Genome analysis and prevalence of SARS-CoV-2 Indonesian variants and the correlation with the outbreak timeline

Maria Prevyolita Indra Muliawan  
Swiss German University  
https://orcid.org/0000-0003-0463-7853

Timotius Christopher Tantokusumo  
Swiss German University

Kholis Abdurachim Audah  
Swiss German University  
<kholis.audah@sgu.ac.id>

Research Article

Keywords: COVID-19, Coronavirus, Mutation, Genome analysis

Posted Date: January 18th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1272031/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Genome analysis and prevalence of SARS-CoV-2 Indonesian variants and the correlation with the outbreak timeline

Maria Prevyolita Indra Mulawan¹, Timotius Christopher Tantokusumo¹, Kholis Abdurachim Audah¹,²,*

¹Department of Biomedical Engineering, Faculty of Life Sciences and Technology, Swiss German University, Tangerang 15143 Banten, Indonesia
²Academic Research and Community Service, Swiss German University, Tangerang 15143 Banten, Indonesia

Abstract

Since its first appearance in China, in the end of 2019, severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has mutated into several variants. The genome analysis of the spike of SARS-CoV-2 is necessary to identify the mutations in each variant. In this study, we obtained the whole genome sequences of SARS-CoV-2 Wild Type and variants from GenBank (National Center of Biotechnology Information, USA), and GISAID EpiCoV database (Germany). We analyzed the spike glycoprotein gene sequences using the Basic Local Alignment Search Tool (BLAST), and identified the changes in nucleotide and amino acid. Additionally, we also discuss the prevalence and correlation between the submission time of SARS-CoV-2 variants on the GISAID database with the outbreak timeline. The result of our analysis showed that there are 9 amino acid changes in SARS-CoV-2 alpha, beta and delta variants, and 3 amino acid changes in gamma variant. There were 8,861 submissions of Indonesian variants on the GISAID database per November 21, 2021. The correlation between submission time of SARS-CoV-2 variants and outbreak timeline showed that SARS-CoV-2 delta (B.1.617.2) variant potentially caused the sudden increase of COVID-19 confirmed cases from July to September 2021.

Keywords: COVID-19, Coronavirus, Mutation, Genome analysis

*Corresponding author.
E-mail: kholis.audah@sgu.ac.id (Audah K)

December 24, 2021
1. Introduction

In the end of year 2019, China reported its first case of infectious disease in Wuhan, which is now known as the coronavirus disease (2019) or COVID-19. In Wuhan Jinyintan Hospital, three bronchoalveolar lavage samples were collected from a patient who had pneumonia with unknown etiology (or cause). Real time polymerase chain reaction (PCR) was performed on these three samples, and the assays showed positive results for pan-Betacoronavirus. The infectious disease is later known to be caused by severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2). The disease quickly spread in other countries, causing a pandemic. [1] [2]

SARS-CoV-2 is the seventh of the coronaviruses that is known to infect humans, with the other six being MERS-CoV, SARS-CoV-1, HCoV-NL63, CoV-HKU1, HCoV-OC43, and HCoV-229E. SARS-CoV-2 is known to be closely related with the SARS-CoV-1 virus, a virus that originated from Guangdong, China in late 2002. [3, p. 6] As of November 2021, there are over 255 millions confirmed cases of COVID-19, with over 5 millions deaths worldwide. [4] While in Indonesia, there are over 4.25 million confirmed cases of COVID-19 with over 144 thousand deaths. [5]

SARS-CoV-2 is a single-stranded RNA virus [6]. Its entire genome has been characterized using RNA-based metagenomic next generation sequencing approach, resulting in 29,881 bp in length (GenBank no. MN908947), which encodes 9,860 amino acids [7]. The gene fragments express structural and non-structural proteins. Structural proteins are encoded by the S, E, M, and N genes. While non-structural proteins are encoded by the ORF region. [8]

The surface of SARS-CoV-2 is covered by a large number of glycosylated S proteins, which bind to the host cell receptor angiotensin-converting enzyme 2 (ACE2) and thus mediates viral cell entry. The binding of the S protein, widely known as spike protein, to the receptor promotes virus entry, allowing the viral RNA to be released and replicated in the host cell [9]. This explains how the spike protein of SARS-CoV-2 plays an important role in the infection process.

