Proofreading function and mutations in SARS-CoV-2 and their impact on viral infectivity

Tiffany Lin

Abstract

SARS-CoV-2 is a positive-sense single-stranded RNA virus that causes the respiratory disease, COVID-19 (coronavirus disease 19) and is a member of the severe acute respiratory syndrome-related coronavirus species in the Coronaviridae family. The virus has led to more than 590 million infections and more than 6.4 million deaths worldwide as of the end of August 2022. The coronavirus family also includes SARS-CoV, which is the virus that led to the 2003 SARS epidemic, affecting 26 countries at the time. Mutations in SARS-CoV-2 have accumulated since the emergence of the virus in 2019, and new variants are increasingly observed. SARS-CoV-2 has a unique proofreading function that has been shown to be carried out by a protein complex that corrects mismatched base pairs during replication. This paper aims to address the degree that the SARS-CoV-2's proofreading mechanism results in fewer protein mutations in its spike glycoprotein compared to other viruses. The proofreading function limits the number of mutations that survive RNA replication for the SARS-CoV-2 virus. Examining the proofreading function of the SARS-CoV-2 virus will reveal important insights into its mutation rates and changes in viral infectivity, suggesting pathways for treatments in the future.

Introduction

SARS-CoV-2 is a single-stranded positive-sense RNA virus, with a viral diameter of 69-140 nanometers, that is able to infect humans and cause a family of symptoms commonly referred to as "COVID-19", named as an abbreviated form for **co**rona **vi**rus **d**isease that originated in 20**19**. SARS-CoV-2 can spread rapidly through the projection of respiratory droplets from the mouth or nose when breathing, coughing, or sneezing. Most people infected with the virus will experience mild to moderate respiratory illness and recover without requiring special treatment. However, there is the possibility that some people will become seriously ill and require immediate medical attention. Nearly 90% of people hospitalized for COVID-19 have underlying conditions according to the CDC's Morbidity and Mortality Weekly Report; the most common underlying conditions include hypertension (49.7%), obesity (48.3%), chronic lung disease (34.6%), diabetes mellitus (28.3%), and cardiovascular disease (27.8%). One study showed that 74.5% of those hospitalized due to coronavirus were age 50 or older, with the highest rates among those over 65 (Brenza, 2020). This shows that older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illnesses (WHO, 2021). Current vaccinations for COVID-19 include the Pfizer-BioNTech vaccine called Comirnaty, Spikevax by ModernaTX, the Novavax vaccine also known as Nuvaxovid, and the Johnson & Johnson vaccine called Janssen. Studies and current data show that antibodies made in the human body after vaccination and a booster shot recognize and protect humans against a few COVID-19 variants, such as Omicron. Though studies also show that the protection and vaccination effects (original COVID-19 booster shots) can wane substantially over time (National Institutes of Health, 2022).

COVID-19 originated in Wuhan, China, and has spread to over 228 countries and territories including the most severely infected countries such as the United States, India, France, and

Brazil (Worldometer, 2022). COVID-19 is part of the coronavirus family, including common viruses that cause a variety of diseases from the head or chest cold to more severe but rare diseases, like severe acute respiratory syndrome (SARS) and Middle East respiratory syndrome (MERS)(CDC, 2021). The severe acute respiratory syndrome (SARS) pandemic originated in 2002 and was eventually brought under control in July 2003; during the period of infection, there were 8,098 reported cases of SARS and 774 deaths (NHS, 2019). The Middle East respiratory syndrome (MERS) originated in 2012 and had a total of 2,591 laboratory-confirmed cases and 882 associated deaths by the end of July 2022 (WHO, 2022). The SARS-CoV-2 virus has led to a worldwide pandemic, causing over 590 million infections and 6.4 million deaths as of the end of August 2022.

Viruses are infectious agents of small size which usually vary in diameter from 20 nanometers to 250-400 nanometers (Britannica, 2021) and simple composition including a nucleoid, capsid, envelope, and enzymes, and can only multiply in living cells of animals, and plants, or bacteria (Krug, 2022). Viruses can infect the human body by recognizing and attaching to host cells; binding to cellular receptors to facilitate their attachment to the host cell. Once a virus attaches to the host cell, it is engulfed through the cellular membrane and subsequently enters the cell's cytoplasm. Inside the cell, the virus will remove its viral coat into smaller cellular vesicles, thereby releasing its genetic material into the cytoplasm for its replication (Robertson, 2022). All viruses consist of short sequences of nucleic acid, which can be in the form of ribonucleic acid (RNA) or deoxyribonucleic acid (DNA) as their genetic material. Compared to most other organisms where DNA is always a double-stranded structure, viruses are unique because they can be differentiated based on how they store their genomic information; their DNA or RNA material can be either single-stranded or double-stranded. Viruses that contain DNA as their genetic material are called 'DNA viruses', whereas the viruses containing RNA as their genetic material are called 'RNA viruses'. Regardless of the type of genetic material contained by a virus, the DNA or RNA is covered by a protein capsid, with some viruses containing an additional envelope covering the capsid.

