Journal of Taibah University Medical Sciences (2024) 19(1), 70–81

Taibah University

Journal of Taibah University Medical Sciences

www.sciencedirect.com

Original Article

SARS-CoV-2 genetic variation and bacterial communities of naso-oropharyngeal samples in middle-aged and elderly COVID-19 patients in West Java, Indonesia

Muhammad M. Adzdzakiy, M.Si. [a], Sutarno Sutarno, Ph.D. [a], Isnaini Z. Asyifa, M.Biomed [b], Alvira R. Sativa, M.Sc. [c], Ahmad R.A. Fiqri, S.Si. [b], Azzania Fibriani, Ph.D. [c], Ryan B. Ristandi, M.D. [d], Ratih A. Ningrum, Ph.D. [e], Syam B. Iryanto, M.Kom. [f], Anggia Prasetyoputri, Ph.D. [g], Anik B. Dharmayanthi, Ph.D. [h] and Sugiyono Saputra, Ph.D. [g,i,*]

[a] *Graduate School of Bioscience, Faculty of Mathematics and Natural Sciences, Universitas Sebelas Maret, Jl. Ir. Sutami 36A Surakarta, Central Java, Indonesia*
[b] *Master Program in Biomedical Science, Faculty of Medicine, Universitas Indonesia, Jl. Salemba Raya No.6, Jakarta, Indonesia*
[c] *School of Life Science and Technology, Institut Teknologi Bandung, Jl. Ganesa 10, Bandung, West Java, Indonesia*
[d] *West Java Health Laboratory, Jl. Sederhana No. 3-5, Pasteur, Sukajadi, Bandung, West Java, Indonesia*
[e] *Research Center for Genetic Engineering, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia*
[f] *Research Center for Computation, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia*
[g] *Research Center for Applied Microbiology, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia*
[h] *Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia*
[i] *Research Center for Applied Zoology, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia*

Received 24 February 2023; revised 21 July 2023; accepted 1 September 2023; Available online 13 September 2023

الملخص

أهداف البحث: تعكس عدد حالات كورونا في إندونيسيا خطورة الأمراض وانتشارها بسرعة. تم إجراء الرصد الجينومي لفيروس كورونا المستجد والتحقيق في المجتمعات البكتيرية للجهاز التنفسي العلوي في غرب جاوة كرد فعل للتهديد المتزايد حيث يمكن أن يؤثر تفاوت التوازن في الميكروبات في الجهاز التنفسي العلوي بشكل سلبي على حالة المرضى.

طرق البحث: استخدمنا منصة تسلسل النانوبور لتحليل 43 عينة من التغيرات الجينية لفيروس كورونا المستجد و11 عينة مختارة لتسلسل الجينات عبر الريبوزوم 16، تم جمعهم في الفترة من مايو إلى أغسطس 2021.

النتائج: خمسة أنواع فيروسية في السكان (أ.ي.23؛ أ.ي.24؛ أ.ي.26؛ أ.ي.42؛ ب.1.1.7) كانت تهيمن على وجود النوع أ.ي.23 (>82%). الجزء الأكثر تحورا في الجينوم الخاص بفيروس كورونا المستجد كان بروتين ال (أس) (>20%). لم يكن هناك ارتباط بين الأنواع الفيروسية لفيروس كورونا المستجد، وتكرار التحور، وملف المريض، وحالات انتشار كورونا بسرعة. لم يكن هناك ارتباط للوفرة النسبية للبكتيريا، والتنوع الألفا والبيتا، وتحليل الفروق في التوازن البيئي البكتيري مع ملف المريض وحالات الانتشار السريع. يظهر تحليل تسلسل الميتاجينوم ثمانية أنواع ذات وفرة متفاوتة في ملفات المرضى الفردية، بما في ذلك البكتيريا الزرقاء الزائفة والبكتيريا الهيموفيلوس بارا إنفلونزا.

الاستنتاجات: وفقا للبيانات، يظهر غرب جاوة سيادة ذات صلة بنوع أ.ي.23 (المتحور دلتا) مصحوب بالعدوى العالمية خلال تلك الفترة الزمنية، مما يدعم أهمية برنامج الرصد في تعقب تقدم المرض. قد يساهم العملية المتعددة العوامل

* Corresponding address: Research Center for Biosystematics and Evolution, National Research and Innovation Agency (BRIN), Jl. Raya Jakarta Bogor Km 46, Bogor, West Java, Indonesia.
E-mail: sugiyono.saputra@brin.go.id (S. Saputra)

Peer review under responsibility of Taibah University.

Production and hosting by Elsevier

1658-3612 © 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). https://doi.org/10.1016/j.jtumed.2023.09.001

في تقدم المرض الذي تسببت فيه البكتيريا في كل مريض، كما هو مبين في النتائج المتنوعة للمجتمعات البكتيرية.

الكلمات المفتاحية: تسلسل الريبوسوم 16؛ المجتمع البكتيري؛ انتشار كورونا بسرعة؛ تكنولوجيا أكسفورد نانوبور؛ متغيرات فيروس كورونا المستجد

Abstract

Objective: The number of COVID-19 cases in Indonesia reflects the disease severity and rapid dissemination. In response to the mounting threat, SARS-CoV-2 genomic surveillance and the investigation of naso-oropharyngeal bacterial communities in West Java were conducted, as dysbiosis of the upper respiratory tract microbiota might adversely affect the clinical condition of patients.

Methods: We utilized the Oxford Nanopore sequencing platform to analyze genetic variation of 43 samples of SARS-CoV-2 and 11 selected samples for 16S rRNA gene sequencing, using samples collected from May to August 2021.

Results: The prevalence of AY.23 (>82%) predominated among five virus lineages in the populations (AY.23, AY.24, AY.26, AY.42, B.1.1.7). The region in the SARS-CoV-2 genome found to have the highest number of mutations was the spike (S) protein (>20%). There was no association between SARS-CoV-2 lineages, mutation frequency, patient profile, and COVID-19 rapid spread-categorized cases. There was no association of bacterial relative abundance, alpha-beta diversity, and linear discriminant analysis effect size analysis with patient profile and rapid spread cases. MetagenomeSeq analysis showed eight differential abundance species in individual patient profiles, including *Pseudomonas aeruginosa* and *Haemophilus parainfluenzae*.

Conclusions: The data demonstrated relevant AY.23 dominance (the Delta variant) in West Java during that period supporting the importance of surveillance program in monitoring disease progression. The inconsistent results of the bacterial communities suggest that a complex multifactor process may contribute to the progression of bacterial-induced disease in each patient.

Keywords: 16S rRNA sequencing; Bacterial community; COVID-19 rapid spread; Oxford Nanopore Technologies; SARS-CoV-2 variants

© 2023 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

The Coronaviridae family has caused several outbreaks in recent years, including severe acute respiratory syndrome (SARS) (2002), Middle East respiratory syndrome (MERS) (2012), and recently, the COVID-19 pandemic that emerged from Wuhan, China, in December 2019.[1] The first reported SARS-CoV-2 infection in Indonesia was recorded in March 2020 and since then has rapidly spread in the community with more than six million total cases (10/09/2022), massively threatening public health, economic and social aspects.[2,3] Middle-aged and elderly patients are considered vulnerable COVID-19 groups in the community, so SARS-CoV-2 surveillance and characterization of the patient's naso-oropharyngeal bacterial community are critical to mitigate potential disease exacerbation in these vulnerable groups. Furthermore, the region of West Java accounts for 16% of national cases, making it the province with the second highest number of cumulative cases in the country.[4–6]

The pandemic impact was more threatening due to the high rates of mutation in the virus, yielding more lineage diversity. Mutations can induce selective pressure on the SARS-CoV-2 genome that could affect its pathological characteristics. A change in amino acid sequence can alter the virus transmissibility, replication efficiency, and ability to evade the immune system response. Reduced vaccine efficacy is reportedly caused by mutations in B.1.617.2 variants (Delta) compared to B.1.1.7 variants (Alpha) in patients who only received their first dose.[4,7,8] SARS-CoV-2 genomic surveillance provides various data, including its characteristics, mutation profile, and lineage diversity.[9] It is essential to monitor the currently circulating variants in the community to ensure the efficiency of developed diagnostic kits and vaccination programs.

