Relationship Between Structure and Pathogenicity

Xinyu Yu

The Hong Kong Polytechnic University, Hong Kong, China

21097273d@connect.polyu.hk

Abstract. The global COVID-19 crisis caused by the SARS-CoV-2 virus has led to the emergence of multiple variants, capturing substantial international attention. This extensive study offers a comprehensive examination of the structural attributes and disease-causing potential of the original SARS-CoV-2 strain and its primary variants, namely Alpha, Beta, Delta, Gamma, and Omicron.

Keywords: COVID-19, variants, pathogenicity.

1. Introduction

In 2019, identified a series of pneumonia cases of unidentified etiology originating from seafood markets in South China. Subsequent investigations established these cases as acute respiratory infections attributable to Novel Coronavirus 2019, presently recognized as COVID-19. The confirmation of the initial case in Wuhan marked the beginning of an accelerated spread of the pandemic. Later, COVID-19 expeditiously disseminated to all provinces in China, instilling apprehension among the public due to its highly infectious characteristics. Coincidentally, nations worldwide encountered an unparalleled escalation in the severity of the pandemic. Since its initial identification, COVID-19 has swiftly diffused across the world, giving rise to a pervasive pandemic with far-reaching ramifications.

The evolution of different variants has been a significant concern. The Alpha variant B.1.1.7 was first identified in the UK. It exhibited higher transmissibility compared to the original strain and quickly spread to various countries. The Beta variant has shown potential resistance to certain monoclonal antibody treatments. The Gamma variant, or P.1, originated in Brazil and shares some characteristics with the Beta variant, including potential antibody resistance. The Delta variant, or B.1.617.2, has raised concerns due to its increased transmissibility and potential to cause disease. Lastly, the Omicron variant of SARS-CoV-2, characterized by numerous spike protein mutations, has raised concerns due to its high transmissibility and potential resistance to existing vaccines.

Although the current pandemic appears to be nearing its conclusion, the complete cessation of the COVID-19 pandemic remains difficult to ascertain due to the continuous emergence and the gradual relaxation of vigilance among the public. At present, numerous uncertainties persist, warranting careful consideration. This article aims to provide a review of the fundamental aspects of COVID-19, encompassing its basic information and epidemiology. Furthermore, an overview of the various variants of the virus, including their emergence, genetic characteristics, and implications, will be presented.

2. Basic information of Covid-19

2.1. Brief Information

The WHO classified the pandemic brought on by the new coronavirus as COVID-19. The Coronaviridae family, a member of the Cornidovirineae suborder, is where SARS-CoV-2 is categorized.

The Coronaviridae family encompasses a different group of enveloped RNA viruses with genomes spanning approximately 30,000 nucleotides. These viruses have a predilection for infecting mammals and avian species [1]. A distinctive characteristic shared by all members of the Coronaviridae family is the presence of Spike proteins embedded in a lipid envelope surrounding the viral particle [1].

Among the known coronaviruses, SARS-CoV-2 stands out as the seventh identified coronavirus with the capacity to invade and replicate within human cells [1]. While three of these coronaviruses typically induce mild upper-respiratory infections akin to common influenza, the infections caused by MERS-CoV, SARS-CoV, and the recently discovered SARS-CoV-2 can manifest as upper-respiratory illnesses, potentially culminating in fatal outcomes [1].

2.2. Structure

SARS-CoV-2 is a ssRNA virus with an envelope and roughly thirty kilobytes. SARS-CoV-2 also encodes non-structural proteins, such as the 2-3C-like main protease, the membrane protein, and the membrane-associated protein with a molecular weight of 180 to 200 kilodaltons [2].

An important factor in regulating viral variety and host tropism is the SARS-CoV-2 Spike protein (S), a trimeric glycoprotein that extends from the viral surface [2]. The name "Coronaviruses" derives from the unusual crown-like shape of these glycoproteins [2]. The S is made up of two subunits that are determined by protease enzymes [2]. The Spike proteins are efficiently hidden from immune monitoring by polysaccharide molecules, allowing them to avoid detection throughout this entrance phase [2].

The coronavirus M glycoprotein exhibits three transmembrane domains, contributing to its structural characteristics [2]. Within the Golgi apparatus, the M proteins undergo glycosylation, a post-translational modification process [2]. Both the fusion of the virus with the host cell and the protein's acquisition of antigenic characteristics depend on the structural alteration of the M protein, and both processes are facilitated by this glycosylation event [2]. The M protein's presence is essential for the intracellular regeneration of viruses [2]. Interactions between the M protein and the genomic RNA occur in ERGIC, where the M protein associates with the genomic RNA, facilitating the assembly and production of virions [2]. Furthermore, the M protein is vital in stimulating the host cell by engaging in a receptor-dependent process, leading to the activation of the Interferon beta pathway [2].

