International Journal of
Life Sciences
International Peer Reviewed Open Access Refereed Journal

Int. J. of Life Sciences, 2022; 10 (4): 301-307
ISSN:2320-7817(p) | 2320-964X(o)

Research Article Open Access

Fundamentals of SARS CoV-2 Proteome and Variants

Nirjar Bhattacharya

Dulles High School, TX, USA
Email: njarbinx@gmail.com

Manuscript details:
Received: 04.09.2022
Accepted: 05.11.2022
Published: 31.12.2022

Cite this article as:
Nirjar Bhattacharya (2022) Fundamentals of SARS CoV-2 Proteome and Variants., *Int. J. of Life Sciences*, 10 (4): 301-307.

Available online on http://www.ijlsci.in
ISSN: 2320-964X (Online)
ISSN: 2320-7817 (Print)

ABSTRACT

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is a novel coronavirus that is responsible for the COVID-19 pandemic. The currently ongoing COVID-19 pandemic has had devastating effects on human health, resulting in 219 million recorded cases and 4.55 million deaths worldwide. Beyond its direct impact on healthcare, this pandemic has also disrupted the daily lives of millions and will likely continue to do alter lifestyles for years through its negative impact on the economy. This paper aims to simplify the SARS-CoV-2 virus so that anybody can understand its molecular function. This paper also explains how this virus is changing with time and how these changes may affect the world's response to this virus. The information regarding SARS-CoV-2 and its structure is important since knowing this information allows people around the world to learn about this seemingly indomitable virus, to understand how the virus works, and how these variations each have a different impact on their body. All this information will increase the overall public knowledge of the virus and help them avoid misinformation and possible threats from SARS-CoV-2.

Keywords: *SARS-CoV-2, biology, viruses, proteins, mutations.*

INTRODUCTION

SARS-CoV-2 is a betacoronavirus, the same genera as other clinically important coronaviruses SARS-CoV and MERS-CoV. These viruses are enveloped positive-sense RNA viruses which infect mammals. Coronaviruses are characterized by club-like spikes on the surface, which appear as a corona, or crown, around the virion in electron microscopy images. Coronaviruses enter cells via binding of the viral spike (S) protein to the cellular receptor angiotensin-converting enzyme 2 (ACE2) (Mittal *et al.,* 2020). Upon entry, virions replicate their (+) ssRNA genome in double-membrane vesicles formed of restructured endoplasmic reticulum (ER) membranes. Host membranes are also used for assembly and envelopment of the SARS-CoV-2 virion prior to egress out of the cell. Like other beta coronaviruses, The SARS-CoV-2 virus is comprised of nearly 29

proteins that are used to hijack the host cell and promote viral replication. In this review, we are studying SARS-CoV-2 proteins and how they interact with host cells with the aim of better understanding why this virus is more transmissible than past coronaviruses. Furthermore, by studying and summarizing the molecular underpinnings of SARS-CoV-2 replication we hope to further reveal the pathogenesis, or the development of a disease, of COVID-19. Lastly, this paper seeks to explain how the immune system reacts to SARS-CoV-2 and how mutations alter this response. An initial infection must first bypass the innate immune system, which in addition to providing physical barriers to infection, also activates general defense mechanisms. The innate immune system then triggers the adaptive, or more targeted, immune system. The adaptive immune system provides a more specified response through neutralizing antibodies and cytotoxins produced by B and T adaptive cells which target the invading pathogen. These cells are also used to create a memory of infection so that secondary infections can be quickly stopped. Overall, this paper is used the increase the public knowledge of this novel virus.