Bacterial infection in COVID-19 patients is associated with increased risk of morbidity and mortality.[10] The incidences of microbiota coinfection and/or secondary infection in COVID-19 patients have been highly variable throughout studies, but bacterial infection has consistently been more frequently detected than fungus, archaebacteria, and other viral infections.[11,12] Moreover, studies have reported different microbiomes present in COVID-19 patients, those who have recovered from COVID-19, and the healthy control group.[13–16] Potentially pathogenic bacteria in the patients' upper respiratory tract bacterial community can be detected by fully sequencing the 16S ribosomal RNA (rRNA) gene employing the Oxford Nanopore Technologies (ONT) platform.[17]

The microbiota community study in COVID-19 patients in Indonesia remains understudied.[18] To the best of our knowledge, there is no published study of bacterial community analyses of middle-aged and elderly COVID-19 patients in West Java, Indonesia. SARS-COV-2 genetic variants and bacterial community analyses might be correlated with patients' clinical records and COVID-19 rapid spread cases, which will be elaborated upon if any association is found between those parameters.

Materials and Methods

Sample preparation

Naso-oropharyngeal clinical swabs from three health institutions (Dinas Kesehatan Kabupaten Bogor, Laboratorium Kesehatan Jawa Barat, Laboratorium Kesehatan Daerah Karawang) from the West Java region were collected in viral transport media (VTM) tubes from May to August 2021. A total of 43 samples for SARS-CoV-2 genetic variant

analyses and 11 samples for bacterial community analyses were selected based on the low cycle threshold (Ct) value (<25) and middle-aged elderly (≥40 years old) patient criteria. Written informed consent was obtained before using the samples. The Health Research Ethics Committee, University of Indonesia, and Cipto Mangunkusumo Hospital (HREC-FMUI/CMH; 20-10-1321_EXP) approved the study protocol for the donated samples for research purposes.

Nucleic acid isolation

Genomic SARS-CoV-2 RNA was isolated with the Viral Nucleic Acid Extraction Kit II (Geneaid Biotech Ltd., New Taipei City, Taiwan). Before PCR amplification, 43 VTM samples were treated with the LunaScript RT SuperMix Kit (New England Biolabs [NEB], Ipswich, MA, USA) to construct cDNA from isolated RNA. Bacterial DNA was isolated using the QIAamp DNA Mini Kit (Qiagen, Hilden, Germany). All of the procedures followed the instructions suggested by the manufacturers and were carried out in Biosafety Level-2 and Biosafety Level-3 Laboratories, Cibinong Science Center, National Research and Innovation Agency (BRIN).

PCR amplification

For SARS-CoV-2 genetic variant analyses, modified V3 and V4 primers from ARTIC network nCoV-2019 sequencing protocols[19] were used to amplify the SARS-CoV-2 genome. A modification in V3 protocol was applied by adding 74F and 74R primers (0.3 μL at a final concentration of 2.5 pmol in pool B).[6] Multiplex PCR was conducted with 30 s of heat activation at 98 °C; and 30 cycles of 15 s of denaturation at 98 °C and 5 min of combined annealing–elongation at 65 °C. The PCR product was run on a 1% agarose gel to qualitatively check the overlapping 400 bp amplicons covering ~30 Kbp SARS-CoV-2 genomes.

For bacterial community analyses, the primers for amplification of the V1–V9 bacterial 16S rRNA gene were 27F 5'-AGAGTTTGATCCTGGCTCAG-3' and 1492R 5'-GGTTACCTTGTTACGACTT-3'. PCR was conducted using the following protocols: 5 min pre-denaturation at 95 °C; 35 cycles of 1 min denaturation at 94 °C, 1 min annealing at 55 °C, 1 min elongation at 72 °C; and 5 min post-elongation at 72 °C. The 1500 bp-sized amplicons were qualitatively checked on a 1% agarose gel.

Library preparation and sequencing

SARS-CoV-2 variant analyses

PCR products were purified with AMPure XP beads (Beckman Coulter, Brea, CA, USA) in a 1:1 ratio. The cleaned amplicons were subsequently quantified using the Qubit 2.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA). The end prep reaction was briefly subjected to the NEBNext Ultra II End Prep Kit (NEB), and the native barcoding was processed using the NB Expansion Kit 1-24 (ONT, Oxford, UK). In a single run, the samples that would be sequenced were 23 barcoded samples and one negative control. The 24 barcoded mixtures were then pooled and purified using AMPure XP beads and quantified with Qubit. The AMII Adapter Mix (ONT) was ligated to the barcoded mixtures with T4 DNA ligase (NEB) before another cleanup with AMPure XP beads and Qubit quantification. The final library was briefly put on an ice rack before use, with only 20 ng of the final library loaded into the R9.4.1 flow cell on the MinION Mk1b/Mk1c or PromethION sequencer machine (ONT). The sequencing process stopped when the sequenced coverage achieved >99% with the reference genome (Wuhan-Hu-1 MN908947.3).

Bacterial community analyses

Library preparation of bacterial community samples was carried out following the 16S Barcoding Kit 1-24 SQK-16S024 (ONT) protocol. The concentration of PCR products was quantified using the Qubit Fluorometer. PCR amplification was performed as the following protocol: 1 min of pre-denaturation at 95 °C; 25 cycles of 20 s of denaturation at 95 °C, 30 s of annealing at 55 °C, 2 min of elongation at 65 °C; and 5 min of final elongation at 65 °C. The crude amplicons were cleaned with AMPure XP beads and washed with 70% ethanol. Before being pooled, the samples were quantified with Qubit. The final library was kept cold before being loaded into the R9.4.1 flow cell on the MinION Mk1b/Mk1c or PromethION sequencer (ONT).

Data analyses

Sequencing was carried out using MinKNOW (v20.06.4; ONT) with a fast base-calling option and was monitored with RAMPART (v2.1.0). The sequence consensus was generated by ARTIC (v.1.1.0), and the base-calling process was performed using Guppy (v4.0.14; ONT). The fastq reads were aligned to the reference sequence, and the primer sequences were trimmed with minimap2 (v2.10-r761) or Geneious Prime (v2022.1.1). Medaka (v1.0.3) workflows and bcftools (v.1.10.2) were used to classify the variant, polish the sequence, and build a consensus sequence.

For SARS-CoV-2 genetic variant analyses, the SARS-CoV-2 lineages were identified using PANGOLIN (v3.1.16), and pangoLEARN (v1.2.105). MAFFT[20] was used to align the sequences. IQ-TREE with UFBoot and SH-aLRT 1000 replications and ModelFinder[21–24] were employed to construct a phylogenetic tree (maximum likelihood [ML]; TIM+F evolutionary model).

For bacterial community analyses, 16S sequence data were analyzed using EPI2ME (v3.4.2; ONT). All data were statistically tested and visualized using MicrobiomeAnalyst web service.[25–27] Alpha diversity analyses measuring the diversity of the intragroup was performed with the Shannon Index.[28] The statistical significance of alpha diversity was determined using the Mann–Whitney and Kruskal–Wallis H tests. Beta diversity analysis was performed to estimate the degree of diversity in the intergroup. The beta diversity matrix distance was calculated with Bray–Curtis measurement and was visualized by

performing principal coordinates analysis (PCoA). Its significance and effect size were measured using permutational multivariate analysis of variance (PERMANOVA).[28,29] Linear discriminant analysis effect size (LEFSe) and metagenomeSeq algorithms were employed to identify differentially abundant bacteria between middle-aged versus elderly patients, male versus female patients, and COVID-19 rapid spread versus nonrapid spread cases.

Results

SARS-CoV-2 genetic variation

The AY.23 lineage dominated COVID-19 cases in all patient groups (Figure 1). A total of 86% and 83% of middle-aged and elderly patients, respectively, were infected with the AY.23 lineage, and so were 86% of female and male patients. Other lineages were also observed, including AY.24, AY.26, and AY.42 with a total of <18% of middle-aged, elderly, female, and male patients.

The region having the highest number of mutations in the virus genome was the spike (S) protein with 20% cumulated amino acid substitution frequency (Figure 2). Other regions with relatively high mutation frequency were the nonstructural protein (NSP) 3 (NSP3) (18%), NSP12 (9%), N protein (7%), and NSP15 (5%).