The pentameric membrane protein E has a complex shape that is divided into a few structural domains [3]. Notably, the hydrophobic transmembrane region has garnered significant research attention due to its self-interaction capacity, resulting in a pentameric ion channel that exhibits minimal or negligible ion selectivity [3]. As a result, this specific region possesses the capability to act as a viroporin, which is a viral protein that can modify the permeability of cellular membranes, particularly ion channels, and disrupt normal cellular functions. This viroporin-like function is believed to be significant during the stages of viral infection or release [3]. Through topological analysis, it has been discovered that the N-terminus is situated within intracellular organelles [3]. On the other hand, the C-terminus of the protein is positioned in the cytoplasm [3].

Exhibiting a helical structure with flexibility, the N protein is a phosphoprotein comprised of 349 to 470 amino acids. It functions as an RNA-binding protein, capable of binding to the viral genomic RNA [2]. In coronaviruses, the N protein assumes a critical role across multiple facets, encompassing virus replication, and transcription [2]. Shedding light on the functional characteristics of the N protein, crystallographic studies have been conducted on this protein derived from Avian infectious bronchitis virus and SARS coronaviruses. The research has revealed that the N-terminal portion is responsible for attaching to the viral genomic RNA, whereas the C-terminal section assists in the formation of larger complexes during the assembly of the virus [2]. Furthermore, the C-terminal region of the N protein interacts with another viral protein called the M protein [2]. Importantly, when both the viral M and N proteins are expressed together using mammalian expression vectors, they interact within the cell and create structures that resemble complete virus particles [2]. Furthermore, the N protein serves as an interferon antagonist, impeding the immune system's endeavors to eliminate the virus [2].

3. Evolution of COVID-19

3.1. Wild Type

The MRCA genome serves as the originator of all global SARS-CoV-2 infections, hereafter referred to as proCoV2 [4]. It is highly plausible that this MRCA genome was transmitted to the initial human case, marking the beginning of the COVID-19 pandemic. The presence of this genetic entity predates or aligns with December 24, 2019, a significant date when the genomic sequence of the SARS-CoV-2 infection, became available to researchers. By comparing the proCoV2 genome with the Wuhan-1 genome, researchers have identified three variations occurring at 49 specific positions [4]. Notably, the Wuhan-1 strain has undergone evolutionary changes through a succession of three α mutations within the progenitor, namely α1, α2, and α3 [4].

From January 2020 onwards, a critical mutation, referred to as D614G, emerged in the wild-type variant of SARS-CoV-2 [4]. In regions where SARS-CoV-2 is prevalent, isolates carrying a D614G mutation within the viral spike (S) protein exhibit predominance [5]. This observation suggests that this specific mutation potentially enhances the transmission capability of the virus. A team of scientists from Florida, USA, has provided the subsequent explanation for this phenomenon. The ACE2 serves as the receptor for SARS-CoV-2, as well as for the previously identified SARS-CoV virus [5]. ACE2 is a metalloprotease with the ability to bind to the virus and facilitate its entry into target cells [5]. This binding event primarily takes place via a distinct subdomain located within the S1 region, referred to as the RBD [5]. Notably, the D614G mutation has been implicated in potentially increasing the number of available binding sites by constraining the shedding of the S1 subdomain [5].

Within a mere three-month period, an overwhelming majority of approximately 90% of COVID-19 cases were found to be with the G614 variant [5]. This variant, characterized by the substitution of D with G, has exhibited a substantial increase in its ability to bind to human ACE2 receptors [5]. This heightened affinity for ACE2 receptors potentially contributes to various sequelae, such as the manifestation of hyposmia (reduced sense of smell) [5].

3.2. Alpha Type

The B.1.1.7 emerges as a subsequent lineage following the SARS-CoV-2, initially identified in England and characterized by 17 significant mutations [6]. Researchers have reported that the B.1.1.7 variant exhibits a heightened transmissibility ranging over 40% when compared to the original wild type strain [7]. This increased transmissibility can be attributed to specific mutations [7].

