virus family subdivided into two subfamilies, *coronavirinae* and *torovirinae*

The subgroups of coronaviruses family are alpha (α), beta (β), gamma (γ) and delta (δ) coronavirus.

Alpha coronaviruses (Alpha-CoV): are the first of the four genera of Coronaviruses,

Species: Human coronavirus 229E, Human coronavirus NL63, bat coronavirus.

*The novel strain 229E were subsequently imaged by electron microscopy in 1967 by Scottish virologist June Almeida at St. Thomas Hospital in London.

Beta coronaviruses (β -CoVs or Beta-CoVs):

Species: SARS-CoV, MERS-CoV, SARS-CoV-2/COVID-19.

Gamma coronavirus:

Species: Avian coronavirus (IBV), Beluga whale coronavirus SW1

Delta coronavirus:

Species: Bulbul coronavirus HKU11

History of previously occurred Human Coronaviruses:

Coronavirus disease was first described in 1931, with the first coronavirus (HCoV-229E) isolated from humans in 1965. the outbreak of severe acute respiratory syndrome in late 2002, only two **Human Coronaviruses** (HCoV) were known – HCoV-229E and HCoV-OC43.

The SARS coronavirus (SARS-CoV) has been identified after other human coronaviruses were identified, therefore there are three human coronavirus groups that still exist: **group 1** (HCoV-229E and HCoV-NL63), **group 2** (HCoVOC43 and HCoV-HKU1), and SARS-CoV is an outline between group 2 and **group 3**.

1. SARS-CoV: In 2003, the Chinese population was infected with a virus causing Severe Acute Respiratory Syndrome (SARS) in Guangdong province. The virus was confirmed as a member of the Beta-coronavirus subgroup and was named SARS-CoV [7]. The infected patients exhibited pneumonia symptoms [31] with a diffused alveolar injury which lead to acute respiratory distress syndrome (ARDS).

Viral genome analyses revealed that SARS-CoV belongs to *Group B*.

SARS initially emerged in Guangdong, China and then spread rapidly around the globe with more than 8000 infected persons and 776 deceases.

A decade later in 2012, a couple of Saudi Arabian nationals were diagnosed to be infected with another coronavirus. The detected virus was confirmed as a member of Coronaviruses and was named as Middle East Respiratory Syndrome (MERS-CoV).

The SARS-related coronavirus (SARS-CoV) was one of several viruses identified by the World Health Organization (WHO) in 2016 as **a likely cause of a future epidemic**.

2. MERS-CoV: In 2012, Saudi Arabian nationals were diagnosed to be infected with it.

MERS-CoV is a member beta-coronavirus subgroup and phylogenetically diverse from other human-CoV.

The infection of MERS-CoV initiates from a mild upper respiratory injury while progression leads to severe respiratory disease [8] . Similar to SARS-coronavirus, patients infected with MERS-coronavirus suffer pneumonia, followed by (ARDS) and renal failure. MERS-CoV belongs to Group C.

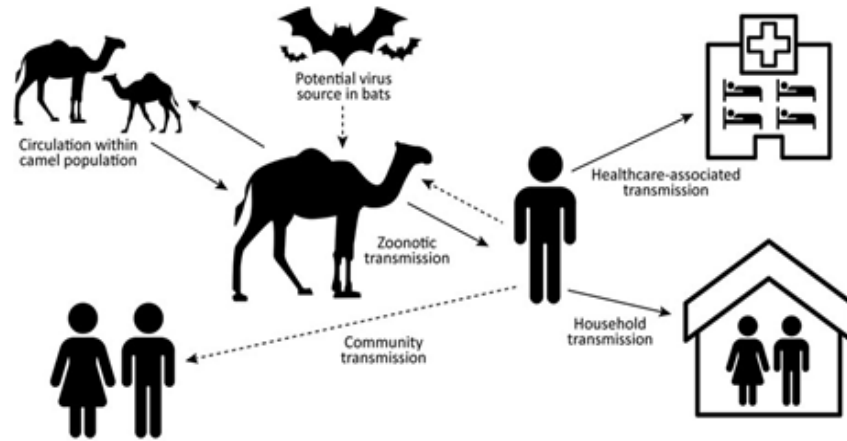


Figure3. Graphical representation of MERS-CoV and its spread.

Primary Hosts of Coronaviruses:

The source of origination and transmission is important in order to understand the strategies to be applied in prevention from infection.

- SARS-CoV:** Initially researchers focused on raccoon dogs and palm civets as a key reservoir of transmission, the samples were isolated from the civets at the food market and it showed positive results for viral RNA detection, suggesting that civets are the Secondary hosts.

In 2001, the samples were isolated from the healthy persons of Hong Kong and the molecular assessment showed 2.5% frequency rate of anti-bodies against SARS-coronavirus. These indications suggested that SARS-coronavirus may be circulating in humans before the outbreak.