Phylogenetic reconstruction of 43 SARS-CoV-2 sequenced was performed by ML based on the TIM + F parameter model (Figure 3). A SARS-CoV-2 genome (NC_045512.2) reference sequence labeled as EPI_ISL_402124 was used as an outgroup to define the phylogenetic tree root. AY.24 and AY.23 have an equal distance from the referenced Wuhan sequence. AY.26 and AY.42

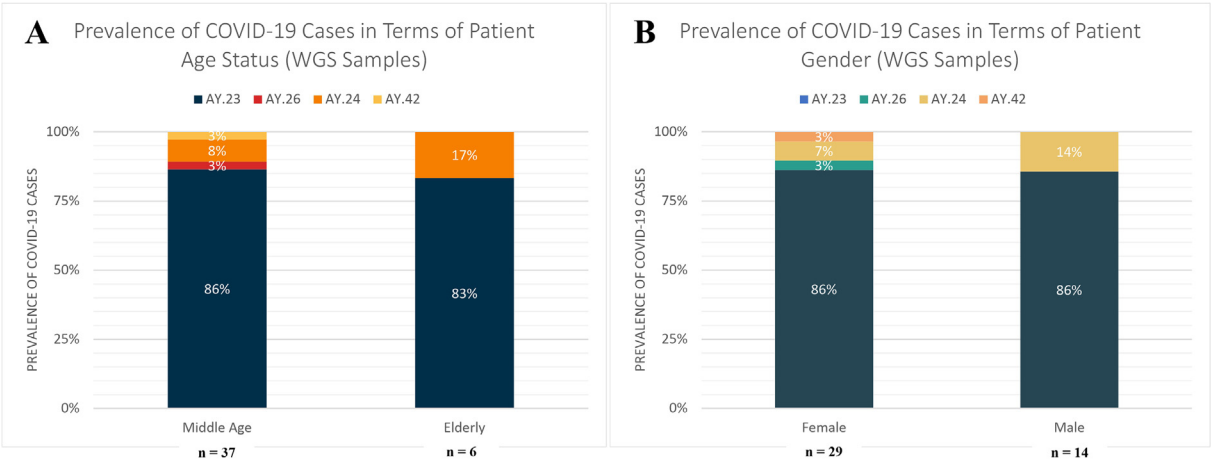

Figure 1: Visualization of SARS-CoV-2 lineage composition in the middle-aged versus elderly patient group (A) and female versus male patient group (B). The AY.23 lineage dominated all patient groups with a greater than 83% prevalence compared with other lineages, including AY.26, AY.24, and AY.42. A higher proportion of lineage AY.24 was observed in elderly and male patients compared to middle-aged and female patients.

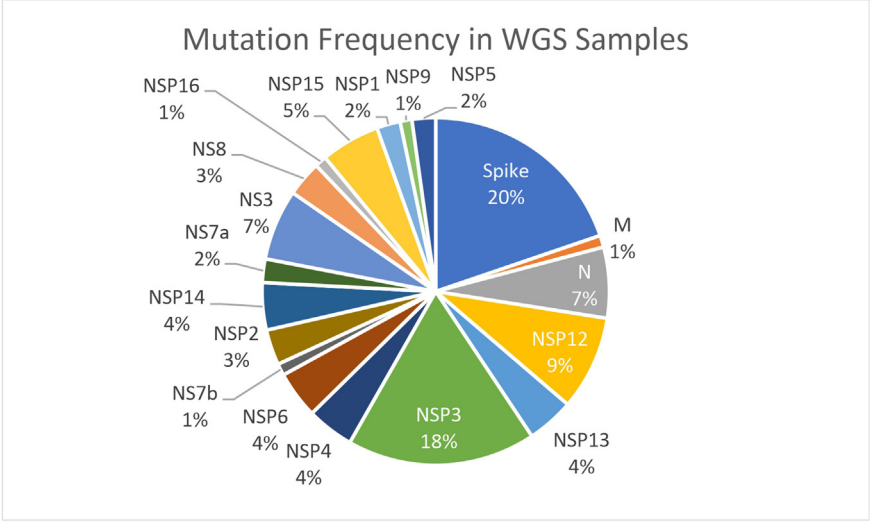

Figure 2: Visualization of detected amino acid substitution frequency in the SARS-CoV-2 genomes. Only the top 19 substitutions with the highest frequency are shown. The highest mutation occurrence was spotted in the S protein region with 20% frequency.

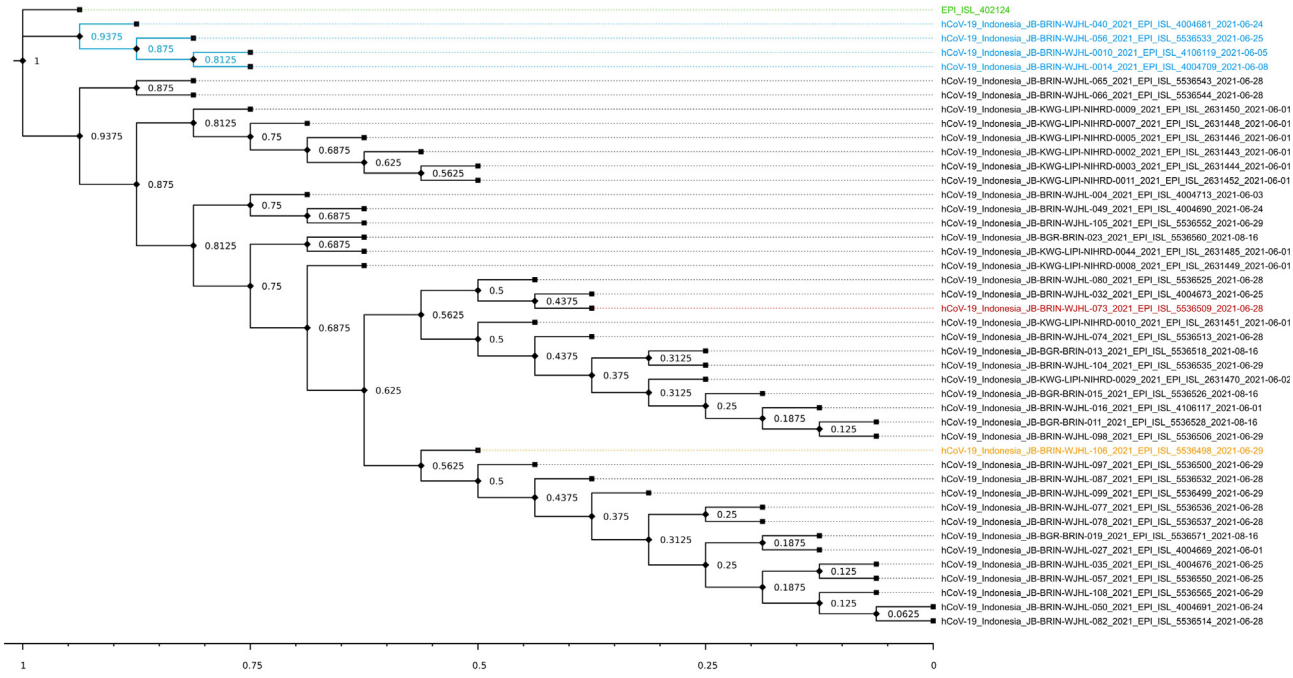

Figure 3: Phylogenetic tree of SARS-CoV-2 identified lineages from 43 sample pools representing four lineages (PANGO lineage). The grouping of certain lineages was based on the extent of variations in the genomes. The detected SARS-CoV-2 lineage was AY.23 (written in black), which was found to be more closely related to AY.26 (red) and AY.42 (orange) than to AY.24 (blue) lineage. The used outgroup was the Wuhan referenced sequence (NC_045512.2). This tree was built using MAFFT and IQ-TREE with 1000 UFBoot and SH-aLRT replications.

lineages were more closely related to the AY.23 lineage than AY.24 and the reference sequence (Figure 3).

