Genomic characterization of reinfection cases in the first wave of the COVID-19 pandemic in Indonesia

Erike Anggraini Suwarsono, Laifa Annisa Hendarmi, Chris Adhiyanto

ABSTRACT

Background: Coronavirus disease 2019 (COVID-19) has become a significant disease causing a pandemic worldwide. Many variants were detected within the globe and may have an important role in transmission capability or reinfection conditions. The ‘unknown’ level of immunity’s protection from previous infection and the more virulent virus could lead to reinfection. Using the whole genome sequence, we can analyze the genomic insight within the first and second infections to help understand the reinfection pattern of COVID-19.

Methods: In this study, we investigated three COVID-19 cases with reinfection for primary and secondary infection each. The SARS-CoV-2 confirmation was conducted using qPCR with multiplex gene targets Orf1ab and E. Confirmation of reinfection was done by genomic analysis. Library preparation and sequencing were done using ARTIC protocol using nanopore technology.

Results: There was a wide variety of mutations between two episodes of infection and different clinical manifestations in each episode.

Conclusions: The result concluded that reinfection happened in worsened clinical symptoms and distinct genomic variation in each episode of infection.

Keywords: COVID-19, reinfection, genomic sequence.


INTRODUCTION

COVID-19 has become a significant disease that is causing a pandemic around the world. In almost 2 years, the world has been growing together with the SARS-CoV-2. Many variants were detected within the globe and may have an important role in transmission capability or reinfection conditions. A limited study of COVID-19 reinfection might contribute to scarce evidence of this re-infection case itself. The ‘unknown’ level of immunity protects the host from previous infection and more virulent viruses, which could lead to emerging reinfection. Antibody in the other coronavirus infection can still be detected within 1-3 years. However, several studies on COVID-19 showed a shorter period of protection. An early study in the UK, from 6614 participants, showed that 83% of antibodies could protect against COVID-19 reinfection in five months. Another study by Mauad T et al. showed that antibody protection in COVID-19 can last within 3-6 months of follow-up. A single case report by Tillet RL et al. reported that reinfection occurred in a 25-year-old 25-year-old man three months from the previous infection with a distinct genome sequence.

As the definition of reinfection is the condition of a second episode of infection caused by a different genome sequence from the first episode, we conducted a whole genome sequence for each episode of infection to analyze the genomic insight of the first and second encounter to help better understand of reinfection in COVID-19. All the reinfection cases were occurred in around first wave of COVID-19 in Indonesia, before vaccination program implemented by our government. This study aimed to capture the genomic characterization of SARS-CoV2 during reinfection events.

METHODS

Our laboratory is one of the Indonesian COVID-19 molecular networks; therefore, we receive samples from hospitals and healthcare facilities around Banten province. The inclusion criteria of this study were all subjects who tested in our laboratory during the year 2020 with positive PCR and already tested negative for PCR in the meantime of the first attempt of COVID-19 infection. However, they were positive again after recovering. In this study, we analyzed three paired cases consecutively of COVID-19 infection in Indonesian adults who have not yet been vaccinated and met the criteria. Those samples were chosen for this study because both paired samples (first and second infection) had a CT value under 30, as this number was a pre-requisite for the quality of whole genome sequencing.

Sample collection and management

All those specimens were taken from nasopharyngeal (NP) and oropharyngeal (OP) swabs in one viral transport medium at 4-8°C during transportation to our laboratory. Each specimen was submitted...
to our laboratory, with informed consent from the patient and clinical data from the referred hospital. We conducted RNA extraction and Real-time PCR (RT-PCR) test on the same day after receiving, then kept the aliquot RNA at -80°C for whole genome sequencing at different times. Ethical clearance of this research was issued by the Ethics Committee of the Faculty of Medicine UIN Syarif Hidayatullah Jakarta (No. B- 005/F12/KPK/TL.00/02/2021).

RNA extraction and RT-PCR
The NP-OP swab in one viral transport medium (VTM) was extracted manually using Accu-RNA Biosensor SD (Biosensor SD, Republic of Korea). From 200μl VTM, we eluded 50μl RNA elution, and the RNA extraction procedures were conducted manually from the factory, thus aliquoted for RT-PCR and stored in a freezer -80°C for whole genome sequencing (WGS). We used a Nanodrop® spectrophotometer to assess the RNA purity before RT-PCR; the purity ratio was 2.0 (260/280 nm absorbance). We conducted multiplex real-time PCR targeted in genes Orf1ab and E (Standard M Biosensor SD, Republic of Korea). The PCR machine was Roche LC 480 with thermal cycling conditions per Standard M Biosensor’s manual protocol.