Later on conducting study, Rhinolophus bats were also found to have anti-SARS-CoV antibodies suggesting the bats as a source of viral replication.

And it was identified as, Bats are a major reservoir of many strains of SARS-related coronaviruses, and several strains have been identified in palm civets, which were likely ancestors of SARS-CoV.

- MERS-CoV:** it first emerged in 2012 in Saudi Arabia. MERS-coronavirus also pertains to have camels as a zoonotic source or primary host.

Later on detail study it was found that, the detected virus was confirmed as a member of coronaviruses and named as the Middle East Respiratory Syndrome Coronavirus (MERS-CoV). The World health organization reported that MERS-coronavirus infected more than 2428 individuals and 838 deaths [8].

MERS-CoV is a member beta-coronavirus subgroup and phylogenetically diverse from other human-CoV

MERS-coronavirus was also detected in Pipistrellus and Perimyotis bats, proffering that bats are the key host and transmitting medium of the virus.

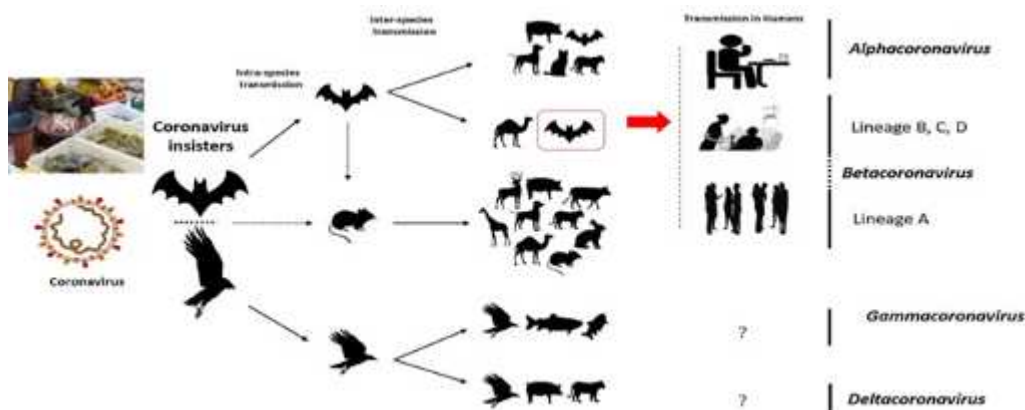


Figure4. The key reservoirs and mode of transmission of coronaviruses.

(suspected reservoirs of SARS-CoV-2 are red encircled); only α and β coronaviruses have the ability to infect humans, the consumption of infected animal as a source of food is the major cause of animal to human transmission of the virus and due to close contact with an infected person, the virus is further transmitted to healthy persons. Dotted black arrow shows the possibility of viral transfer from bat whereas the solid black arrow represents the confirmed transfer.

ABOUT

Coronavirus Disease 2019 (COVID-19)

1. BASIC STRUCTURE:

It is a positive-stranded RNA viruses with a crown-like appearance under an electron microscope, the prefix is derived from a Latin term (*coronam* – crown) due to the presence of spike glycoproteins on the envelope.

It is also reported that Spike glycoprotein of the Wuhan coronavirus is modified via **homologous recombination**.

The spike glycoprotein of SARS-CoV-2 is the mixture of bat SARS-CoV and a not known Beta-CoV.

The spike protein of SARS-CoV-2 contains a 3-D structure in the RBD region to maintain the van der Waals forces. [32]

2. GENOME

Recent studies show that SARS-CoV-2 has a similar genomic organization to other coronaviruses, the genome has the usual 5' methylated cap and a 3' Poly Adenylated tail, consisting of a 5'-3' untranslated region (UTR), Open reading frame (orf1) encoding non-structural proteins, a spike protein (S) gene, envelope protein (E) gene, a membrane protein (M) gene, a Nucleocapsid protein (N) gene.

There are 265 nucleotides in the 5'UTR and 342 nucleotides in 3'UTR.

The 5' methylated cap and 3' Poly Adenylated tail allows the positive-sense RNA genome to be directly translated by the host cell's ribosome on viral entry.

Studies also highlighted that SARS-CoV-2 genes share <80% nucleotide identity and 89.10% nucleotide similarity with SARS-CoV genes.

Recent studies have indicated notable variations in SARS-CoV and SARS-CoV-2 such as the absence of 8a protein and fluctuation in the number of amino acids in 8b and 3c protein in SARS-CoV-2.