Bacterial community profile

Bacterial relative abundances were explored in 11 selected samples belonging to middle-aged and elderly patient groups, female and male patient groups, and COVID-19 rapid spread cases and nonrapid spread cases. We found a total of 1459 operational taxonomic units (OTUs) assigned to five phyla, 12 classes, 26 orders, 38 families, 72 genera, and 233 species (Figure 4). Collectively, Firmicutes (48%) were the most abundant phyla found, followed by Bacteroidetes (28%), Actinobacteria (18%),

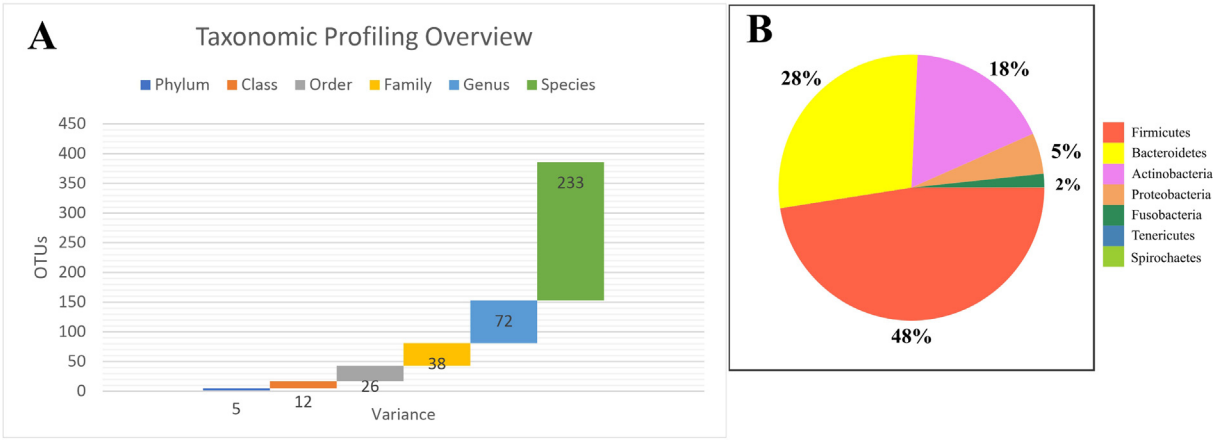

Figure 4: The taxonomic profile of the naso-oropharyngeal bacterial community from middle-aged and elderly COVID-19 patient groups. (A) The taxonomic profile overview of various OTU hierarchies showed a total of 1459 OTUs that were classified into five phyla, 12 classes, 26 orders, 38 families, 72 genera, and 233 species. (B) The abundance of the identified phyla in the sample population was dominated by Firmicutes (48%), Bacteroidetes (28%), and Actinobacteria (18%). The abundance of other phyla was only detected less than 10%.

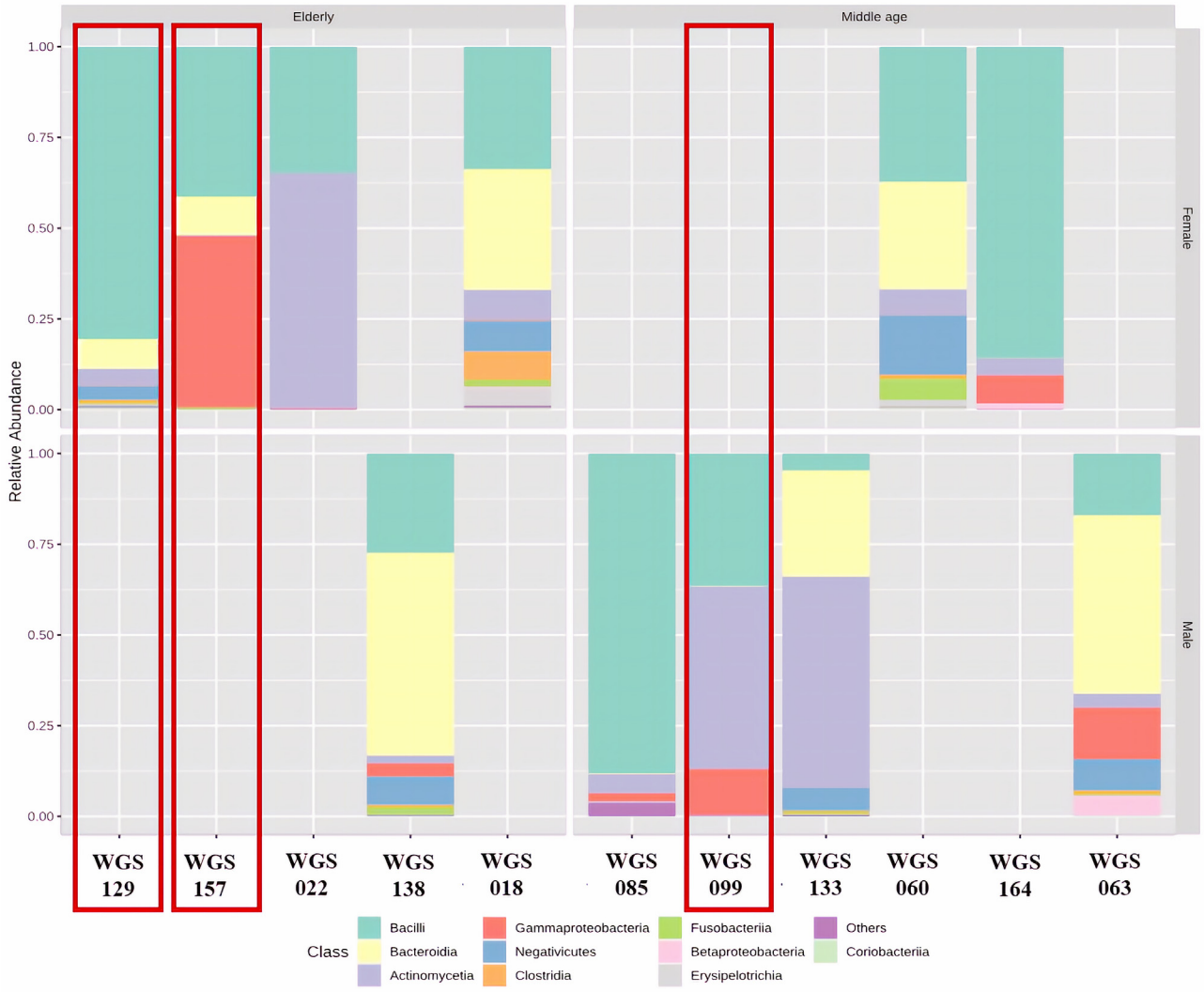

Figure 5: The relative abundance percentage comparison of bacterial community classes based on the patients' age status and sex. Only the top 10 classes are shown. WGS 129, WGS 157, and WGS 099 are marked with red squares to represent the COVID-19 nonrapid spread cases.

and Proteobacteria (5%) (Figure 4). The Bacilli class dominated all patient groups including middle-aged, elderly, female, and male patients (Figure 5). There was no associated pattern between bacterial abundance and the rapid spread-labeled cases as shown in Figure 5.

Alpha diversity results were not statistically significant for all age-based patient groups (middle-aged and elderly) and all sex-based groups (female and male), with p-values of 0.93 and 0.54, respectively. Shannon indices for middle-aged patients (2.88) were slightly higher than elderly patients (2.76), and female patients had lower indices (2.61) compared to the male patient (3.05), although this difference was not significant (Figure 6). The COVID-19 rapid spread cases were not different to nonrapid spread cases.

Beta diversity results showed an identical degree of dissimilarity (51.9%) in the comparison of middle-aged versus elderly patients and female versus male patients (Figure 7). However, none of these age- and sex-based patient groups had real meaning as the statistical tests were insignificant with [PERMANOVA] $F = 0.64$ and 0.92, $R2 = 0.07$ and 0.09, $p < 0.77$ and < 0.46, respectively, for each group. The bacterial diversity in elderly patients was found to be higher than that found in middle-aged patients. Similarly, the bacterial diversity in male patients was also higher than that in female patients, as indicated by the wider covered areas of the ellipsoid pattern (Figure 7). The nonrapid spread cases were found to have no specific plotted features associated with certain patient groups as shown in Figure 7.

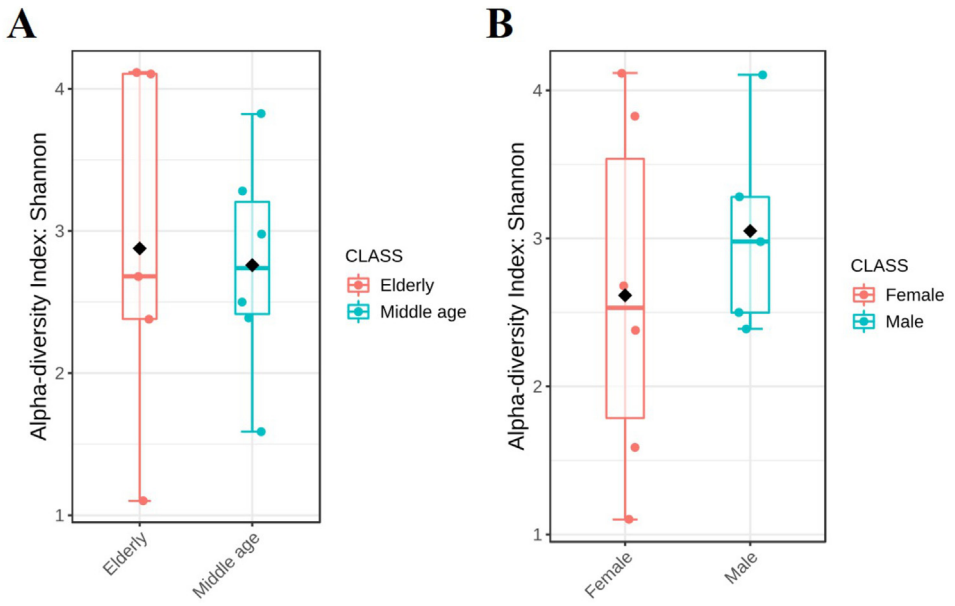

Figure 6: Alpha diversity utilizing Shannon indices results were statistically nonsignificant in all groups. (A) Elderly patients had average Shannon indices of 2.88, which was not significantly different from middle-aged patients at 2.76 ($p = 0.93$). (B) Female patients had lower Shannon indices of 2.61 compared to male patients at 3.05, which was also not significantly different ($p = 0.54$).