2. Structural Proteins:

Of the 29 SARS-CoV-2 proteins, four are structural proteins that make up the physical virion that surrounds and protects the viral genome. This class of viral proteins consists of Spike (S), membrane/matrix (M), envelope (E), and nucleocapsid (N) proteins (Fig 1A). The spike protein, displayed on the viral envelope, is commonly targeted by host antibodies and serves as the key that the virus uses to bind to the locks found on the host cell's surface. In both SARS-CoV and SARS-CoV-2, the unique crown structure of the S protein allows for binding and unlocking of the host ACE2 receptor to facilitate viral entry. The M protein is the most abundant protein in the virus, and it defines the shape of the viral envelope by interacting with and stabilizing other viral proteins, such as spike. M protein is a glycoprotein that is also essential for viral assembly. In contrast, The E protein is the least abundant of the major structural proteins and participates in viral assembly and egress in host cells. Furthermore, the E protein plays a role in virus pathogenesis and may act as a virulence factor. Finally, the N protein is the only structural protein that binds to the RNA genome and is also involved in viral

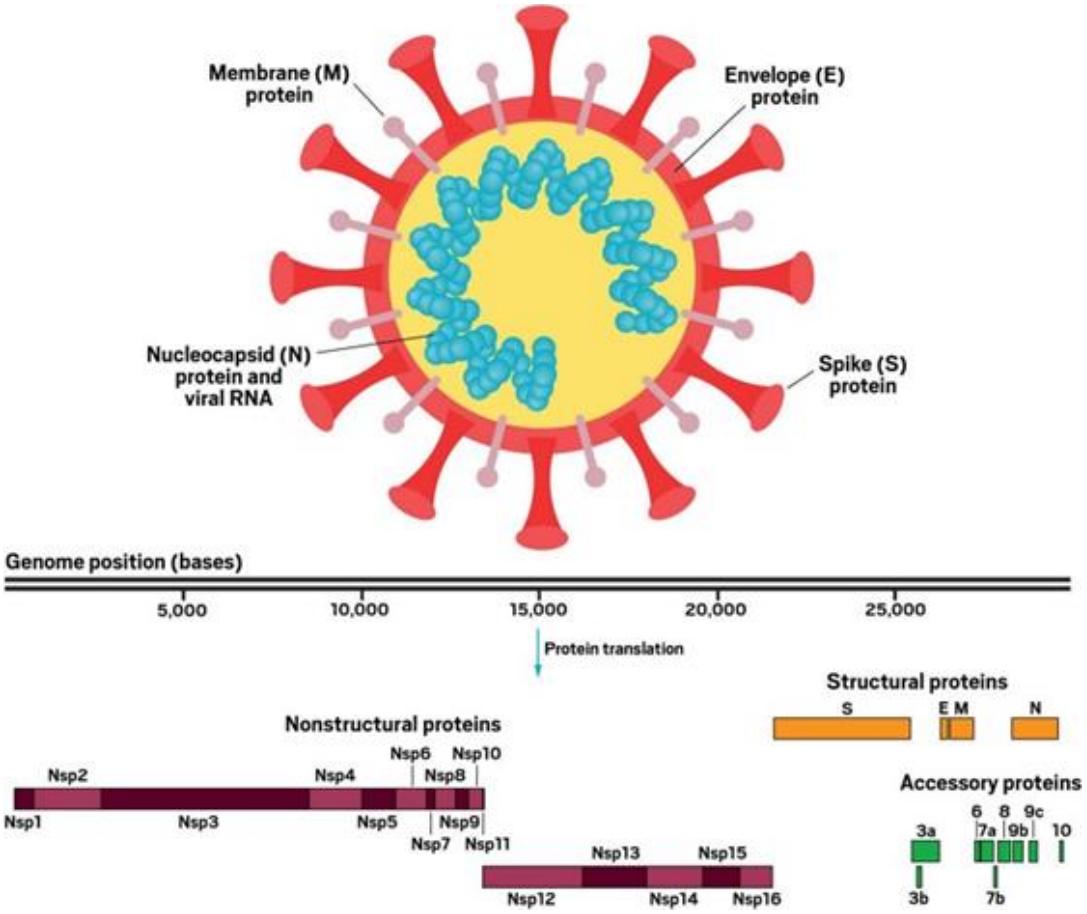

Figure 1 - Protein Breakdown of COVID-19 Virus-

assembly, budding, and genome packaging. The N protein does this by recognizing and associating with the genomic RNA before self-associating into an oligomer, or a polymer with relatively few repeating units, to form the capsid containing the genetic information (Mousavizadeh and Ghasemi, 2020).