Major differences between DNA and RNA viruses include the process of replication, size of the genome, and mechanism of protein synthesis. One of the main differences between DNA and RNA viruses is in how their genetic material is replicated. Firstly, the cellular compartment where DNA or RNA is replicated has major differences. Upon infecting a host, the replication of the viral DNA occurs inside the nucleus of the host cell, as opposed to viral RNA which must first undergo transcription and then it is replicated in the cytoplasm. Additionally, DNA and RNA viruses also have different size genomes caused by the accuracy of replication. DNA viruses can contain large genomes due to the high accuracy of DNA replication, resulting from nucleotide selectivity, DNA proofreading, and mismatch repair; whereas many RNA viruses contain small genomes due to their error-prone replication (Dey, 2022). During protein synthesis, viral DNA needs to be transcribed into RNA then mRNA first, before it is translated into viral proteins; on the other hand, RNA viruses can bypass the step of transcription during protein synthesis as they are already RNA which may help the virus increase the rate of protein synthesis (Lakna, 2017).

Mutations can be caused by one or more mismatched base pairs in DNA or RNA, or the presence of uncomplimentary bases in double-stranded DNA or RNA replication. Even though

SARS-CoV-2 is a single-stranded RNA virus, it produces a double-stranded viral DNA during its replication processes using an RNA-dependent RNA polymerase (RNA replicase) to convert viral RNA into a complementary strand of DNA, then copying it to produce a double-stranded replicative form of viral DNA. This means that an error in matching base pairs may occur in the replication stage of the single-stranded RNA virus (Barr & Fearns, 2022). Examples of mismatched base pairs include an A/C or G/T pairing, as opposed to their natural A/T and G/C pairing. These mismatches may appear most commonly due to errors in replication, though can arise from spontaneous deamination of cytosine and/or adenine, or during homologous recombination (U.S. National Library of Medicine, 2022). Failure to repair a mismatched base pair could result in many downstream events within the virus including changes in the protein sequence and structure; possibly affecting the host, though, sometimes these mutations can go unnoticed or have negligible effects.

One mechanism that organisms have adapted to ameliorate the impact of random mutations is through a process called proofreading. Proofreading occurs in a wide array of organisms, including human cells and viruses. In the DNA replication of human cells, proofreading occurs to prevent mutations or permanent changes in the DNA sequence. Although replication errors and DNA damages occur naturally in the cells of our bodies frequently, it is unlikely to cause mutations as they are mostly detected and fixed by DNA proofreading and other repair mechanisms. When a DNA polymerase detects an incorrectly paired nucleotide, if the proofreading action proceeds correctly, the mismatched nucleotide is removed and replaced before continuing with DNA synthesis. One of these proofreading mechanisms is called ‘mismatch repair’ and, in human cells, begins when the heterodimers MutSa (MSH2/MSH6) and MutSβ (MSH2/MSH3) protein complexes recognize the base-base mismatches and latter insertion-deletion loops. Then, the heterodimer MutLa (MLH1/PMS2) forms a ternary complex with one of the MutS complexes and promotes the repair process (Hong et al., 2008) by excising the incorrect nucleotide and replacing the missing section with correct nucleotides; with the enzyme DNA ligase sealing the gap (DAV, 2020). Thus, having a proofreading function may allow a different rate of mutations, thereby impacting a virus’s rate of infectivity.

Viral infectivity is defined as the capacity of viruses to invade a host cell and exploit all its resources to replicate, which may lead to infection and subsequent diseases in the human host (Rodríguez et al., 2013). One of the ways RNA viruses increase or decrease their level of infectivity is through a mutation. A mutation is a change in the nucleotide sequence of the DNA or RNA strand and the effect of a mutation depends on the region in which the sequence of genetic material has been changed (Biology Dictionary, 2017). Most commonly, mutations occur when there is a copying error in DNA or RNA replication, commonly referred to as mismatched base pairs, or when nucleotides are altered, inserted, or deleted. Mutations tend to occur more often in RNA because RNA polymerase, the enzyme responsible for RNA synthesis, lacks the same accuracy as DNA polymerase when transcribing the genetic code from DNA (Tatomir, 2022). All mutations (including viral mutations) occur randomly and may or may not continue to propagate via genetic selection for viruses to better adapt to their surroundings and more effectively transfer from host to host. Some beneficial functions of viral mutations could include helping the virus gain traits that help with its speed of reproduction or to better adhere to the surface of human cells (Sobhanie, 2021). However, because mutations are random, there is the possibility that some viral mutations will not affect a virus’ ability to infect or cause disease and