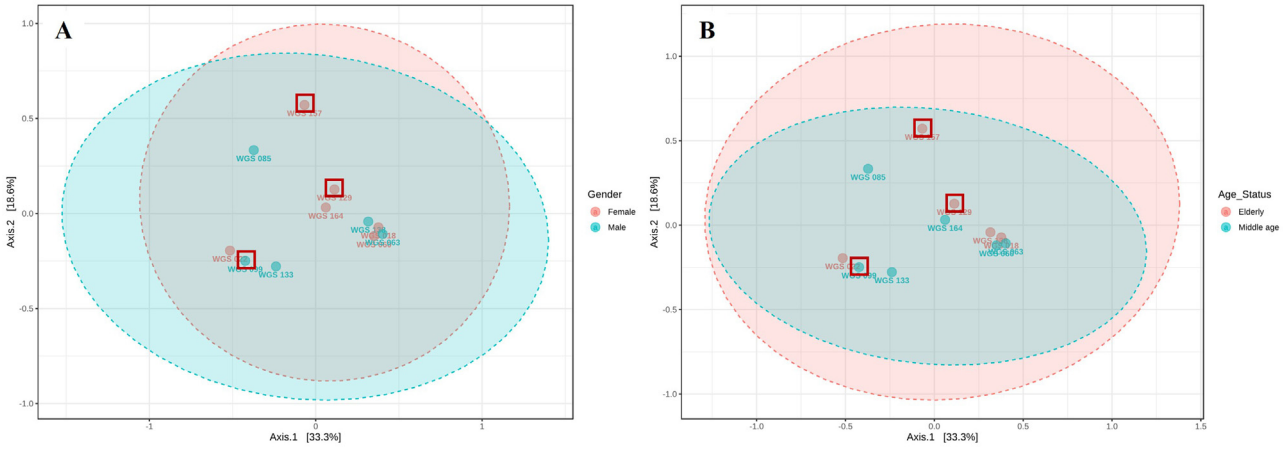

Figure 7: Beta diversity was visualized by PCoA utilizing PERMANOVA with Bray–Curtis distance results being statistically insignificant in all sample groups. (A) The comparison of middle-aged versus elderly patients showed a degree of dissimilarity of 51.9% ([PERMANOVA] $F = 0.64$, $R2 = 0.07$, $p < 0.77$). (B) The comparison of female versus male patients showed an identical result of 51.9% ([PERMANOVA] $F = 0.92$, $R2 = 0.07$, $p < 0.77$). The nonrapid spread cases of WGS 129, WGS 157, and WGS 099 are marked by red squares.

Table 1: MetagenomeSeq analysis results of eight bacteria species based on FDR score (<0.05).

Bacteria Species	p-values	FDR	Associated Patient Group
Limosilactobacillus fermentum	1.47E-06	0.000364	Elderly
Ligilactobacillus salivarius	3.79E-05	0.004695	Elderly
Pseudomonas aeruginosa	0.00025611	0.018879	Elderly
Cutibacterium acnes	0.00030449	0.018879	Middle aged
Corynebacterium propinquum	0.000209	0.035799	Female
Staphylococcus caprae	0.000289	0.035799	Male
Haemophilus parainfluenzae ATCC 33392	0.00057	0.035919	Female
Prevotella loescheii	0.000579	0.035919	Male

Note: p-value = probability value, FDR = false discovery rate.

The differentially abundant species were not found to be significant when LEFSe analysis was employed. Interestingly however, eight bacteria species were associated with certain groups based on metagenomeSeq analysis (Table 1). The eight differentially abundant species were *Cutibacterium acnes* in middle-aged patients; *Limosilactobacillus fermentum*, *Pseudomonas aeruginosa*, and *Ligilactobacillus salivarius* in elderly patients; *Corynebacterium propinquum* and *Haemophilus parainfluenzae* ATCC 33392 in female patients; and *Staphylococcus caprae* and *Prevotella loescheii* in male patients.

Discussion

Genomic surveillance of SARS-CoV-2 clinical samples has been essential to monitor virus evolution, enabling analysis of the association between SARS-CoV-2 genetic variation and the patients' bacterial metagenomic data.[4,6] This study, which characterized the SARS-CoV-2 genetic variation and bacterial community from COVID-19 patients with Nanopore sequencing, has three main findings. First, we observed the similarity between the detected lineages of circulating viruses in West Java and the national and global infection waves. Delta variant (AY.23 lineage) was found to dominate Indonesia's second wave. Second, the parameters being interrogated, including bacterial prevalence, relative abundance, alpha diversity, and beta diversity, were not associated with certain patient profiles. Third, *L. fermentum, L. salivarius*, and *P. aeruginosa* were found to be associated with elderly patients; *C. acnes* with middle-aged patients; *C. propinquum* and *H. parainfluenzae* ATCC 33392 with female patients; and *S. caprae* and *P. loescheii* with male patients.

A shift in SARS-CoV-2 variant domination during Indonesia's first and second waves of COVID-19 was previously reported.[6] Our study also confirmed that Indonesian lineages were the most prevalent during the first pandemic wave, whereas the Delta variants were dominant during the second wave. High prevalence of the AY.23 lineage, as one of the defined Delta variants, was observed in all patient groups (Figure 1), indicating that the transmission rate of this virus lineage was high and did not discriminate against any patient's age and sex. All Delta variant lineages, including AY.23, AY.24, AY.26, and AY.42, have increased transmissibility and induced poorer outcomes.[7] The mortality risk of Indonesian COVID-19 patients was associated with higher age, male sex, and pre-existing comorbidities.[29] A meta-analysis study[30] found that elderly patients (≥75 years old) have a higher morbidity and mortality risk than middle-aged patients (60–74 years old) in China, but these risks are not associated with the sex of the patient. Other studies also concluded that SARS-CoV-2 infection did not correlate with the patient's sex.[6,9,18] Although it should be noted that all of the aforementioned studies looked at SARS-CoV-2 infections in general, without considering any particular lineages .

In this study, lineage AY.23 was found to be the dominant circulating lineage in West Java (Figure 3) and is therefore relevant to the global infection cases.[3,9,10] First documented in India in September 2020 before being detected in Indonesia in January 2021, the Delta variants have since been dominating COVID cases globally.[3] Other than its high transmissibility, the high prevalence of the Delta variant in the second wave of COVID-19 in Indonesia was contributed by sociodemographic factors (i.e., the homecoming tradition before celebrating Eid Al-Fitr). Additionally, many recreational sites were still operating with health and safety measures despite the enforced restrictions on community activities (PPKM) in Indonesia,[9] which may have contributed to the increased rate of COVID-19 infections during May to August 2021.