Mutations are changes or small errors in genetic sequence. Mutation is the cause of diversity among organisms [10]. Genetic mutation is also the reason why humans exist until today. Even though viruses are not living organisms, they are also known for their mutation ability as they infect their host’s cells. The outcome of genetic mutation in viruses may cause them to infect a host more easily, or even evade the immune system defense [11]. It is known that RNA viruses are able to mutate at high rates [12]. There are several factors that determine viral mutation rates, which include: polymerase infidelity, template secondary structure, cellular microenvironment, replication mechanisms, proofreading, sequence context, and access to post-replicative repair [13]. SARS-CoV-2 virus typically accumulates two single-letter mutations in the genome each month [14].

According to WHO, there are several classifications of SARS-CoV-2 variants, including the: variants of concern (VOC), variants of interest (VOI), and variants under monitoring.
Variants of interest are the variants that have genetic mutations that can affect virus characteristics (such as transmissibility, disease severity, immune escape, diagnostic or therapeutic escape) and can cause significant community transmission or multiple COVID-19 clusters. Variants of concern are those that meet the criteria of variants of interest, and is associated in either increase in transmissibility or detrimental change in COVID-19 epidemiology; or increase in virulence or change in clinical disease presentation; or decrease in effectiveness of public health or available diagnostics, therapeutics or vaccines. While variants under monitoring are SARS-CoV-2 variants that have genetic mutations that are suspected to affect virus characteristics can pose a future risk, however the evidence of phenotypic or epidemiological impact is still unclear [15]

As of November 26 2021, the variants of concern according to WHO [15] are:
- Alpha variant, originated from United Kingdom, September 2020
- Beta variant, originated from South Africa, May 2020
- Gamma variant, originated from Brazil, November 2020
- Delta variant, originated from India, October 2020

The newly detected variant of concern is the omicron variant, which originated from South Africa and Botswana in November 2021 [15] [16].

In this study, we analyzed the change in nucleotide and amino acid in spike glycoprotein gene sequence of SARS-CoV-2 alpha, beta, delta and gamma mutations in Indonesia. Additionally, we also discuss the prevalence of SARS-CoV-2 variants in Indonesia; and the correlation between submission time of SARS-CoV-2 variants on the GISAID database with the outbreak timeline.

2. Results

Table 1. Submission of SARS-CoV-2 variants in Indonesia on GISAID database with time range

| Time Range            | Types of Variants | Total of submission |
|-----------------------|-------------------|---------------------|
| 01 March 2020 - 01 May 2020 | Alpha             | 0                   |
|                       | Beta              | 0                   |
|                       | Gamma             | 0                   |
|                       | Delta             | 0                   |
|                       | Wild Type         | 6                   |
| Total submitted: 6    |                   |                     |
| 02 May 2020 - 01 July 2020 | Alpha           | 0                   |
|                       | Beta              | 0                   |
|                       | Gamma             | 0                   |
|                       | Delta             | 0                   |
|                       | Wild Type         | 20                  |
| Total submitted: 20   |                   |                     |

December 24, 2021
| Date Range                        | Alpha | Beta | Gamma | Delta | Wild Type |
|----------------------------------|-------|------|-------|-------|-----------|
| 02 July 2020 - 01 September 2020 | 0     | 0    | 0     | 0     | 8         |
| Total submitted: 8               |       |      |       |       |           |
| 02 September 2020 - 01 November 2020 | 0     | 0    | 0     | 0     | 80        |
| Total submitted: 80              |       |      |       |       |           |
| 02 November 2020 - 01 January 2021 | 0     | 0    | 0     | 0     | 26        |
| Total submitted: 26              |       |      |       |       |           |
| 02 January 2021 - 01 March 2021  | 0     | 0    | 0     | 0     | 328       |
| Total submitted: 328             |       |      |       |       |           |
| 02 March 2021 - 01 May 2021      | 13    | 1    | 0     | 2     | 752       |
| Total submitted: 768             |       |      |       |       |           |
| 02 May 2021 - 01 July 2021       | 34    | 9    | 0     | 382   | 886       |
| Total submitted: 1,311           |       |      |       |       |           |
| 02 July 2021 - 01 September 2021 | 18    | 7    | 0     | 0     | 18        |
| Total submitted: 3,299           |       |      |       |       |           |
| Date Range: 02 September 2021 - 21 November 2021 | Gamma | Delta | Wild Type |
|-----------------------------------------------|-------|-------|-----------|
| Total submitted: 3,015                         | 2     | 1,889 | 1,383     |
| Alpha                                         | 12    |       |           |
| Beta                                          | 5     |       |           |
| Gamma                                         | 0     |       |           |
| Delta                                         | 2,707 |       |           |
| Wild Type                                     | 291   |       |           |

