

## A review of current aspects of sars family, genome, database, drug, vaccine and its pathogenic member SARS-CoV-2

Rawaa AlChalabi<sup>1</sup>, Basma Kh. Alani<sup>1</sup>, Taisir K. Ibrahim<sup>2</sup>, AhmedAbdul Jabbar Suleiman<sup>3</sup>

<sup>1</sup>Al-Nahrain University, College of Biotechnology, Baghdad, Iraq

<sup>2</sup>Tikrit University, College of Science, Biology Department, Saladin, Iraq

<sup>3</sup>Biotechnology Department, College of Science, University of Anbar, Ramadi, Anbar, Iraq

Received:

August 01, 2022

Accepted:

October 11, 2022

Published Online:

November 30, 2022

### Abstract

The coronaviruses (CoV) show great diversity with respect to their genomic structures and the ability to infect animals and human beings. Multiple omics studies have been conducted to understand and identify the genome organization, immunological responses, and molecular markers for vaccine and drug development. However, due to constant mutational changes in the CoV-2 genome, the drug and vaccine development is becoming a challenge to target new variants. A tremendous amount of research is being carried out for vaccine development by utilizing bioinformatics and immunoinformatics, however, most of the developed vaccines are under trial. In this review, SARS family members are systematically illuminated for their differences, infections and mechanism of action along with recommendations to counter the new variants of CoV-2 through multi-omics, machine learning and structural bioinformatics techniques.

**Keywords:** SARS, COVID-19, SARS-CoV-2, Pandemic, Coronaviruses, RNA viruses, MERS-CoV

### How to cite this:

AlChalabi R, Alani BK, Ibrahim TK and Suleiman AAJ. A review of current aspects of sars family, genome, database, drug, vaccine and its pathogenic member SARS-CoV-2. Asian J. Agric. Biol. 2023(2). 2022106 DOI: <https://doi.org/10.35495/ajab.2022.106>

\*Corresponding author email:  
ahmed.suleiman@uoanbar.edu.iq

This is an Open Access article distributed under the terms of the Creative Commons Attribution 3.0 License. (<https://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

### Introduction

Coronaviruses (CoV) are very distinct types of viruses that have great diversity and the ability to mutate (Cui et al., 2019). They infect many animals and induce respiratory syndrome in human beings. CoVs are positive-sense single-stranded (ss) RNA viruses belonging to a family called coronaviridae (Cui et al., 2019). After the phylogenetic analysis and genome organization studies, CoV is grouped into the sub-family Coronaviridae which has four genera (Cui et al., 2019), including Alphacoronavirus, Deltacoronavirus,

Gammacoronavirus, Betacoronavirus. The origin of these genera was revealed through evolutionary analysis and it was found that  $\alpha$ CoV and  $\beta$ CoV originated from rodents and bats (Kumar et al., 2020). The CoV transmission is species independent and spreads rapidly in hosts. The focus of this review is  $\beta$ CoV, which caused Middle East respiratory syndrome CoV (MERS-CoV) and severe acute respiratory syndrome CoV (SARS-CoV) (Fung et al., 2021). In 2002, SARS-CoV emerged, whereas MERS-CoV originated in 2012 (Cui et al., 2019). CoV is a zoonotic disease, and is responsible for fatal respiratory disease in humans that has made



coronavirus a general public concern. By the end of 2019, an outbreak of a new betacoronavirus named SARS-CoV-2 was identified in Wuhan, China, and it spread throughout the World as a pandemic. It caused COVID-19 (Wu et al., 2020). It has surpassed MERS and SARS-CoV due to its rapid spread among people (Deng and Peng, 2020; Han et al., 2020).

### **Emergence and Spread**

In the middle of November 2002 and July 2003, the first outbreak of SARS was reported in southern China. This outbreak resulted in a total of 8273 cases in China while approximately 775 in the rest of the world. The major impact of this outbreak was in Hong Kong which has a fatality rate of 9.6% and this SARS then spread from Hong Kong to 37 different countries in the early months of 2003. This outbreak was controlled and exterminated by January of next year. Phylogenomic analysis of the virus helped to identify its origin, that is, bats. It was identified that through bats, the disease prevailed to human beings in the Chinese animal market (Platto et al., 2020).

The Second outbreak occurred in December 2019, in which a unique coronavirus strain was identified, named SARS-CoV-2. Initially, the mode of SARS-CoV-2 transmission in human beings was not known but it was observed that most of the infected people had exposure to large seafood or have visited the animal market. A zoonotic origin was hypothesized based on the information. The spread of this SARS variant had the same mode of transmission, as of SARS-CoV. It has an animal origin and then it was migrated to human beings from animal reservoirs to the environment (Stokes, 2020). Also, it is anticipated that the SARS-CoV-2 has intermediate hosts that allow it to be more potent and pathogenic. Later, SARS-CoV-2 spread to multiple other countries leading to multiple cases. The World Health Organization declared a pandemic in March 2020, caused by SARS-CoV-2. There is a higher risk of acquiring disease in the older people and the ones with medical conditions like heart problems, diabetes, and compromised immune system (Abdelrahman et al., 2020).

### **Morphology and Genome Organization of SARS Family**

#### **Morphology**

The coronavirus-specific morphological analysis of

SARS-CoV-2 was performed using electron microscopy. The SARS-CoV has a size of 70-90 nm (Park et al., 2020). The higher sequence identity between both viruses suggested the same structure of both virus types. (Kumar et al., 2020). Coronaviruses have viral protein spikes on their surface and the viral RNA is encapsulated by the lipid bilayer embedded inside the host membrane (Yang et al., 2020). The research on understanding the structure of viral spikes has opened more opportunities for therapeutic treatments.