Different lineages are defined based on the mutation set detected on the SARS-CoV-2 genomes. The highest frequency of amino acid substitutions was found in the S protein region (Figure 2).[3,31,32] Other regions with relatively high mutation frequency were NSP3, NSP12, N protein, and NSP15 (Figure 2). These findings are in accordance with the top globally detected mutation-containing regions in the GISAID database and other studies.[3,6,32] Mutations on the RNA-dependent RNA polymerase (RdRp) region, including NSP3 and NSP12, might affect the virus replication and immunogenic evasive response.[18]

Among all substitutions detected, the D614G substitution was found in 100% samples. This particular substitution was frequently reported in the surge of Delta variants domination and correlated with the increase in SARS-CoV-2 transmissibility and viral titer load in the respiratory tract.[18,33] Another two mutations of interest were detected in nearly all samples, namely L452R (100%) and P681R (98%). L452R was reported to be a key mutation that increases the stability of the S protein, promotes viral replication, and indicates decreased binding to monoclonal antibodies and subsequently may impact their neutralization potential.[32,34,35] Similarly, P681R substitution along with L452R, N501Y, and P681H were found to correlate with an increase in virus transmissibility and virus fusogenicity to S protein, pathogenicity, and antibody neutralization.[32,35–37]

The phylogenetic tree (Figure 3) revealed that the AY.23 lineage genomes were more closely related to the AY.26 and AY.42 lineages than to the AY.24 lineage. It is important to note that the lineage assignment by PANGOLEARN (v1.10.2) in the PANGOLIN system (v3.1.30) is dynamic with an average recall value for designated lineages of 95.8%, so the detailed lineage classification remains subject to change. A sequence may be assigned to a different lineage later in the future if there is a redefinition of one lineage and/or when the lineage is declared inactive/nonrelevant in the current spatiotemporal host population.[38–40]

In this study, the relative abundances of bacteria at the species level were considered low (Figure 4A), particularly when compared to another study that found 919 species in COVID-19 patients, 675 species in recovered patients, and 1476 species in the healthy control group.[16] Differences in bacterial abundance in those three groups suggests a dysbiosis progression in COVID-19 patients. Another study found that the diversity of microbiome in COVID-19 patients, recovered patients, and healthy control groups did not differ at the phyla level, which were dominated by Firmicutes, Bacteroidetes, and Proteobacteria (Figure 4B).[41] On the other hand, some studies have found varying dominance

of certain phyla. Either Actinobacteria being the most dominant and followed by Firmicutes, and Proteobacteria,[18] or Proteobacteria was found to be the major phyla followed by Firmicutes, and Actinobacteria.[42] Nevertheless, all of those studies are in agreement with this study, which revealed no association between certain phyla or classes with patient profiles (Figures 4B and 5).

This study found that the alpha diversities of the samples were insignificant (Figure 6) with any patient profile, which is similar to the findings of a previous study.[15] Rapid and nonrapid spread-labeled cases were also not associated with certain patterns or data. A more prevalent bacterial infection in elderly patients (≥60 years old) was previously shown, eventhough it was not elaborated if the case was also associated with a higher diversity of infecting bacteria.[43] Alpha diversity of microbiota in COVID-19 patients was lowest compared to the recovered patients and healthy people.[16]

The beta diversity measured in this study was found insignificant to patient profiles as found in similar study,[16] albeit that same study reported significantly diverse microbiota among COVID-19 patients, recovered patients, and healthy groups. Rapid spread cases were shown to not differ from nonrapid spread cases, as indicated by the randomly distributed sample plot of those mentioned cases (Figure 7). The microbiome of COVID-19 patients was found to be more diverse than that of patients with upper respiratory tract infection and chronic obstructive pulmonary disease.[41]

LEFSe analysis is widely used to determine the differential abundance of bacteria species in select observed groups.[15,28] However, a few studies have noted the weakness of this method, which is relatively insensitive, has low discriminative power, and tends to detect false-positive results.[44,45] The use of multiapproach in differential abundance analysis is essential as there is no gold-standard method agreed upon on this topic, and each published method has its weakness and superiority.[45] We managed this problem by applying a false discovery rate-corrected p-value cutoff of 0.05 in LEFSe analysis and using a comparative alternate method with metagenomeSeq. MetagenomeSeq is supposed to detect false-positive results in high replication data (>10),[46] which is opposed to the mentioned consistency of this method in other studies.[44,45]

Interestingly, there were differences in outcomes between the two employed approaches. While LEFSe did not find any single differential abundant species, metagenomeSeq found eight different abundant species associated with certain patient groups. This outcome could be impacted by the different characteristics of statistical tests used in the respective approach (LEFSe vs. metagenomeSeq).[45]

L. fermentum, P. aeruginosa, and *L. salivarius* were associated with elderly patients, while *C. acnes* was associated with middle-aged patients. The findings of *L. fermentum* and *L. salivarius* are interesting because these two species are generally classified as probiotic bacteria in the human body and not normally discovered as commensal bacteria in the upper respiratory tract.[47,48] Meanwhile, *P. aeruginosa* is an opportunistic pathogen that is classified as one of the ESKAPE pathogens group because it has multidrug resistance[12] and was also reported to be discovered in low prevalence (9%) in other studies[11,49] *C. acnes* is a normal microbiota on human skin, oral cavity, digestive tract, and the urogenital tract. However, this bacterium is also an opportunistic pathogen capable of invasively causing pneumonia, infection in the respiratory tract, and pleuropulmonary in rare cases.[50,51] Despite this different niche between the four bacteria, all of them share a similarity in that some of their strains have a potential antibiotic resistance,[12,48,50,52] so their existence in the upper respiratory tract should be given more attention.

C. propinquum and *Haemophilus influenzae* ATCC 33392 were associated with female patients, while *S. caprae* and *P. loescheii* were associated with male patients. *C. propinquum* is a commensal bacterium that is also reported as differentially abundant bacteria in COVID-19 patients.[41] *H. parainfluenzae* and *S. caprae* are classified as commensal bacteria but have also been reported to be opportunistic pathogens. *Haemophilus* spp. have multiple drug-resistance genes and have been associated with an increased prevalence of acute respiratory infection.[53,54] Additionally, *P. loescheii* has been associated with impacting the hematological indices of COVID-19 patients.[15]

All of these differently abundant bacteria might contribute to the disease progression. The microbial community could affect the clinical manifestation, morbidity, and mortality risks of COVID-19 patients with the increasing probability of pathogen coinfection, secondary infection, and or infection resulting in opportunistic pathogen infections while being in a complex interaction with the human immune system.[11,13,55,56] In contrast with our study, a few studies did not find real differences in bacterial communities in COVID-19 patients versus non-COVID-19 patients with respect to the age and sex of the patient.[18,57] The different bacterial communities observed in our study versus a few other studies might be affected by complex multifactors, including the patient's age and sex, geographical factor, medical treatment, pollutant exposure, and genotype.[12,13,18]

Our study presented a domination of Delta variant and naso-oropharyngeal bacterial communities in all patient groups, albeit no association was found between the above-mentioned factors. The relatively insufficient clinical data and small sampling region were a few limitations of this study. The collected data may not represent the entire West Java region due to additional inclusion criteria for whole genome sequencing. This study was focused on the West Java region because of its high prevalence of COVID-19 infection cases at that time.[6] Nevertheless, the relatively small clinical data were accompanied by in-depth analyses of the existing literature and metagenomic analyses. Our metagenomic analyses of the naso-oropharyngeal bacterial community in middle-aged and elderly COVID-19 patients in West Java is the first published work in this time frame.

Future studies are needed by collecting comprehensive clinical data, including asymptomatic cases; designing more stringent surveillance in longitudinal studies to observe the dynamics of the bacterial community in the patients; and adding another analysis parameter by analyzing the gut–lung axis bacterial community in COVID-19 patients. These considerations are important for better genomic and metagenomic surveillance in Indonesia, specifically in the West Java region. Nevertheless, we successfully obtained data on

SARS-CoV-2 genetic variations and the bacterial community of the COVID-19 patients of certain groups. Thus, this study adds to the body of literature on the much-needed information on SARS-CoV-2 evolution, the bacterial community profiles in those patients, and also the association among the virus, the bacterial community, and patients' profile in West Java, Indonesia.

Conclusions

The AY.23 lineage (Delta variant) dominated the COVID-19 cases in West Java in the second wave of the pandemic in Indonesia (May to August 2021). The highly mutated genome region in SARS-CoV-2 was the S protein region. We noticed a wide variety of naso-oropharyngeal bacterial communities in all of the patients, indicating a unique microenvironment in the patients' naso-oropharyngeal tract. However, there was no association between SARS-CoV-2 genetic variation and bacterial community in COVID-19 patients with the COVID-19 rapid spread cases on clinical samples from middle-aged and elderly COVID-19 patients in West Java.

Source of funding

This work was supported by the Riset Produktif dan Inovatif (RISPRO) by Educational Fund Management Institution (LPDP) Ministry of Finance of the Republic of Indonesia for COVID-19 Funding Program (07/FI/P-19.2B3/X/2020) under the project name "Surveilans Genom SARS-CoV-2" (VENOMCoV) LIPI/BRIN. The funders had no role in the study design, data collection and analysis, publication decisions, or manuscript preparation.

Conflict of interest

The authors have no conflict of interest to declare.

Ethical approval

This work was approved by the Komisi Etik Penelitian Kesehatan from FKUI/RSCM (20-10-1321_EXP).