*December 24, 2021*
Figure 1. Graph of submitted SARS-CoV-2 Indonesian variants on GISAID database with time range

Table 2. Cumulative confirmed COVID-19 cases and deaths in Indonesia

| Date             | Number of confirmed case | Number of confirmed death |
|------------------|--------------------------|---------------------------|
| 02 March 2020    | 2                        | 0                         |
| 02 May 2020      | 10,843                   | 831                       |
| 02 July 2020     | 59,394                   | 2,987                      |
| 02 September 2020| 180,646                  | 7,616                      |
| 02 November 2020 | 415,402                  | 14,044                     |
| 02 January 2021  | 758,473                  | 22,555                     |
| 02 March 2021    | 1.35 million             | 36,518                     |
| 02 May 2021      | 1.68 million             | 45,796                     |
| 02 July 2021     | 2.23 million             | 59,534                     |
| 02 September 2021| 4.10 million             | 133,676                    |
Table 3. Total submission of SARS-CoV-2 variants in Indonesia on GISAID database until 21 November 2021

| Types of Variants | Total of viruses | Percentages |
|-------------------|------------------|-------------|
| VOC Delta GK/478K.V1 (B.1.617.2+AY.x) first detected in India | 4,980 | 56.20% |
| Wuhan wildtype | 3,780 | 42.66% |
| VOC Alpha 202012/01 GRY (B.1.1.7+Q.x) first detected in the UK | 77 | 0.87% |
| VOC Beta GH/501Y.V2 (B.1.351+B1.351.2+B.1.351.2) first detected in South Africa | 22 | 0.25% |
| VOC Gamma GR/501T.V3 (P.1+P.1.x) first detected in Brazil/Japan | 2 | 0.02% |

Total viruses: 8,861

December 24, 2021
### Table 4. Spike glycoprotein gene sequence result of Wild Type vs Delta (B.1.617.2.55) variant

| Accession ID       | Location from Wild type |
|--------------------|-------------------------|
| NC_045512.2        | 21618 21987 22028 22227 22917 22995 23073 23403 23604 |
| EPI_ISL_1824604     | 23074 |

| Location from Delta Variant |
|-----------------------------|
| 21594 21963 22004 22203 22893 22971 23049 23050 23379 23580 |

*Format: nucleotide(amiino acid)*

### Table 5. Spike glycoprotein gene sequence result of Wild Type vs Alpha (B.1.1.7) variant

| Accession ID       | Location from Wild type |
|--------------------|-------------------------|
| NC_045512.2        | 21765 21770 23063 23271 23403 23604 23709 24506 24914 |
| EPI_ISL_1169047     | 23023 23231 23363 23564 23669 24466 24874 |

| Location from Alpha Variant |
|-----------------------------|

*Format: nucleotide(amiino acid)*
Table 6. Spike glycoprotein gene sequence result of Wild type vs Beta (B.1.351) Variant

| Location from Wild type | Accession ID | 21801 | 22206 | 22280 | 22813 | 23012 | 23063 | 23149 | 23403 | 23664 |
|------------------------|--------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| **EPI_ISL_1824605 (Beta)** | A (D) | A(D) | A(T) | G(K) | G(E) | A(N) | G(K) | A(D) | C(A) |
|                        | C (A) | G(G) | A(P) | T(N) | A(K) | T(Y) | A(K) | G(G) | T(V) |

| Location from Beta Variant | 21777 | 22182 | 22258 | 22789 | 22988 | 23039 | 23125 | 23379 | 23640 |