Mutation N501Y significantly impacts the receptor binding domain, leading to an augmented affinity of the SARS-CoV-2 virus for the human ACE2 receptor [6]. The mutation P681H exerts notable influence on the infection and transmission dynamics of the virus. It has been associated with altered viral characteristics that affect the spread and severity of the disease [6]. Another noteworthy modification is the deletion ΔH69/ΔV70 within the spike protein, which has independently arisen in various lineages of SARS-CoV-2 [7]. This deletion has been linked to evading the immune response in individuals with uncompromised immune systems, potentially leading to more severe disease outcomes [7]. Additionally, the presence of this deletion poses challenges for certain commercial testing kits, as they may fail to detect the spike glycoprotein gene, thereby affecting the accuracy of diagnostic measures.

In addition to its heightened transmissibility, the Alpha variant of SARS-CoV-2 also exhibits a greater level of virulence compared to the original strain. According to a study conducted by researchers from Peking University, 60% patients infected with the Alpha variant experienced squeals such as serve fatigue [8]. Despite the increased infectivity and virulence of the B.1.1.7 variant, vaccines developed for the original strain of SARS-CoV-2 still demonstrate efficacy in providing protection [8]. For instance, the Novavax vaccine, which exhibited approximately 96% effectiveness against the Wuhan wild-type strain, retains an approximate 86% effectiveness against the Alpha variant [8].

3.3. Beta Type

The B.1.351 of SARS-CoV-2 brings forth a consequential lineage marked by specific mutations, including E484K, K417N, and N501Y [9]. Notably, the E484K mutation is not exclusive to the Beta variant but is also observed in subsequent generations of the Alpha variant, contributing to immune evasion [9]. These variants have been identified in multiple countries across different continents.

Preliminary modeling studies indicate that the lineage known as 501Y.V2, which encompasses these mutations, may exhibit approximately 50% higher transmissibility compared to previously circulating variants. Experimental investigations utilizing deep mutation scanning and mouse models have shed light on the functional implications of these mutations. The N501Y substitution, for instance, has been shown to enhance the affinity of the virus to human ACE2 receptors, thereby facilitating viral entry into host cells [10]. Additionally, there is evidence suggesting that the E484K substitution may also increase the binding affinity to human ACE2 receptors [10]. When combined with the N501Y mutation, the effect on binding affinity is further augmented [10].

In addition, the mutation E484K should shoulder the responsibility of immune escape [11]. Class 2 antibodies specifically bind to E484 on the SARS-CoV-2 S protein [11]. Notably, empirical observations have demonstrated that the presence of the E484K substitution within this region imparts increased resilience to the neutralizing effects exerted by class 2 antibodies. This evidence suggests that the E484 region assumes a critical role in orchestrating the immune response against the virus and serves as a pivotal target for neutralizing antibodies [11].

The immune evasion capability of the Beta variant is approximately 40 times higher than that of the original strain and 20 times higher than that of the Alpha variant, following the administration of two doses of the Pfizer vaccine [12]. Additionally, another study suggests that the effectiveness of the Pfizer vaccine decreases to 75% when confronted with the Beta variant, despite the administration of two doses, whereas the efficacy remains at 89.5% against the Alpha variant [12].

3.4. Delta Type

Emerging in May 2021, the Delta variant, scientifically designated as B.1.617.2, has rapidly gained substantial prevalence on a global scale, prompting an inundation of reports pertaining to its epidemiological significance. Notably, this variant has exhibited an astonishingly swift rise to dominance, particularly within the geographical context of England, where it surpassed the prevalence of other variants, including the previously prominent B.1.1.7 strain, within a remarkably short timeframe of approximately one month. Extensive global research efforts have been undertaken to investigate the Delta variant, yielding significant insights into its distinct characteristics.

Firstly, the viral load associated with B.1.617.2 infections, as revealed through polymerase chain reaction (PCR) analysis, exhibits a remarkable increase of approximately 1000-fold compared to the Wuhan wild type [13]. This stark contrast suggests a wider dissemination of the variant, characterized by accelerated transmission dynamics and prolonged viral persistence within the host [13].

Secondly, extensive investigations utilizing data from the United Kingdom and India have led researchers to estimate that the Delta variant possesses a heightened infectious capacity ranging from 40% to 80% [14]. Notably, when utilizing the New England region of the United States as a reference, the transmissibility of the Delta variant appears to be even higher [14]. Additional calculations provide further evidence, indicating that the newly identified variant is estimated to be between 60% and 120% more transmissible [14].

Thirdly, in addition to its heightened transmissibility, the Delta variant has demonstrated an elevation in virulence when compared to previously circulating variants. The figures displayed in this visually informative chart presents a clear and concise overview of the relevant data.

