Full-genome sequencing and mutation analysis of SARS-CoV-2 isolated from Makassar, South Sulawesi, Indonesia

Muhammad Nasrum Massi[1,2,3], Rufika Shari Abidin[3], Abd-ElAziem Farouk[4], Handayani Halik[3,5], Gita Vita Soraya[6], Najdah Hidayah[3], Rizalinda Sjahril[1,2], Irda Handayani[7], Mohamad Saifudin Hakim[8], Faris Muhammad Gazali[9], Vivi Setiawaty[10] and Tri Wibawa[8]

[1] Department of Clinical Microbiology, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[2] Microbiology Laboratory, Hasanuddin University Hospital, Makassar, South Sulawesi, Indonesia
[3] Hasanuddin University Medical Research Center Laboratory, Faculty of Medicine, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[4] Department of Biotechnology, Faculty of Science, Taif University, Taif City, Al-Hawiyya, Saudi Arabia
[5] Mega Rezky University, Makassar, South Sulawesi, Indonesia
[6] Department of Biochemistry, Faculty of Medicine, Hasanuddin University, Makassar, South Sulawesi, Indonesia
[7] Clinical Pathology Laboratory, Wahidin Sudirohusodo Hospital, Makassar, South Sulawesi, Indonesia
[8] Department of Microbiology, Faculty of Medicine, Public Health and Nursing, Gadjah Mada University, Yogyakarta, Indonesia
[9] Master Program in Biotechnology, Postgraduate School, Gadjah Mada University, Yogyakarta, Indonesia
[10] National Institute for Health Research and Development, Ministry of Health, Jakarta, Indonesia

Corresponding author
Muhammad Nasrum Massi, nasrumm2000@yahoo.com, nasrumm@unhas.ac.id

ABSTRACT

Introduction. A global surge in SARS-CoV-2 cases is occurring due to the emergence of new disease variants, and requires continuous adjustment of public health measures. This study aims to continuously monitor and mitigate the impact of SARS-CoV-2 through genomic surveillance, to determine the emergence of variants and their impact on public health.

Methods. Data were collected from 50 full-genome sequences of SARS-CoV-2 isolates from Makassar, South Sulawesi, Indonesia. Mutation and phylogenetic analysis was performed of SARS-CoV-2 from Makassar, South Sulawesi, Indonesia.

Results. Phylogenetic analysis showed that two samples (4%) were of the B.1.319 lineage, while the others (96%) were of the B.1.466.2 lineage. Mutation analysis of the spike (S) protein region showed that the most common mutation was D614G (found in 100% of the sequenced isolates), followed by N439K (98%) and P681R (76%). Several mutations were also identified in other genomes with a high frequency, including P323L (nsp12), Q57H (ns3-orf3a), and T205I (nucleoprotein).

Conclusion. Our findings highlight the importance of continuous genomic surveillance to identify new viral mutations and variants with possible impacts on public health.

INTRODUCTION

The first cases of coronavirus disease 2019 (COVID-19) were detected in Wuhan, China, on 31st December 2019 (*Zhou et al., 2020*; *Oley et al., 2021*). It was subsequently confirmed that these cases were due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (*Coronaviridae Study Group of the International Committee on Taxonomy of Viruses, 2020*), a respiratory virus that is highly transmissible between humans, and spread rapidly globally. The outbreak was declared a pandemic by the World Health Organization on 30th January 2020 (*WHO, 2020b*).

Phylogenic studies indicate that the virus first emerged in late November 2019, and has an evolutionary rate of approximately 1.1×10^{-5} subs/site/year (*Duchene et al., 2020*). Compared to other RNA viruses, coronaviruses (including SARS-CoV-2) have proofreading activity encoded by non-structural protein (nsp) 14, which reduces RNA replication errors (*Sevajol et al., 2014*; *Smith et al., 2013*; *Wu et al., 2020*). However, the highly transmissible nature of SARS-CoV-2 increases the likelihood of viral mutations and the rise of new variants in each infected individual.

Furthermore, variants that confer advantageous traits for viral survival (such as higher receptor binding capacity and immune evasion) are more likely to dominate the viral population, as observed in the case of the D614G variant (*Korber et al., 2020*; *Lauring & Hodcroft, 2021*). Although this variant was already present in China in sequences collected in late January to early February 2020, the mutation was first reported in the UK in early March 2020 (*Korber et al., 2020*). By June/July 2020, the D614G mutation was present in 63% of the total submitted sequences worldwide (*Koyama, Platt & Parida, 2020*). Viruses with the D614G variant have a higher viral load, but are not associated with disease severity (*Korber et al., 2020*). Over time, D614G diverged, leading to new variants in different parts of the world, including the B.1.1.7, B.1.351, P.1, B.1.617.2, and B.1.1.529 variants from the UK, South Africa, Brazil, India, and South Africa respectively. These variants are currently classified as variants of concern (VOC) by the World Health Organization. VOC are those associated with an observed increase in viral load, moderate immune evasion, and disease severity (*WHO, 2021b*; *Robert Koch Institut, 2021*; *ECDC, 2022*). The emergence of new impactful variants highlights the importance of the containment viral spread, genomic surveillance, and the development of broadly cross-protective vaccines (*Walensky, Walke & Fauci, 2021*; *Abdool Karim & de Oliveira, 2021*).

Indonesia announced the first two confirmed cases of COVID-19 on 2nd March 2020 (*Alwan, 2007*). By 1st December 2021, there were 4,256,687 (278 new within December 2021 alone) confirmed cases, 143,480 (10 new within December 2021 alone) deaths, and 4,104,964 recoveries, from 510 districts across all 34 Indonesian provinces (*WHO, 2020a*). However, the incidence of disease varies between areas, due to the archipelagic nature of the country and the spread of the virus. Genomic surveillance plays an important role in the containment of the virus, and in mitigating the emergence of new variants. From March 2020 to April 2021, 1,210 SARS-CoV-2 sequences were identified, accounting for 0.073% of the 1,668,360 infections recorded in Indonesia, with a monthly average of 140 days between specimen collection and sequence upload to Global Initiative on Sharing All

Influenza Data (GISAID) (*Pradipta et al., 2022*). By 10th December 2021, a total of 10,094 SARS-CoV-2 genome sequences had been submitted to GISAID from Indonesia (*GISAID, 2021*).

Makassar is the capital city of the South Sulawesi province, which lies at the center of Indonesia. It has the largest economy in eastern Indonesia, due in particular to the Makassar Megapolitan region at its center, which acts as a busy hub for transportation throughout the country. As of 19th April 2021, a total of 50 SARS-CoV-2 whole-genome sequencing data collected from both the Wahidin Sudirohusodo and Universitas Hasanuddin hospitals had been submitted to GISAID (*Ministry of Health of Republic Indonesia, 2021*).