*Format: nucleotide(abbreviation)*

Table 7. Spike glycoprotein gene sequence result of Wild type vs Gamma (B.1.470) Variant

| Location from Wild type | Accession ID | 21597 | 22286 | 23403 |
|------------------------|--------------|-------|-------|-------|
| **EPI_ISL_2854770 (Gamma)** | C(S) | C(L) | A(D) |
|                        | T(F) | T(F) | G(G) |

| Location from Gamma Variant | 21550 | 22239 | 23356 |

*Format: nucleotide(abbreviation)*

December 24, 2021
Table 8. Amino acid mutations in SARS-CoV-2 alpha, beta, delta and gamma variant

| Amino acid mutations | SARS-CoV-2 variants |
|----------------------|--------------------|
|                      | Alpha | Beta | Delta | Gamma |
| Δ68-69               | ✔     |      | ✔     |       |
| A222V                |        | ✔    | ✔     |       |
| A570D                | ✔     |      | ✔     |       |
| A701V                |        | ✔    | ✔     |       |
| D1118H               | ✔     |      | ✔     |       |
| D215G                | ✔     |      | ✔     |       |
| D614G                | ✔     | ✔    | ✔     | ✔     |
| D80A                 |        | ✔    | ✔     | ✔     |
| E156K                |        |      | ✔     |       |
| E484K                | ✔     |      | ✔     |       |
| G142D                |        | ✔    | ✔     |       |
| G504X                | ✔     |      | ✔     | ✔     |
| K417N                |        | ✔    | ✔     |       |
| K529K                | ✔     |      | ✔     |       |
| L242F                |        | ✔    | ✔     |       |
| L452R                |        | ✔    | ✔     |       |
| N501Y                | ✔     | ✔    | ✔     |       |
| P681H                | ✔     |      | ✔     |       |
| P681R                |        | ✔    | ✔     |       |
| S12F                 |        | ✔    | ✔     |       |
| S982A                | ✔     |      | ✔     |       |
| T19R                 |        | ✔    | ✔     |       |
| T240P                |        | ✔    | ✔     |       |
| T478K                |        | ✔    | ✔     |       |
| T716I                | ✔     |      | ✔     |       |
| V70I                 | ✔     |      | ✔     |       |
| Total mutated amino acids | 9     | 9    | 9     | 3     |
3. Discussion

3.1 Analysis of SARS-CoV-2 Indonesian variants spike protein sequence

In the alpha (B.1.1.7) variant, based on Table 5, the mutations in amino acids of the spike glycoprotein were: Δ68-69, V70I, N501Y, A570D, D614G, P681H, T716I, S982A, and D1118H. The SARS-CoV-2 B.1.1.7 variant has the Δ68-69 mutation, which is an amino-terminal domain deletion. The mutation N501Y in SARS-CoV-2 B.1.1.7 variant spike glycoprotein strengthens its binding to the receptor ACE2 [18]. Additionally, the mutation N501Y reduces neutralization by a small proportion of RBD (receptor binding domain) antibodies [19]. The mutation D614G in B.1.1.7 increases infectivity of the virus, which is correlated with S1 shedding and more incorporation of S protein into the pseudovirion [18][3].

According to Table 6, the mutations in amino acids of the spike glycoprotein of the beta (B.1.351) variant were: D80A, D215G, T240P, K417N, E484K, N501Y, K529K, D614G, and A701V. The existence of mutations K417N and E484K suggest that the SARS-CoV-2 B.1.351 variant may overcome polyclonal antibody response by reducing neutralization from class 1 and class 2 RBD-specific antibodies; with K417N playing an important role in viral escape [20]. Mutation N501Y of the spike glycoprotein in the SARS-CoV-2 B.1.351 variant allows the virus to strengthen its bond with the receptor ACE2 [18]. The infectivity of the SARS-CoV-2 beta variant is also increased by the mutation D614G [3, p. 6].

The mutations in amino acids of gamma (B.1.470) variant’s spike glycoprotein as seen in Table 7 were: S12F, L242F, and D614G. Just like in SARS-CoV-2 alpha (B.1.1.7) and beta (B.1.351) variant spike glycoprotein sequence analysis, D614G mutation in SARS-CoV-2 B.1.470 variant also increases infectivity of the virus [3, p. 6]. The S12F mutation has a surface exposure, and might have an impact on cell interaction [21].