Authors contributions

MMA was responsible for choosing the topic, conducting the literature search, extracting and analyzing the data, interpreting results, updating reference lists, writing the manuscript, and reviewing the manuscript. S was responsible for reviewing the manuscript. IZA was responsible for extracting and analyzing the data. ARS was responsible for extracting and analyzing the data. ARAF was responsible for extracting and analyzing the data. AF was responsible for extracting and analyzing the data and reviewing the manuscript. ABD was responsible for designing the study, extracting and analyzing the data, reviewing the data, interpreting results, and reviewing the manuscript. RAN was responsible for providing research materials and funding acquisition. AP was responsible for designing the study, funding acquisition, extracting and analyzing the data, reviewing the data, and reviewing the manuscript. SBI was responsible for designing the data analysis, analyzing, and reviewing the data. RBR was responsible for extracting and analyzing the data and reviewing the manuscript. SS was responsible for designing the study, funding acquisition, providing research materials, extracting and analyzing the data, reviewing the data, interpreting results, and reviewing the manuscript. All authors have critically reviewed and approved the final draft and are responsible for the content and similarity index of the manuscript.

Acknowledgment

MMA, IZA, ARS, and ARAF were the recipients of the LIPI Fellowship Program 2021 for the VenomCoV Project. Our greatest gratitude goes to all members of the West Java Health Laboratory for collecting samples and the LIPI COVID-19 Diagnostic Team and VENOMCoV Project team for assisting with the research materials.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jtumed.2023.09.001.

References

1. Sayed AM, Khattab AR, AboulMagd AM, Hassan HM, Rateb ME, Zaid H, et al. Nature as a treasure trove of potential anti-SARS-CoV drug leads: a structural/mechanistic rationale. **RSC Adv 2020**; 10: 19790–19802.
2. UNICEF UNDP, Prospera SMERU. *Analysis of the social and economic impacts of COVID-19 on households and strategic policy recommendations for Indonesia*. Jakarta: UNICEF; 2021.
3. GISAID. *Phylodynamics of pandemic coronavirus in Indonesia*; 2022 https://www.gisaid.org/phylodynamics/indonesia/. [Accessed 10 September 2022].
4. Funk CD, Laferriere C, Ardakani A. A snapshot of the global race for vaccines targeting SARS-CoV-2 and the COVID-19 pandemic. **Front Pharmacol 2020**; 11: 937. https://doi.org/10.3389/fphar.2020.00937.
5. Soebandrio A, Kusumaningrum T, Yudhaputri FA, Oktavianthi S, Malik SG, Myint KSA. Characteristics of children with confirmed SARS-CoV-2 infection in Indonesia. **J Clin Virol Plus 2021**; 1:100027. https://doi.org/10.1016/j.jcvp.2021.100027.
6. Prasetyoputri A, Dharmayanthi AB, Iryanto SB, Andriani A, Nuryana I, Wardiana A, et al. The dynamics of circulating SARS-CoV-2 lineages in Bogor and surrounding areas reflect variant shifting during the first and second waves of COVID-19 in Indonesia. **PeerJ 2022**; 10:e13132. https://doi.org/10.7717/peerj.13132.
7. CDC. *SARS-CoV-2 variant classifications and definitions*; 2021 https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Consequence. [Accessed 31 July 2021].
8. Moore JP, Offit PA. SARS-CoV-2 vaccines and the growing threat of viral variants. **JAMA 2021**; 325(9): 821–822. https://doi.org/10.1001/jama.2021.1114.
9. Fibriani A, Stephanie R, Alfiantie AA, Siregar ALF, Pradani GAP, Yamahoki N, et al. Analysis of SARS-CoV-2 genomes from West Java, Indonesia. **Viruses 2021**; 13: 2097. https://doi.org/10.3390/v13102097.
10. Prasetyoputri A. Detection of bacterial coinfection in COVID-19 patients is a missing piece of the puzzle in the COVID-19 management in Indonesia. **ACS Infect Dis 2021**; 7: 203–205. https://doi.org/10.1021/acsinfecdis.1c00006.

11. Langford BJ, So M, Raybardhan S, Leung V, Westwood D, MacFadden DR, et al. Bacterial coinfection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis. **Clin Microbiol Infect 2020**; 26: 1622–1629. https://doi.org/10.1016/j.cmi.2020.07.016.
12. Hoque MN, Akter S, Mishu ID, Islam MR, Rahman MS, Akhter M, et al. Microbial coinfections in COVID-19: associated microbiota and underlying mechanisms of pathogenesis. **Microb Pathog 2021**; 156:104941. https://doi.org/10.1016/j.micpath.2021.104941.
13. Ren L, Wang Y, Zhang J, Li X, Xiao Y, Li J, et al. Dynamics of the upper respiratory tract microbiota and its association with mortality in COVID-19. **Am J Respir Crit Care Med 2021**; 204(12): 1379–1390. https://doi.org/10.1164/rccm.202103-0814OC.
14. Mozota M, Castro I, Gomez-Torres N, Arroyo R, Lailla Y, Somada M, et al. Administration of *Ligilactobacillus salivarius* MP101 in an elderly nursing home during the COVID-19 pandemic: immunological and nutritional impact. **Foods 2021**; 10(2149): 1–12. https://doi.org/10.3390/foods10092149.
15. Liu J, Liu S, Zhang Z, Lee X, Wu W, Huang Z, et al. Association between the nasopharyngeal microbiome and metabolome in patients with COVID-19. **Synth Syst Biotechnol 2021**; 6: 135–143. https://doi.org/10.1016/j.synbio.2021.06.002.
16. Hoque MN, Sarkar MMH, Rahman MS, Akter S, Banu TA, Goswami B, et al. SARS-CoV-2 infection reduces human nasopharyngeal commensal microbiome with inclusion of pathobionts. **Sci Rep 2021**; 11:24042. https://doi.org/10.1038/s41598-021-03245-4.
17. Zhu X, Yan S, Yuan F, Wan S. The applications of nanopore sequencing technology in pathogenic microorganism detection. **Hindawi Can J Infect Dis Med Microbiol 2020**; 6675206: 1–8. https://doi.org/10.1155/2020/6675206.
18. Rhoades NS, Pinski AN, Monsibais AN, Jankeel A, Doratt BM, Cinco IR, et al. Acute SARS-CoV-2 infection is associated with an increased abundance of bacterial pathogens, including Pseudomonas aeruginosa in the nose. **Cell Rep 2021**; 36:109637. https://doi.org/10.1016/j.celrep.2021.109637. 1–13.
19. Quick J. *nCoV-2019 sequencing protocol v3 (LoCost) V.3*; 2020 https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye. [Accessed 11 January 2022].
20. Katoh K, Rozewicki J, Yamada KD. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief **Bioinform 2019**; 20(4): 1160–1166. https://doi.org/10.1093/bib/bbx108.
21. Nguyen LT, Schmidt HA, Haeseler AV, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. **Mol Biol Evol 2015**; 32: 268–274. https://doi.org/10.1093/molbev/msu300.
22. Trifinopoulos J, Nguyen LT, von Haeseler A, Minh BQ, W-IQ-TREE. A fast online phylogenetic tool for maximum likelihood analysis. **Nucleic Acids Res 2016**; 44(W1): W232–W235. https://doi.org/10.1093/nar/gkw256.
23. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. **Mol Biol Evol 2017**; 35: 518–522. https://doi.org/10.1093/molbv/msx281.
24. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. **Nat Methods 2017**; 14: 587–589. https://doi.org/10.1038/nmeth.4285.
25. Dhariwal A, Chong J, Habib S, King IL, Agellon LB, Xia J. MicrobiomeAnalyst: a web-based tool for comprehensive statistical, visual and meta-analysis of microbiome data. **Nucleic Acids Res 2017**; 45: W181–W188. https://doi.org/10.1093/nar/gkx295.
26. Chong J, Liu P, Zhou G, Xia J. Using MicrobiomeAnalyst for comprehensive statistical, functional, and meta-analysis of microbiome data. **Nat Protoc 2020**; 15: 799–821. https://doi.org/10.1038/s41596-019-0264-1.
27. Lu Y, Zhou G, Ewald J, Pang Z, Shiri T, Xia J. MicrobiomeAnalyst 2.0: comprehensive statistical, functional and integrative analysis of microbiome data. **Nucleic Acids Res 2023**. https://doi.org/10.1093/nar/gkad407.
28. Yip S, Dehcheshmeh MM, Mclelland DJ, Boardman WSJ, Saputra S, Ebrahimie E, et al. Porphyromonas spp., *Fusobacterium* spp., and *Bacteroides* spp. dominate microbiota in the course of macropod progressive periodontal disease. **Sci Rep 2021**; 11:17775. https://doi.org/10.1038/s41598-021-97057-1.
29. Surendra H, Elyazar IRF, Djaafara BA, Ekawati LL, Saraswati K, Adrian V, et al. Clinical characteristics and mortality associated with COVID-19 in Jakarta, Indonesia: a hospital-based retrospective cohort study. **Lancet Reg Health West Pac 2021**; 9:100108. https://doi.org/10.1016/j.lanwpc.2021.100108.
30. Gao T, Shen Q, Guo W, He W, Li J, Zhang Y, et al. Clinical characteristics of elderly patients with COVID-19 in Hunan Province, China: a multicenter, retrospective study. **Gerontology 2020**; 66: 467–475. https://doi.org/10.1159/000508734.
31. Khare S, Gurry C, Freitas L, Schultz MB, Bach G, Diallo A, et al. GISAID's role in pandemic response. **China CDC Wkly 2021**; 3(49): 1049–1051. https://doi.org/10.46234/ccdcw2021.255.
32. Gangavarapu K, Latif AA, Mullen J, Alkuzweny M, Hufbauer E, Tsueng G, et al. Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. **medRxiv 2022**. https://doi.org/10.1101/2022.01.27.22269965.
33. Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking changes in SARS-CoV-2 spike: evidence that D614G increases infectivity of the COVID-19 virus. **Cell 2020**; 182L: 812–817. https://doi.org/10.1016/j.cell.2020.06.043.
34. Motozono C, Toyoda M, Zahradnik J, Saito A, Nasser H, Tan TS, et al. SARS-CoV-2 spike L452R variant evades cellular immunity and increases infectivity. **Cell Host Microbe 2021**; 29: 1124–1136. https://doi.org/10.1016/j.chom.2021.06.006.
35. Cherian S, Potdar V, Jadhav S, Yadav P, Gupta N, Das M, et al. SARS-COV-2 spike mutations, L425R, T478K, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. **Microorganisms 2021**; 9(1542): 1–11. https://doi.org/10.3390/microorganisms9071542.
36. Saito A, Irie T, Suzuki R, Maemura T, Nasser H, Uriu K, et al. Enhanced fusogenicity and pathogenicity of SARS-CoV-2 delta P681R mutation. **Nature 2021**; 602: 300–325. https://doi.org/10.1038/s41586-021-04266-9.
37. Liu Y, Liu J, Johnson BA, Xia H, Ku Z, Schindewolf C, et al. Delta spike P681R mutation enhances SARS-CoV-2 fitness over alpha variant. **bioRxiv 2021**. https://doi.org/10.1101/2021.08.12.456173.
38. Rambaut A, Holmes EC, O'Toole A, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. **Nat Microbiol 2020**. https://doi.org/10.1038/s41564-020-0770-5.
39. O'Toole A, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tools. **Virus Evol 2021**. https://doi.org/10.1093/ve/veab064.
40. O'Toole A, Hill V, Pybus OG, Watts A, Bogoch II , Khan K, et al. Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.352/501Y-V2 with grinch. **Wellcome Open Res 2021**. https://doi.org/10.12688/wellcomeopenres.16661.1.
41. Hoque MN, Rahman MS, Ahmed R, Hossain MS, Islam MS, Islam T, et al. Diversity and genomic determinants of the microbiomes associated with COVID-19 and non-COVID