#### **Genome organization**

A genome similarity of 79% between SARS and CoV and 50% between MERS and CoV, has been reported (Deslandes et al., 2020). Coronavirus genomes constitute 6 to 100 open reading frames (ORFs). Sixteen non-structural proteins (nsps) are encoded by the first ORF which covers about 67% of the genome, while the rest of the ORFs code for structural and accessory proteins (Guo et al., 2020). SARS-CoV-2 has two untranslated regions (UTRs). The 5' untranslated region end is of 265 nucleotides length whereas the 3' untranslated region end is of 358 nucleotides. There has been no substantial sequence variation between the two viruses. SARS-CoV-2 genome consists of two viral cysteine proteases that include papain-like protease (nsp3) and main protease (nsp5). Several other NSPs are present such as helicase (nsp13) that plays its role in replication (Chan et al., 2020). Besides nsps, the ORFs encode four major structural proteins that include accessory proteins, spike surface glycoprotein (S), envelope (E), nucleocapsid protein (N). The M protein contains an N-terminal glycosylated domain, along with other domains including a long C-terminal CT domain, and three transmembrane domains (TM) (Kumar et al., 2020).

In coronavirus, M and E proteins are required for morphogenesis, assembly, and budding. The S protein is a two subunit (S1 and S2) protein containing fusion viral protein. The bat SARS-CoV S1 subunit has been reported to have a 70% sequence identified with human CoV. The S1 subunit has a signal peptide and two domains, receptor binding (RBD) and N terminal (NTD) (Walls et al., 2020).



**Table-1: Pathogenesis, Genomic Characterization and Characteristics of Coronaviruses**

Characteristics	SARS-CoV	MERS-CoV	SARS-CoV-2
<b>Origin</b>	China	Saudi Arabia	China
<b>Origin Year</b>	2003	2012	2019
<b>Hosts</b>	Bats	Bats	Bats
<b>Reservoir (Origin)</b>	Himalayan palm civet	Dromedary camels	Pangolins
<b>Infections Caused</b>	>8000	2428	More than 529 million
<b>Deaths Caused</b>	776	838	6.29 million (as of 30th May 2022)
<b>Fatality rate</b>	>10%	35%	5.9%
<b>Transmission region</b>	Global	Regional	Global
<b>Reproduction Number (R<sub>0</sub>)</b>	3	<1	2–2.5
<b>Receptor</b>	ACE2	DPP4	ACE2
<b>Replication Efficiency</b>	High	Higher	Highest
<b>Organ</b>	Liver, kidney, respiratory tract	Gastrointestinal tract respiratory tract kidney, liver, urinary tract and neurons	Kidney, liver and respiratory tract
<b>Genome Size</b>	29,727 bp	30,119 bp	29,903 bp
<b>Structural</b> (4 (nucleocapsid spike, envelope, membrane))	Yes	Yes	Yes
<b>Number of non-structural and accessory proteins</b>	5, 8	16, 5	15, 6

The main difference was found between the external subdomain which is specifically involved in association with the ACE2 receptor. The spike protein's ectodomain was crystallized to understand the glycoprotein structure. The glycoprotein is similar in structure to the SARS-CoV spike protein whereas the receptor-binding domain has a high percentage of structural differences (Wrapp et al., 2020). The bat SARS-CoV S2 subdomain shares a 99% sequence identity with human SARS-CoV (Coutard et al., 2020). The comparison of genome organization among three coronaviruses is given in Table 1.

## Pathogenesis

### Endocytosis and replication

The endocytosis occurs when the glycoprotein of the viral spike binds to the receptor of the cell, and the host protease stimulates the S protein. For internalization, both viruses use ACE2 receptors and S-protein priming is done by TMPRSS2 serine proteases (Hoffmann et al., 2020). A higher expression of ACE2 receptors leads to a higher risk of COVID-19. The affinity of SARS-CoV-2 spikes protein binding to the ACE2 receptor is about 20 times higher than the SARS-CoV (Wrapp et al., 2020). Once the spike protein binds to the ACE2 receptor, a conformational change takes place. The fusion of the viral envelope protein in the cell membrane of the host allows the entry of the virus through the endosomal pathway

(Coutard et al., 2020). After undergoing translation, the virus produces replicase polyproteins pp1a and pp1b. Proteinases produced by the virus cleave these proteins into smaller proteins. During the viral replication phase, multiple copies of sub-genomic RNA by irregular transcription are produced along with full-length genomic production due to ribosomal frameshifting in translation. When the viral RNA, endoplasmic reticulum and Golgi complex form an interaction, the virion assembly takes place and is released out of the cells with the help of vesicles.

### Symptoms

The SARS incubation period usually remains under 2-7 days, and it can be extended to almost 10 days. The symptoms of coronavirus appear generally with a fever (>38°C) which is often quite high accompanied by rigors and chills (Kumar et al., 2020). The other symptoms include, headache, malaise, muscle pain, mild respiratory issues and diarrhea in some cases.

After 3-7 days, a dry cough and eventually shortness of breath is observed. These can further progress to hypoxemia i.e., low level of oxygen. In about 10-20% of infected patients, the respiratory issues are severe enough and require ventilation. In many cases, chest radiographs appear normal during the illness phase but not for all patients. In some patients, white blood cell count is also decreased, and it can result in a low level of platelets at the peak of the disease that causes



further complications (Kumar et al., 2020).

### **Omicron Aspect of SARS family**

#### **Phylogenetic analysis**

Phylogenetic analysis is a very important technique to determine the relationship of a new micro-organism with the existing one. This allows us to put a new strain on a specific family and subfamily and then formulate preventive measures based on the information about the family.