Meanwhile, from Table 3, the mutations in amino acids of delta (B.1.617.2.55 or AY.55) variant’s spike glycoprotein were: T19R, G142D, E156K, A222V, L452R, T478K, G504X, D614G, and P681R. The delta variant mutations in T19R and G142D may cause disruption in binding some anti-NTD neutralizing antibodies obtained from the wild-type’s spike [22]. The A222V mutation was present in the Delta Plus and Delta variants, at the prevalence of 58% and 9% respectively [23]. The L452R mutation caused structural change resulting in reduction of intramolecular and intermolecular contacts towards ACE2 binding [24]. In addition, another study reported that 14 out of 35 RBD-specific mAbs, including three clinical stage mAbs, had their neutralizing activity decreased or eliminated due to the L452R mutation [25]. Moreover, another report revealed that the L452R mutation allows viruses to evade HLA-24-restricted cellular immunity while also increasing viral infectivity, potentially encouraging viral reproduction [26]. The T478K mutation in the RBD is specific to the Delta B.1.617.2.55 variation and occurs in the epitope region, and is powerful in neutralizing ‘Class 1’ monoclonal antibodies [27]. The D614G mutation alters the spike protein to be more stable, allowing it to penetrate the host cell more efficiently [17]. The P681R mutation in the furin cleavage site
enhancing the basicity of the poly-basic stretch might aid in greater rate of membrane fusion, internalization and hence improved transmissibility [24].

3.2 Analysis of cumulative confirmed cases and deaths and submitted variants on GISAID database

Based on Table 2, the increase in the number of cumulative confirmed COVID-19 cases went steady from September 2020 to January 2021, in which the confirmed cases increased by around 300,000 every 3 months. However from January to March 2021, the number of confirmed cases increased by around 590,000 from 758,473 cases on January 2, 2021 to 1.35 million cases on March 2, 2021. The number of confirmed cases increased sharply from July to September 2021, which went from 2.23 million on July 2 2021 to 4.10 million on September 2 2021; with the number of deaths caused by COVID-19 also increased sharply from 59,534 to 133,676 deaths. By November 21 2021, the cumulative cases of COVID-19 reached the number of 4.25 million with 143,739 deaths caused by COVID-19. The increase in the number of confirmed COVID-19 cases and deaths in Indonesia can be seen in Figure 2.

Based on Table 1, from March-May 2021 to May-July 2021 there was a sudden increase in the number of submitted Indonesian delta variants which went from 2 submissions to 382 submissions on the GISAID EpiCoV database. The number increased to 1,889 submissions in July-September 2021, and kept increasing to 2,707 submissions in September-November 2021. The increasing number of Indonesian delta variant submissions as well as other submitted Indonesian variants on the GISAID EpiCoV database can be seen in Figure 1.

To analyze the correlation between the cumulative COVID-19 cases and deaths, with the cumulative submissions of SARS-CoV-2 Indonesian variants on the GISAID database, direct comparison of the data is performed. The direct comparison can be seen in Table 9 below.

Table 9. Cumulative submission of SARS-CoV-2 Indonesian variants, COVID-19 cases and deaths

| Date          | SARS-CoV-2 Indonesian variant | Case       | Death     |
|---------------|------------------------------|------------|-----------|
|               | Alpha | Beta | Gamma | Delta | WT |           |            |
| 01 May 2020   | 0      | 0    | 0     | 0     | 6  | 10,843    | 831        |
| 01 July 2020  | 0      | 0    | 0     | 0     | 26 | 59,394    | 2,987      |
| 01 September 2020 | 0   | 0    | 0     | 0     | 34 | 180,646    | 7,616      |
| 01 November 2020 | 0   | 0    | 0     | 0     | 114| 415,402    | 14,044     |
| 01 January 2021 | 0    | 0    | 0     | 0     | 140| 758,473    | 22,555     |
| 01 March 2021 | 0      | 0    | 0     | 0     | 468| 1.35 million | 36,518     |