Preliminary data evaluation showed that the N439K variant with a D416G background dominated the SARS-CoV-2 viral population in Indonesia (*Cahyani et al., 2021*). This variant contains an N439K mutation in the receptor-binding motif (RBM), leading to enhanced binding affinity for the human angiotensin I-converting enzyme 2 (hACE2) receptor and resistance to several neutralizing monoclonal and polyclonal antibodies. These properties allow it to infect cells more easily and evade the immune system (*Thomson et al., 2021*). This study aims to: (1) report the full genome sequences of SARS-CoV-2 isolated from COVID-19 patients in Makassar, South Sulawesi, Indonesia, throughout the period of January to April 2021; and (2) perform mutation and phylogenetic analysis of SARS-CoV-2 from Makassar, South Sulawesi, Indonesia.

MATERIALS & METHODS

Clinical sample collections

This study was conducted in Makassar, the capital city of the South Sulawesi province of Indonesia. SARS-CoV-2 RNA samples from COVID-19-positive patients were collected at the Wahidin Sudirohusodo and Hasanuddin University Hospitals, which are type A and B COVID-19 referral hospitals, respectively. The participants were recruited using total sampling. SARS-CoV-2 patients were grouped based on the severity of their illness according to the National Institutes of Health, Treatment Guidelines Panel Coronavirus Disease 2019 (COVID-19) as described previously in *National Institutes of Health (2021)* and *El-Raey et al. (2021)*.

All nasopharyngeal swab samples were collected in viral transport media (Kang Jian Viral Transport Media; Jiangsu Kang Jian Medical Apparatus Co. Ltd., Jiangyan Taizhou, China) and transported to the Hasanuddin University Medical Research Centre (HUM-RC) Laboratory, Faculty of Medicine, Universitas Hasanuddin. SARS-CoV-2 was detected using real-time quantitative PCR (qRT-PCR) with a mBiocov-19 RT-PCR Kit (Biofarma, Bandung, Indonesia) and CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA). For whole-genome sequencing, the samples were selected based on a CT value below 30. Clinical and other laboratory data were obtained from the patient forms, which were sent to the laboratory.

Full-genome sequencing

The total viral RNA was purified from 50 original nasopharyngeal swab samples using a Viral Nucleic Acid Extraction Kit II (Geneaid Biotech Ltd., New Taipei City, Taiwan).

The purified RNA samples were used for whole-genome sequencing (WGS) using the Oxford Nanopore's GridlON sequencer methodology, operated with MinION version 20.06.9 and MinION core version 4.0.11. High accuracy base calling was conducted using Guppy (*Oxford Nanopore Technologies, 2021*) version 4.0.11, with all generated reads assembled using the EPI2ME Labs platform employing ARTIC workflow. This provided the whole genome sequences of the collected samples, in addition to data on amino acid mutations in each isolate.

Genome annotation and phylogenetic analysis

The full-length genomes of the SARS-CoV-2 samples isolated in this study were annotated using the reference genome of hCoV-19/Wuhan/Hu-1/2019 (NC_045512.2). For phylogenetic analysis, a dataset of available SARS-CoV-2 genomes from different countries, including those from other regions in Indonesia, was retrieved from GISAID, representing both VOC and non-VOC. The final dataset consisted of 50 original samples derived from this study, the NCBI reference sequence (NC_045512.2), and 30 sequences retrieved from GISAID, consisting of B.1, B.1.319, B.1.466.2, Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), Delta (B.1.617.2), and Omicron (B.1.1.529). The studies from which this data was sourced are provided in Table S1. The accession IDs of our isolates are provided in Table S2.

A total of 81 sequences were aligned using MAFFT (v7.471) with the L-INS-i algorithm (*Katoh & Standley, 2013*). The aligned dataset was used to construct the maximum likelihood (ML) tree, using RaXML v.8.2.10 with 1,000 bootstraps and the GTR+G model as the best-fitting substitution model (*Stamatakis, 2014*). The ML tree was viewed and annotated using MEGA (v11.1.1) (*Tamura, Stecher & Kumar, 2021*). The final tree was rooted in the ancestor virus hCoV-19/Wuhan/Hu-1/2019 to visualize the evolutionary relationship between our original samples and the other SARS-CoV-2 isolates deposited in GISAID.

Mutation visualization

All mutations were visualized in a 3D protein model using Swiss-PdbViewer (ver. 4.1.0). The protein model was reconstructed using the Wuhan reference sequence (NC_045512.2) against the S Cryo-EM structure (PDB ID: 6XR8; Resolution: 2.90 Å), to observe mutations located at the furin cleavage site.

Protein modelling was performed using the Swiss-Model web server (https://swissmodel.expasy.org/). The quality of the protein model was ensured by observing the QMEAN (>-4.0).

Ethical approval

This study obtained ethical approval from the Health Research Ethical Committee of the Faculty of Medicine, Universitas Hasanuddin, Makassar, South Sulawesi, Indonesia (No. 845/UN4.6.5.31/PP36/2020). Written informed consent was obtained from all patients prior to participating in this study.

RESULTS

Demographics and clinical severity of COVID-19 patients

Over half of the patients in this cohort were male (56%), with the most severely affected patients between the ages of 30–39 years (30%), followed by those aged 50–64 years (20%) and 40–49 years (19%) (see Table 1). Eleven (22%) cases were asymptomatic, while 17 (34%) patients were categorized as mild, 17 (34%) as moderate, and the remainder (10%) as severe. Patients with comorbidities were found in a moderately high proportion (60% of 10 samples) in this study, as shown in Table 1.

Full genome characterization of SARS-CoV-2 isolated from Makassar, South Sulawesi

The phylogenetic tree of our original isolates ($n = 50$) and GISAID sequences ($n = 31$) is shown in Fig. 1. All isolates originating from Makassar were clustered in the main lineage of B.1, where 48 samples were identified from the B.1.466.2 lineage, and two (hCoV-19/Indonesia/SN-MKS_17/2021 and hCoV-19/Indonesia/SN-MKS_36/2021) from the B.1.319 lineage. However, 38 sequences of our isolates formed new clades which were different from other samples from the GISAID reference lineage of B.1.466.2 (EPI_ISL_877456; EPI_ISL_768689; EPI_ISL_728184). All isolates from Makassar were distantly related to other VOC lineages (Alpha/B.1.1.7, Beta/B.1.351, Gamma/P.1, Delta/B.1.617.2, Omicron/B.1.1.529).

Mutation analysis of the S protein

Mutation analysis of the S protein region showed that the most frequently occurring mutations in our study were D614G (50/50, 100%), followed by N439K (49/50, 98%) and P681R (38/50, 76%) (Fig. 2A and Table 2). Other mutations were found at lower frequencies, including L5F, V143F, Y144del, D253G, S255F, A348S, Q677H, A1070S, and P1263L. Mapping of the identified mutations showed they were predominantly in the S1 subunit (10 mutations), while two were located in the S2 subunit, and two others (A348S and N439K) were found in the receptor-binding domain (RBD). D641G, P681R, and Q677H mutations were located close to the furin cleavage site (FCS) region (Fig. 2B). Four mutations (V143F, Y144del, D253G, and S255F) were identified in the N-terminal domain (NTD) of the S1 subunit (Fig. 2A). The distribution of these specific mutations in our collected samples are provided in Table S2.