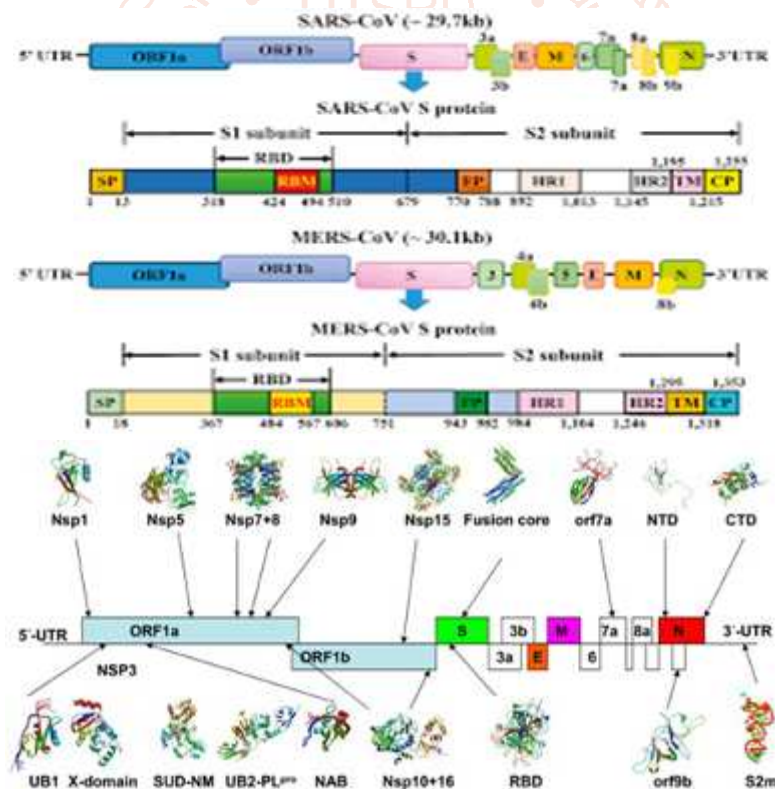


Figure5. Schematic representation of the genome organization and functional domains of S protein for SARS-CoV and MERS-CoV. The single-stranded RNA genomes of SARS-CoV and MERS-CoV encode two large genes, the ORF1a and ORF1b genes, which encode 16 non-structural proteins (nsp1–nsp16) that are highly conserved throughout coronaviruses. The structural genes encode the structural proteins, spike (S), envelope (E), membrane (M), and nucleocapsid (N), which are common features to all coronaviruses. The accessory genes (shades of green) are unique to different coronaviruses in terms of number, genomic organization, sequence, and function. The structure of each S protein is shown beneath the genome organization. The S protein mainly contains the S1 and S2 subunits. The residue numbers in each region represent their positions in the S protein of SARS and MERS, respectively. The S1/S2 cleavage sites are highlighted by dotted lines. SARS-CoV, severe acute respiratory syndrome coronavirus; MERS-CoV, Middle East respiratory syndrome coronavirus; CP, cytoplasm domain; FP, fusion peptide; HR, heptad repeat; RBD, receptor-binding domain; RBM, receptor-binding motif; SP, signal peptide; TM, transmembrane domain.

This polypeptide is proteolytically cleaved to generate various proteins. The proteolytic processing is mediated by **papain-like protease (PLpro)** and **3-chymotrypsin-like protease (3CLpro)**.

The **3CLpro** cleaves the polyprotein at 11 distinct sites to generate various non-structural proteins that are important for viral replication, it also plays a critical role in the replication of virus particles and unlike structural/accessory protein-encoding genes, [8] It is located at the 3' end which exhibits excessive variability [30].

Structure-based activity analyses and high-throughput studies have identified potential inhibitors for SARS-CoV and MERS-CoV 3CL^{pro} [10].

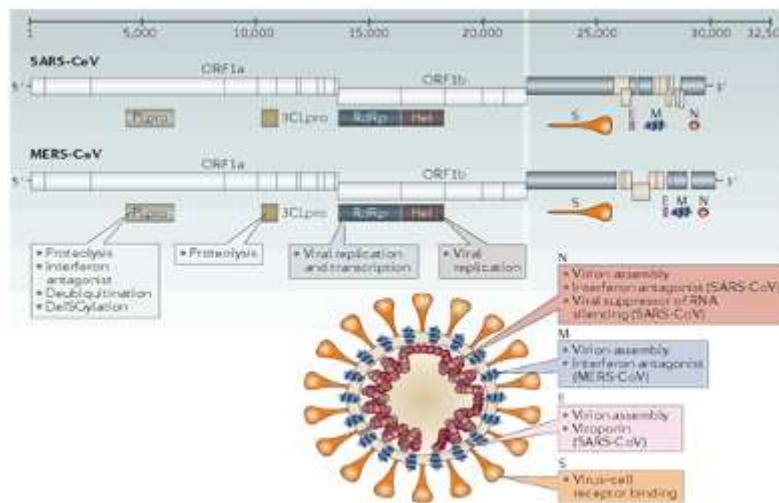


Figure6. Genomic structure of SARS-CoV and MERS-CoV.