December 24, 2021
| Date               | Submission | Total | Cases   | Deaths | Cases DE  | Deaths DE |
|--------------------|------------|-------|---------|--------|----------|-----------|
| 01 May 2021        | 13         | 1     | 0       | 2      | 1,220    | 1.68 million | 45,796 |
| 01 July 2021       | 47         | 10    | 0       | 384    | 2,106    | 2.23 million | 59,534 |
| 01 September 2021  | 65         | 17    | 2       | 2,273  | 3,489    | 4.10 million | 133,676 |
| 21 November 2021   | 77         | 22    | 2       | 4,980  | 3,780    | 4.25 million | 143,739 |

In Table 9 above, the notable changes have been highlighted with pale yellow color. The notable change occurs from July 1 2021 to September 1 2021. On July 1 2021, the number of submissions of the SARS-CoV-2 delta variant in Indonesia on the GISAID database was 384, which then increased to 2,273 on September 1 2021. Meanwhile within the same time range, drastic increase of cumulative COVID-19 cases and deaths also occurred. On July 1 2021, the number of cumulative COVID-19 cases and deaths in Indonesia was respectively 2.23 million and 59,534, which then increased to 4.10 million cumulative cases and 133,676 deaths on September 1 2021. This analysis could show the correlation of the increasing number of delta variant of SARS-CoV-2 with the sudden increase of cumulative COVID-19 cases and deaths in Indonesia. This can be related with the clinical manifestations of the SARS-CoV-2 delta (B.1.617.2) variant, which are: (1) increased transmissibility, (2) potential vaccine effectiveness reduction against symptomatic COVID-19, and (3) potential reduction in monoclonal antibody therapies neutralization. Moreover, B.1.617.2 has a potential of increasing severity, judging from hospitalization rate [28]. However, the limitations to this analysis include: (1) lack of proof due to lack of data (especially about the precise number of COVID-19 cases caused by each variant of SARS-CoV-2), (2) lack of significant analysis that can lead to conclusive result, and (3) few mismatch of date between submissions and cases/deaths due to different sources.

4. Conclusion

As a conclusion, this research studied the genome analysis and prevalence of SARS-CoV-2 Indonesian variants, and the correlation between the submissions of SARS-CoV-2 Indonesian variants: alpha, beta, delta and gamma on the GISAID database with the outbreak timeline. From the genome analysis, the changes of amino acids in the SARS-CoV-2 variants’ spike glycoprotein were obtained. The analysis of the submission time and outbreak timeline, supported by analysis of clinical manifestation caused by the mutations, showed that SARS-CoV-2 delta (B.1.617.2) variant might have caused the sudden increase of COVID-19 confirmed cases and deaths from July to September 2021.

5. Materials and Methods

5.1. Whole genome sequences of SARS-CoV-2

In this study, the SARS-CoV-2 whole genome sequences (WGS) were obtained from GISAID (the Global Initiative on Sharing All Influenza Data) EpiCoV database and GenBank (National
Center for Biotechnology Information). Registration to GISAID was required to access the database.

5.2. SARS-CoV-2 submission timeline analysis
The submission time of SARS-CoV-2 variants on the GISAID database was obtained by accessing the GISAID EpiCoV database. The time range used in this analysis was per 3 months (quarterly) starting from 01 March 2020 until 21 November 2021. The location was set to “Indonesia”, and the variant was set as one of the options below:
- VOC Delta GK/478K.V1 (B.1.617.2+AY.x) first detected in India
- VOC Alpha 202012/01 GRY (B.1.1.7+Q.x) first detected in the UK
- VOC Beta GH/501Y.V2 (B.1.351+B.1.351.2+B.1.351.3) first detected in South Africa
- VOC Gamma GR/501Y.V3 (P.1+P.1.x) first detected in Brazil/Japan
- Empty (to search all submissions within the time range)

5.3. Analysis of cumulative confirmed COVID-19 cases and deaths in Indonesia
The number of cumulative confirmed COVID-19 cases and deaths in Indonesia was obtained by accessing the ‘Our World in Data’ website via the link https://ourworldindata.org/coronavirus-data?country=~IDN. The country was set to ‘Indonesia’ and the data was recorded every 3 months.

5.4. Genetic Composition Analysis
The genetic composition of SARS-CoV-2 spike glycoprotein sequences were analyzed based on their nucleotide variants and amino acid mutations. In this study, the complete genome of SARS-CoV-2 isolate Wuhan-Hu 1 was used as a reference gene [17]. Afterwards, the results were compared to the previous studies and research journals.

