CURRENT SCIENCE AND TECHNOLOGY (CST)
VOL. 2, ISSUE 2, 1 – 7
DOI: https://doi.org/10.15282/cst.v2i2.8859

ORIGINAL ARTICLE

On-Site Sequencing Of Sars-Cov-2 Genomes Identify Two Variant Of Concern Clusters In Pahang, Malaysia

Ummu Afeera Zainulabid[1], Wan Nurliyana Wan Ramli[1,*], Raihanah Haroon[2], Ayesha Aslam[3], Mushtaq Hussain[3], and Hajar Fauzan Ahmad[3*]

[1]Department of Internal Medicine, Kulliyyah of Medicine, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia.
[2]Department of Radiology, Kulliyyah of Medicine, Jalan Sultan Ahmad Shah, Bandar Indera Mahkota, 25200 Kuantan, Pahang, Malaysia.
[3]Bioinformatics and Molecular Medicine Research Group, Dow Research Institute of Biotechnology and Biomedical Sciences, Dow College of Biotechnology, Dow University of Health Sciences, Karachi, Pakistan.
[4]Faculty of Industrial Sciences and Technology, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Pahang, Malaysia.

ABSTRACT – Timely variant identification is required in respond to effective public health measures. It has been demonstrated that comprehending genetic epidemiology can be aided by a combination of public health expertise and through viral genomic sequencing using Oxford Nanopore Technologies (ONT). The purpose of this study is to determine the utility of using Oxford Nanopore Sequencing to elucidate the genetic epidemiology in two significant clusters in Pahang, Malaysia. The combined oropharyngeal and nasopharyngeal swabs of clinical specimens from two significant clusters in Pahang, Malaysia, were retrieved for long-read sequencing. Results: We identify Beta and Delta variants as 2 variants of concern from our analysis. We found that the B.1.351 (ß) and B.1.617 (Δ) variants were responsible for the Taman Tanah Putih Baru and Pasar Kemunting clusters, respectively. In conclusion, ONT long-read sequencing tools offer a practical solution with a number of advantages. ONT devices are compact, affordable, and need little technical hands-on or laboratory equipment to prepare samples. They can also be used to quickly and adaptably execute sequencing analyses and to understand genomic epidemiology.

ARTICLE HISTORY
Received: 09th Nov 2022
Revised: 21st Dec 2022
Accepted: 08th June 2023

KEYWORDS
SARS-CoV-2
Beta variant
Delta variant
Covid-19
Nanopore

INTRODUCTION

The coronavirus has been responsible for multiple worldwide outbreaks throughout human history, such as H1N1 in 2009 and the Middle East Respiratory Syndrome (MERS) caused by MERS-CoV in 2012. In the latter month of 2019, the WHO China Country Office received reports of pneumonia cases with unclear causes, ultimately linked to a novel coronavirus identified as SARS-CoV-2 that causes COVID-19 [1]. The World Health Organization has declared 627,573,579 confirmed cases of COVID-19, including 6,570,363 deaths as of October 2022 [2]. The pandemic is regarded as a worldwide crisis creating severe disruptions across the economy and health system, including Malaysia.

Malaysia was one of the first countries in Southeast Asia to implement the Movement Control Order (MCO) to prevent the spread of the coronavirus during the early outbreak [3]. Malaysia eventually faced a much more difficult task in combating the COVID-19 pandemic in its third wave, which began on 8th September, 2020, due to the Benteng LD cluster in Sabah [4, 5]. Since then, the highest lineage contributor appeared to be from Variant of Concerns (VOCs) and Variant of Interests (VOIs) with D614G mutations in the spike protein [6]. Two large clusters reported in Pahang, Malaysia, includes Taman Tanah Putih Baru and Pasar Kemunting cluster. Taman Tanah Putih Baru, Pahang, was previously under an Enhanced Movement Control Order (EMCO) by the National Security Council (NSC) of Malaysia on 5th May 2021 after 11 residents tested positive for COVID-19. The state government had to 'appeal' for free screening tests after 200 residents ran away from their houses before the EMCO. A total of 4795 surveillance test for COVID-19 was done, and a total of 385 individual was reported positive from this cluster. The Pasar Kemunting cluster, on the other hand, is a workplace cluster linked to a market. The cluster started with one case reported on 21st June 2021 in Kuantan District. The name of the cluster refers to the locality where the outbreak is suspected of having occurred at the workplace located at Jalan Seri Kemunting 2, Kemunting, Kuantan. The index case for this cluster is a Malaysian citizen working there. The index case started experiencing fever and sore throat symptoms on 14th June 2021. The results of the investigation found that 5 other colleagues tested positive for COVID-19. The transmission of the infection is suspected of having occurred in the workplace due to social gatherings, and there were violations of SOPs. As of 20th August 2021, a total of 5135 individuals had been screened in this cluster, 1123 cases were detected positive for COVID-19, and 16 deaths had been reported.