3. Nonstructural Proteins:

The non-structural SARS-CoV-2 proteins regulate and coordinate viral genome replication as well as immune evasion. These nonstructural proteins (NSPs) are expressed as two large polyproteins, pp1a and pp1ab, which are subsequently cleaved into smaller functional proteins by the viral proteins Papain-like protease (Plpro) and 3C-like protease (3CLpro) (Fig 1B). These cleavage products form important NSPs such as helicase and RNA-dependent RNA polymerase (RdRp). RdRp catalyzes the replication of RNA from an RNA template, a function that is unique to viral proteins and is therefore used to replicate the SARS-CoV-2 genome using the initial viral genome as a template. Another key NSP is RNA helicase which helps with gene expression by unwinding RNA and clamping protein complexes to strands of RNA. NSPs are generally characterized as enzymes or transcription factors, which the virus uses to replicate. For example, RNA replicase, is an enzyme encoded by SARS-CoV-2 which catalyzes RNA replication and alters the host's functions by regulating the replication of genetic information within the nucleus, nucleolus, and other host structures. However, despite the importance of this class of proteins, many of the virus's NSPs are still poorly understood (Yadav *et al.,* 2021).

4. Accessory Proteins:

The least understood class of SARS-CoV-2's proteins are the 9 accessory proteins. These proteins are not structural or used in viral entry or replication but instead are important to other parts of viral pathogenesis and interference of immune signaling. This indicates that they can aid in viral maturation as well as help to impair the host's immune response. Specifically, several of these proteins, such as ORF8 discussed below, seek to disrupt interferon and cytokine signaling, which are two key axes in host anti-viral signaling and immune activation. Accessory proteins are currently being researched to gain a deeper understanding of how the difference in the structure of these proteins affects their function (Redondo *et al.,* 2021).

5. Other SARS-CoV-2 Strains:

Due to the high error rate of RdRp, mistakes often occur during viral replication resulting in the production of virions with varying genomes. Thus, like there are many types of proteins in viruses, there are several types of strains or variants. There are multiple variants, or a virus bearing a mutation that leads to a significant change from the original virus, of the SARS-CoV-2 genome that are different from the version first detected in China. Among these are the beta and alpha strains from the original waves of the pandemic. These strains are much less infective than the delta strain. The delta strain is a newer variant that is a "variant of concern" by the CDC because it is 40% to 60% more infectious than other major strains. The delta strain is more infectious due to its increased ability to evade the immune system. Another significant difference between the strains is in the amino acid sequence of the S protein. These mutations help the virus avoid the possible dangers from immune cells such as white blood cells and memory cells, cells that record memories of the virus to allow for a faster immune response upon secondary infection. Essentially, these genetic variants are emerging around the world driving the pandemic forward with their mutated proteins, altered structure, and thus altered functions (Pachetti *et al.,* 2020).

6. Spike (S) Protein Structure, Functions, and Mutations:

The spike (S) protein is the sole viral membrane protein responsible for cell entry. It binds to the receptor on the target cell and mediates subsequent virus-cell fusion and entry into host cells. With a size of 180–200 Kilodaltons (KDa), the coronavirus spike protein is a multifunctional molecular machine that consists of an extracellular N-terminus, a transmembrane (TM) domain anchored in the viral membrane, and a short intracellular C-terminal segment (Fig 3A). The spikes are also coated with polysaccharide molecules to camouflage them, evading surveillance of the host immune system during entry. Spike proteins start the viral replication process by first binding to the ACE2 receptors on the host cell surface through its S1 subunit and then fuses viral and host membranes through its S2 subunit which leads to attachment to the host cell (Fig. 2). The spike protein exists in two structurally distinct states, prefusion and postfusion. The transition from the prefusion to the postfusion state of the spike protein must be triggered

by host proteins, leading to membrane fusion (Duan *et al.,* 2020).