can even have a negative impact on virus survival, causing it to die without having long-term effects on the human body. For mutations that positively alter viral transmission and function, there could be severe consequences to the viral host. In humans, for example, viral mutations could allow them to become better at invading human immune systems, or if they gradually overcome or evade the available vaccinations and treatments. The SARS-CoV-2 virus has mutated over the past 3 years, resulting in genetic variations in the viral population. This genetic variation has impacted SARS-CoV-2's many properties such as transmission or the severity of symptoms of infected individuals.

In order to maintain the integrity of the SARS-CoV-2 viral genome and prevent lethal mutagenesis, the virus has also developed a proofreading mechanism designed to remove incorrect nucleotides or correct mismatched base pairs from newly synthesized viral RNA (Lin et al., 2022). As SARS-CoV-2 and other coronaviruses have a proofreading function, it causes them to mutate less frequently compared to other viruses. In this review, I will be discussing how this proofreading mechanism prevents mutations in SARS-CoV-2 and how that impacts the level and rate of infectivity in humans.

The Proofreading Complex in SARS-CoV-2

In order to minimize the number of genomic mutations, SARS-CoV-2 uses a proofreading mechanism. Proofreading is the process by which mismatched base pairs are detected and removed before the correct bases are added. The proofreading mechanism in SARS-CoV-2 RNA synthesis and repair involves a proofreading complex that is composed of nonstructural protein 14 (nsp14) and nonstructural protein 10 (nsp10). The RNA proofreading function of nsp14 is critical for viral viability and genome maintenance and is primarily conducted via its N-terminal 3'-to-5' exonuclease domain (ExoN). The ExoN domain is responsible for cleaving the nucleic acid chain. The sequence and structure of the nsp14 ExoN domain can be highly divergent from most cellular exonuclease enzymes but previous research has shown that nsp14 is highly conserved amongst coronaviruses. The nsp14 enzyme is strongly stimulated by its interaction with nsp10 (Fig. 1). Studies have also demonstrated that the ExoN domain of nsp14 relies on its negatively charged amino acids (generally three aspartic acids and one glumatic acid) accompanied by bivalent metal ions. Specifically, nsp14 has been shown to rely on Mg^{2+} for its catalytic activity, which is required for the cleavage step of the proofreading process (Rona et al., 2022). The nsp14 enzyme also contains a C-terminal guanine-N7-methyltransferase (N7-MTase) domain responsible for viral mRNA capping, which is an essential step for viral replication as the mRNA-cap can be used by viruses to translate their mRNAs into proteins. Additionally, the cap also provides protection of the viral RNA from degradation by cellular nucleases and prevents viral RNA recognition by immunity mechanisms (Bouvet et al., 2012).

As previously mentioned, nsp14's enzyme activity is dependent on binding to accessory protein nsp10 (Ma et al., 2021). Nsp10 is one of the sixteen nonstructural proteins encoded by the SARS-CoV-2 genome and it interacts with nsp14 to perform 3'-to-5' exoribonuclease activities, which is required for cleaving the strand containing the mismatched base. The nsp10 protein is synthesized as a polyprotein whose cleavage generates many non-structural proteins, including nsp7 and nsp8. The structure of nsp10 is composed of two zinc-binding sites, responsible for

stabilizing the C-terminus of the nsp10 protein, along with two anti-parallel helices which are stacked against an irregular beta-sheet that comprises a beta-subdomain (Lu, 2020).

Nsp14 association with nsp10 has been shown to promote the ability of the nsp14 cleavage domain to excise mismatched nucleotides. Through purification and in vitro assembly of nsp14/nsp10, it has been observed that binding to nsp10 enhances protein stability as it provides a 260-fold increase in the exoribonucleolytic activity of nsp14 (Riccio et al., 2022). Mutations in nsp14 that weaken proofreading function have been shown to critically reduce viral fitness, therefore emphasizing the importance of nsp14 function in SARS-CoV-2 survival and infectivity. Thus, inhibition of nsp14 is expected to be detrimental to coronavirus replication, which could make it a possible therapeutic target.