Mutation analysis of other SARS-CoV-2 genome regions

Mutations occurring with high frequency (>10%) in other regions of the SARS-CoV-2 genome are provided in Table 2 and Table S2. In the structural protein, we found that the T205I mutation of the N protein was present in all collected samples in this study. Only one mutation (2%) occurred in the envelope (E) protein –L73I. Two mutations were found in the membrane (M) protein at low frequency –I24T (2%) and L46F (4%) (Table S2).

In non-structural proteins, three dominant mutations were identified in the NS3 protein of ORF1ab –S126L (76%), T350I (100%), and P822L (98%). In the nsp12 protein of ORF1ab, all samples had a mutation in P323L. The Q57H mutation in the NS3 protein

Table 1 The clinical characteristics of patients with COVID-19 in Makassar, South Sulawesi, Indonesia, from January to April 2021.

Variable	N (%)
Gender	
Male	28 (56)
Female	22 (44)
Age (year)	
0–4	4 (8)
5–17	1 (2)
18–29	9 (8)
30–39	15 (30)
40–49	5 (19)
50–64	10 (20)
65–74	4 (8)
75–84	2 (4)
Clinical severity	
Asymptomatic	11 (22)
Mild	17 (34)
Moderate	17 (34)
Severe	5 (10)
Comorbidities[*]	
Hypertension	1 (10)
Coronary Artery Disease	1 (10)
AV Block	1 (10)
Coagulopathy	2 (10)
Malnutrition	1 (10)
Elevated Liver Enzyme	1 (10)
Anxiety Disorder	1 (10)

Notes.
[*]Data available in 10 samples.

of ORF3a was found in 49 (98%) of our collected samples. No mutations were identified in nsp4, nsp7, nsp9, or nsp10 of the ORF1ab and ORF10 proteins.

DISCUSSION

The COVID-19 pandemic continues to present a pressing challenge to global health. Despite possessing a proofreading enzyme, SARS-CoV-2 mutates and adapts rapidly, resulting in new variants with enhanced fitness. By 29th March 2022, the World Health Organization had identified a total of five VOCs, including previously circulating VOCs (Alpha, Beta, Gamma) and VOCs in current circulation (Delta and Omicron) (*WHO, 2021b*). VOCs tend to demonstrate increased transmissibility, and mitigation measures, such as vaccines, therapeutics, diagnostics, and other public or social health measures, are generally less effective for them (*Thomson et al., 2021*). Although the parental SARS-CoV-2 strains overwhelmingly affected older age groups in particular, our results show that the virus also affects adolescents and young adults in comparable proportions. Global genomic
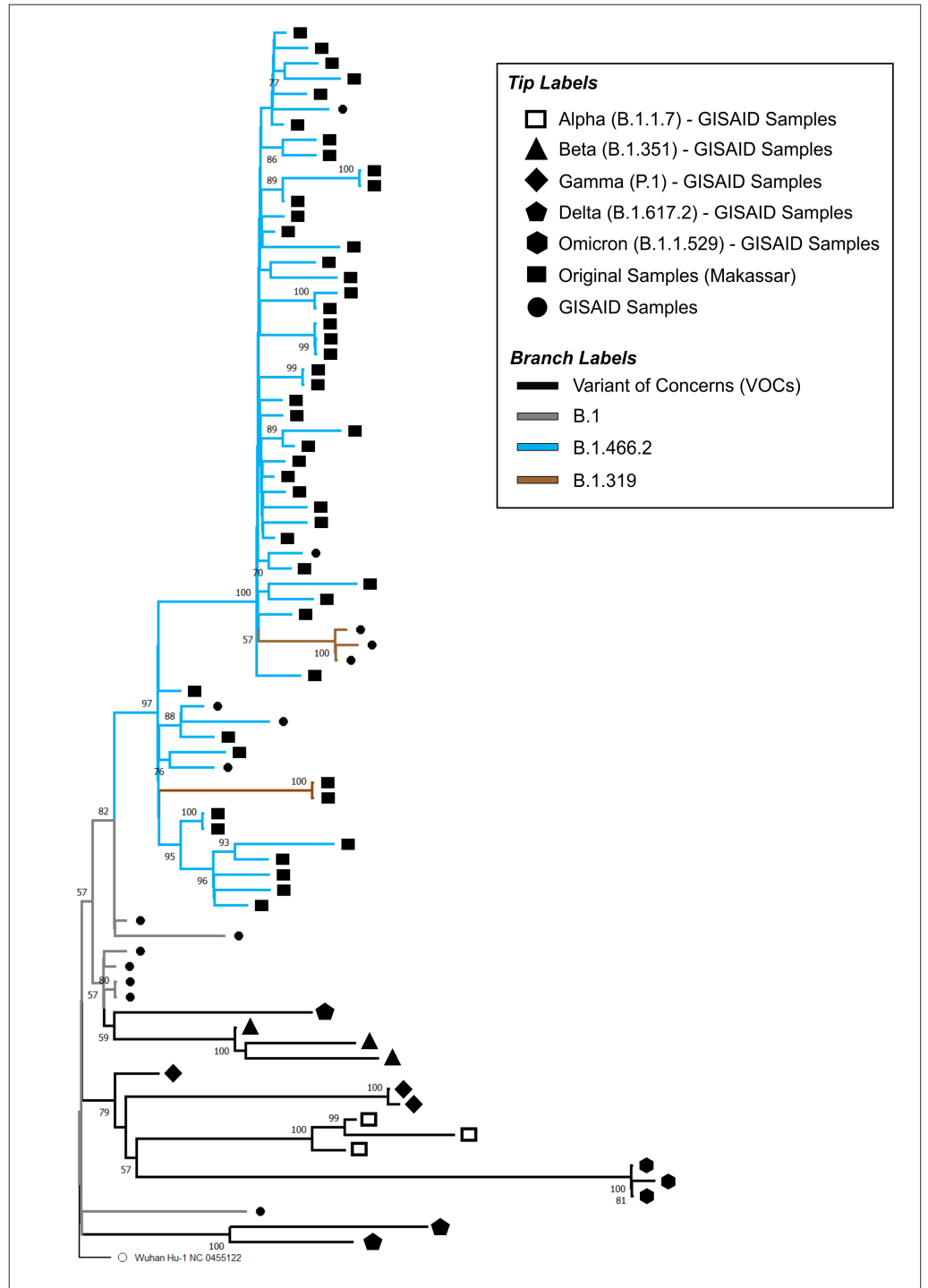
Figure 1 **Phylogenetic analysis of SARS-CoV-2 from Makassar and other countries.** A phylogenetic tree was constructed from the full-length genome of SARS-CoV-2 using the maximum likelihood statistical method, with 1,0000 bootstrap replications and the best DNA substitution model for the dataset (GTR+G). Virus isolates originating from Makassar are indicated in green. The tree was rooted on the ancestor virus hCoV-19/Wuhan/Hu-1/2019.
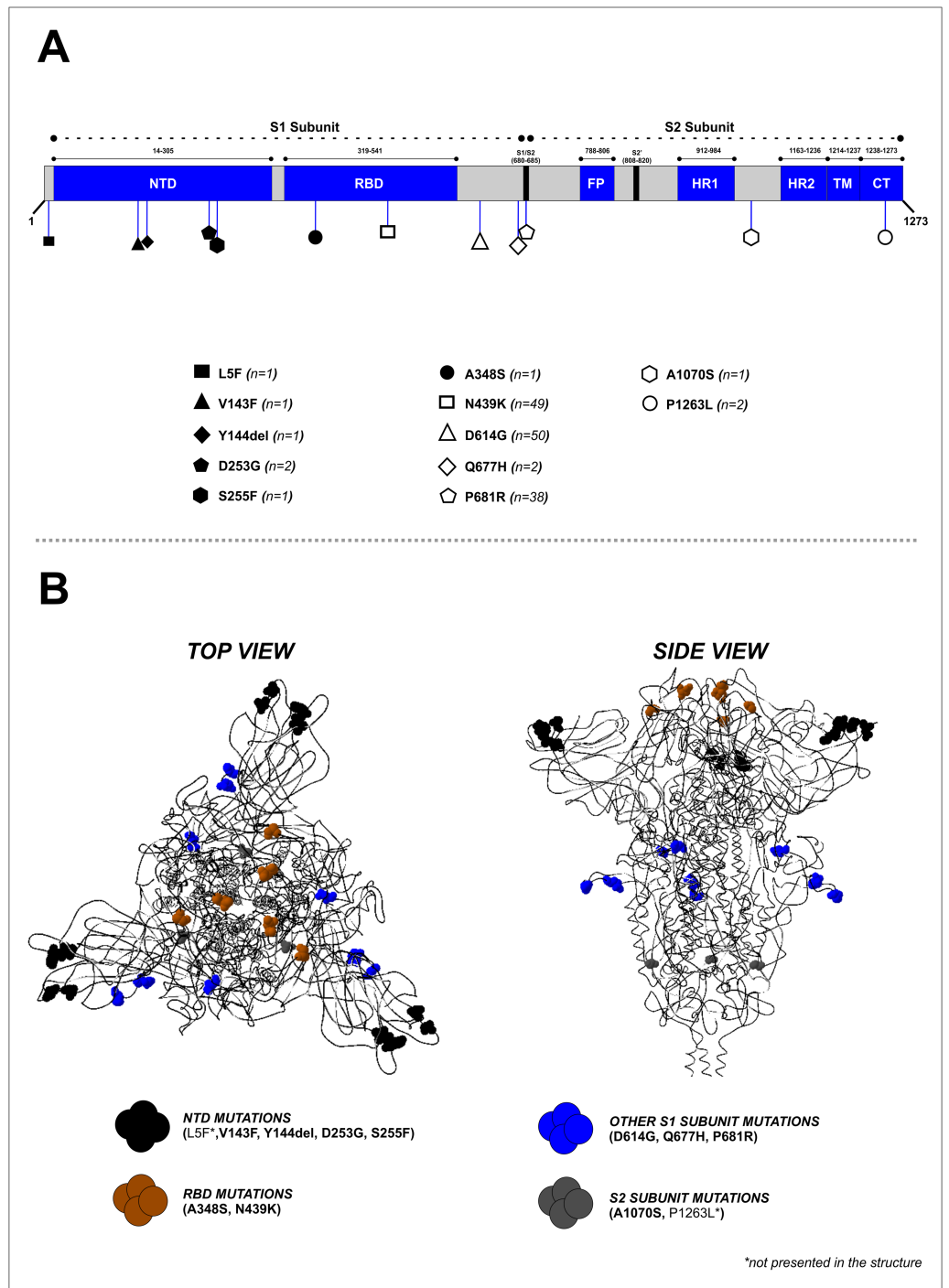
Figure 2   Mapping of the S protein mutations identified in this study. (B) Visualization of the 3D structure of the S protein containing identified mutations, in top and side views. The mutations were found predominantly in the S1 subunit—10 mutations were found in this region, compared to only two in the S2 subunit.