Among the features under consideration, the transmissibility of the B.1.617.2 variant holds paramount significance when compared to previous variants. The successful evolution of this variant can be attributed to multiple factors that contribute to its exceptional transmissibility. Notably, two key factors play a crucial role in this regard. Firstly, the B.1.617.2 variant possesses a rapid cell-cell fusion capacity, thereby expediting the infection process within individuals. This heightened fusion

capacity can be attributed to the mutation of P681R, which accelerates both the attachment and penetration procedures of the variant [15]. As a result, the fusion activity rate of the Delta variant is observed to nearly double compared to other variants, including the Gamma, Alpha, and Beta types [15]. This accelerated fusion activity enhances the variant's ability to establish infection more swiftly within host cells, leading to a faster overall infection process. Secondly, the B.1.617.2 variant exhibits a shortened generation time, indicating a quicker interval between successive infections from one patient to another. Prior to the emergence of the Delta variant, dominant variants typically exhibited generation times ranging from approximately 4.0 to 5.4 days [16]. In contrast, the B.1.617.2 variant demonstrates a notably reduced generation time of approximately 3.2 days [16]. This shortened generation time signifies a more rapid transmission of the variant between individuals, facilitating its efficient spread within populations. The available vaccines have demonstrated their enduring effectiveness in preventing infection. Dr. Sanchez further underscored that the vaccines continue to exhibit substantial efficacy, particularly in mitigating the risk of severe disease and mortality. Even in the face of this more contagious variant, the vaccines have proven to be formidable allies in safeguarding individuals from the detrimental consequences of COVID-19.

3.5. Omicron Type

The Omicron variant has surfaced with a discernible molecular trait termed an SGTF on PCR assay, originating from a specific mutation referred to as 69-70del. This genetic modification stands out as a pivotal characteristic shared by the Omicron variant along with its sub-lineages. Notably, the Omicron variant has exhibited genetic diversity, resulting in the identification of distinct subvariants. Notably, BA.1 and BA.2 share numerous common mutations, suggesting a close genetic relationship between these subvariants. Furthermore, recent investigations have identified two additional subvariants, BA.4 and BA.5, in South Africa, further contributing to the genetic diversity observed within the Omicron lineage.

The Omicron variant has garnered attention due to several notable attributes. Firstly, it exhibits a high transmissibility, indicating a heightened ability to spread rapidly within populations [17]. Despite its elevated transmissibility, studies have suggested that the Omicron variant may be associated with a relatively low severity of infection [17].

The spike protein and its RBD in the Omicron variant have undergone a comprehensive set of genetic alterations, resulting in the accumulation of approximately 30 mutations [17]. Within this array of mutations, specific ones such as N501Y and Q498R have been identified as key determinants influencing the transmissibility of the variant [17].

Moreover, mutations H655Y and N679K, located in close proximity to the furin cleavage site, have been observed to promote increased S1/S2 cleavage [18]. This augmented cleavage, alongside the P681H mutation which also enhances S1/S2 cleavage, leads to a higher level of transmissibility exhibited [18]. The increased efficiency of S1/S2 cleavage contributes to the enhanced viral replication and dissemination observed in Omicron-infected individuals [18].

In addition to its heightened transmissibility, the Omicron variant possesses a set of mutations that confer partial resistance to vaccines. These mutations, in conjunction with those found in previous variants, contribute to the distinctive characteristics of the Omicron variant. This alteration in the immunogenicity of the variant may have implications for the efficacy of existing vaccines in recognizing and neutralizing the Omicron variant.

Furthermore, investigations have revealed an overall decrease in the flexibility of the s protein in the Omicron variant. This decrease in flexibility may have consequences for the stability and function of the protein itself. The altered structural dynamics of the s protein could potentially affect the ability of vaccines to effectively recognize and target the viral protein.

However, an intriguing observation has emerged regarding patients infected with the Omicron variant, indicating that they have lower chances of experiencing long-term complications stemming from COVID-19 when compared to patients infected with previous variants. Preliminary evidence suggests that patients infected with the Omicron variant exhibit a reduced frequency of long-term

cardiorespiratory symptoms, potentially indicating a less severe impact compared to previous variants. However, it is important to note that neurological symptoms continue to prevail even in cases of Omicron infection. Among the frequently reported neurological symptoms, insomnia and impaired balance have been highlighted, although no obvious differences have been observed in their prevalence between different variants.