More in-depth about the entry process, the initial attachment of the virion to the host cell is initiated by interactions between the S protein and the host's ACE2 receptor. The S-protein–receptor interaction is the primary determinant for how the coronavirus infects a host. Many coronaviruses utilize ACE2 as their cellular receptor to enter mammalian cells. Following receptor binding, the virus must next enter the host cell and its cytosol or inner fluid. This is generally accomplished by cleavage of S protein by a protease. The SARS-CoV-2 S protein is cleaved and activated by two host proteases, TMPRSS2 and Furin. This is followed by the fusion of the viral and cellular membranes. The S protein cleavage occurs at two sites within the S2 portion of the protein. The first cleavage is important for separating the RBD and the fusion domains of the S protein, while the second is important for exposing the fusion peptide, which assists with the previously mentioned fusion of host and viral membranes (V'kovski *et al.,* 2021). This fusion occurs at the plasma membrane where the virus joins two series of 7 repeating amino acids, to form a protein that mediates membrane fusion (Fig 3B). The formation of these bundles allows for the mixture of viral and cellular membranes, resulting in fusion and ultimately release of the viral genome into the cytoplasm and eventually the double-membrane vesicles (DMV) to hijack the host cell. This is how the S protein facilitates viral entry into a host cell. The next step is the translation of the replicase gene from the virion genomic RNA and then the translation and assembly of the viral replicase complexes. Following replication and genomic RNA synthesis, encapsulation of the virus occurs resulting in the formation of the mature virus. Following assembly, virions are transported to the cell surface in vesicles and released by exocytosis (Huang *et al.,* 2020).