Table 2 **The dominant mutations in SARS-CoV-2 proteins identified in this study.** Mutations with a frequency of > 10% in our collected samples are included.

No	Region	Mutation	Frequency [n (%)]
1	nsp3-ORF1ab	S126L	38 (76)
		T350I	50 (100)
		P822L	49 (98)
2	nsp6-ORF1ab	L75F	38 (76)
		V149F	6 (12)
3	nsp12-ORF1ab	P323L	50 (100)
4	nsp13-ORF1ab	S259L	38 (76)
5	Spike	N439K	49 (98)
		D614G	50 (100)
		P681R	38 (76)
6	NS3-ORF3a	Q57H	49 (98)
7	Nucleoprotein	T205I	50 (100)

surveillance of SARS-CoV-2 plays an important role in mitigating and monitoring the emergence and spread of new variants, as well as in providing information for public health measures (*WHO, 2021*; *WHO, 2021a*).

This study reports SARS-CoV-2 genomic surveillance using whole-genome sequencing in Makassar, the capital city of the South Sulawesi province of Indonesia, from January to April 2021. We performed phylogenetic and mutation analysis for in-depth characterization of the SARS-CoV-2 variants circulating in this region.

Phylogenetic analysis showed that most isolates originated from the B.1.466.2 lineage, while the others (two samples) belonged to the B.1.319 lineage. These results were higher than those of another study carried out in West Java, another province in Indonesia, between September 2020 to June 2021, which found that 27% of isolates were of the B.1.466.2 lineage (*Fibriani et al., 2021*). B.1.466.2 variants are more locally dispersed in Indonesia, and have been detected at lower rates in other Southeast Asian countries (*Cahyani et al., 2021*) B.1.466.2 is one of the Indonesian lineages identified by PANGO, and although it has not been considered a cause for concern on a global scale, it has been designated on alert for further monitoring by the World Health Organization (*WHO, 2021b*).

The results of the mutation analysis show that ORF1ab contains the greatest number of amino acid mutations, followed by the ORF3a, Spike, and N. The ORF1ab combines codes of a polyprotein that are cleaved into 16 NSPs, which play a variety of roles in viral replication (*Yoshimoto, 2020*). Amino acid mutations are also able to affect the structure of proteins, which can impact their function in viral replication. The ORF1ab consists of two amino acid mutations (T350I and P323L) observed in 100% of the genome sequences in the cohort. The *ORF3a* gene, which encodes an ion channel protein related to NLRP3 inflammasome reactivation (*Yoshimoto, 2020*), was found to have a total of 16 amino acid mutations in the cohort, with the Q57H mutation present in 49 (98%) sequences. The nucleocapsid (N) protein encoded by the *N* gene encapsulates the RNA genome of

SARS-CoV-2 (*Yoshimoto, 2020*), and all 50 sequences collected in this cohort were found to contain the T205I mutation.

