Whole-genome-scale analysis of circulating SARS-CoV-2 during the first COVID-19 wave in an international tourist destination, Bali, Indonesia

Sri Masyeni, Duwi Sumohadi, Krisna Duta, Ida Ayu Mahayani, Komang Parwata, Agus Eka Darwinata, Harapan

ABSTRACT

Background: The emergence of severe acute respiratory syndromes coronavirus 2 (SARS-CoV-2) variants such as B.1.17 (alpha), B.1.351 (beta), and B.1.617.2 (delta) have caused a significant rise in coronavirus disease 2019 (COVID-19) cases worldwide. In response, the Indonesian government imposed restrictions on international and domestic travel, especially in areas with a high number of COVID-19 cases including Java Island and Bali. The aim of this study was to investigate the mutation patterns of indigenous SARS-CoV-2 strains in Bali.

Methods: We conducted whole-genome sequencing on isolates collected from COVID-19-confirmed patients at the end of the first wave (January – March 2021). The sequencing was carried out at the 1st Base Pte Ltd Laboratory (Singapore) using Oxford Nanopore Platform (GridION) using IDT ARTIC nCoV-2019 V3 Panel. The phylogenetic tree was constructed on MEGA 7.0 with hCoV-19/Wuhan/WIV04/2019 as the genomic reference.

Results: Our data revealed that none of the isolates were variants of concern. Among the five lineages tracked, the viral isolates were most likely from B.1.466.2 (n=4). All viruses had D614G mutation in the spike protein with clade GH being most predominant (n=10) and followed by GR (n=1) and O (n=1). Additionally, mutations at NS3, NSP3, NSP12, and NSP6 were found.

Conclusion: At the end of the first COVID-19 wave in Bali, variants of concern were not detected which could be attributed to the heavy mobility restrictions.

Keywords: Bali, B.1.466.2, D614G, genomic surveillance, Indonesia.


INTRODUCTION

In 2019, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) emerged and have caused coronavirus disease 2019 (COVID-19) pandemic and associated with significant health problem in Indonesia.\(^1\)\(^2\) In comparison to the previous two coronaviruses (MERS-CoV and SARS-CoV), the infectivity of SARS-CoV-2 is the highest but has a lower mortality rate.\(^3\) The clinical spectrum of COVID-19 ranging from being asymptomatic to developing acute respiratory distress and systemic involvements.\(^4\) Treatment of this disease was and remains a major challenge, attributed to the lack of hospital infrastructure and absence of specific antiviral.\(^5\)\(^6\) Countries in tropical regions, such as Indonesia, is prone to COVID-19 coinfecion with dengue increasing its health burden.\(^7\) Hence, strict COVID-19 regulations coupled with genomic surveillance of SARS-CoV-2 are required to suppress the pandemic waves in Indonesia.

The first laboratory-confirmed COVID-19 cases in Indonesia were officially reported on March 2, 2020, transmitted from a Japanese traveler.\(^8\) A few weeks later on March 20, 2020, the cases were no longer solely from importation, but also local transmission indicating the SARS-CoV-2 outbreak in Indonesia.\(^8\) Despite these data, earlier importation of COVID-19 cases might occur since two cities in Indonesia, Jakarta and Denpasar, were found among global top high-risk cities which ranked 29\(^{th}\) and 11\(^{th}\), respectively.\(^8\) Denpasar is located in a tourism hotspot province of Indonesia, Bali, which receives intense frequency of international tourists.
Following the active and local human-to-human transmission of SARS-CoV-2, Indonesian government closed its border for foreign travelers while nationally imposing the COVID-19 restrictions to curb the pandemic.\textsuperscript{10} These restrictions were then temporarily lifted. Steady rise of confirmed COVID-19 cases along with the emergence of new and more infectious SARS-CoV-2 strains, Indonesia closed its border once again from January to March 2022.\textsuperscript{11}