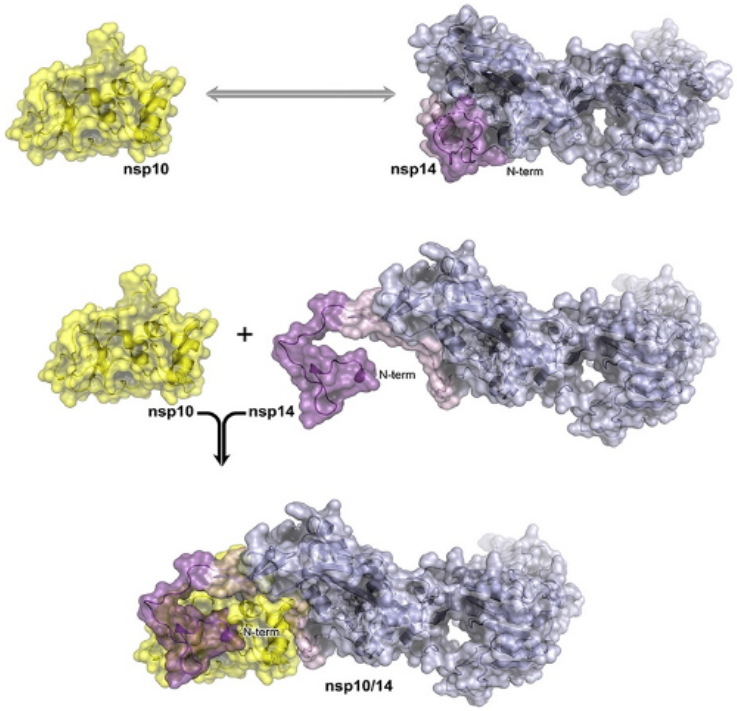

Figure 1. Binding of nonstructural protein 10/14 (from (Czarna et al., 2022))

Spike Protein Mutations in SARS-CoV-2

Despite SARS-CoV-2 containing a proofreading mechanism, mutations can still arise. The first critical step of viral infection is catalyzed by the spike protein, which is the main antigenic component of SARS-CoV-2. The spike protein on the surface of the SARS-CoV-2 virus interacts with the human angiotensin-converting enzyme 2 (ACE2) receptor to gain entry into a cell and initiate viral infections that lead to various levels of COVID-19 disease severity (Magazine et al., 2022). This spike glycoprotein is a type I membrane protein, which forms a trimer, anchored to the viral membrane by its transmembrane segment (Zhang et al., 2021). Positive evolutionary selection of mutations within the spike protein has led to the genesis of new SARS-CoV-2 variants with greatly enhanced overall fitness, which led to higher rates of viral infectivity. Thus,

the spike protein represents an important therapeutic target and is a critical antigen in host immune responses. The original variant of SARS-CoV-2's spike protein is the D614 variant, and 13 other mutated variants including the D614G variant, through to omicron (InvivoGen, 2022).

a. D614G variant

The D614G mutation in the viral spike protein occurred at the early stage of the coronavirus pandemic, emerging in late January to early February 2020. The D614G mutation is a missense mutation where an alteration in a single DNA base pair causes the substitution of aspartic acid (single-letter code: D) with glycine (single-letter code: G) at position 614 in the protein encoded by the mutated gene. This SARS-CoV-2 strain became the dominant form of the virus globally, replacing the initial strain identified in China, by June 2020. Glycine is a non-polar amino acid with a single hydrogen atom as its side chain; whereas aspartic acid is a polar amino acid with an acidic side chain. The codons that code for aspartic acid and glycine are GAC and GGC, respectively. Thus, a single mutation in the RNA codon causing the A to G shift can mutate the aspartic acid to glycine shift in the peptide sequence of the target protein. Using bioinformatic methods, scientists have found that D614G mutation in the viral spike protein is a frequently occurring mutation across many geographical locations (Dutta, 2021).

The D614G mutation variant has been shown to enhance the infectivity of pseudoviruses in the laboratory. Pseudoviruses are kinds of viruses used in the lab that help researchers obtain conclusions on a positive or negative influence of a mutation. A pseudovirus contains fragments of host-cell DNA without containing any of the nucleic acid components of the infectious virus to which they are related; and can be naturally produced during an infection or artificially produced in a laboratory for research purposes (Cuffari, 2021). As pseudotyped viruses, G614 variants have considerably higher infectious concentrations of an antibody (titers) than D614 variants. This indicates that the spike D614G mutation makes SARS-CoV-2 more infectious and transmissible from person to person, though it is not associated with increased disease severity. A study conducted by a team at New York University involved putting the D614G mutation into the human lung, liver, and colon cells, also using the original Wuhan version of the virus (also known as the "wild type") for clear comparisons. They found that the spike mutation definitely increased transmissibility, as much as eight times higher than the original version of COVID-19 (Open Access Government, 2021). Although there is no research showing that the D614G mutation is associated with increased COVID-19 severity, a study using more than 4000 coronavirus genomes claimed that viruses containing D614G mutation are more virulent, and thus, are associated with higher disease-related mortality (Dutta, 2021).