The S protein, encoded by the *S* gene, is a SARS-CoV-2 surface glycoprotein of particular interest, as it plays an essential role in viral entry into the host cell by binding to cell surface angiotensin-converting enzyme 2 (ACE2) proteins. The S protein is a target for antibody neutralization; thus, amino acid mutations in antibody binding sites within the S protein allow for the evasion of antibody-based host defense mechanisms (*Walls et al., 2020*). In addition, the RNA sequence of the S region is employed as the target of the primer used in the PCR viral detection method. Consequently, nucleic acid mutations in these areas will result in false negative PCR results (*Wang et al., 2020*; *Vogels et al., 2020*). All genomic sequences collected in this cohort contain the background D614G (B.1. lineage) mutation, and 98% of them contain the N439K mutation.

The D614G mutation quickly became prominent in SARS-CoV-19 strains (*Korber et al., 2020*; *Zhou et al., 2022*; *Viedma et al., 2021*), and remains present in VOCs circulating currently. D614G was detected in early cases of COVID-19 in several Europe regions, and quickly became dominant (55–85% by the end of April 2020) (*Viedma et al., 2021*; *Tang et al., 2020*). In East Asia, the transition from D to G was slower, with almost 90% of viral samples from Chinese patients still displaying D614 at the end of April 2020 (*Tang et al., 2020*). In Indonesia, D614G was reported to be dominant (94.2%) until June 2021 (*Cahyani et al., 2021*), and this is consistent with our findings in Makassar city during our study period of January–April 2021. In other Indonesian provinces (namely Yogyakarta and Central Java) the D614G variant circulating since May 2020 finally replaced the D614 variant completely as of September 2020 (*Wijayanti et al., 2022*). Several studies have attempted to examine the effects of this mutation, and results suggest that it is a crucial to the FCS region. Structural and biochemical research on a full-length G614 spike trimer revealed that there are connections in D614G not seen in D614, which prevent the S1 subunit from prematurely losing its binding to the ACE2 (*Zhang et al., 2021*). Other reports found that the change of amino acid at position 614 from aspartic acid (D) to glycine (G) results in a greater ability to penetrate ACE2-expressing cells and replicate in human respiratory cells (*Plante et al., 2021*; *Zhang et al., 2020*). This ability essentially increases the number of spikes that can allow infection to occur, and indicates that the D614 viral spike is less stable than the G614 variant. The G614 mutation is thought to confer a moderate degree of infectivity and transmissibility, which could account for its persistent circulation worldwide (*Korber et al., 2020*; *Plante et al., 2021*; *Zhang et al., 2020*).

Similarly to B.1.466.2, the N439K mutation was found in moderately high frequency in this study. The N439K mutation was first detected in Scotland in March 2020 and increased in prevalence until June 2020, before disappearing entirely from the country following stringent public health measures. Despite this, the mutation has reappeared independently several times in other parts of the world. It was first sampled in Romania in May 2020, and as of January 2021 had been detected in 34 countries as the second most observed receptor-binding domain (RBD) mutation worldwide (*Thomson et al., 2021*). The N439K mutation was most dominant in Indonesia between January and April 2021, and formed its own characteristic lineage (B.1.466.2) in the Indonesian population (*Rambaut et al., 2020*;

O'Toole et al., 2021a; *O'Toole et al., 2021b*), as reflected in the samples obtained for this study. Experiments involving recombinant N439K, hACE2 proteins, and X-ray structural analysis have shown that the N439K mutation results in a 2-fold increase in hACE2 binding affinity (*Thomson et al., 2021*). The effect of this on transmissibility and infectivity measures requires further evaluation. The clinical outcomes and viral replication of N439K mutants in cell tissue culture were comparable to those of N439 SARS-CoV-2 (*Thomson et al., 2021*). The N439K mutants demonstrated more significant 2-fold reduction in response to polyclonal antibody in a small proportion (6.8%) of sera collected from 442 individuals who recovered from COVID-19. A reduction in binding of more than 2-fold was also demonstrated in 16.7% of 140 mAbs isolated from recovered individuals (*Thomson et al., 2021*). In addition, *Zhou et al. (2021)* found that the N439K RBD was significantly resistant to the SARS-CoV-2 neutralizing antibody REGN10987, which could potentially result in neutralization failure. This information is important for ensuring the efficacy of vaccines and immunotherapy. Amino acid changes appear to accumulate gradually over time. N439K is the most prominent mutation in the RBD domain, and should be used preferentially to determine the effect of mutations on pathogen evolution.

Another amino acid mutation that should be highlighted is found in the spike protein–P681R. The frequency of P681R was found to be moderately high in Makassar during the study period, compared to data from the rest of Indonesia (42.6%) and the world (2.1%). The potential causes of this, including natural conditions, lifestyle, genetic factors of hosts, and adherence to health protocols, should be further investigated. Studies on viruses with P681R have revealed that the mutation could result in a reduction in neutralizing antibodies. Furthermore, it has been shown to increase cell–cell fusion by promoting S protein cleavage, mediated by furin (*Saito et al., 2021*). Interestingly, P681R was found frequently in the Delta variant, based on data obtained from GISAID from June to August 2021 (92.7% of all data submitted to GISAID and 99.6% of all Delta variants submitted to GISAID). This mutation may have benefitted from a cumulative fitness advantage over other variants, due to its high frequency in the Indonesian B.1.466.2 lineage.

These findings highlight the necessity of ongoing surveillance to identify novel COVID-19 mutations. The ability to detect new variants and their spread is critical for appropriate public health response and policy making.

CONCLUSION

We report the full-genome sequences of SARS-CoV-2 collected from Makassar, South Sulawesi, Indonesia, between January and April 2021. Lineage B.1.466.2 represents the most dominant (96%) of our isolates, and none of these variants are classified as VOC. All SARS-CoV-2 isolates were found to contain the D614G mutation, while the N439K mutation was also found in high proportions (98%) compared to the lineage in the total Indonesian SARS-CoV-2 sequences submitted to GISAID. Our study therefore highlights the importance of continuous genomic surveillance of SARS-CoV-2, particularly in the South Sulawesi region, as well as in other parts of Indonesia, to monitor the emergence and spread of novel mutations and variants with possible impacts on global health.

ACKNOWLEDGEMENTS

The authors would like to thank Muhammad Faruk, MD, for his contributions in editing this manuscript, and express their gratitude to all individuals who participated in this study.

Funding

This work is supported by the Ministry of Research and Technology/National Research and Innovation Agency of Republic Indonesia (Ristek-BRIN) in collaboration with Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan/LPDP) (Grant no. 13/FI/P-KCOVID-19.2B3/X/2020). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
The Ministry of Research and Technology/National Research and Innovation Agency of Republic Indonesia (Ristek-BRIN).
Indonesia Endowment Fund for Education (Lembaga Pengelola Dana Pendidikan/LPDP): 13/FI/P-KCOVID-19.2B3/X/2020.