The phylogenetic analysis clusters the SARS-CoV-2 inside beta coronaviruses including SARS-CoV which led the researchers to its induction in the subgenus Sarbecovirus of genus betacoronavirus. SARS-CoV-2 is categorized in a separate lineage that contains other viruses RaTG13, RmYN02, ZC45, and ZXC21 taken from the four horseshoe bat coronaviruses along with the one identified in pangolins (Gorbalenya et al., 2020; Yang et al., 2020).

#### **SARS-CoV-2 functional genomics analysis**

The source, spread and intermediate host can be determined using the SARS-CoV-2 sequences. During the early research, 96% identical sequences were observed against RaTG13 confirming the presence of an ACE2 receptor in both (Zhou et al., 2020). Therefore, it was suggested that the pangolins were the carriers of SARS-CoV-2 (Zhang et al., 2020). The variant analysis discovered important mutations and their associated risk (Plante et al., 2021). Several mutation analysis studies have reported an in-depth molecular understanding of the genome functionality of SARS-CoV-2.

The researchers found three crucial mutations that were associated with geographical distributions associated with each of the three subtypes (Zhou et al., 2020). It relies significantly on the binding affinity of ACE2 with the spike glycoprotein for transmission and replication kinetics (Zhu et al., 2022). A study showed the binding affinity of N439K S protein to be quite high with ACE2 receptors in human beings. These N439K were similar in all the aspects like symptoms, infection cycle, and replication to the wild type (Thomson et al., 2021). In addition to this, Benton with his colleagues found out that the G614 open

conformation increases infectivity and predominance of SARS-CoV-2 (Benton et al., 2021), and Hou with his coworkers carried out research on the D614G virus variant for the viral pathogenesis and transmission. These variants were found with function to enhance the infectivity and high transmission of the virus in human and mouse model cells (Hou et al., 2020). Most of the research work demonstrated the functional effect of the mutation, and it was observed that the virus variants have an improved capacity to spread than its pathogenicity (Zhou et al., 2021). Furthermore, the clinical data proposed that the alteration of D614G has no substantial association with the severity of the disease and it would not change the effectiveness of vaccine targets under development (Korber et al., 2020). It was also found that N501Y contributes to a higher rate of transmissions, ranges from 40 to 70% (Gu et al., 2020), 484K and K417N induce an immunogenic response to the antibodies (Wibmer et al., 2021).

Furthermore, two coronavirus variants E484Q and L452R have been found in India and reported to be the main cause of the surge in April 2021 cases (Zhu et al., 2022). Three other SARS-CoV-2 variants emerged in the UK (B.1.1.7 (501Y.V1)), South Africa (B.1.351 (501Y.V2)), and Brazil (B.1.1.28.1 (P.1)) (Planas et al., 2021). A new variant, B.1.617.2 (delta variant) has been reported to spread faster as compared to other coronaviruses (Fontanet et al., 2021).

The mutations mostly happen due to changes in spikes, and they were analyzed using the bioinformatics tools to understand the functional changes associated with the mutations. Several bioinformatics tools have been employed to predict the effect of these mutations on the structural stability of viral proteins (Pereira et al., 2020). The variants N439K, N501Y, and K417N have been reported to have a neutral impact on the stability of the protein. PPA Pred web server (that uses mutational data to predict the effect of mutations on binding between the two proteins) has also been used for the prediction of ACE-spike protein binding (Brielle et al., 2020). All the mutations were found to increase the ability, except for N439K and N501Y, causing an increased viral infection (Zhu et al., 2022).



**Table-2: Integration of SARS-CoV-2 Specific Databases and Genomics Resources**

Resource Name	Web Link	Data
GISAID	<a href="https://www.gisaid.org/">https://www.gisaid.org/</a>	Strains
Nextstrain	<a href="https://nextstrain.org/sars-CoV-2/">https://nextstrain.org/sars-CoV-2/</a>	Strains
GESS	<a href="https://wan-bioinfo.shinyapps.io/GESS/">https://wan-bioinfo.shinyapps.io/GESS/</a>	Strains
NCBI SARS-CoV-2 Resources	<a href="https://www.ncbi.nlm.nih.gov/sars-CoV-2/">https://www.ncbi.nlm.nih.gov/sars-CoV-2/</a>	Genomics Data
National Genomics Data Center	<a href="https://bigd.big.ac.cn/">https://bigd.big.ac.cn/</a>	Genomics Data
European Nucleotide Archive	<a href="https://www.ebi.ac.uk/ena/browser/home">https://www.ebi.ac.uk/ena/browser/home</a>	Genomics Data
COVID-19 data portal	<a href="https://www.Covid19dataportal.org/">https://www.Covid19dataportal.org/</a>	Genomics Data
The UCSC SARS-CoV-2 Genome Browser	<a href="https://genome.ucsc.edu/Covid19.html">https://genome.ucsc.edu/Covid19.html</a>	Genome Resources
The COVID-19 corpus	<a href="https://www.semanticscholar.org/cord19">https://www.semanticscholar.org/cord19</a>	Literature
LitCoVid	<a href="https://www.ncbi.nlm.nih.gov/research/coronavirus/">https://www.ncbi.nlm.nih.gov/research/coronavirus/</a>	Literature
BioRxiv & MedRxiv	<a href="https://connect.biorxiv.org/relate/content/181">https://connect.biorxiv.org/relate/content/181</a>	Preprints
Drug bank	<a href="https://go.drugbank.com/Covid-19">https://go.drugbank.com/Covid-19</a>	Drug Database
COVID19 Drug Repository	<a href="http://www.Covid19.md.biu.ac.il/">www.Covid19.md.biu.ac.il/</a>	Drug Database
Coronavirus3D	<a href="http://www.coronavirus3d.org/">www.coronavirus3d.org/</a>	Protein Structure Data
CoV3D	<a href="http://www.Cov3d.ibbr.umd.edu/">www.Cov3d.ibbr.umd.edu/</a>	Protein Structure Data
RCSB PDB	<a href="http://www.rcsb.org/">www.rcsb.org/</a>	Protein Structure Data
UniProt	<a href="http://www.uniprot.org/">www.uniprot.org/</a>	Proteome Data