Genomic epidemiology data could be used to assess the impact of government’s responses toward the SARS-CoV-2 transmission. For example, a year of genomic surveillance in Africa revealed major importation of variants of concern (VOC) and variants of interest (VOI) was due to the lack of mobility restriction.\textsuperscript{12} Moreover, whole genome sequencing could inform the mutation dynamics of SARS-CoV-2. Unfortunately, available published literatures reporting whole genome sequencing of SARS-CoV-2 in Indonesia only covered Java Island and Sulawesi Island, or immunocompromised population.\textsuperscript{13-15} Herein, we performed whole-genome sequencing on SARS-CoV-2 positive samples collected from Bali from January to March 2021. The isolates were categorized into clades following the nomenclature Global Initiative on Sharing All Influenza Data (GISAID) proposed. The SARS-CoV-2 virus reference strain is assigned to clade L, while others were assigned to clades S, V, G, GH, GR, and GV.\textsuperscript{16,17} Clade S members have L84S mutation at NS8, while clade V members have L37F mutation at NSP6 and G251 mutation at NS3.\textsuperscript{16,17} As for SARS-CoV-2 strains with D614G mutation in the spike protein, they belong to clade G.\textsuperscript{16,17} Additional mutations other than D614, namely NS3-Q57H, N-G204R and G251 mutation at NS3. PCR conditions were as follows: 12.5 µL Q5 Hot Start High-Fidelity 2x Master Mix (M0493L, New England BioLabs, Ipswich, MA, USA). The mixture contained 16 µL LunaScript SuperMix and 4 µL RNA samples.

Multiplex PCR was performed using the cDNA samples. Two PCR reactions (Pool 1 and Pool 2) were used per sample using primers from IDT ARTIC nCoV-2019 sequencing protocol v3 (LoCost)). The RNA samples were directly used in cDNA synthesis using LunaScript™ RT SuperMix Kit (E3010L, New England BioLabs, Ipswich, MA, USA). The mixture was then quantified using Qubit dsDNA HS kit (Q32851, ThermoFisher, Carlsbad, CA, USA) and 100 ng was taken forward to library preparation using SQK-LSK109 with EXP-NBD104 (1-12) & EXP-NBD114 (13-24) for multiplexing samples using different barcode. Details protocols can be found elsewhere.\textsuperscript{18}

Phylogenetic analysis

Data of SARS-CoV-2 sequences available until March 2021 were downloaded from GISAID, where they were subsequently aligned on MAFFT v.7.309, as implemented in Geneious v.10.1.3.\textsuperscript{19,20} Phylogenetic tree was then constructed to depict the relationship between the SARS-CoV-2 viruses from Bali and other places with the maximum likelihood-based phylogenetic tree was estimated using Randomized Axelerated Maximum Likelihood (RAxML) v.7.2.8.\textsuperscript{21,22} Reconstructions of phylogenetic tree were performed on MEGA 7.0. visualization using FigTree v1.4.3 was used for the resultant phylogenies.

RESULTS

SARS-CoV-2 obtained from all samples along with mutations are presented in Table 1. Among the 12 patients, 3 patients were above 60 years old, and one patient was 4 years old, with male being the most predominant (n=10). Most of the samples were collected in March (n=9) then in February (n=1) or January (n=2). From the most to the least frequent, the viruses were found to be in these following clades: GH (n=10), GR (n=1), and O (n=1). As many as 4 viruses were from B.1.466.2 lineage, whilst the others were from B.1.470 (n=2), B.1.375 (n=2), B.1.1.269 (n=1), B.1.468 (n=1), and B.1 (n=1). All viruses had mutations on the D614G protein. Other common mutations were found in NSP3, NSP6, NSP12, and NS3. None of the viruses was indicated to have VOC mutations.

The phylogenetic tree among the isolated viruses is presented in Figure 1. The viruses had established a distant relationship with the overall common ancestor suggesting a rapid mutation of this virus. The reference genome of Wuhan isolate (hCoV-19/Wuhan/2004/2019)
<table>
<thead>
<tr>
<th>Virus name</th>
<th>Accession ID</th>
<th>Date of sample collection</th>
<th>Patient's Characteristics</th>
<th>Clade</th>
<th>Lineage</th>
<th>Mutations</th>
</tr>
</thead>
</table>
**Virus name**  
**Accession ID**  
**Date of sample collection**  
**Patient’s Characteristics**  
**Clade**  
**Lineage**  
**Mutations**

<table>
<thead>
<tr>
<th>Virus name</th>
<th>Accession ID</th>
<th>Date of sample collection</th>
<th>Patient’s Characteristics</th>
<th>Clade</th>
<th>Lineage</th>
<th>Mutations</th>
</tr>
</thead>
</table>

**Figure 1.** Phylogenetic tree of SARS-CoV-2 from samples collected in Bali at the end of the first COVID-19 wave (January to March 2021).

had the closest relationship with the ancestor and grouped along with two Bali isolates (hCoV-19/Indonesia/BA-Unwar-KL20/2021 and hCoV-19/Indonesia/BA-Unwar-TB16/2021). The isolate collected on 8 January 2021 had the least mutation, meanwhile the isolate with the highest number of mutations was from sample collected later on 16 March 2021 (Figure 1).