Competing Interests

The authors declare there are no competing interests

Author Contributions

- Muhammad Nasrum Massi conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Rufika Shari Abidin conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Abd-ElAziem Farouk performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Handayani Halik analyzed the data, prepared figures and/or tables, and approved the final draft.
- Gita Vita Soraya conceived and designed the experiments, performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Najdah Hidayah conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Rizalinda Sjahril conceived and designed the experiments, prepared figures and/or tables, and approved the final draft.
- Irda Handayani performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

- Mohamad Saifudin Hakim conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Faris Muhammad Gazali performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Vivi Setiawaty conceived and designed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Tri Wibawa conceived and designed the experiments, performed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.

Human Ethics

The following information was supplied relating to ethical approvals (i.e., approving body and any reference numbers):

The study has obtained ethical approval from the Health Research Ethical Committee of Faculty of Medicine, Universitas Hasanuddin, Makassar, South Sulawesi, Indonesia (No. 845/UN4.6.5.31/PP36/2020).

Microarray Data Deposition

The following information was supplied regarding the deposition of microarray data:

https://www.epicov.org/epi3/frontend#252fd0

Data Availability

The SARS-CoV-2 genome sequences used in this study are available via GISAID (https://www.epicov.org/epi3/frontend#252fd0):

EPI_ISL_1247019; EPI_ISL_1247873; EPI_ISL_1248934; EPI_ISL_1249610; EPI_ISL_1250222; EPI_ISL_1258423; EPI_ISL_1258424; EPI_ISL_1258463; EPI_ISL_1258464; EPI_ISL_1258465; EPI_ISL_1263319; EPI_ISL_1263320; EPI_ISL_1263321; EPI_ISL_1263322; EPI_ISL_1263323; EPI_ISL_1630277; EPI_ISL_1630278; EPI_ISL_1630279; EPI_ISL_1630280; EPI_ISL_1630281; EPI_ISL_2691912; EPI_ISL_2692137; EPI_ISL_2692142; EPI_ISL_2692143; EPI_ISL_2692144; EPI_ISL_2692169; EPI_ISL_2692170; EPI_ISL_2692171; EPI_ISL_2692179; EPI_ISL_2692205; EPI_ISL_2692206; EPI_ISL_2692531; EPI_ISL_2692533; EPI_ISL_2692706; EPI_ISL_2692880; EPI_ISL_2692882; EPI_ISL_2692883; EPI_ISL_2692884; EPI_ISL_2692886; EPI_ISL_2692887; EPI_ISL_2692889; EPI_ISL_2692891; EPI_ISL_2692892; EPI_ISL_2692896; EPI_ISL_2692897; EPI_ISL_2692899; EPI_ISL_2692900; EPI_ISL_2692902; EPI_ISL_2692903; EPI_ISL_2692905.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj.13522#supplemental-information.

REFERENCES

Abdool Karim SS, de Oliveira T. 2021. New SARS-CoV-2 variants—clinical, public health, and vaccine implications. *New England Journal of Medicine* **384**:1866–1868 DOI 10.1056/nejmc2100362.

Alwan A. 2007. World Health Organization. *Disaster Medicine and Public Health Preparedness* **1**:7–8 DOI 10.1097/DMP.0b013e3180676d32.

Cahyani I, Putro EW, Ridwanuloh AM, Wibowo SHB, Hariyatun H, Syahputra G, Akbariani G, Utomo AR, Ilyas M, Loose MW, Kusharyoto W, N. 6#, Susanti S. 2021. Genome profiling of SARS-cov-2 in indonesia, ASEAN, and the neighbouring east asian countries: features, challenges, and achievements. *BioRxiv* **7**:2021.07.06.451270 DOI 10.1101/2021.07.06.451270.

Coronaviridae Study Group of the International Committee on Taxonomy of Viruses. 2020. The species Severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. *Nature Microbiology* **5(4)**:536–544 DOI 10.1038/s41564-020-0695-z.

Duchene S, Featherstone L, Haritopoulou-Sinanidou M, Rambaut A, Lemey P, Baele G. 2020. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evolution* **6**:veaa061 DOI 10.1093/ve/veaa061.

ECDC. 2022. SARS-CoV-2 variants of concern as of 5 January 2022. *Available at https://Www.Ecdc.Europa.Eu/En/Covid-19/Variants-Concern*.

El-Raey F, Alboraie M, Youssef N, Yousef A, Abdelmoaty AA, Hassan E, Hassany SM, Abd-Elsalam S, Elsharkawy R, Farrag K, Elbadry M, Tag-Adeen M, Abdeen N, Elshaarawy O, Haydara T, Ahmed Hassan A, El-Sayed M, Fouad Y, Abdel Baki A, Zaky S. 2021. Predictors for severity of SARS-cov-2 infection among healthcare workers. *Journal of Multidisciplinary Healthcare* **14**:2973–2981 DOI 10.2147/JMDH.S335226.

Fibriani A, Stephanie R, Alfiantie AA, Siregar ALF, Pradani GAP, Yamahoki N, Purba WS, Alamanda CNC, Rahmawati E, Rachman RW, Robiani R, Ristandi RB. 2021. Analysis of SARS-CoV-2 genomes from West Java, Indonesia. *Viruses* **13(10)**:2097 DOI 10.3390/v13102097.

GISAID. 2021. Submission tracking: Indonesia, GISAID. *Available at https://www.gisaid.org/* (accessed on 29 March 2022).

Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution* **30**:772–780 DOI 10.1093/molbev/mst010.

Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, Hengartner N, Giorgi EE, Bhattacharya T, Foley B, Hastie KM, Parker MD, Partridge DG, Evans CM, Freeman TM, de Silva TI, Angyal A, Brown RL, Carrilero L, Green LR, Groves DC, Johnson KJ, Keeley AJ, Lindsey BB, Parsons PJ, Raza M, Rowland-Jones S, Smith N, Tucker RM, Wang D, Wyles MD, McDanal C, Perez LG, Tang H, Moon-Walker A, Whelan SP, LaBranche CC, Saphire EO, Montefiori DC, Angyal A, Brown RL, Carrilero L, Green LR, Groves DC, Johnson KJ, Keeley AJ, Lindsey

BB, Parsons PJ, Raza M, Rowland-Jones S, Smith N, Tucker RM, Wang D, Wyles MD. 2020. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**:812–827 DOI 10.1016/j.cell.2020.06.043.

Koyama T, Platt D, Parida L. 2020. Variant analysis of SARS-cov-2 genomes. *Bulletin of the World Health Organization* **98**:495–504 DOI 10.2471/BLT.20.253591.

Lauring AS, Hodcroft EB. 2021. Genetic variants of SARS-CoV-2—what do they mean? *Journal of the American Medical Association* **325(6)**:529–531 DOI 10.1001/jama.2020.27124.

Ministry of Health of Republic Indonesia. 2021. SARS-CoV-2 genomic surveillance Indonesia: update of whole genome sequence (WGS) in Indonesia (as of 7 August 2021).

National Institutes of Health. 2021. Treatment guidelines panel. Coronavirus disease 2019 (COVID-19), NIH. 2019. 1–243*Available at* https://www.covid19treatmentguidelines.nih.gov/.%0Ahttps://www.covid19treatmentguidelines.nih.gov/.