respiratory diseases. **Gene Rep 2021**; 23:101200. https://doi.org/10.1016/j.genrep.2021.101200.
42. Saha OFT, Dubourg G, Yacouba A, Bossi V, Raoult D, Lagier JC. Profile of the nasopharyngeal microbiota affecting the clinical course in COVID-19 patients. **Front Microbiol 2022**; 13:871627. https://doi.org/10.3389/fmicb.2022.871627.
43. Liu Y, Mao B, Liang S, Yang JW, Lu HW, Chai YH, et al. Association between age and clinical characteristics and outcomes of COVID-19. **Eur Respir J 2020**; 55:2001112. https://doi.org/10.1183/13993003.01112-2020.
44. Paulson JN, Stine OC, Bravo HC, Pop M. Differential abundance analysis for microbial marker-gene surveys. **Nat Methods 2013**; 10: 1200−1202. https://doi.org/10.1038/nmeth.2658.
45. Nearing JT, Douglas GM, Hayes MG, MacDonald J, Desai DK, Allward N, et al. Microbiome differential abundance methods produce different results across 38 datasets. **Nat Commun 2022**; 13(342): 1−16. https://doi.org/10.1038/s41467-022-28034-z.
46. McMurdie PJ, Holmes S. Waste not, want not: why rarefying microbiome data is inadmissible. **PLoS Comput Biol 2014**; 10(4):e1003531. https://doi.org/10.1371/journal.pcbi.1003531.
47. Duar RM, Lin XB, Zheng J, Martino ME, Grenier T, Perez-Munoz ME, et al. Lifestyles in transition: evolution and history of the genus Lactobacillus. **FEMS Microbiol Rev 2017**; 41(1): S27−S248. femsre/fux030.
48. Du T, Lei A, Zhang N, Zhu C. The beneficial role of probiotic Lactobacillus in respiratory diseases. **Front Immunol 2022**; 13(908010): 1−14. https://doi.org/10.3389/fimmu.2022.908010.
49. Bassetti M, Kollef MH, Timsit JF. Bacterial and fungal superinfections in critically ill patients with COVID-19. **Intensive Care Med 2020**; 46: 2071−2074. https://doi.org/10.1007/s00134-020-06219-8.
50. Cobo F, Borrego J, Rodriguez-Granger J, Sampedro A, Navarro-Mari JM. A rare case of pleural infection due to *Propionibacterium acnes* (*Cutibacterium acnes*). **Rev Esp Quimioter 2017**; 31(2): 173−174.
51. Boman J, Nilson B, Sunnerhagen T, Rasmussen M. True infection or contamination in patients with positive Cutibacterium blood cultures—a retrospective cohort study. **Eur J Clin Microbiol Infect Dis 2022**; 41: 1029−1037. https://doi.org/10.1007/s10096-022-04458-9.
52. Klein G. Antibiotic resistance and molecular characterization of probiotic and clinical Lactobacillus strains in relation to safety aspects of probiotics. **Foodborne Pathog Dis 2011**; 8(2): 267−281. https://doi.org/10.1089/fpd.2010.0672.
53. Stefanini I, De Renzi G, Foddai E, Cordani E, Mognetti B. Profile of bacterial infections in COVID-19 patients: antimicrobial resistance in the time of SARS-CoV-2. **Biology 2021**; 10(822): 1−16. https://doi.org/10.3390/biology10090822.
54. Hurst JH, McCumber AW, Aquino JN, Rodriguez J, Heston SM, Lugo DJ, et al. Age-related changes in the nasopharyngeal microbiome are associated with severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection and symptoms among children, adolescents, and young adults. **Clin Infect Dis 2022**; 75(1): e928−e937. https://doi.org/10.1093/cid/ciac184.
55. Jain S, Batra H, Yadav P, Chand S. COVID-19 vaccines currently under preclinical and clinical studies, and associated antiviral immune response. **Vaccines 2020**; 8(649): 1−16. https://doi.org/10.3390/vaccines8040649.
56. Cohen CA, Li APY, Hachim A, Hui DSC, Kwan MYW, Tsang OTY, et al. SARS-CoV-2 specific T cell responses are lower in children and increase with age and time after infection. **Nat Commun 2021**; 12L: 4678. https://doi.org/10.1038/s41467-021-24938-4.
57. Saeed NK, Al-Khawaja S, Alsalman J, Almusawi S, Albalooshi NA, Al-Biltagi M. Bacterial coinfection in patients with SARS-CoV-2 in the Kingdom of Bahrain. **World J Virol 2021**; 10(4): 168−181. https://doi.org/10.5501/wjv.v10.i4.168.

How to cite this article: Adzdzakiy MM, Sutarno S, Asyifa IZ, Sativa AR, Fiqri ARA, Fibriani A, Ristandi RB, Ningrum RA, Iryanto SB, Prasetyoputri A, Dharmayanthi AB, Saputra S. SARS-CoV-2 genetic variation and bacterial communities of naso-oropharyngeal samples in middle-aged and elderly COVID-19 patients in West Java, Indonesia. J Taibah Univ Med Sc 2024;19(1):70−81.