Whole genome sequencing
The extracted RNA was prepared for whole genome sequencing based on previous protocols published previously. Briefly, we synthesized cDNA from 50ng RNA from each SARS-CoV2 positive sample. A 100 ng of purified double-stranded cDNA was amplified, and post-amplification, the amplicons were pooled and purified using 1x AMPure beads (AMPure XP, Beckman Coulter, Cat.No. A63881). The purified sample (200ng) was prepared for library preparation, native barcode ligation and adapter ligation; thus, sequencing was done using ARTIC protocol that targeted whole genome sequencing. Libraries were pooled in multiplexes of 24 per flow cell and sequenced on the GridION platform (Oxford Nanopore Technologies, UK).

Phylogenetic analysis
All samples were assessed for read quality, continued to multiple sequence analysis for alignment build and checked the positional homology between the amino acids of each sequence. We constructed the phylogeny tree using MEGAX with maximum likelihood phylogenetic models.

RESULTS
We conducted whole genome sequencing for those cases, including primary and secondary infection. From the genome analysis, we observed that the secondary infection had more genome mutations and a wider range of amino acid substitutions in all cases. The detailed clinical and genomic insight can be seen in Table 1. Most of the SARS-CoV-2 in these cases were GH clade, and B.1.459 lineage, and all have amino acid substitution in D614G. However, as we analyzed case by case, we found a different clade and lineage. For instance, in the first case, we discovered

Table 1.  Clinical and genomic insight of COVID-19 reinfection cases of first and second encounter

<table>
<thead>
<tr>
<th>Variables</th>
<th>First infection</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>Male</td>
<td>Male</td>
<td>Male</td>
<td>Female</td>
</tr>
<tr>
<td>Age</td>
<td>31 y.o</td>
<td>55 y.o</td>
<td>36 y.o</td>
<td></td>
</tr>
<tr>
<td>Background</td>
<td>Healthcare worker</td>
<td>Lecturer</td>
<td>Healthcare worker</td>
<td></td>
</tr>
<tr>
<td>Main clinical symptoms</td>
<td>Fever, malaise,</td>
<td>Mild fever, nasal congestion</td>
<td>Asymptomatic (being tested as contact tracker)</td>
<td></td>
</tr>
<tr>
<td>Suspected transmission</td>
<td>Family cluster</td>
<td>Travelling to East Java</td>
<td>Positive colleague encounter</td>
<td></td>
</tr>
<tr>
<td>Radiology result</td>
<td>no main radiological finding</td>
<td>No data</td>
<td>No data</td>
<td></td>
</tr>
<tr>
<td>Hospitalization</td>
<td>Self-isolation</td>
<td>Self-isolation</td>
<td>Self-isolation</td>
<td></td>
</tr>
<tr>
<td>Real time PCR result</td>
<td>Positive date</td>
<td>September 19th, 2020</td>
<td>October 23th, 2020</td>
<td>October 27th, 2020</td>
</tr>
<tr>
<td></td>
<td>Negative date*</td>
<td>September 25th and 26th, 2020</td>
<td>October 27th and 28th, 2020</td>
<td>November 2nd and 3rd, 2020</td>
</tr>
<tr>
<td>GISAID Clade</td>
<td>G</td>
<td>GH</td>
<td>GH</td>
<td></td>
</tr>
<tr>
<td>Lineage</td>
<td>B.1.459</td>
<td>B.1.459</td>
<td>B.1.459</td>
<td></td>
</tr>
<tr>
<td>Months need for second infection</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variables</th>
<th>Second infection</th>
<th>First infection</th>
<th>Patient A</th>
<th>Patient B</th>
<th>Patient C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Clinical symptoms</td>
<td>Shortness of breath and dizziness</td>
<td>Severe pneumonia</td>
<td>high fever, shortness of breath and anosmia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Suspected transmission</td>
<td>Family cluster</td>
<td>Family cluster</td>
<td>Family cluster</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiology result</td>
<td>ground glass appearance in CT thorax</td>
<td>Bronchopneumonia in thorax photo</td>
<td>ground glass appearance in CT thorax hospitalization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hospitalization</td>
<td>hospitalization</td>
<td>hospitalization</td>
<td>hospitalization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real time PCR result</td>
<td>Positive date</td>
<td>November 17th, 2020</td>
<td>January 6th, 2021</td>
<td>January 12th, 2021</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Negative date</td>
<td>December 18th and 19th, 2020</td>
<td>Februari 8th, and 9th 2021</td>
<td>January 25th and 26th, 2021</td>
<td></td>
</tr>
<tr>
<td>GISAID Clade</td>
<td>GH</td>
<td>GH</td>
<td>GH</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lineage</td>
<td>B.1.459</td>
<td>B.1.470</td>
<td>B.1.466.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* A true negative PCR result should be validated in 24 hours in sequence.
and ancestor with the wider mutation in secondary infection (Figure 2). However, the virus that infected for the second time in patients B and C had different roots compared to the first infection, although the time interval of infection was not too long. The same figure shows that all the virus is close to Wuhan rather than another variant circulated as the cases found in the first wave of the pandemic. Based on the result in Figure 2, we can see that D614G amino acid mutation in Spike protein was found in all cases, as well as NSP12 P323L. However, the second infection had a wider amino acid mutation than the first infection.