The typical coronavirus (CoV) genome is a single-stranded, non-segmented RNA genome, which is approximately 26–32 kb. It contains 5'-methylated caps and 3'-polyadenylated tails and is arranged in the order of 5', replicase genes, genes encoding structural proteins (spike glycoprotein (S), envelope protein (E), membrane protein (M) and nucleocapsid protein (N)), polyadenylated tail and then the 3' end. The partially overlapping 5'-terminal open reading frame 1a/b (ORF1a/b) is within the 5' two-thirds of the CoV genome and encodes the large replicase polyprotein 1a (pp1a) and pp1ab. These polyproteins are cleaved by papain-like cysteine protease (PLpro) and 3C-like serine protease (3CLpro) to produce non-structural proteins, including RNA-dependent RNA polymerase (RdRp) and helicase (Hel), which are important enzymes involved in the transcription and replication of CoVs. The 3' one-third of the CoV genome encodes the structural proteins (S, E, M and N), which are essential for virus–cell-receptor binding and virion assembly, and other non-structural proteins and accessory proteins that may have immunomodulatory effects.

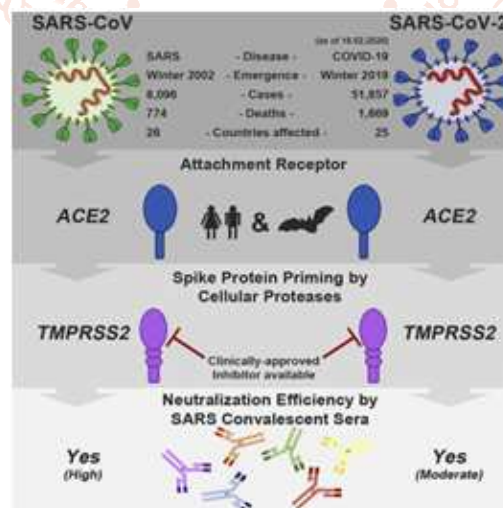


Figure7. graphical abstract of SARS-CoV2 Cell entry:

SARS-CoV-2 uses the SARS-CoV receptor ACE2 for host cell entry

The spike protein of SARS-CoV-2 is primed by TMPRSS2

Antibodies against SARS-CoV spike may offer some protection against SARS-CoV-2

The single N501T mutation in SARS-CoV-2's Spike protein may have significantly enhanced its binding affinity for ACE2.

The 394 glutamine residue in the MPR region of SARS-CoV-2 is recognized by the critical lysine 31 residue on the human ACE2 receptor [33].

In a fluorescent study, SARS-CoV-2 uses the SARS-CoV receptor ACE2 for entry and the serine protease TMPRSS2 for S protein priming.

3. SEQUENCE ANALYSIS

In order to identify similar sequences and key/conserved residues, and to infer phylogeny, multiple sequence alignment of SARS-CoV-2 3CL^{pro} followed by phylogenetic tree analyses were performed using T-Coffee [15] and the alignment Figure was generated using ESPript3 [16]. Physicochemical parameters of SARS-CoV-2 3CL^{pro} including isoelectric point, instability index, grand average of hydropathicity (GRAVY), and amino acid and atomic composition were investigated using the ProtParam tool of ExPASy [14].

4. STRUCTURAL ANALYSIS

To probe the molecular architecture of SARS-CoV-2 3CL^{pro}, comparative homology modelling was performed using Modeller v9.11 [17]. To select closely-related templates for modelling, PSI-BLAST was performed against all known structures in the protein databank (PDB) [18]. Chimera v1.8.1 [19] and PyMOL educational version [20] were used for initial quality estimation, energy minimisation, mutation analyses, and image processing.

5. LIGAND AND DATABASE PREPARATION:

A comprehensive medicinal plant library containing 32,297 potential anti-viral phytochemicals and traditional Chinese medicinal compounds was generated from our previously collected data and studies [13, [21], [22], [23]], and screened against the predicted SARS-CoV-2 3CL^{pro} structure. Molecular operating environment (MOE) [24] was used for molecular docking, ligand-protein interaction and drug likeness analyses. All analyses were performed using the same protocols that are already described in our previous studies [13, 25, 26]. The qualitative assessment of absorption, deposition, metabolism, excretion and toxicity (ADMET) profile of selected hits were predicted computationally by using ADMET sar server [27].

6. MOLECULAR DYNAMICS AND SIMULATIONS:

Explicit solvent molecular dynamics (MD) simulations were performed to verify docking results and to analyze the binding and stability of potential compounds by using the predicted SARS-CoV-2 3CL^{pro} homology model.

GROMACS v5.1.4, GROMOS96 and the PRODRG server were employed to run 50 ns MD simulations [28, 29] following the same protocol as described in our previous studies [13].

7. SEQUENCE AND STRUCTURAL ANALYSIS RESULTS:

Multiple sequence alignment results revealed that 3CL^{pro} was conserved, with 100% identity among all SARS-CoV-2 genomes.