4. Summary

In conclusion, the SARS-CoV-2 virus has undergone several significant mutations, leading to the emergence of different variants with varying characteristics. The G614 variant, prevalent in 90% of cases, exhibited higher affinity for ACE2 receptors and potential symptoms like reduced sense of smell. The B.1.1.7 originated in England. Mutations N501Y, P681H, and ΔH69/ΔV70 contributed to enhanced ACE2 receptor binding, altered infection dynamics, and potential immune evasion. Vaccines retained efficacy against the Alpha variant, albeit with slightly reduced levels. The B.1.351 was characterized by specific mutations such as E484K, K417N, and N501Y. The Beta variant exhibited higher immune escape capability and slightly reduced effectiveness of vaccines compared to the original strain and the Alpha variant. The B.1.617.2 rapidly became globally dominant, showing significantly increased viral load, transmissibility, and virulence. Its rapid cell-cell fusion capacity and shortened generation time contributed to its exceptional transmissibility. The Omicron variant of SARS-CoV-2 exhibited unique genetic traits, including S-gene target failure and distinct subvariants. It showed high transmissibility and potential resistance to existing vaccines due to multiple spike protein mutations. However, it may be associated with lower disease severity and reduced long-term complications compared to previous variants, although neurological symptoms persisted. Overall, the emergence of these variants highlights the ongoing evolution of SARS-CoV-2 and its ability to adapt. Monitoring and understanding these variants are crucial for effective public health interventions, and controlling the spread of COVID-19.

References

[1] Huang, Y., Yang, C., Xu, Xf. et al. Structural and functional properties of SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19. Acta Pharmacol Sin 41, 1141–1149 (2020).

[2] Hu, B., Guo, H., Zhou, P. et al. Characteristics of SARS-CoV-2 and COVID-19. Nat Rev Microbiol 19, 141–154 (2021).

[3] Trobec, Tomaž. The role of the SARS-CoV-2 envelope protein as a pH-dependent cation channel. The Journal of physiology vol. 599, 14 (2021).

[4] Kumar, Sudhir et al. An Evolutionary Portrait of the Progenitor SARS-CoV-2 and Its Dominant Offshoots in COVID-19 Pandemic. Molecular biology and evolution vol. 38, 8 (2021): 3046-3059.

[5] Zhang, Lizhou et al. The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity. bioRxiv: the preprint server for biology 2020.06.12.148726.

[6] Meng, Bo et al. Recurrent emergence of SARS-CoV-2 spike deletion H69/V70 and its role in the Alpha variant B.1.1.7. Cell reportsvol. 35, 13 (2021): 109292.

[7] Davies, Nicholas G et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science (New York, N.Y.) vol. 372, 6538 (2021): eabg3055.

[8] Mahase, Elisabeth. Covid-19: Novavax vaccine efficacy is 86% against UK variant and 60% against South African variant. BMJ (Clinical research ed.) vol. 372 n296. 1 Feb. 2021.

[9] Tegally, Houriiyah et al. Detection of a SARS-CoV-2 variant of concern in South Africa. Nature vol. 592, 7854 (2021): 438-443.

[10] Harvey, William T et al. SARS-CoV-2 variants, spike mutations and immune escape. Nature reviews. Microbiology vol. 19, 7 (2021): 409-424.

[11] Baum, Alina et al. Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. Science (New York, N.Y.) vol. 369, 6506 (2020): 1014-1018.

[12] Garcia-Beltran, Wilfredo F et al. Multiple SARS-CoV-2 variants escape neutralization by vaccine-induced humoral immunity. Cell vol. 184, 9 (2021): 2372-2383.e9.

[13] Li, Baisheng et al. Viral infection and transmission in a large, well-traced outbreak caused by the SARS-CoV-2 Delta variant. Nature communications vol. 13, 1 460. 24 Jan. 2022

[14] Earnest, Rebecca et al. "Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA." medRxiv: the preprint server for health sciences 2021.10.06.21264641. 7 Oct. 2021.

[15] Zhang, Jun et al. Membrane fusion and immune evasion by the spike protein of SARS-CoV-2 Delta variant. bioRxiv: the preprint server for biology 2021.08.17.456689. 17 Aug. 2021.

[16] Hart, William S et al. Generation time of the alpha and delta SARS-CoV-2 variants: an epidemiological analysis. The Lancet. Infectious diseases vol. 22, 5 (2022): 603-610.

[17] Lupala, Cecylia S et al. "Mutations on RBD of SARS-CoV-2 Omicron variant result in stronger binding to human ACE2 receptor." Biochemical and biophysical research communications vol. 590 (2022): 34-41.

[18] Vu, Michelle N et al. Loss-of-function mutation in Omicron variants reduces spike protein expression and attenuates SARS-CoV-2 infection. bioRxiv: the preprint server for biology 2023.04.17.536926. 10 Jul. 2023.