**DISCUSSION**

In Indonesia, the isolates appeared to be included in clades L, G, GR, GH, and O. Herein, the results revealed that high prevalence of SARS-CoV-2 belonged to clade GH, where only one that belonged to clade GR. We also found one isolate that could not be included in the seven major clades (G, GH, GR, GV, S, V, and L), hence included into clade O. In a continental scale, based on a report published on 19 April 2021, Asia was reported to be predominated by clade GR which happened to similarly occur on South America, Africa, and Oceania. In a more specific region, a study analyzing samples from four patients in Java Island, Indonesia, revealed the predominance of GH (n=3). A study from a neighboring country, Malaysia, employing 1000 participants using data deposited until July 19, 2021, revealed the similar predominance of clade GH (n=311). Meanwhile in Singapore, within January to March 2021, clades G and GH were predominant compared with others. Furthermore, in this present study, the most dominant tracked lineage was B.1.466.2 of the Balinese isolates. In a nationwide scale, the lineage occupied 50% of daily infection, where Java and Sumatra had been suggested as its epicenters of transmission. This SARS-CoV-2 lineage was also observed in Malaysia and Singapore, but not in other ASEAN countries. These suggest that the prevalence of SARS-CoV-2 clades is more likely to be influenced by the nearby cities or countries, and the clade distribution could be significantly different when distantly located countries are included in the analysis.

This present study revealed that all Balinese isolates had D614G mutation. Other mutations occurred predominantly at NS3, NSP3, NSP6, and NSP12. Previously, the number of D614G mutations were only occurred among 36/60 and 23/40 isolates, collected in 2020. Later findings (January-April 2021) from Sulawesi Island revealed 100% D614G mutations, which is similar with the results herein. The foregoing studies in Java and Sulawesi Islands also found the mutations at NS3, NSP3, and NSP12. Mutations at NSP3 were also observed in high frequency by a study using GISAID metadata of B.1.466.2 strains isolated from...
Indonesia.26 It is also worth-noting that the isolate from a 78-years-old male patient in this present study had the most frequent mutations with NSP6 being the major mutation site. Mutation locations have been found to be linked with the disease severity and viral replication and stability, suggesting the importance of monitoring SARS-CoV-2 mutations especially during the pandemic.28,29 Moreover, phylogenetic tree herein highlights that circulating SARS-CoV-2 strains in Balinese have sequence variations compared with the early ancestors.

In regards with VOCs, Indonesia has reported alpha and beta variants in January 2021.27 However, in early 2021, among South-east Asian countries, Indonesian COVID-19 cases were dominated by non-VOCs.27 In this present study, we reported the absence of VOCs at the end of the first COVID-19 wave based on the whole genome sequencing of 12 patients from January to March 2021. This is in line with the government's report that a single case of alpha variants was found only in the beginning of the second wave (June 2021).30 During January to March 2021, a strict lockdown policy was imposed by the government due to a significant increase of COVID-19. The policy mostly covered restrictions on people's mobility and face-to-face social interaction; such as but not limited to, 75% working office activities should be performed from home, online learning for all schools, and 50% limit for worship house occupancy. At the same time, Indonesian government also imposed domestic and international travel bans.31 Taken altogether, the COVID-19 restrictions had successfully prevent the entry of VOCs into Bali during the first wave of the pandemic.

CONCLUSIONS
All collected isolates from Bali had multiple mutations from their early ancestors. No VOCs were found from the Balinese isolates at the end of the first COVID-19 wave in Indonesia which could indicate the effectiveness of COVID-19 restrictions. As the restrictions have been loosened in Bali, we recommend the continuation of genomic surveillance on SARS-CoV-2 mutations.

ETHICS APPROVAL
The study protocol was approved by the Universitas Warmadewa Ethical Review Board (No. 65/Unwar/FKIK/EC-KEPK/VII/2021).

ACKNOWLEDGMENT
We would like to all patients participate in this present study.

DISCLOSURE OF CONFLICTS OF INTEREST
No conflict of interest.

FUNDING
This study was funded by Faculty of Medicine and Health Sciences, Universitas Warmadewa.

AUTHORS CONTRIBUTION
The authors contributed to the research process, including preparation, conceptualization, data collection and analysis, drafting, and approval to publish.

REFERENCES