The SARS-CoV-2 3CL^{pro} protein sequence was compared with its closest homologs (Bat-CoV, SARS-CoV, MERS-CoV, Human-CoV and Bovine-CoV).

The results revealed that SARS-CoV-2 3CL^{pro} clustered with bat SARS-like coronaviruses and sharing 99.02% sequence identity (Fig1A).

Furthermore, it shares 96.08%, 87.00%, 90.00% and 90.00% sequence identity with SARS-CoV, MERS-CoV, Human-CoV and Bovine-CoV homologs, respectively (Fig. 1B).

These findings were consistent with an initial study reporting that SARS-CoV-2 is more similar to SARS-CoV than MERS-CoV, and shares a common ancestor with **bat** coronaviruses.

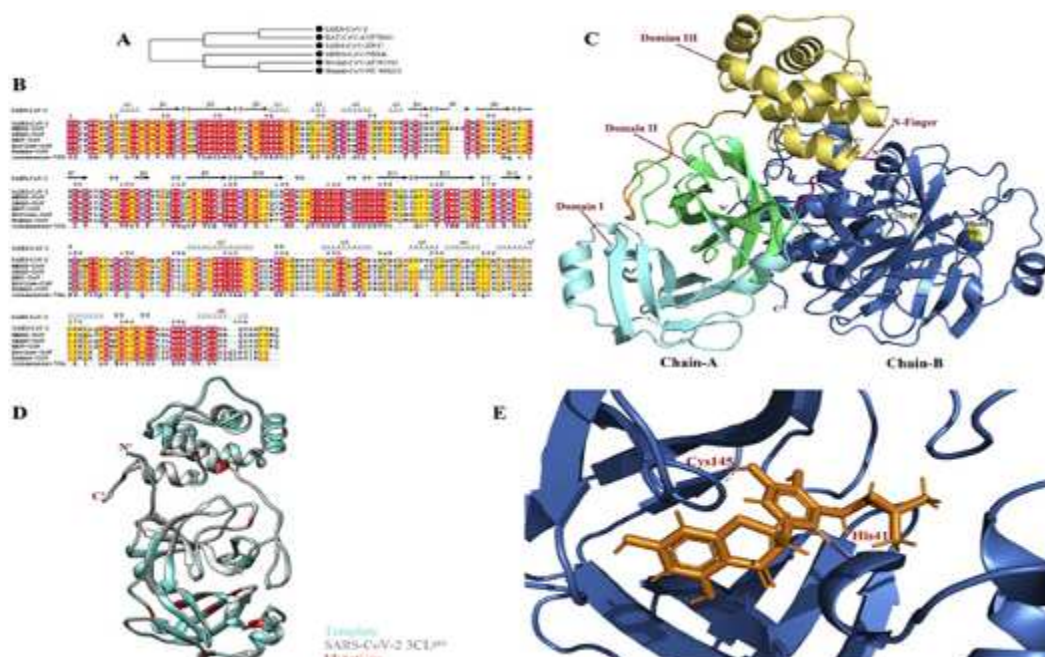


Figure8. Phylogenetic tree inferred from closest homologs of SARS-CoV-2.

The maximum likelihood method was used to construct this tree. (B) Multiple sequence alignment of closest homologs of SARS-CoV-2 3CLpro sharing $\geq 70\%$ sequence identity. (C) Cartoon representation of the SARS-CoV-2 3CLpro homodimer. Chain-A (protomer-A) is in multicolor and Chain-B (protomer-B) is in dark blue. The N-finger that plays an important role in dimerization maintaining the active conformation is shown in hot pink, domain I is colored cyan, domain II is shown in green, and domain III is colored yellow.

The N- and C-termini are labelled. Residues of the catalytic dyad (Cys-145 and His-41) are highlighted in yellow and labelled. (D) Cartoon representation of the 3CLpro monomer model (chain/protomer-A) of SARS-CoV-2 superimposed with the SARS-CoV 3CLpro structure. The SARS-CoV 3CLpro template is colored cyan, the SARS-CoV-2 3CLpro structure is colored grey, and all identified mutations are highlighted in red. (E) Docking of 5,7,3',4'-tetrahydroxy-2'-(3,3-dimethylallyl) flavone inside the receptor-binding site of SARS-CoV-2 3CLpro, showing hydrogen bonds with the catalytic dyad (Cys-145 and His-41). The 3CLpro structure is colored dark blue, the 5,7,3',4'-tetrahydroxy-2'-(3,3-dimethylallyl) flavone is orange, and hydrogen colored maroon.