Herein, we report 2 variants of concern – ß and Δ variants in these 2 large clusters in Pahang, Malaysia. The representative sample of Taman Tanah Putih Baru cluster patient genome surveillance showed that he contracted with

*CORRESPONDING AUTHOR | HF Ahmad | ✉ fauzanahmad@ump.edu.my

B.1.351 lineage, which had been identified previously in October 2021 in Eastern Cape Province, South Africa. This cluster ended on 23rd May 2021. In addition, the representative sample of Pasar Kemunting cluster patient genome surveillance showed to be related to B.1.617 lineage, which was first detected in India in late 2020, and this cluster ended on 20th August 2021. Moreover, the utility of long-read sequencing devices from Oxford Nanopore Technologies (ONT) was tested, showing as a promising diagnostic tool in investigating COVID-19 for disease surveillance.

MATERIALS AND METHODS

Sample Collection, Patients Background and Ethical Approval

Clinical specimens were collected directly from combined oropharyngeal and nasopharyngeal swab specimens from two individuals with diverse clinical presentations representing both clusters. Both were COVID-19 positive and diagnosed *via* reverse transcriptase PCR (RT-PCR) with a threshold cycle (C_T) value below 30. The study was approved by the International Islamic University Malaysia Research Ethics Committee (IREC 2021-080).

DNA Extraction, Nanopore Library Preparation and Sequencing

The genomic RNA of SARS-CoV-*2* was isolated using Maxwell HT simplyRNA kit (Promega, USA) according to the manufacturer's recommended procedures with some modifications. The concentration and quality of gDNA were measured by Nanodrop spectrophotometer (Thermo Scientific, USA) and agarose gel electrophoresis, respectively, for quality and control measurement prior library preparation. The library preparation was performed using NEBNext® ARTIC SARS-CoV-2 Companion Kit (ONT®, UK) according to the manufacturer's instructions. Briefly, the extracted total RNA was converted into cDNA using SuperScript™ IV Reverse Transcriptase (Invitrogen, US) with some modifications, a hexamer annealing and extension step of 25ºC for 2 minutes was performed followed by cDNA synthesis at 42ºC for 50 minutes. The purified PCR products were quantified using DeNovix® dsDNA High Sensitivity Assay (DeNovix Inc., US) and The prepared CDNA library was then loaded onto a MINION flow cell and a 72 hr sequenced for 72 hours.

Nanopore Basecalling and Data Analysis

The generated *.fast5* files were basecalled on Guppy (ONT, ver. 4.4.1) using high accuracy mode followed by demultiplexing using Guppy barcoder. Basecalled reads were individually classified using Kraken2 (http://ccb.jhu.edu/software/kraken2/) and MiniKraken2_v2 library (ftp://ftp.ccb.jhu.edu/pub/data/kraken2_dbs/old/minikraken2_v2_8GB_201904.tgz). The Kraken2 reports were visualized on Krona Plots using KronaTools (https://github.com/marbl/Krona/wiki). The SARS-CoV-2 genomes were reconstructed from the raw reads using a combination of bioinformatic tool as listed in https://github.com/CDCgov/SARS-CoV-2_Sequencing/tree/master/protocols/BFX-UT_ARTIC_Illumina. Briefly, the raw reads were aligned to the reference strain WuHan-Hu-1 genome (GenBank accession number: MN908947) using bwa mem 0.7.17-r1188 and subsequently trimmed to remove primer binding region and a consensus genome was generated from the filtered alignment using iVar v1.2.