b. Omicron variant

The Omicron variants of SARS-CoV-2 were first discovered in Botswana in late November 2021, and various Omicron lineages have arisen since (Alex, 2022). The Omicron spike structure revealed an unusually tightly packed receptor binding domain (RBD) organization with long-range impacts that were not observed in other spike variants (Gobeil et al., 2022). It also has 15 receptor-binding domain amino acids, which is much more compared to other variants suggesting that cell entry mechanisms and subsequent replication rates may be different than previous variants (Marcos, 2022). The Omicron variant spike protein has 37 mutations, while

other variant spike proteins have had fewer. This is significant because small mutations on the spike protein have potentially great implications for the increased transmission of the virus, slower immune response from the human body, or reduced efficacy of vaccinations; as shown by the study mentioned in the D614G variant. Several mutations in the Omicron variant include R493, S496, and R498, which create new salt bridges and hydrogen bonds between the spike protein and the human cell receptor (ACE2). This appears to increase binding affinity, while other mutations in Omicron such as K417N, decrease the strength of this bond (Goldhawk, 2021).

The Omicron spike protein is far better at evading antibodies that are commonly used as treatments than other variants, as well as evading the immunity produced by both vaccines and natural infection. The Omicron spike protein's strong binding with human cells and increased antibody evasion are likely contributing factors to the increased transmissibility of the Omicron variant; which is why the variant spreads so rapidly and often becomes the dominant variant of SARS-CoV-2 (Goldhawk, 2021). From previous advanced research analysis using cryo-electron microscopy and X-ray crystallography, it has been shown that the Omicron variant bears more mutations in its spike protein due to the new salt bridges and hydrogen bonds as mentioned before. This means that there is a direct connection between mutations of this variant and new chemical interaction sites, though previous studies have shown the Omicron variant to have decreased its pathogenicity as opposed to its increased transmissibility (Bálint et al., 2022).

Proofreading Mechanism & Mutations in Other Viruses

Many organisms have a proofreading system, but many RNA viruses, like influenza for example, do not, and therefore lack the ability to fix mistakes in replication like SARS-CoV-2; and create a crowd of mutants around them. Coronaviruses are the only known organisms that encode an RNA-dependent, RNA-proofreading system (Denison et al., 2011). From previous research on frequencies of mutations accumulated, it proved an estimated spontaneous mutation rate of SARS-CoV-2 being $1.3x10^{-6} \pm 0.2x10^{-6}$ per base per infection cycle despite having proofreading capabilities (Amicone et al., 2022).

The Influenza virus is one of many viruses that do not have a proofreading mechanism. Influenza viruses are enveloped viruses of the Orthomyxoviridae family, which are classified into four genera, which include influenza virus A-D (IAV, IBV, ICV, and IDV). With regards to human health, IAVs and IBVs are of main concern; ICVs are endemic and only cause mild diseases in humans, and IDVs primarily cause infection in cattle (Flerlage et al., 2021). Mutations in influenza viruses occur frequently because the virus' replication machinery does not have a proofreading mechanism. Frequent mutational events explain the occurrence of seasonal influenza epidemics that may differ in severity and age groups affected, due to the lack of proofreading mechanisms (WHO, 2014). Through previous research, it is shown that Influenza viruses have a mutation rate of $2.0x10^{-6} \pm 0.6x10^{-6}$ per base per infection cycle(Nobusawa & Sato, 2006). Though both the Influenza virus and SARS-CoV-2 infect human respiratory systems through surface proteins and show similar symptoms, only SARS-CoV-2 has a proofreading mechanism, which results in mutations nearly half as often as in Influenza (Manzanares-Meza & Medina-Contreras, 2020).

Rhinovirus is another virus that does not have a proofreading mechanism. Rhinoviruses are non-enveloped, single-stranded positive-sense RNA viruses belonging to the Picornaviridae family and are divided into three species, including Rhinovirus A, -B, and -C, which are antigenically diverse (Waman et al., 2014). Rhinoviruses are a major cause of the common cold and other severe respiratory illnesses, including asthma and pneumonia. The genetic diversity of the Rhinovirus is attributed to a high mutation rate, which is due to the lack of proofreading activities and its error-prone RNA polymerase. The error rate of Picornavirus (which includes Rhinoviruses) RNA polymerases is estimated to range between 10^{-3} and 10^{-4} errors per nucleotide of a cycle of replication. A research group also conducted an experiment estimating an in vivo mutation rate of $3.4x10^{-4}$ per base per infection cycle during a five-day acute infection period, which is over 100x larger than the SARS-CoV-2 mutation rate (Tapparel et al., 2011).