8. MOLECULAR DOCKING:

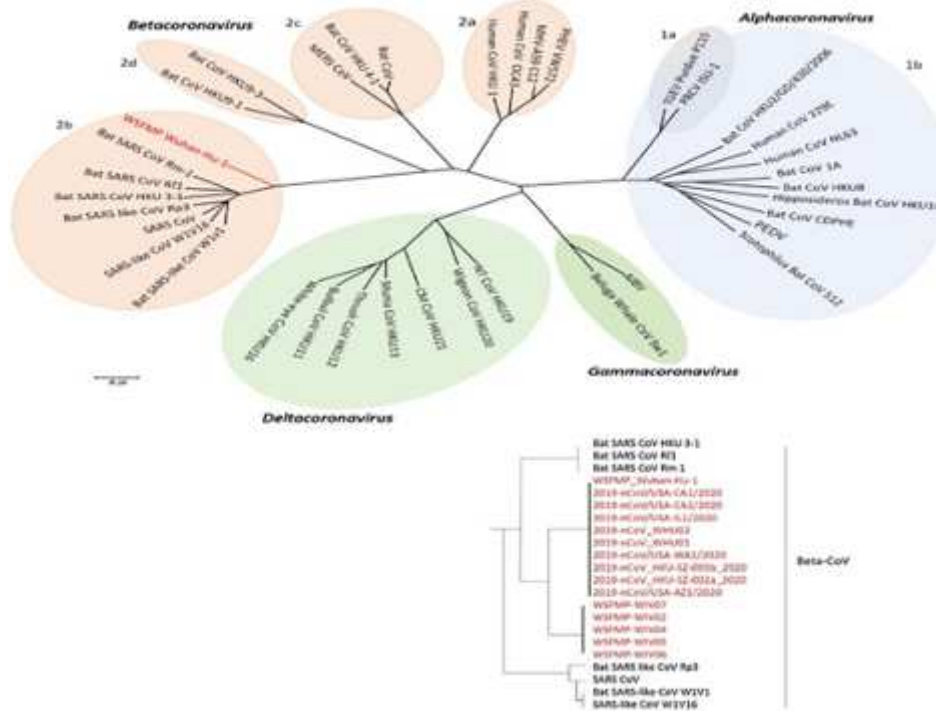
To test this hypothesis, sequences were docked (R)-N-(4-(tert-phenyl)-N-[2-(tert-butylamino)-2-oxo-1-(pyridin-3-ethyl furan-2-carboxamide)], a potential noncovalent inhibitor of SARS-CoV 3CLpro named ML188 [35], with the SARS-CoV-2 3CLpro homology model.

The inhibitor target is to bind Cys-His catalytic dyad (Cys-145 and His-41) with the other residues, and increase as a result the docking score ($S = -12.27$) relatively high.

However, surprisingly, ML188 did not show significant binding to the catalytic dyad (Cys-145 and His-41) of SARS-CoV-2, and the docking score ($S = -8.31$) was considerably lower.

These results indicated that the 12-point mutation is identified which may disrupt important hydrogen bonds and alter the receptor binding site.

9. PHYLOGENETIC TREE:



The phylogenetic tree shows the relationship of Wuhan-Hu-1 (denoted as red) to selected coronavirus is based on nucleotide sequences of the complete genome. The viruses are grouped into four genera.

This tree is based on the published trees of Coronaviruses and is reconstructed with sequences of the complete RNA-dependent RNA polymerase- coding region of the representative novel coronaviruses.

Severe Acute Respiratory Syndrome Coronavirus (SARS- CoV), The Middle East respiratory syndrome coronavirus (MERS-CoV); Porcine Enteric Diarrhea Virus (PEDV); Wuhan seafood market pneumonia (Wuhan-Hu-1).

Genetic analysis has revealed that the coronavirus genetically clusters with the genus Beta coronavirus, in subgenus *Sarbecovirus* (lineage B) together with two bat-derived strains. It is 96% identical at the whole genome level of other bat coronavirus samples

Recent studies have indicated notable variations in SARS-CoV and SARS-CoV-2 such as the absence of 8a protein and fluctuation in the number of amino acids in 8b and 3c protein in SARS-CoV-2 [34] (Fig. 4). It is also reported that Spike glycoprotein of the Wuhan coronavirus is modified via homologous recombination.

(SARS-Cov). The results suggested that **homologous recombination** may have occur and contributed to the 2019-nCoV **cross-species transmission**.

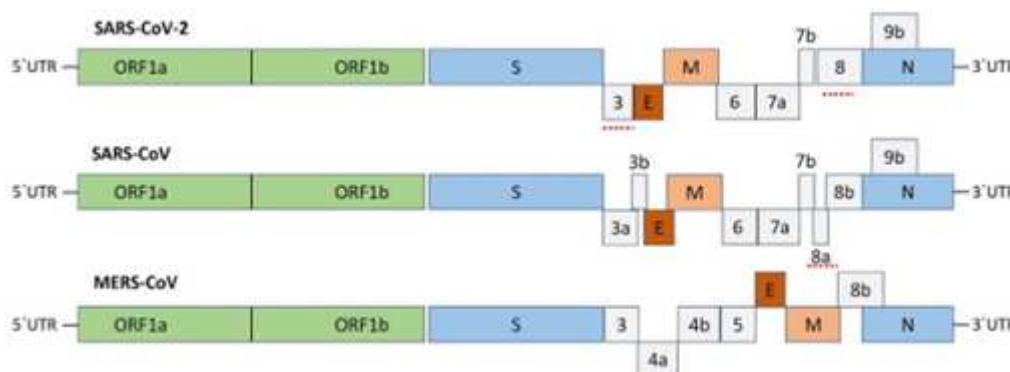


Figure9. Betacoronaviruses genome organization;