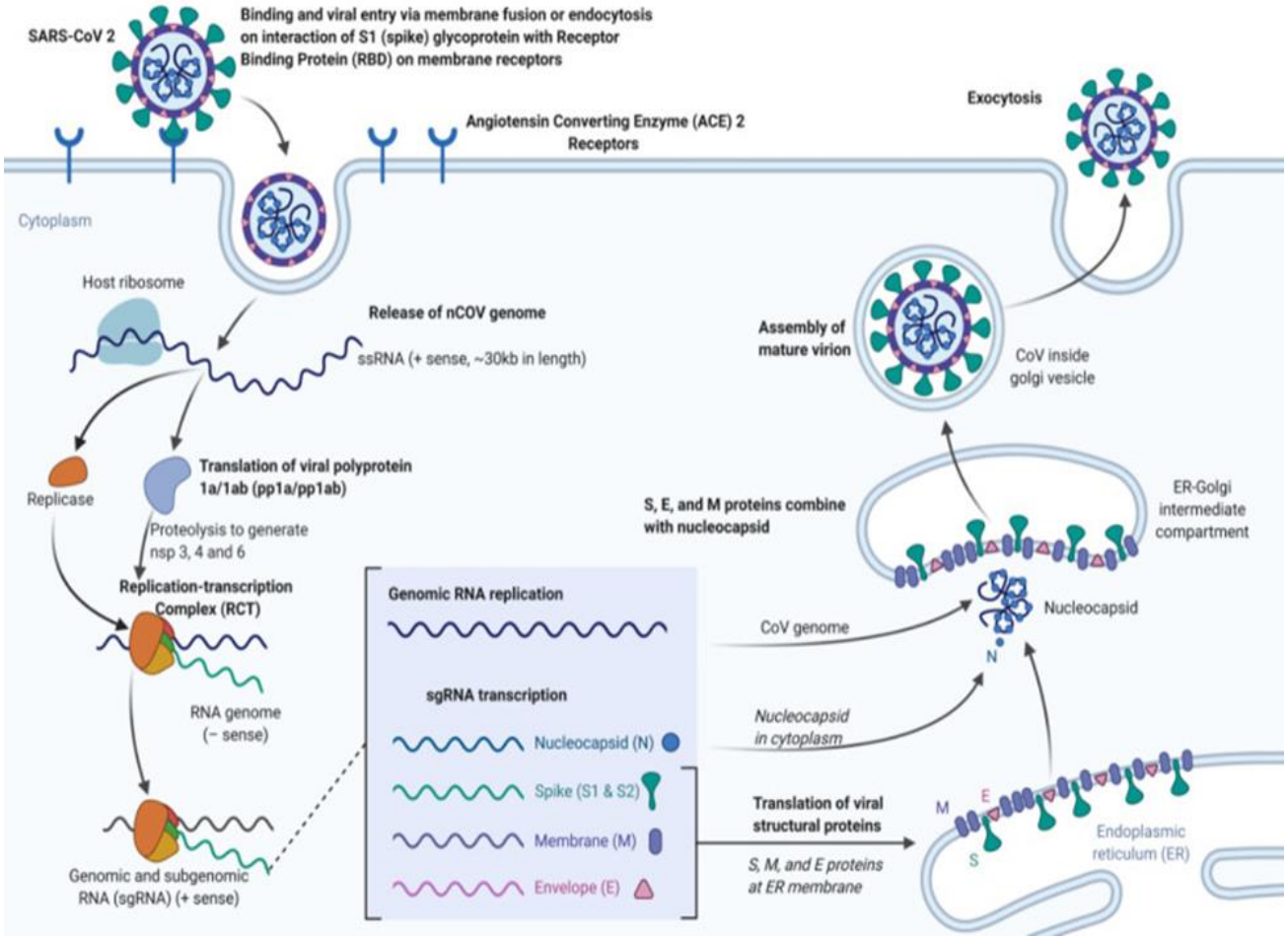

Figure 2 - Viral Replication Breakdown-

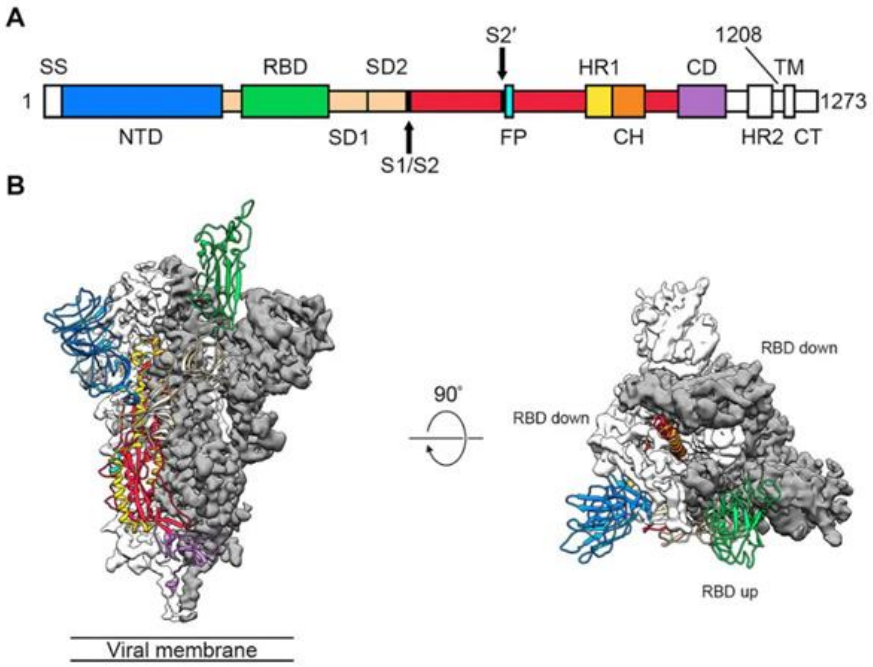

Figure 3 : Spike Protein Breakdown

Since the Spike protein is a key protein for the initiation of viral replication and entry, changes to the S protein may significantly alter the virus' functions or methodology for entering the host cell. Therefore, mutations within the S protein, specifically the RBD in the S1 subunit, often result in variant formation. Thus, S1 is considered a hotspot for mutations that may have high clinical relevance in terms of virulence, transmissibility, and host immune evasion. One dangerous S1 mutation is K484E, which replaces the positively charged lysine K with negatively charged glutamic acid E. This mutation results in significantly increased ACE2 binding affinity with decreased affinity to neutralizing antibodies, potentially indicating how new variants, like the delta strain, have higher transmissibility and virulence. Another mutation at D614G is carried by over 99% of major variants since early 2020. This mutation does not change the binding affinity to ACE2 or neutralizing antibodies for the virus. Instead, it may lead to an increase in spike density by preserving the integrity of the spike and avoiding S1 degradation. This allows for more functional spikes to be available and results in viruses with the D614G variants being armed with increased infectivity. Another common type of mutation in the spike proteins are deletions in the RBD which wipe out epitopes, or the part of an antigen molecule to which an antibody attaches itself to start the immune response, which eventually allows the virus to evade the host's immune response and causing certain memory cells to become irrelevant or defective (Khateeb, et. Al., 2021). Due to its important functions, the S protein is commonly targeted by neutralizing antibodies from the host's immune system and in vaccines against SARS-CoV-2. Therefore, understanding the function of spike mutations and predicting novel mutations is critical for developing lasting therapeutics. While these are the most common and dangerous characterized mutations there are many other mutations, to the spike protein which have altered its infectivity rate and binding affinity to antibodies and host cell receptors or helped in evading the immune response (Harvey, et. al., 2021)

7. Mutations in NSPs and Accessory Proteins:

Two other mutation hot spots are in the nonstructural protein NSP1 and the accessory protein ORF8. Both proteins have been found to be related to SARS-CoV-2 virulence and transmissibility. NSP1 antagonizes type I interferon (IFN) induction in the host to dampen immune activation. SARS-CoV-2 also uses NSP1 to suppress cellular, but not viral, protein synthesis by plugging the ribosome mRNA entry channel as well as for more unknown mechanisms (Zhang, et. al., 2021). Similarly, ORF8 is known as an immune-evasive protein that downregulates major histocompatibility