Oley MH, Oley MC, Kepel BJ, Tjandra DE, Langi F, Herwen , Aling DMR, Joicetine Wagiu AM, Faruk M. 2021. ICAM-1 levels in patients with covid-19 with diabetic foot ulcers: a prospective study in southeast asia. *Annals of Medicine and Surgery* **63**:102171 DOI 10.1016/j.amsu.2021.02.017.

O'Toole Á, Hill V, Pybus OG, Watts A, Bogoch II, Khan K, Messina JP, COVID-19 Genomics UK (COG-UK) consortium, Network for Genomic Surveillance in South Africa (NGS-SA), Brazil-UK CADDE Genomic Network, Tegally H, Lessells RR, Giandhari J, Pillay S, Tumedi KA, Nyepetsi G, Kebabonye M, Matsheka M, Mine M, Tokajian S, Hassan H, Salloum T, Merhi G, Koweyes J, Geoghegan JL, De Ligt J, Ren X, Storey M, Freed NE, Pattabiraman C, Prasad P, Desai AS, Vasanthapuram R, Schulz TF, Steinbrück L, Stadler T, Swiss Viollier Sequencing Consortium, Parisi A, Bianco A, García de Viedma D, Buenestado-Serrano S, Borges V, Isidro J, Duarte S, Gomes JP, Zuckerman NS, Mandelboim M, Mor O, Seemann T, Arnott A, Draper J, Gall M, Rawlinson W, Deveson I, Schlebusch S, McMahon J, Leong L, Lim CK, Chironna M, Loconsole D, Bal A, Josset L, Holmes E, St. George K, Lasek-Nesselquist E, Sikkema RS, Oude Munnink B, Koopmans M, Brytting M, Sudha Rani V, Pavani S, Smura T, Heim A, Kurkela S, Umair M, Salman M, Bartolini B, Rueca M, Drosten C, Wolff T, Silander O, Eggink D, Reusken C, Vennema H, Park A, Carrington C, Sahadeo N, Carr M, Gonzalez G, SEARCH Alliance San Diego, National Virus Reference Laboratory, SeqCOVID-Spain, Danish Covid-19 Genome Consortium (DCGC), Communicable Diseases Genomic Network (CDGN), Dutch National SARS-CoV-2 surveillance program, Division of Emerging Infectious Diseases (KDCA), de Oliveira T, Faria N, Rambaut A, Kraemer MUG. 2021a. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. *Wellcome Open Research* **6**:121 DOI 10.12688/wellcomeopenres.16661.2.

O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, Colquhoun R, Ruis C, Abu-Dahab K, Taylor B, Yeats C, du Plessis L, Maloney D, Medd N, Attwood SW, Aanensen DM, Holmes EC, Pybus OG, Rambaut A. 2021b. Assignment of

epidemiological lineages in an emerging pandemic using the pangolin tool. *Virus Evolution* **7**:veab064 DOI 10.1093/ve/veab064.

Oxford Nanopore Technologies. 2021. MinION Mk1B IT requirements. *Available at* https://bit.ly/3EDRhLs.

Plante JA, Liu Y, Liu J, Xia H, Johnson BA, Lokugamage KG, Zhang X, Muruato AE, Zou J, Fontes-Garfias CR, Mirchandani D, Scharton D, Bilello JP, Ku Z, An Z, Kalveram B, Freiberg AN, Menachery VD, Xie X, Plante KS, Weaver SC, Shi PY. 2021. Spike mutation D614G alters SARS-CoV-2 fitness. *Nature* **592**:116–121 DOI 10.1038/s41586-020-2895-3.

Pradipta A, Kumaheri MA, Wahyudi LD, Susanto AP, Agasi HI, Shankar AH, Sudarmono P. 2022. Accelerating detection of variants during COVID-19 surges by diverse technological and public health partnerships: a case study from Indonesia. *Frontiers in Genetics* **13**:1–4 DOI 10.3389/fgene.2022.801332.

Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nature Microbiology* **5**:1403–1407 DOI 10.1038/s41564-020-0770-5.

Robert Koch Institute. 2021. Bericht zu Virusvarianten von SARS-CoV-2 in Deutschland. Berlin: Robert Koch Institute. *Available at* https://www.rki.de/DE/Content/ InfAZ/N/Neuartiges_Coronavirus/DESH/Bericht_VOC_2021-06-23.pdf .

Saito A, Irie T, Suzuki R, Maemura T, Nasser H, Uriu K, Kosugi Y, Shirakawa K, Sadamasu K, Kimura I, Ito J, Wu J, Iwatsuki-Horimoto K, Ito M, Yamayoshi S, Loeber S, Tsuda M, Wang L, Ozono S, Butlertanaka EP, Tanaka YL, Shimizu R, Shimizu K, Yoshimatsu K, Kawabata R, Sakaguchi T, Tokunaga K, Yoshida I, Asakura H, Nagashima M, Kazuma Y, Nomura R, Horisawa Y, Yoshimura K, Takaori-Kondo A, Imai M, Chiba M, Furihata H, Hasebe H, Kitazato K, Kubo H, Misawa N, Morizako N, Noda K, Oide A, Suganami M, Takahashi M, Tsushima K, Yokoyama M, Yuan Y, Tanaka S, Nakagawa S, Ikeda T, Fukuhara T, Kawaoka Y, Sato K. 2021. Enhanced fusogenicity and pathogenicity of SARS-CoV-2 Delta P681R mutation. *Nature* **602**:300–306 DOI 10.1038/s41586-021-04266-9.

Sevajol M, Subissi L, Decroly E, Canard B, Imbert I. 2014. Insights into RNA synthesis, capping, and proofreading mechanisms of SARS-coronavirus. *Virus Research* **194**:90–99 DOI 10.1016/j.virusres.2014.10.008.

Smith EC, Blanc H, Vignuzzi M, Denison MR. 2013. Coronaviruses lacking exoribonuclease activity are susceptible to lethal mutagenesis: evidence for proofreading and potential therapeutics. *PLOS Pathogens* **9**:e1003565 DOI 10.1371/journal.ppat.1003565.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**:1312–1313 DOI 10.1093/bioinformatics/btu033.

Tamura K, Stecher G, Kumar S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. *Molecular Biology and Evolution* **38**:3022–3027 DOI 10.1093/molbev/msab120.

Tang L, Schulkins A, Chen CN, Deshayes K. 2020. The SARS-CoV-2 spike protein D614G mutation shows increasing dominance and may confer a structural advantage

to the furin cleavage domain. *Preprints* **May 24**:1–14
DOI 10.20944/preprints202005.0407.v1.