The Betacoronavirus for human (SARS-CoV-2, SARS-CoV and MERS-CoV) genome comprises of the 5'-untranslated region (5'-UTR), open reading frame (orf) 1a and 1a/b (green box) encoding non-structural proteins (nsp) for replication. structural proteins including spike (blue box), envelop (maroon box), membrane (pink box), and nucleocapsid (cyan box) proteins, accessory proteins (light gray boxes) such as orf 3, 4a, 4b 6, 7a, 7b, 8b and 9b in the SARS-CoV-2 genome, and the 3'-untranslated region (3'-UTR).

The dotted underlined in red are the protein which shows key variation between SARS-CoV-2 and SARS-CoV.

SARS-CoV Showed high sequence identity to **SARS-CoV-2**.

CONCLUSION:

The novel coronavirus originated from the Hunan seafood market at Wuhan, China where bats, snakes, raccoon dogs, palm civets, and other animals are sold, has rapidly spread up to 109 countries. The zoonotic source of SARS-CoV-2 is not confirmed, however, **sequence-based analysis suggested, bats as the key reservoir**.

DNA recombination was found to be involved at spike glycoprotein which assorted SARS-CoV with the Receptor Binding Domain (RBD) of another Beta Coronavirus, thus could be the reason for *cross-species transmission* and rapid infection. According to phylogenetic trees, **SARS-CoV2 is closer to SARS-CoV**.

Upon deep analysis it was also found that, *3CL^{pro} is conserved in SARS-CoV-2*.

It is highly similar to bat SARS-like coronavirus 3CL^{pro}, with some differences from other beta-coronaviruses.

SARS-CoV-2 3CL^{pro} is conserved, share 99.02% sequence identity with SARS-CoV 3CL^{pro} and together with **12 point-mutations**.

These 12 point Mutations also disrupt hydrogen bonds and alter the receptor binding site of SARS-CoV-2 3CL^{pro}.

Taking advantage of all the genomic information currently available, researchers constructed a phylogenetic tree including representatives of other coronaviridae members, such as Bat coronavirus (BCoV) and severe acute respiratory syndrome (SARS). It was confirmed that high sequence similarity (>99%) is observed between all sequenced 2019-nCoV genomes available, with the closest BCoV

sequence sharing 96.2% sequence identity, confirming the notion of a zoonotic origin of 2019-nCoV genome.

Finally, a full proteomic comparison with other coronaviridae members, by identifying key amino acids and its differences are considered for deriving its relationship from previously coronavirus approaches.

REFERENCES:

- [1] X. Xu, P. Chen, J. Wang, et al. **Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission** Sci. China Life Sci, 63 (2020), pp. 457-46
- [2] W. Ji, W. Wang, X. Zhao, et al. **Cross-species transmission of the newly identified coronavirus 2019-nCoV**
- [3] J. Med. Virology., 92 (2020), pp. 433-44 J. Cui, F. Li, Z.-L. **Origin and evolution of pathogenic coronaviruses** Nat Rev Microbiology 17 (3) (2019), pp. 181-192
- [4] C.-C. Lai, T.-P. Shih, W.-C. K, H.-J. Tang, P.-R. **Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and corona virus disease-2019 (COVID-19): the epidemic and the challenges**.
- [5] Organization WH. **Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected human cases: international guidance**, 2 March 2020. World Health Organization.
- [6] **"The woman who discovered the first coronavirus"**.
- [7] J. Pier, **Severe acute respiratory syndrome** Nat Med, 10 (12) (2004), pp. S88-S97
- [8] Rahman, A. Sarkar **Risk factors for fatal middle east respiratory syndrome coronavirus infections in**