complex class I (MHC-I) and type I IFN signaling pathways in host cells. By doing this both NSP1 and ORF8 inhibit the production of well-known stimulators of antiviral genes and the adaptive immune system and inhibitors of viral replication in infected cells. Thus, variants with partial deletion of NSP1 have substantially limited virus viability and thus contribute to milder infections. Likewise, several mutations in SARS-CoV-2 ORF8 proteins have been identified. For example, the alpha variant contains a premature stop codon at position 27 of ORF8 (Benedetti, et. al., 2020). These mutations often cause truncated NSP1 and ORF8 and account for less than 5% of infections worldwide. This might be due to the virus attempting to not kill the host quickly and rather replicate as much as possible. Other more common mutations of these proteins have the opposite effect where the proteins have enhanced functions and therefore antagonize IFN and other immune responses more strongly to hinder the eradication of SARS-CoV-2. Some examples of these mutations are in the S24L and W45L mutations in the ORF8 protein (Rashid, et. al., 2021). These proteins are some of the fastest evolving viral proteins in beta coronaviruses which suggests that these proteins have a direct correlation to infectivity and clinical outcomes.

CONCLUSION

Increased mutation rates allow viruses to adapt to changing environments indicating why understanding these mutations is important to understanding and neutralizing new virus variants. This paper goes into the structure and mutations of the S protein, NSP1 protein, and ORF8 protein. These viral proteins are critical in viral replication due to their roles in viral entry or immune evasion. Knowing the structure of these viral proteins has permitted us to recognize mutations, where they occur, and how they impact protein function. This is important since the protein mutations referenced have impacted the infectivity and functionality of SARS-CoV-2 to increase transmissibility. This further understanding of these proteins will allow for greater predictions for the ongoing pandemic and future pandemics. For example, mutational analysis has already been put into therapeutic use with the analysis of NSP1, where scientists have shown that variants with mutations that result in partial deletion of NSP1's amino acid chain have resulted in weaker variants. This has shown that NSP1 could be a future therapeutic target for drugs to suppress SARS-CoV-2.