5.5. Similarity Analysis
In order to analyze the similarity between the reference and the query sequences, a nucleotide Basic Local Alignment Search Tool (BLAST) provided by NCBI was used. A pairwise with dots for identities alignment view was utilized to locate the mutations in the nucleotides. In addition, the Coding Region Sequences (CDS) feature was used to analyze the isolates’ amino acids sequence.
Authors’ Contribution

K.A.A. had overseen the project conceptualization process, acquired research funding for the study, designed the methodology for the study, and helped writing as well as reviewing the manuscript draft alongside M.P.I.M. and T.C.T. T.C.T. provided the whole genome sequence of SARS-CoV-2 Wild Type, while M.P.I.M. provided the whole genome sequences of the variants. M.P.I.M. provided the data of submission of SARS-CoV-2 variants in Indonesia on GISAID database and cumulative confirmed COVID-19 cases and deaths in Indonesia with time range, and the analysis of the correlation between the provided data. T.C.T. performed the genome sequence similarity analysis between Wild Type and each variant and provided the data of amino acid mutations in each variant. T.C.T. and M.P.I.M. provided analysis of each variant’s mutation based on the data of amino acid mutations provided by T.C.T. All authors have read and approved the final version of the manuscript.

Competing Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgments

We thank the Swiss German University for funding this research through the Faculty Research Fund granted to Kholis Abdurachim Audah.

References

[1] Mayo Clinic, “Coronavirus disease 2019 (COVID-19) - Symptoms and causes,” Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/coronavirus/symptoms-causes/syc-20479963 (accessed Dec. 05, 2021).

[2] WHO, “Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19).” WHO. [Online]. Available: https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf#:~:text=People%20with%20COVID%2D19,mild%20disease%20and%20recover.

[3] L. Zhang et al., “The D614G mutation in the SARS-CoV-2 spike protein reduces S1 shedding and increases infectivity,” bioRxiv, p. 2020.06.12.148726, Jun. 2020, doi: 10.1101/2020.06.12.148726.

[4] WHO, “WHO Coronavirus (COVID-19) Dashboard.” https://covid19.who.int (accessed Dec. 05, 2021).

[5] H. Ritchie et al., “Coronavirus Pandemic (COVID-19),” Our World Data, Mar. 2020. Accessed: Dec. 05, 2021. [Online]. Available: https://ourworldindata.org/coronavirus-data

[6] R. Lu et al., “Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding,” The Lancet, vol. 395, no. 10224, pp. 565–574, Feb. 2020, doi: 10.1016/S0140-6736(20)30251-8.

[7] Y. Huang, C. Yang, X. Xu, W. Xu, and S. Liu, “Structural and functional properties of

December 24, 2021
SARS-CoV-2 spike protein: potential antivirus drug development for COVID-19,” *Acta Pharmacol. Sin.*, vol. 41, no. 9, pp. 1141–1149, Sep. 2020, doi: 10.1038/s41401-020-0485-4.

[8] J. F.-W. Chan *et al.*, “Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan,” *Emerg. Microbes Infect.*, vol. 9, no. 1, pp. 221–236, Jan. 2020, doi: 10.1080/22221751.2020.1719902.

[9] A. R. Fehr and S. Perlman, “Coronaviruses: An Overview of Their Replication and Pathogenesis,” in *Coronaviruses*, vol. 1282, H. J. Maier, E. Bickerton, and P. Britton, Eds. New York, NY: Springer New York, 2015, pp. 1–23. doi: 10.1007/978-1-4939-2438-7_1.

[10] L. Loewe, “Genetic Mutation,” *Nat. Educ.*, vol. 1, no. 1, p. 113, 2008.

[11] M. Wei-Haas, “The coronavirus is mutating—but what determines how quickly?,” *Science: Coronavirus Coverage*, Feb. 06, 2021. https://www.nationalgeographic.com/science/article/the-coronavirus-is-mutating-but-what-determines-how-quickly

[12] S. Duffy, “Why are RNA virus mutation rates so damn high?,” *PLOS Biol.*, vol. 16, no. 8, p. e3000003, Aot 2018, doi: 10.1371/journal.pbio.3000003.