### Integration of database resources for COVID-19

As the omics analysis on the SARS-CoV-2 increased, genome sequencing for several types of research was carried out. It resulted in the production of tones of data, which needed to be reserved in databases for current and future references. GISAID is the first database that has free access and contains genome sequence data of the first virus. GISAID supports multiple studies on viral genome sequence analysis. Furthermore, services like sequence alignment, 3D protein models, diagnostic primers, phylogenetic analysis, and drug targets are offered here. So, it is safe to say that GISAID is one of the high-quality database resources that provide us information on viruses related to regional analysis, mutation analysis, and sequence mutations (Zhu et al., 2022). Some of the databases which contains the sequence information with their websites are given in Table 2.

### Transcriptomics data and immune response

The transcriptomics data analysis allows us to understand the quantitative RNA expression of the organism (Castro et al., 2022). RNA sequencing data taken from infected COVID-19 patients has revealed that the viral infection causes reduced antiviral defense accompanied by high production of cytokines. Through different studies, the immunological

characteristics are also identified which are associated with blood, lung, and bronchoalveolar lavage fluid, showing an association between SARS-CoV infection and its link to increased cytokine production (Blanco-Melo et al., 2020).

Currently, the immunological response of patients to viral infection by using single cell technology is a very emerging field of research. Immune samples taken from the infected COVID-19 patients have revealed diverse blood profiles. A diverse immune response landscape was demonstrated and the pathway of immune response to COVID is revealed by the use of single cell technology (Huang et al., 2021).

The immune abnormalities caused by COVID-19 also can cause bacterial infections. Thus, many studies were performed to understand the role of microbiome in COVID-19 patients (Cao et al., 2021). These studies also recognized the interaction between the virus and microbiota. These studies might be beneficial in the prediction of diagnosis, treatment, and prognosis. Besides this, tremendous amounts of research is being conducted on the use of probiotics for adjunctive therapy for alleviation and prophylaxis of symptoms (Bottari et al., 2021).

The transcriptomic data of COVID 19 was also analyzed using clinical bioinformatics and it was found that SARS-CoV-2 can affect the pattern of gene

expression for many genes that includes the immune-related molecules, like IL6, IFIH1, DDX58, STAT1, IRF7, ISG15, IFIT3, TNF, IRF9, and MX1. The studies have shown that the alterations in the expression of IL6 and TNF are closely associated with an increased level of plasma cytokines level in coronavirus-infected patients (Darif et al., 2021). The enrichment analysis also demonstrated that the expression level changes are also closely linked to the neurological symptoms, cardiac dysfunction, and coagulation dysfunction in SARS-CoV-2 infected patients (Zhang et al., 2021).

Most of other evidence revealed that disturbances in the signaling pathway like TL4 (toll like receptor) signaling pathway cascade can affect immune cells such as B and T cells which causes the generation of Th17 releasing cytokines leading to an illness associated with acquired and innate immune systems. In a study by Li and his colleagues, it was revealed that the molecules like NF- $\kappa$ B and IL 6 induce COVID-19 innate immune antagonism (Attiaq et al., 2021). Also, SARS can increase the cell apoptosis to MAPK and STAT3 signaling pathways (Hemmat et al., 2021). Thus, SARS-CoV-2 can affect many types of immune-associated molecules and signaling cascades. These types of SARS-CoV-2 infection cases possess a distinct impact on immune cell proliferation, differentiation, chemotaxis, and migration (Zhang et al., 2021). Therefore, there is a need for further studies to elucidate the underlying molecular mechanisms and factors.

### **Drug, Vaccine, and Therapeutics for SARS family**

#### **Drug and therapeutics**

Target selection and identification are very important for the drug development pathway. This allows us to identify the target with high specificity with existing drugs and treatments that can then be used to cure COVID-19 (Chen et al., 2020; Hodgson, 2020).

#### **Vaccines**

In the last decade, significant development has been made for distinct types of vaccines through techniques such as subtractive genomics, recombinant protein vaccines, whole virus vaccine, structural vaccinology, subunit vaccine. With the outbreak of COVID-19 pandemic, the main target of the scientific community was to develop a vaccine that is not only biologically safe but also has few side effects and curtails its follow-up for waves (Kaur and Gupta, 2020). Due to the genetic similarity of SARS-CoV-2 with SARS-

CoV and MERS CoV, their vaccines were considered potentially useful against COVID-19 (Ahmed et al., 2020). Furthermore, SARS and MERS-CoV infection is effectively controlled by neutralizing the antibodies produced against S-protein and it was also considered as a prospective approach to control COVID-19 (Zhou et al., 2020). In genome comparison of two viruses, it was identified that the spike proteins of CoV-2 were quite variable in their S1 subunit, thus, antibody neutralization is ineffective against COVID-19 (Ishack and Lipner, 2021).

Currently, several teams are working for developing vaccines. In one of the approaches, the researchers are working to develop inactive, protein subunit, vector-based, virus-like particles, and RNA-DNA-based vaccines against COVID-19 (Grifoni et al., 2020). Previously, for SARS-CoV, different types of vaccines were developed which were able to protect animals against infection but their use caused lung damage and eosinophil infiltration in model mice and infections in the liver in ferrets (Abdelmageed et al., 2020). Typically, a new vaccine follows research and development (R&D), clinical trials and final approval from a regulatory authority. This process mostly requires 12-18 months and sometimes even more (Ishack and Lipner, 2021). Chen and his colleagues categorized vaccines for COVID-19 into three broad categories (Ishack and Lipner, 2021); Subunit Vaccine, Nucleic Acid Vaccine and Whole Virus Vaccine.