**DISCUSSION**

All reinfection cases have a dominant D614G mutation in spike (S) protein, as this type is a predominated mutation worldwide in the first wave of the pandemic. Zhang L et al. conducted a study that concluded the D614G mutation can increase viral infectivity. The substitution from S protein D614 to G614 makes the virus more stable, so it transmits efficiently. Therefore, this type of mutation is dominant in Indonesia and other countries at the beginning of the pandemic. Consistent with our findings, we can find the mutation in all cases. The possibility of reinfection in SARS-CoV-2 may be because of the suboptimal antibody responses during the first infection. Many factors affect the protection condition of the antibodies post-natural infection. However, the different points of mutation of the second virus might have a major influence.

Furthermore, our genome sequence showed a wider mutation and genomic variation in the secondary infection, mostly in non-structural (NS) and S proteins. As we know, S protein mediates receptor recognition and binding during infection, playing an important part in viral transmission. However, it is still unclear to what extent immune responses against a previous variant will protect against reinfection by another variant. A certain mutation on the receptor-binding domain, such as D614G, confers resistance to commonly produced antibodies during SARS-CoV-2 infection in vitro. Overall, these reinfections are
symptoms were worsened in secondary infection. Some studies have demonstrated that severe clinical manifestation will be associated with higher neutralizing antibody responses. This could be why the patients got re-infected, as the first infection was in mild manifestation.  
These reinfections are potentially a consequence of divergent mutations enhancing virus transmission and virulence from first to second infection. Even in the same lineage, the mutated genome of the viruses probably has different immunogenic responses and an adequate virulence expression in the body that may also contribute to vulnerable conditions to a second infection.

Another possibility contributing to reinfection is the polymorphism in the immune system, genetic, or another risk factor from the host. Some people probably have the chance to get reinfection, and some are not. During this pandemic, the second subject already got another reinfection (the third infection was not sequenced) with mild pneumonia. However, he didn't have underlying diseases such as diabetes mellitus, hypertension or autoimmune.

Overall, the reinfection case of COVID-19 still needs more studies and evidence to understand, as COVID-19 itself is a rapidly evolving disease. However, it is still necessary to perform WGS in cases of reinfection in subjects who have received the first vaccination to booster to compare their genomic patterns.

CONCLUSION
Our data conclude that reinfection could occur in COVID-19 with distinct genomic sequences. The two cases occurred in young individuals with worsened clinical symptoms of secondary infection. The possibility of reinfection in SARS-CoV-2 may be because of the suboptimal antibody responses during the first infection. Many factors affected the protection condition of the antibodies post-natural infection. However, the different points of mutation of the second virus might have a major influence. The infection of SARS-CoV-2 may produce an early humoral response. However, the information on protection duration and the appearance of neutralizing antibodies after infections to reduce reinfection is still limited. Previous study shows that within 17-19 days after onset, the virus-specific IgG can be detected, and virus-specific IgM can be detected in 94% of cases. Thus, the titer of IgG dan IgM is higher in severe clinical conditions. Moreover, serum concentration of specific IgA protection is decreased over time. As we can see in all first infection cases, the clinical symptoms were mild or asymptomatic; thus, the symptoms were worsened in secondary infection. Some studies have demonstrated that severe clinical manifestation will be associated with higher neutralizing antibody responses. This could be why the patients got re-infected, as the first infection was in mild manifestation.

CONFLICT OF INTEREST
The authors declare that they have no competing interests.

ETHICAL CONSIDERATION
The ethical clearance of this study was issued by the Ethical Committee of Bali Medical Journal 2023; 12(3): 3141-3145 | doi: 10.15562/bmj.v12i3.4713

Figure 2. Heatmap diagram of amino acid mutation in first and second infection of each case.
the Faculty of Medicine UIN Syarif Hidayatullah Jakarta with the reference number B-005/F12/KPK/TL.00/02/2021.

**FUNDING**

The authors are grateful to the Center of Research and publication of UIN Syarif Hidayatullah Jakarta for funding this study.

**AUTHORS’ CONTRIBUTIONS**

EAS did the conceptualization, design, data collection and analysis, and also had a major contribution in writing the manuscript. LAH, CA, had a contribution to the design, data analysis and writing the manuscript.

**ACKNOWLEDGMENTS**

The authors acknowledged the Molecular Division of COVID-19 Laboratory of UIN Syarif Hidayatullah Jakarta.

**REFERENCES**