[13] R. Sanjuán and P. Domingo-Calap, “Mechanisms of viral mutation,” *Cell. Mol. Life Sci. CMLS*, vol. 73, no. 23, pp. 4433–4448, Dec. 2016, doi: 10.1007/s00018-016-2299-6.

[14] E. Callaway, “The coronavirus is mutating — does it matter?,” *Nature*, vol. 585, no. 7824, pp. 174–177, Sep. 2020, doi: 10.1038/d41586-020-02544-6.

[15] WHO, “Tracking SARS-CoV-2 variants.” https://www.who.int/emergencies/emergency-health-kits/trauma-emergency-surgery-kit-who-tesk-2019/tracking-SARS-CoV-2-variants (accessed Dec. 05, 2021).

[16] ECDC, “SARS-CoV-2 variants of concern as of 3 December 2021,” *European Centre for Disease Prevention and Control*, Dec. 03, 2021. https://www.ecdc.europa.eu/en/covid-19/variants-concern (accessed Dec. 05, 2021).

[17] F. Wu *et al.*, “A new coronavirus associated with human respiratory disease in China,” *Nature*, vol. 579, no. 7798, pp. 265–269, Mar. 2020, doi: 10.1038/s41586-020-2008-3.

[18] W. T. Harvey *et al.*, “SARS-CoV-2 variants, spike mutations and immune escape,” *Nat. Rev. Microbiol.*, vol. 19, no. 7, pp. 409–424, Jul. 2021, doi: 10.1038/s41579-021-00573-0.

[19] D. A. Collier *et al.*, “Sensitivity of SARS-CoV-2 B.1.1.7 to mRNA vaccine-elicited antibodies,” *Nature*, vol. 593, no. 7857, pp. 136–141, May 2021, doi: 10.1038/s41586-021-03412-7.

[20] C. K. Wibmer *et al.*, “SARS-CoV-2 501Y.V2 escapes neutralization by South African COVID-19 donor plasma,” *Nat. Med.*, vol. 27, no. 4, pp. 622–625, Apr. 2021, doi: 10.1038/s41591-021-01285-x.

[21] L. F. S. M. Timmers *et al.*, “SARS-CoV-2 mutations in Brazil: from genomics to putative clinical conditions,” *Sci. Rep.*, vol. 11, no. 1, p. 11998, Jun. 2021, doi: 10.1038/s41598-021-91585-6.

[22] Y. Liu *et al.*, “The SARS-CoV-2 Delta variant is poised to acquire complete resistance to wild-type spike vaccines,” Aug. 2021. doi: 10.1101/2021.08.22.457114.

[23] S. R. Kannan *et al.*, “Evolutionary analysis of the Delta and Delta Plus variants of the SARS-CoV-2 viruses,” *J. Autoimmun.*, vol. 124, p. 102715, Nov. 2021, doi: 10.1016/j.jaut.2021.102715.
[24] S. Cherian et al., “Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India,” May 2021. doi: 10.1101/2021.04.22.440932.

[25] M. McCallum et al., “SARS-CoV-2 immune evasion by variant B.1.427/B.1.429,” BioRxiv Prepr. Serv. Biol., p. 2021.03.31.437925, Apr. 2021, doi: 10.1101/2021.03.31.437925.

[26] C. Motozono et al., “An emerging SARS-CoV-2 mutant evading cellular immunity and increasing viral infectivity,” Apr. 2021. doi: 10.1101/2021.04.02.438288.

[27] D. Planas et al., “Reduced sensitivity of SARS-CoV-2 variant Delta to antibody neutralization,” Nature, vol. 596, no. 7871, pp. 276–280, Aug. 2021, doi: 10.1038/s41586-021-03777-9.

[28] UpToDate, “SARS-CoV-2 Variants of Concern,” 2021. https://www.uptodate.com/contents/image?imageKey=ID%2F131216 (accessed Dec. 05, 2021).
Figure 1

Graph of submitted SARS-CoV-2 Indonesian variants on GISAID database with time range
Cumulative Confirmed COVID-19 Cases and Deaths in Indonesia

Figure 2

Graph of cumulative confirmed COVID-19 cases and deaths in Indonesia