Effects of Viral Transmissibility & Infectivity due to Proofreading Mechanisms

Table 1. Transmissibility of different viruses from their basic reproductive rate (R_0) and severity of illnesses for infected individuals (data from (Spencer et al., 2020)).

	SARS-CoV-2	Influenza	Rhinovirus
Transmissibility, R_0 (reproductive rate)	4.18	1.68	1.88
Incubation period, days (mean)	5.07	2.61	2.36
Infectious period, days (mean)	15.20	4.58	9.40
Hospitalization period, days (mean)	4.96	6.36	1.19

Transmissibility is defined as the quality of a disease being able to be passed on from one person to another (Oxford Languages, 2022). Thus, the transmissibility of respiratory viruses can be estimated by their basic reproduction number (R_0). The reproductive number is defined as the average number of successful transmissions per infectious individual in a population, thus the reproductive number is significant in reflecting the capacity of a virus to be transmitted (Leung, 2021). As transmissibility is mainly determined by the infectivity of the specific virus, the results above show that SARS-CoV-2 has the highest transmissibility with the reproductive rate of “4.18” among other viruses compared in the table; therefore suggesting that SARS-CoV-2 is more infective than other mentioned viruses (Table 1). This could possibly imply that although viruses with a proofreading function experience fewer rates of mutation, successful mutations are more severe in infectivity and illnesses. On the other hand, viruses without a proofreading

function tend to have quicker rates of mutation but those successful variants are less severe, as shown by the infectious period and hospitalization period for all viruses compared in the table.

Discussion

The proofreading mechanism in SARS-CoV-2 plays a significant role as it minimizes the frequency and rate of viral mutations. Without this mechanism, SARS-CoV-2 would have more frequent spontaneous mutations which have the possibility of positively affecting the reproduction rate of viruses and may help the virus to better adhere to the surface of human host cells. If SARS-CoV-2 did not have a proofreading mechanism it may threaten human health more severely as we have observed each new mutation variant of SARS-CoV-2 to possess higher transmissibility, disease severity, and ability to evade vaccine-induced and natural immunity. Thus, the possibility of increased mutational events could have highly detrimental impacts with regard to human health.

From the evaluation of mutation rates of SARS-CoV-2, Influenza, and Rhinoviruses, results suggest that the Influenza virus and the Rhinovirus have a much higher mutation rate than SARS-CoV-2 due to their lack of proofreading functions. Previous research showed that SARS-CoV-2 has a mutation rate of $1.3x10^{-6} \pm 0.2x10^{-6}$ per base per infection cycle, while Influenza and Rhinoviruses have a much higher mutation rate of $2.0x10^{-6} \pm 0.6x10^{-6}$ and $3.4x10^{-4}$, respectively. Thus, supporting the hypothesis that the proofreading mechanism of viruses are closely connected with the rate of viral mutation. Results also indicated that SARS-CoV-2 has a higher transmissibility among other viruses brought to comparison. The infectious period of SARS-CoV-2 is also much higher when compared to Influenza and Rhinoviruses. This suggests that despite a reduced mutation rate, SARS-CoV-2 is still more infectious. Two possible explanations for this difference could be that (1) SARS-CoV-2 disease severity would be much higher without proofreading or (2) limiting the number of viral mutations via proofreading increases viral infectivity, perhaps by limiting competition between variants.

Despite SARS-CoV-2 having a proofreading function, SARS-CoV-2 is still considered a highly infectious virus and is more infectious than similar viruses that do not have proofreading activity. We do not currently have evidence that proves whether or not proofreading activity is responsible for or involved in this difference in viral infectivity, though perhaps proofreading has decreased the infectivity of a potentially more dangerous virus. This can be of further investigation for scientists in the future to provide evidence to see if SARS-CoV-2 would be more dangerous without its proofreading mechanism.

Conclusion

To conclude, proofreading can become a possible therapeutic target by further investigating how SARS-CoV-2's level of infectivity might change if it did not have a proofreading mechanism to clarify the effect of proofreading on SARS-CoV-2, which could be potentially applicable to other viruses. From the increased viral infectivity observed by mutation variants of SARS-CoV-2 that enhance spike protein affinity for human cells, it is vital that the scientific community continue to share knowledge about viral infectivity for emerging dangerous variants. Additionally, awareness should be raised for the active implementation of protective and preventive measures that may

help reduce the transmissibility of SARS-CoV-2. Thereby lowering the risks of individuals being infected and developing serious illnesses with side effects detrimental to health. Though appropriate vaccination and therapeutic development should occur in a timely manner, as prevention is key and will be more useful than cure.

Acknowledgements

I would like to thank Polygence for providing a platform to conduct my review paper with the help of my mentor Lauren Sundby.