Saudi Arabia analysis of the WHO Line List, 2013-2018

- [9] D. Needle, G.T. Lountos, D.S. Waugh **Structures of the Middle East respiratory syndrome coronavirus 3C-like protease reveal insights into substrate specificity** Acta Crystallogr. D Biol. Crystallogr., 71 (2015), pp. 1102-1111
- [10] A.K. Ghosh, K. Xi, K. Ratia, **Design and synthesis of peptide mimetic severe acute respiratory syndrome chymotrypsin-like protease inhibitors.**
- [11] V. Kumar, K.P. Tan, Y.M. Wang **Identification, synthesis and evaluation of SARS-CoV and MERS-CoV 3CLpro** Med. Chem., 24 (2016), pp. 3035-3042
- [12] T. Pillaiyar, M. Manickam, V. Namasivayam, **An overview of severe acute respiratory syndrome-coronavirus (SARS-CoV) Peptides and small molecule chemotherapy** J. Med. Chem., 59 (2016), pp. 6595-6628
- [13] M. Tahir ul Qamar, A. Maryam, I. Muneer. **Computational screening of medicinal plant phytochemicals to discover serotype inhibitors** Sci. Rep., 9 (2019), pp. 1-16
- [14] E. Gasteiger, A. Gattiker, C. Hoogland, **ExpASY: The proteomics server for in-depth protein knowledge and analysis** Nucleic Acids Res., 31 (2003), pp. 3784-3788
- [15] C. Notredam, D.G. Higgins, J. Heringa **T-COFFE: A novel method for fast and accurate multiple sequence alignment** J. Mol. Biol., 302 (2000), pp. 205-217
- [16] P. Gouet, E. Courcelle, D.I. Stuart, *et al* **ESPrpt: analysis of multiple sequence alignments in PostScript** Bioinformatics, 15 (1999), pp. 305-308
- [17] N. Eswar, B. Webb, M.A. Marti-Renom, **Comparative protein structure modeling using Modeller** Bioinformatics, 15 (2006),
- [18] M. Johnson, I. Zaretskaya, Y. Raytselis, **NCBI BLAST: a better web interface** Nucleic Acids Res., 36 (2008), pp. W5-W9
- [19] E.F. Pettersen, T.D. Goddard, C.C. Huang, **UCSF Chimera—a visualization system for exploratory research and analysis** J. comput Chem., 25 (2004), pp. 1605-1612
- [20] W. Delano, **An open-source molecular graphics tool** CCP4 Newsletter on protein crystallography, 40 (2002), pp. 82-92
- [21] A. Mumtaz, U. A. Ashfaq, M. Tahir Ul Qamar, **MPD3: a useful medicinal plants database for drug designing** Nat. Prod. Res., 31 (2017), pp. 1228-1236
- [22] U. A. Ashfaq, A. Mumtaz, M. Tahir Ul Qamar, **MAPS Database: medicinal plant activities, phytochemical and structural database** Bioinformation, 9 (2013), pp. 993-995
- [23] A. Mumtaz, U.A. Ashfaq, M. Tahir ul Qamar, **Addendum**, Nat. Prod. Res. (2020)
- [24] S. Vilar, G. Cozza, S. Moro **Medicinal chemistry and the molecular operating environment (MOE): application of QSAR and molecular docking to drug discovery** Current. Top. Med. Chem., 8 (2008), pp. 1555-1572
- [25] M. Tahir Ul Qamar, S. Saleem, U.A. Ashfaq, **Epitope-based peptide vaccine design and target site depiction against MERS: an immune-informatics study** J. Transl. Med., 17 (2019),
- [26] M. Tahir Ul Qamar, A. Bari, M.M. Adeel, **Peptide vaccine against chikungunya virus: immuno-informatics combined with molecular docking approach** J. Transl. Med., 16 (2018),
- [27] F. Cheng, W. Li, Y. Zhou, **A comprehensive source and free tool for assessment of chemical ADMET properties** J. Chem. Inf. Model., 52 (2012), pp. 3099-3105
- [28] H. J. C. Berendsen, D. van der Spoel, R. van Drunen **GROMACS: a message-passing parallel molecular dynamics implementation** Comp. Phys. 91 (1995), pp. 43-56
- [29] D. M. van Aalten, R. Bywater, J.B. Findlay, **a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules** J. Compt. Aided Mol. Des., 10 (1996), pp. 255-262
- [30] Wu, S. Zhao, B. Yu, Y.-M. Chen, W. Wang, Z.-G. **A new coronavirus associated with human respiratory disease in China** Nature, 1-5 (2020)
- [31] **Pneumonia outbreak associated with a new coronavirus of probable bat origin** Nature, 579 (2020), pp. 270-273
- [32] X. Xu, P. Chen, J. Wang, J. Feng, H. Zhou, X. Li, **Evolution of the novel coronavirus from the ongoing Wuhan outbreak and modeling of its spike protein for risk of human transmission** Science China Life Sciences, 63 (3) (2020), pp. 457-46
- [33] Y. Wan, J. Shang, R. Graham, R.S. Baric, F. Li **Receptor recognition by novel coronavirus from Wuhan: an analysis based on decade-long structural studies of SARS** J Virology (2020)
- [34] Wu, Y. Peng, B. Huang, X. Ding, X. Wang, **Genome composition and divergence of the novel coronavirus (2019-nCoV)** Cell Host Microbe (2020)
- [35] J. Jacobs, S. Zhou, E. Dawson, et al. **Discovery of Non covalent Inhibitors of the SARS Main Proteinase 3CLpro**, National Center for Biotechnology Information (US), Bethesda (MD) (2010)