The vaccine platform for COVID-19 is divided into recombinant protein, RNA-DNA, viral vector-based, inactivated, live, and attenuated vaccines. The main characteristic of the COVID-19 vaccine is the consideration of unwanted immunoprotection minimization, stockpiling suitability, and taking into account hypertension or diabetes in patients (Han et al., 2021).

Most of the vaccine developing organizations for SARS-CoV-2 infection are targeting S protein antibodies in human beings by delivering S antigen through vaccine inoculation. The first vaccine used for COVID-19 was mRNA-1273, utilizing the S protein in a distinct type of lipid nanoparticles for injection (Fang et al., 2022). This is developed by Moderna therapeutics. It is thought that after the injection of this vaccine, the host cell will take in the foreign mRNA, and then the immune system will generate an immune response against the S protein to prevent infection and invasion. Similarly, another mRNA vaccine is under research by Clover biopharmaceuticals, China (Ishack



and Lipner, 2021).

### Role of Bioinformatics in SARS Drug Discovery

In silico studies and analyses have gained a lot of acceptance in the last two decades. It is one of the most rapidly growing disciplines which has applications in almost all the fields and its potential is quite visible in the last decade. One of the important domains of in silico analysis in medical sciences is the development of an efficient way for vaccine and drug discovery by the use of computational tools. Several in silico researches were done to predict the potential drugs and chemical compounds for the treatment and cure of COVID-19 patients (Ray et al., 2021).

Many machine learning techniques were implemented for screening of chemical compounds that are suitable for the target protein.

Molecular docking, virtual screening and molecular dynamic simulation are among the most widely used computational techniques for drug discovery against pathogens including viruses. Molecular docking envisages the region of one molecule to another when they are bound in the form of a stable complex. Molecular docking helps to predict interaction of chemical compounds or ligands to a target biomolecule in order to find the activity and affinity of the potential drug molecule (Broni et al., 2021). On the other hand, molecular dynamic simulation is a machine learning technique for drug discovery. In this method, the structures of drug target compounds are predicted in silico and then they are simulated inside a virtual host environment for stability analysis of in silico structures or molecularly docked compounds. Besides these, homology modeling, distribution, drug-likeness, metabolism, absorption, excretion (ADME) analysis, and network-based identification are used extensively for screening of many potential drug candidates against pathogens (Manne, 2021).

### Recommendations and limitations

This review focused primarily on the multi-omics approaches applied on the CoV-2 genome to elucidate evolutionary behavior of the drug and vaccine targets. This limitation hinders the insights that can be gained through the research focused on proteomics of CoV-2. Therefore, it is recommended that research should be designed and executed to elucidate drug and vaccine targets that remain conserved over time in CoV-2. Structural bioinformatics and machine learning techniques should be applied to predict the probabilities of possible mutations that can arise in the

CoV-2 genome. Such insights can allow researchers to design drugs and vaccines that can target current and future mutations in CoV-2 target proteins.

### Conclusion

The SARS-CoV-2 is a novel pathogenic virus that has caused multiple fatalities in the world. Multiple potential drugs and vaccines have been developed for CoV-2. However, due to rapid mutational changes in the CoV-2, the development of vaccines for new variants of CoV-2 is becoming a challenging task. Therefore, incorporating bioinformatics and immunoinformatics, the vaccine development processes can be enhanced to counter new COV-2 variants.

### Acknowledgement

Authors would like to express deep thanks to the Biotechnology Department, University of Anbar for great support in accessing the source for references of this work.

**Disclaimer:** None.

**Conflict of Interest:** None.

**Source of Funding:** None.

### References

- Abdelmageed MI, Abdelmoneim AH, Mustafa MI, Elfadol NM, Murshed NS, Shantier SW and Makhawi AM, 2020. Design of a Multi-epitope-Based Peptide Vaccine against the E Protein of Human COVID-19: An Immunoinformatics Approach. *BioMed Res. Int.* 2020: 2683286. <https://doi.org/10.1155/2020/2683286>
- Abdelrahman Z, Li M and Wang X, 2020. Comparative Review of SARS-CoV-2, SARS-CoV, MERS-CoV, and Influenza A Respiratory Viruses. *Front. Immunol.* 11: 552909. <https://doi.org/10.3389/fimmu.2020.552909>
- Ahmed SF, Quadeer AA and McKay MR, 2020. Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies. *Viruses.* 12(3): 254. <https://doi.org/10.3390/v12030254>
- Attiq A, Yao LJ, Afzal S and Khan MA, 2021. The triumvirate of NF- $\kappa$ B, inflammation and cytokine storm in COVID-19. *Int. Immunopharmacol.* 101:



108255.  
<https://doi.org/10.1016/j.intimp.2021.108255>
- Benton DJ, Wrobel AG, Roustan C, Borg A, Xu P, Martin SR, Rosenthal PB, Skehel JJ and Gamblin SJ, 2021. The effect of the D614G substitution on the structure of the spike glycoprotein of SARS-CoV-2. *Proc. Nat. Acad. Sci. USA.* 118(9). e2022586118.  
<https://doi.org/10.1073/pnas.2022586118>
- Blanco-Melo D, Nilsson-Payant BE, Liu WC, Uhl S, Hoagland D, Møller R, Jordan TX, Oishi K, Panis M, Sachs D, Wang TT, Schwartz RE, Lim JK, Albrecht RA and tenOever BR, 2020. Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell.* 181(5): 1036-1045.e9.  
<https://doi.org/10.1016/j.cell.2020.04.026>
- Bottari B, Castellone V and Neviani E, 2021. Probiotics and CoVid-19. *Int. J. Food Sci. Nutr.* 72(3): 293–299.  
<https://doi.org/10.1080/09637486.2020.1807475>
- Brielle ES, Schneidman-Duhovny D and Linial M, 2020. The SARS-CoV-2 Exerts a Distinctive Strategy for Interacting with the ACE2 Human Receptor. *Viruses.* 12(5): E497.  
<https://doi.org/10.3390/v12050497>
- Broni E, Kwofie SK, Asiedu SO, Miller WA and Wilson MD, 2021. A Molecular Modeling Approach to Identify Potential Antileishmanial Compounds Against the Cell Division Cycle (cdc)-2-Related Kinase 12 (CRK12) Receptor of *Leishmania donovani*. *Biomolecules.* 11(3), 458.  
<https://doi.org/10.3390/biom11030458>
- Cao J, Wang C, Zhang Y, Lei G, Xu K, Zhao N, Lu J, Meng F, Yu L, Yan J, Bai C, Zhang S, Zhang N, Gong Y, Bi Y, Shi Y, Chen Z, Dai L, Wang J and Yang P, 2021. Integrated gut virome and bacteriome dynamics in COVID-19 patients. *Gut Microbes.* 13(1): 1–21.  
<https://doi.org/10.1080/19490976.2021.1887722>
- Castro ISL, Freitas-Lopes R, do L Ferreira S, de S Maciel TEF, Florez JC, Zambolim EM, Zambolim L and Caixeta ET, 2022. Transcriptome Analysis Uncovers the Gene Expression Profile of *Hemileia vastatrix* (Race XXXIII) during the Interactions with Resistant and Susceptible Coffee. *Agronomy.* 12(2): 444.  
<https://doi.org/10.3390/agronomy12020444>
- Chan JF, Kok KH, Zhu Z, Chu H, To KK, Yuan S and Yuen KY, 2020. *Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan.* *Emerg. Microbes Infect.* 9(1): 221-236.
- Chen WH, Strych U, Hotez PJ and Bottazzi ME, 2020. The SARS-CoV-2 Vaccine Pipeline: An Overview. *Curr. Trop. Med. Rep.* 7(2): 61–64.  
<https://doi.org/10.1007/s40475-020-00201-6>
- Coutard B, Valle C, de Lamballerie X, Canard B, Seidah NG and Decroly E, 2020. The spike glycoprotein of the new coronavirus 2019-nCoV contains a furin-like cleavage site absent in CoV of the same clade. *Antivir. Res.* 176: 104742. doi: 10.1016/j.antiviral.2020.104742
- Cui J, Li F and Shi ZL, 2019. Origin and evolution of pathogenic coronaviruses. *Nature Rev. Microbiol.* 17(3): 181–192. <https://doi.org/10.1038/s41579-018-0118-9>
- Darif D, Hammi I, Kihel A, El Idrissi Saik I, Guessous F and Akarid K, 2021. The pro-inflammatory cytokines in COVID-19 pathogenesis: What goes wrong? *Microb. Pathogen.* 153: 104799.  
<https://doi.org/10.1016/j.micpath.2021.104799>
- Deng SQ and Peng HJ, 2020. Characteristics of and public health responses to the coronavirus disease 2019 outbreak in China. *J. Clin. Med.* 9(2): 575.
- Deslandes A, Berti V, Tandjaoui-Lambotte Y, Alloui C, Carbonnelle E, Zahar JR, Briclher S and Cohen Y, 2020. SARS-CoV-2 was already spreading in France in late December 2019. *Int. J. Antimicrob. Agents.* 55(6): 106006.  
<https://doi.org/10.1016/j.ijantimicag.2020.106006>
- Fang E, Liu X, Li M, Zhang Z, Song L, Zhu B, Wu X, Liu J, Zhao D and Li Y, 2022. Advances in COVID-19 mRNA vaccine development. *Sig. Transd. Target. Therap.* 7(1): 1–31.  
<https://doi.org/10.1038/s41392-022-00950-y>
- Fontanet A, Autran B, Lina B, Kieny MP, Karim SSA and Sridhar D, 2021. SARS-CoV-2 variants and ending the COVID-19 pandemic. *Lancet.* 397(10278): 952–954.  
[https://doi.org/10.1016/S0140-6736\(21\)00370-6](https://doi.org/10.1016/S0140-6736(21)00370-6)
- Fung M, Otani I, Pham M and Babik J, 2021. Zoonotic coronavirus epidemics: Severe acute respiratory syndrome, Middle East respiratory syndrome, and coronavirus disease 2019. *Ann. Allerg. Asthma Immunol.* 126(4): 321–337.  
<https://doi.org/10.1016/j.anai.2020.11.021>
- Gorbalenya A, Baker S, Baric R, de Groot R, Drosten C, Gulyaeva A, Haagmans B, Lauber C, Leontovich A, Neuman B, Penzar D, Perlman S,