Thomson EC, Rosen LE, Shepherd JG, Spreafico R, Da Silva Filipe A, Wojcechowskyj JA, Davis C, Piccoli L, Pascall DJ, Dillen J, Lytras S, Czudnochowski N, Shah R, Meury M, Jesudason N, De Marco A, Li K, Bassi J, O'Toole A, Pinto D, Colquhoun RM, Culap K, Jackson B, Zatta F, Rambaut A, Jaconi S, Sreenu VB, Nix J, Zhang I, Jarrett RF, Glass WG, Beltramello M, Nomikou K, Pizzuto M, Tong L, Cameroni E, Croll TI, Johnson N, Di Iulio J, Wickenhagen A, Ceschi A, Harbison AM, Mair D, Ferrari P, Smollett K, Sallusto F, Carmichael S, Garzoni C, Nichols J, Galli M, Hughes J, Riva A, Ho A, Schiuma M, Semple MG, Openshaw PJM, Fadda E, Baillie JK, Chodera JD, Rihn SJ, Lycett SJ, Virgin HW, Telenti A, Corti D, Robertson DL, Snell G. 2021. Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated immunity. *Cell* **184**:1171–1187 DOI 10.1016/j.cell.2021.01.037.

Viedma E, Dahdouh E, González-Alba JM, González-Bodi S, Martínez-García L, Lázaro-Perona F, Recio R, Rodríguez-Tejedor M, Folgueira MD, Cantón R, Delgado R, García-Rodríguez J, Galán JC. 2021. J. Mingorance, Genomic epidemiology of SARS-CoV-2 in Madrid, Spain, during the first wave of the pandemic: fast spread and early dominance by D614G variants. *Microorganisms* **9**:1–11 DOI 10.3390/microorganisms9020454.

Vogels CBFF, Brito AF, Wyllie AL, Fauver JR, Ott IM, Kalinich CC, Petrone ME, Casanovas-Massana A, Catherine Muenker M, Moore AJ, Klein J, Lu P, Lu-Culligan A, Jiang X, Kim DJ, Kudo E, Mao T, Moriyama M, Oh JE, Park A, Silva J, Song E, Takahashi T, Taura M, Tokuyama M, Venkataraman A, El Weizman O-E, Wong P, Yang Y, Cheemarla NR, White EB, Lapidus S, Earnest R, Geng B, Vijayakumar P, Odio C, Fournier J, Bermejo S, Farhadian S, Dela Cruz CS, Iwasaki A, Ko AI, Landry ML, Foxman EF, Grubaugh ND. 2020. Analytical sensitivity and efficiency comparisons of SARS-CoV-2 RT–qPCR primer–probe sets. *Nature Microbiology* **5**:1299–1305 DOI 10.1038/s41564-020-0761-6.

Walensky RP, Walke HT, Fauci AS. 2021. SARS-CoV-2 variants of concern in the united states-challenges and opportunities. *JAMA The Journal of the American Medical Association* **325**:1037–1038 DOI 10.1001/jama.2021.2294.

Walls AC, Park YJ, Tortorici MA, Wall A, McGuire AT, Veesler D. 2020. Structure, function, and antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* **181**:281–292 DOI 10.1016/j.cell.2020.02.058.

Wang R, Hozumi Y, Yin C, Wei GW. 2020. Mutations on COVID-19 diagnostic targets. *Genomics* **112**:5204–5213 DOI 10.1016/j.ygeno.2020.09.028.

WHO. 2020a. Coronavirus disease 2019 (2019-nCOV) situation report –11. *Available at* https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200131-sitrep-11-ncov.pdf?sfvrsn=de7c0f7_4.

WHO. 2020b. Timeline—COVID-19 27 April 2020. *Available at* https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19.

WHO. 2021. Genomic sequencing of SARS-CoV-2: a guide to implementation for maximum impact on public health. World Health Organization. *Available at https://www.who.int/publications/i/item/9789240018440*.

WHO. 2021a. SARS-CoV-2 genomic sequencing for public health goals, WHO—Interim Guid. 1–20.

WHO. 2021b. Tracking SARS-CoV-2 variants, World Heal. Organ. *Available at https://www.who.int/activities/tracking-SARS-CoV-2-variants* (accessed on 28 March 2022).

Wijayanti N, Gazali FM, Supriyati E, Hakim MS, Arguni E, Daniwijaya MEW, Nuryastuti T, Nuhamunada M, Nabilla R, Haryana SM, Wibawa T. 2022. Evolutionary dynamics of SARS-CoV-2 circulating in Yogyakarta and Central Java, Indonesia: sequence analysis covering furin cleavage site (FCS) region of the spike protein. *International Microbiology* 1–10. Advance online publication DOI 10.1007/s10123-022-00239-8.

Wu A, Peng Y, Huang B, Ding X, Wang X, Niu P, Meng J, Zhu Z, Zhang Z, Wang J, Sheng J, Quan L, Xia Z, Tan W, Cheng G, Jiang T. 2020. Genome composition and divergence of the novel coronavirus (2019-nCoV) originating in China. *Cell Host & Microbe* **27**:325–328 DOI 10.1016/j.chom.2020.02.001.

Yoshimoto FK. 2020. The proteins of severe acute respiratory syndrome coronavirus-2 (SARS CoV-2 or n-COV19), the cause of COVID-19. *The Protein Journal* **39**:198–216 DOI 10.1007/s10930-020-09901-4.

Zhang J, Cai Y, Xiao T, Lu J, Peng H, Sterling SM, Walsh RM, Rits-Volloch S, Zhu H, Woosley AN, Yang W, Sliz P, Chen B. 2021. Structural impact on SARS-CoV-2 spike protein by D614G substitution. *Science* **372**:525–530 DOI 10.1126/science.abf2303.

Zhang L, Jackson CB, Mou H, Ojha A, Peng H, Quinlan BD, Rangarajan ES, Pan A, Vanderheiden A, Suthar MS, Li W, Izard T, Rader C, Farzan M, Choe H. 2020. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nature Communications* **11(1)**:6013 DOI 10.1038/s41467-020-19808-4.

Zhou P, Lou Yang X, Wang X-GG, Hu B, Zhang L, Zhang W, Si HR, Zhu Y, Li B, Huang CL, Chen HD, Chen J, Luo Y, Guo H, Di Jiang R, Liu MQ, Chen Y, Shen XR, Wang X-GG, Zheng XS, Zhao K, Chen QJ, Deng F, Liu LL, Yan B, Zhan FX, Wang YY, Xiao GF, Shi ZL. 2020. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**:270–273 DOI 10.1038/s41586-020-2012-7.

Zhou W, Xu C, Wang P, Anashkina AA, Jiang Q. 2022. Impact of mutations in SARS-COV-2 spike on viral infectivity and antigenicity. *Briefings in Bioinformatics* **23**:1284–1294 DOI 10.1093/bib/bbab375.

Zhou W, Xu C, Wang P, Luo M, Xu Z, Cheng R, Jin X, Guo Y, Xue G, Juan L, Anashkina AA, Nie H, Jiang Q. 2021. N439K variant in spike protein alter the infection efficiency and antigenicity of SARS-CoV-2 based on molecular dynamics simulation. *Frontiers in Cell and Developmental Biology* **9**:1–11 DOI 10.3389/fcell.2021.697035.