Bibliography

1. Alex, S. (2022, June 7). *Omicron lineage spike proteins use ACE2 receptors more efficiently*. News-Medical.

2. Amicone, M., Borges, V., Alves, M. J., Isidro, J., Zé-Zé, L., Duarte, S., Vieira, L., Guiomar, R., Gomes, J. P., & Gordo, I. (2022). Mutation rate of SARS-CoV-2 and emergence of mutators during experimental evolution. *Evolution, Medicine and Public Health*, *10*(1), 142–155. https://doi.org/10.1093/emph/eoac010

3. Bálint, G., Vörös-Horváth, B., & Széchenyi. (2022). Omicron: increased transmissibility and decreased pathogenicity. *Nature*, *151*(7).

4. Barr, J., & Fearns, R. (2022, August 18). *RNA virus*. Wikipedia.

5. Biology Dictionary. (2017, April 28). *Mutation*. Biology Dictionary.

6. Bouvet, M., Ferron, F., Imbert, I., Gluais, L., Selisko, B., Coutard, B., Canard, B., & Decroly, E. (2012). Stratégies de formation de la structure coiffe chez les virus à ARN. In *Medecine/Sciences* (Vol. 28, Issue 4, pp. 423–429). https://doi.org/10.1051/medsci/2012284021

7. Brenza, A. (2020, April 9). *Nearly 90% of People Hospitalized for COVID-19 Have Underlying Conditions, Says CDC*. Health.

8. Britannica. (2021). *Size and shape*. Britannica.

9. CDC. (2021, November 4). *COVID-19*. CDC Centers for Disease Control and Prevention.

10. Cuffari, B. (2021). *What is a Pseudovirus?* News-Medical.

11. Czarna, A., Plewka, J., Kresik, L., Matsuda, A., Karim, A., Robinson, C., O'Byrne, S., Cunningham, F., Georgiou, I., Wilk, P., Pachota, M., Popowicz, G., Wyatt, P. G., Dubin, 12. G., & Pyrć, K. (2022). Refolding of lid subdomain of SARS-CoV-2 nsp14 upon nsp10 interaction releases exonuclease activity. *Structure*, *30*(8), 1050-1054.e2. https://doi.org/10.1016/j.str.2022.04.014

13. DAV. (2020). *DNA proofreading and repair*.

14. Denison, M. R., Graham, R. L., Donaldson, E. F., Eckerle, L. D., & Baric, R. S. (2011). Coronaviruses: an RNA proofreading machine regulates replication fidelity and diversity. *RNA Biology*, *8*(2), 270—279. https://doi.org/10.4161/rna.8.2.15013

15. Dey, A. (2022). *Errors In DNA Replication: 13 Facts Most Beginners Don't Know*. Lambda Geeks.

16. Dutta, S. (2021). *D614G Mutation in SARS-CoV-2 Spike Protein*. News-Medical.

17. Flerlage, T., Boyd, D. F., Meliopoulos, V., Thomas, P. G., & Schultz-Cherry, S. (2021). Influenza virus and SARS-CoV-2: pathogenesis and host responses in the respiratory tract. In *Nature Reviews Microbiology* (Vol. 19, Issue 7, pp. 425–441). Nature Research. https://doi.org/10.1038/s41579-021-00542-7

18. Gobeil, S., Henderson, R., & Stalls, V. (2022). *Structural diversity of the SARS-CoV-2 Omicron spike*. NIH.

19. Goldhawk, B. (2021, December 23). *World's first molecular-level analysis of omicron variant spike protein*. MedicalXpress.

20. Hong, Z., Jiang, J., Hashiguchi, K., Hoshi, M., Lan, L., & Yasui, A. (2008). Recruitment of mismatch repair proteins to the site of DNA damage in human cells. *Journal of Cell Science*, *121*(19).

21. InvivoGen. (2022). *SARS-CoV-2 Structural Genes*. InvivoGen.

22. Krug, R. (2022, August 22). *What is a virus?* Britannica.

23. Lakna. (2017, December 6). *Difference Between DNA and RNA Viruses*. PEDIAA.

24. Leung, N. H. L. (2021). Transmissibility and transmission of respiratory viruses. In *Nature Reviews Microbiology* (Vol. 19, Issue 8, pp. 528–545). Nature Research. https://doi.org/10.1038/s41579-021-00535-6

25. Lin, L., Wang, Y., Li, Q., Hu, M., & Shi, Y. (2022). Novel SARS-CoV-2 therapeutic targets: RNA proofreading complex and virus-induced senescence. In *Cell Death and Differentiation* (Vol. 29, Issue 2, pp. 263–265). Springer Nature. https://doi.org/10.1038/s41418-021-00909-6