- Poon L, Samborskiy D, Sidorov I, Sola I and Ziebuhr J, 2020. The species Severe acute respiratory syndrome-related coronavirus: Classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiol.* 5: 536–544. <https://doi.org/10.1038/s41564-020-0695z>
- Grifoni A, Sidney J, Zhang Y, Scheuermann RH, Peters B and Sette A, 2020. A Sequence Homology and Bioinformatic Approach Can Predict Candidate Targets for Immune Responses to SARS-CoV-2. *Cell Host Microbe.* 27(4): 671–680.e2. <https://doi.org/10.1016/j.chom.2020.03.002>
- Gu H, Chen Q, Yang G, He L, Fan H, Deng YQ, Wang Y, Teng Y, Zhao Z, Cui Y, Li Y, Li XF, Li J, Zhang NN, Yang X, Chen S, Guo Y, Zhao G, Wang X, Luo DY, Wang H, Yang X, Li Y, Han G, He Y, Zhou X, Geng S, Sheng X, Jiang S, Sun S, Qin CF and Zhou Y, 2020. Adaptation of SARS-CoV-2 in BALB/c mice for testing vaccine efficacy. *Science.* 369(6511): 1603–1607. <https://doi.org/10.1126/science.abc4730>
- Guo YR, Cao QD, Hong ZS, Tan YY, Chen SD, Jin HJ, Tan KS, Wang DY and Yan Y, 2020. The origin, transmission and clinical therapies on coronavirus disease 2019 (COVID-19) outbreak—an update on the status. *Mil. Med. Res.* 7(1): 11.
- Han HJ, Nwagwu C, Anyim O, Ekweremadu C and Kim S, 2021. COVID-19 and cancer: From basic mechanisms to vaccine development using nanotechnology. *Int. Immunopharmacol.* 90: 107247. <https://doi.org/10.1016/j.intimp.2020.107247>
- Han Q, Lin Q, Jin S and You LC, 2020. Coronavirus 2019-nCoV: A brief perspective from the front line. *J. Infect.* 80(4):373–377. doi: 10.1016/j.jinf.2020.02.010.
- Hemmat N, Asadzadeh Z, Ahangar NK, Alemohammad H, Najafzadeh B, Derakhshani A, Baghbanzadeh A, Baghi HB, Javadrashid D, Najafi S, Ar Gouilh M and Baradaran B, 2021. The roles of signaling pathways in SARS-CoV-2 infection; lessons learned from SARS-CoV and MERS-CoV. *Arch. Virolog.* 166(3): 675–696. <https://doi.org/10.1007/s00705-021-04958-7>
- Hodgson J, 2020. The pandemic pipeline. *Nature Biotechnol.* 38(5): 523–532. <https://doi.org/10.1038/d41587-020-00005-z>
- Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA, Drosten C and Pöhlmann S, 2020. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell.* 181(2):271–280.
- Hou YJ, Chiba S, Halfmann P, Ehre C, Kuroda M, Dinnon KH 3rd, Leist SR, Schäfer A, Nakajima N, Takahashi K, Lee RE, Mascenik TM, Graham R, Edwards CE, Tse LV, Okuda K, Markmann AJ, Bartelt L, de Silva A, Margolis DM, Boucher RC, Randell SH, Suzuki T, Gralinski LE, Kawaoka Y and Baric RS, 2020. SARS-CoV-2 D614G variant exhibits efficient replication ex vivo and transmission in vivo. *Science.* 370(6523): 1464–1468. <https://doi.org/10.1126/science.abe8499>
- Huang L, Shi Y, Gong B, Jiang L, Zhang Z, Liu X, Yang J, He Y, Jiang Z, Zhong L, Tang J, You C, Jiang Q, Long B, Zeng T, Luo M, Zeng F, Zeng F, Wang S, Yang X and Yang Z, 2021. Dynamic blood single-cell immune responses in patients with COVID-19. *Sig. Transd. Target. Therap.* 6(1): 110. <https://doi.org/10.1038/s41392-021-00526-2>
- Ishack S and Lipner SR, 2021. Bioinformatics and immunoinformatics to support COVID-19 vaccine development. *J. Med. Virol.* 93(9): 5209–5211. <https://doi.org/10.1002/jmv.27017>
- Kaur SP and Gupta V, 2020. COVID-19 Vaccine: A comprehensive status report. *Virus Res.* 288: 198114. <https://doi.org/10.1016/j.virusres.2020.198114>
- Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI; Sheffield COVID-19 Genomics Group, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO and Montefiori DC, 2020. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell.* 182(4): 812–827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>
- Kumar S, Maurya VK, Prasad AK, Bhatt MLB and Saxena SK, 2020. Structural, glycosylation and antigenic variation between 2019 novel coronavirus (2019-nCoV) and SARS coronavirus (SARS-CoV). *Virusdisease.* 31(1): 13–21. <https://doi.org/10.1007/s13337-020-00571-5>
- Manne R, 2021. Machine Learning Techniques in Drug DisCoVery and Development. *Int. J. Appl. Res.* 7(4): 21–28.