Conflict of Interest: None of the authors have any conflicts of interest to disclose. All the authors approved the final version of the manuscript.

REFERENCES

Benedetti, F., Snyder, G. A., Giovanetti, M., Angeletti, S., Gallo, R. C., Ciccozzi, M., Zella, D. (2020, August 31). *Emerging of a SARS-COV-2 viral strain with a deletion in NSP1*. Journal of Translational Medicine. https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-020-02507-5.

Duan, L., Zheng, Q., Zhang, H., Niu, Y., Lou, Y., Wang, H. (1AD, January 1). *The sars-COV-2 spike glycoprotein biosynthesis, structure, function, and antigenicity: Implications for the design of spike-based vaccine immunogens*. Frontiers. https://www.frontiersin.org/articles/10.3389/fimmu.2020.576622/full.

Huang, Y., Yang, C., Xu, X.-feng, Xu, W., Liu, S.W. (2020, August 3). Structural and functional properties of SARS-COV-2 spike protein: Potential antivirus drug development for covid-19. Nature News. https://www.nature.com/articles/s41401-020-0485-4.

Khateeb, J., Li, Y., Zhang, H. (2021, July 12*). Emerging sars-COV-2 variants of concern and potential intervention approaches*. Critical Care. https://ccforum.biomedcentral.com/articles/10.1186/s13054-021-03662-x.

Mittal, A., Manjunath, K., Ranjan, R. K., Kaushik, S., Kumar, S., Verma, V. (n.d.). *Covid-19 pandemic: Insights into structure, function, and hACE2 receptor recognition by SARS-COV-2. PLOS Pathogens*. https://journals.plos.org/plospathogens/article?id=10.1371%2Fjournal.ppat.1008762.

Pachetti, M., Marini, B., Benedetti, F., Giudici, F., Mauro, E., Storici, P., Masciovecchio, C., Angeletti, S., Ciccozzi, M., Gallo, R. C., Zella, D., Ippodrino, R. (2020, April 22). *Emerging sars-COV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. Journal of Translational Medicine*. https://translational-medicine.biomedcentral.com/articles/10.1186/s12967-020-02344-6.

Rashid, F., Suleman, M., Shah, A., Dzakah, E. E., Wang, H., Chen, S., Tang, S. (1AD, January 1). *Mutations in SARS-COV-2 ORF8 altered the bonding network with interferon regulatory factor 3 to evade host immune system*. Frontiers. https://www.frontiersin.org/articles/10.3389/fmicb.2021.703145/full.

Redondo, N., Zaldívar-López, S., Garrido, J. J., Montoya, M. (1AD, January 1*). SARS-COV-2 accessory proteins in viral pathogenesis: Knowns and unknowns*. Frontiers. https://www.frontiersin.org/articles/10.3389/fimmu.2021.708264/full.

Schoeman, D., Fielding, B. C. (2019, May 27). *Coronavirus envelope protein: Current knowledge*. Virology journal. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6537279/.

V'kovski, P., Kratzel, A., Steiner, S., Stalder, H., Thiel, V. (2020, October 28). *Coronavirus Biology and replication: Implications for SARS-COV-2*. Nature News. https://www.nature.com/articles/s41579-020-00468-6.

Zhang, K., Miorin, L., Makio, T., Dehghan, I., Gao, S., Xie, Y., Zhong, H., Esparza, M., Kehrer, T., Kumar, A., Hobman, T.C., Ptak, C., Gao, B., Minna, J.D., Chen, Z., García-Sastre, A., Ren, Y., Wozniak, R.W., Fontoura, B.M.A. (n.d.). *NSP1 protein of SARS-COV-2 disrupts the mrna export machinery to inhibit host gene expression*. Science advances. https://pubmed.ncbi.nlm.nih.gov/33547084/.