26. Lu, S. (2020). *NSP10*. NCBI.

27. Ma, Z., Pourfarjam, Y., & Kim, I. K. (2021). Reconstitution and functional characterization of SARS-CoV-2 proofreading complex. *Protein Expression and Purification*, *185*. https://doi.org/10.1016/j.pep.2021.105894

28. Magazine, N., Zhang, T., Wu, Y., McGee, M. C., Veggiani, G., & Huang, W. (2022). Mutations and Evolution of the SARS-CoV-2 Spike Protein. In *Viruses* (Vol. 14, Issue 3). MDPI. https://doi.org/10.3390/v14030640

29. Manzanares-Meza, L. D., & Medina-Contreras, O. (2020). SARS-CoV-2 and influenza: a comparative overview and treatment implications. *NIH*, *77*(5).

30. Marcos, T. (2022). *Replication Race: Scientists unravel why Omicron spreads faster*. Proteintech.

31. National Institutes of Health. (2022, September 9). *COVID-19 Vaccines*. NIH Covid-19 Research.

32. NHS. (2019, October 24). *SARS (severe acute respiratory syndrome)*. NHS.

33. Nobusawa, E., & Sato, K. (2006). Comparison of the Mutation Rates of Human Influenza A and B Viruses. *Journal of Virology*, *80*(7), 3675–3678. https://doi.org/10.1128/jvi.80.7.3675-3678.2006

34. Open Access Government. (2021, February 18). *The spike mutation makes COVID eight times more infectious*. Open Access Government.

35. OxfordLanguages. (2022). *Transmissibility*. Oxford Dictionary.

36. Riccio, A. A., Sullivan, E. D., & Copeland, W. C. (2022). Activation of the SARS-CoV-2 NSP14 3'-5' exoribonuclease by NSP10 and response to antiviral inhibitors. *Journal of Biological Chemistry*, *298*(1). https://doi.org/10.1016/j.jbc.2021.101518

37. Robertson, S. (2022, February 13). *What is a Virus?* News-Medical.

38. Rodríguez, L., Kovac, & Hernández. (2013). *Virus Infectivity*. Science Direct.

39. Rona, G., Zeke, A., Miwatani-Minter, B., de Vries, M., Kaur, R., Schinlever, A., Garcia, S. F., Goldberg, H. v., Wang, H., Hinds, T. R., Bailly, F., Zheng, N., Cotelle, P., Desmaële, D., Landau, N. R., Dittmann, M., & Pagano, M. (2022). The NSP14/NSP10 RNA repair complex as a Pan-coronavirus therapeutic target. *Cell Death and Differentiation*, *29*(2), 285–292. https://doi.org/10.1038/s41418-021-00900-1

40. Sobhanie, M. (2021, December 14). *How do virus mutations happen, and what do they mean?* Wexnermedical.

41. Spencer, J. A., Shutt, D. P., Moser, S. K., Clegg, H., Wearing, H. J., Mukundan, H., & Manore, C. A. (2020). *Epidemiological parameter review and comparative dynamics of influenza, respiratory syncytial virus, rhinovirus, human coronavirus, and adenovirus*. https://doi.org/10.1101/2020.02.04.20020404

42. Tapparel, C., Cordey, S., Junier, T., Farinelli, L., van Belle, S., Soccal, P. M., Aubert, J. D., Zdobnov, E., & Kaiser, L. (2011). Rhinovirus Genome Variation during Chronic Upper and Lower Respiratory Tract Infections. *PLoS ONE*, *6*(6). https://doi.org/10.1371/journal.pone.0021163

43. Tatomir, J. (2022, February 12). *DNA Mutations vs. RNA Mutations: What is a Mutation in DNA*. Study.Com.

44. U.S. National Library of Medicine. (2022). *Base Pair Mismatch*. U.S. National Library of Medicine.

45. Waman, V. P., Kolekar, P. S., Kale, M. M., & Kulkarni-Kale, U. (2014). Population Structure and Evolution of Rhinoviruses. *Plos One*.

46. WHO. (2014, March 4). *How pandemic influenza emerges*. WHO.

47. WHO. (2021, June 14). *Coronavirus disease (COVID-19)*. WHO (World Health Organization).

48. WHO. (2022, July). *Middle East respiratory syndrome*. WHO EMRO.

49. Worldometer. (2022, September 12). *Countries where COVID-19 has spread*. Worldometer.

50. Zhang, J., Xiao, T., Cai, Y., & Chen, B. (2021). Structure of SARS-CoV-2 spike protein. In *Current Opinion in Virology* (Vol. 50, pp. 173–182). Elsevier B.V. https://doi.org/10.1016/j.coviro.2021.08.010