- <https://doi.org/10.22271/allresearch.2021.v7.i4a.8455>
- Park WB, Kwon NJ, Choi SJ, Kang CK, Choe PG, Kim JY, Yun J, Lee GW, Seong MW, Kim NJ, Seo JS and Oh MD, 2020. Virus Isolation from the First Patient with SARS-CoV-2 in Korea. *J. Korean Med. Sci.* 35(7): e84. <https://doi.org/10.3346/jkms.2020.35.e84>
- Pereira R, Oliveira J and Sousa M, 2020. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. *J. Clin. Med.* 9(1): 132. <https://doi.org/10.3390/jcm9010132>
- Planas D, Veyer D, Baidaliuk A, Staropoli I, Guivel-Benhassine F, Rajah MM, Planchais C, Porrot F, Robillard N, Puech J, Prot M, Gallais F, Gantner P, Velay A, Le Guen J, Kassis-Chikhani N, Edriss D, Belec L, Seve A, Courtellemont L, Péré H, Hocqueloux L, Fafi-Kremer S, Prazuck T, Mouquet H, Bruel T, Simon-Lorière E, Rey FA and Schwartz O, 2021. Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization. *Nature.* 596(7871): 276–280. <https://doi.org/10.1038/s41586-021-03777-9>
- Plante JA, Mitchell BM, Plante KS, Debbink K, Weaver SC and Menachery VD, 2021. The variant gambit: COVID-19's next move. *Cell Host Microbe.* 29(4): 508–515. <https://doi.org/10.1016/j.chom.2021.02.020>
- Platto S, Xue T and Carafoli E, 2020. COVID19: An announced pandemic. *Cell Death Dis.* 11(9): 799. <https://doi.org/10.1038/s41419-020-02995-9>
- Ray M, Sable MN, Sarkar S and Hallur V, 2021. Essential interpretations of bioinformatics in COVID-19 pandemic. *Meta Gene.* 27: 100844. <https://doi.org/10.1016/j.mgene.2020.100844>
- Stokes EK, 2020. Coronavirus Disease 2019 Case Surveillance—United States, January 22–May 30, 2020. *MMWR. Morb. Mort. Week. Rep.* 69. <https://doi.org/10.15585/mmwr.mm6924e2>
- Thomson EC, Rosen LE, Shepherd JG, Spreafico R, da Silva Filipe A, Wojcechowskyj JA, Davis C, Piccoli L, Pascall DJ, Dillen J, Lytras S, Czudnochowski N, Shah R, Meury M, Jesudason N, De Marco A, Li K, Bassi J, O'Toole A, Pinto D, Colquhoun RM, Culap K, Jackson B, Zatta F, Rambaut A, Jaconi S, Sreenu VB, Nix J, Zhang I, Jarrett RF, Glass WG, Beltramello M, Nomikou K, Pizzuto M, Tong L, Cameroni E, Croll TI, Johnson N, Di Iulio J, Wickenhagen A, Ceschi A, Harbison AM, Mair D, Ferrari P, Smollett K, Sallusto F, Carmichael S, Garzoni C, Nichols J, Galli M, Hughes J, Riva A, Ho A, Schiuma M, Semple MG, Openshaw PJM, Fadda E, Baillie JK, Chodera JD; ISARIC4C Investigators; COVID-19 Genomics UK (COG-UK) Consortium, Rihn SJ, Lycett SJ, Virgin HW, Telenti A, Corti D, Robertson DL and Snell G, 2021. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell.* 184(5): 1171-1187.e20. <https://doi.org/10.1016/j.cell.2021.01.037>
- Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT and Veesler D, 2020. Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell.* 181(2):281-292.
- Wibmer CK, Ayres F, Hermanus T, Madzivhandila M, Kgagudi P, Oosthuysen B, Lambson BE, de Oliveira T, Vermeulen M, van der Berg K, Rossouw T, Boswell M, Ueckermann V, Meiring S, von Gottberg A, Cohen C, Morris L, Bhiman JN and Moore PL, 2021. SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma. *Nature Med.* 27(4): 622–625. <https://doi.org/10.1038/s41591-021-01285-x>
- Wrapp D, Wang N, Corbett KS, Goldsmith JA, Hsieh CL, Abiona O, Graham BS and McLellan JS, 2020. Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science.* 367(6483): 1260-1263.
- Wu JT, Leung K and Leung GM, 2020. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet.* 395(10225): 689-697.
- Yan R, Zhang Y, Li Y, Xia L, Guo Y and Zhou Q, 2020. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science.* 367(6485): 1444–1448. <https://doi.org/10.1126/science.abb2762>
- Yang Y, Xiao Z, Ye K, He X, Sun B, Qin Z, Yu J, Yao J, Wu Q, Bao Z and Zhao W, 2020. SARS-CoV-2: Characteristics and current advances in research. *Virology.* 17(1): 117. <https://doi.org/10.1186/s12985-020-01369-z>
- Zhang H, Chen H, Zhang J, Chen X, Guo B, Zhi P, Li Z, Liu G, Yang B, Chi X and Wang Y, 2021. Bioinformatics analysis of SARS-CoV-2 infection-associated immune injury and therapeutic prediction for COVID-19. *Emerg. Crit. Care Med.* 1(1): 20–28. <https://doi.org/10.1097/EC9.000000000000005>



- Zhang T, Wu Q and Zhang Z, 2020. Probable Pangolin Origin of SARS-CoV-2 Associated with the COVID-19 Outbreak. *Curr. Biol.* 30(7): 1346-1351.e2. <https://doi.org/10.1016/j.cub.2020.03.022>
- Zhou B, Thao TTN, Hoffmann D, Taddeo A, Ebert N, Labroussaa F, Pohlmann A, King J, Steiner S, Kelly JN, Portmann J, Halwe NJ, Ulrich L, Trüeb BS, Fan X, Hoffmann B, Wang L, Thomann L, Lin X, Stalder H, Pozzi B, de Brot S, Jiang N, Cui D, Hossain J, Wilson MM, Keller MW, Stark TJ, Barnes JR, Dijkman R, Jores J, Benarafa C, Wentworth DE, Thiel V and Beer M, 2021. SARS-CoV-2 spike D614G change enhances replication and transmission. *Nature.* 592(7852): 122–127. <https://doi.org/10.1038/s41586-021-03361-1>
- Zhou P, Yang XL, Wang XG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Jiang RD, Liu MQ, Chen Y, Shen XR, Wang X, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF and Shi ZL, 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature.* 579(7798): 270–273. <https://doi.org/10.1038/s41586-020-2012-7>
- Zhou Y, Hou Y, Shen J, Huang Y, Martin W and Cheng F, 2020. Network-based drug repurposing for novel coronavirus 2019-nCoV/SARS-CoV-2. *Cell Discov.* 6: 14. <https://doi.org/10.1038/s41421-020-0153-3>
- Zhu Z, Zhang S, Wang P, Chen X, Bi J, Cheng L and Zhang X, 2022. A comprehensive review of the analysis and integration of omics data for SARS-CoV-2 and COVID-19. *Brief. Bioinformat.* 23(1). <https://doi.org/10.1093/bib/bbab446>

#### **Contribution of Authors**

All authors contributed equally in this article.