Whole-Genome Sequencing and Phylogenetic Tree Analysis

The SARS-CoV-2 genome was reconstructed from the raw reads using a combination of several bioinformatic tools enlisted in https://github.com/CDCgov/SARS-CoV-2_Sequencing/tree/master/protocols/BFX-UT_ARTIC_Illumina. Genome sequences from other studies related to humans and animal coronaviruses were mined from the GISAID (https://www.gisaid.org) and NCBI GenBank (https://www.ncbi.nlm.nih.gov/genbank/). The multiple sequence alignment was performed using Clustal X [7] and observed in BioEdit [8] and finalized using MEGA XI [9]. Evolutionary analysis was conducted in MEGA XI by reconstructing bootstrap consensus tree of sequences employing Neighbor-Joining (NJ) method with 1000 bootstrap replicates to represent the evolutionary history of the taxa analyzed.

Data Availability

The sequences deposited in the GISAID database are EPI_ISL_3221234 and EPI_ISL_2622089.

RESULT AND DISCUSSION

Case Presentation

Case 1. A 72-year-old man, an active smoker with underlying hypertension, presented with a fever history and loose stools for 4 days. His symptoms were also associated with a productive cough and loss of appetite. He denied shortness of breath, chest pain, vomiting, anosmia or ageusia. His chest radiograph showed small patchy peripheral consolidations in both lungs, predominantly at the lower zones (Figure 1). Initial blood investigation revealed a white cell count of 7.3 x 109/L, haemoglobin of 14.1 g/dL and platelet of 298 x 109/L.

His renal and liver profile was normal. His real-time PCR showed positive result and he was treated for COVID-19 Category 3.

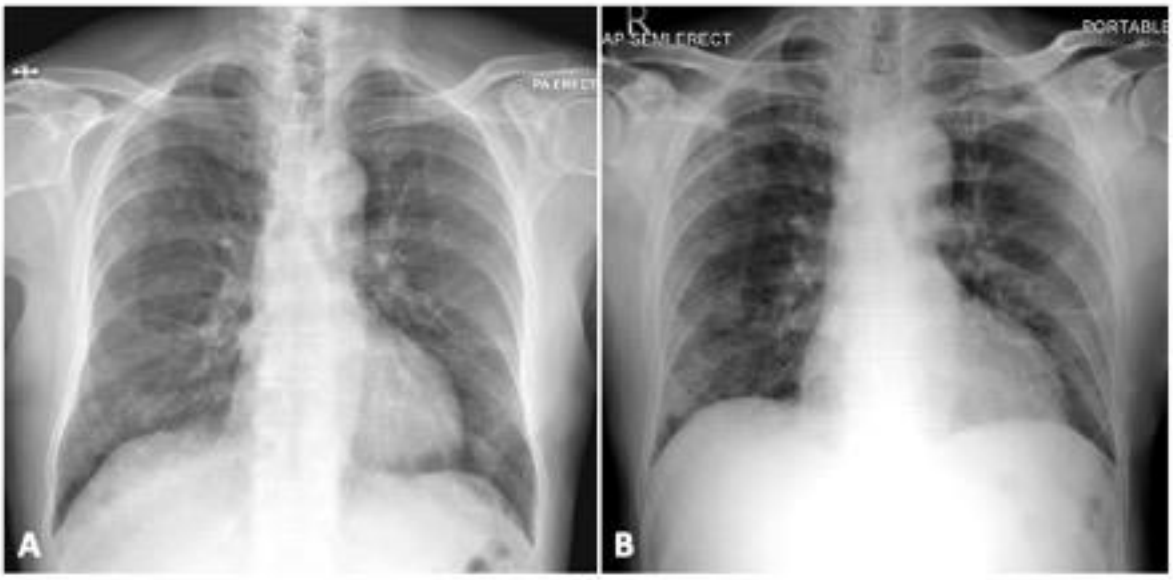

Figure 1: Two serial frontal chest radiographs. (A) - Acquired on the day of admission and (B) - Acquired 4 days afterwards, showing worsening bilateral, almost symmetrical subpleural air space opacities peripherally distributed.

On day 5 of hospitalization, he clinically deteriorated, and he became tachypneic and required oxygen supplementation. His clinical diagnosis was revised to COVID-19 Category 4 and he was started on IV Dexamethasone 6 mg daily. A repeated chest radiograph showed worsening bilateral peripheral lung consolidation during his clinical deterioration, particularly on the right side. Scattered ground glass opacities are also noted within both lungs. A CT-pulmonary angiogram was also ordered to look for the possibility of a pulmonary embolism. However, it was none and showed a typical feature of COVID-19 (Figure 2).

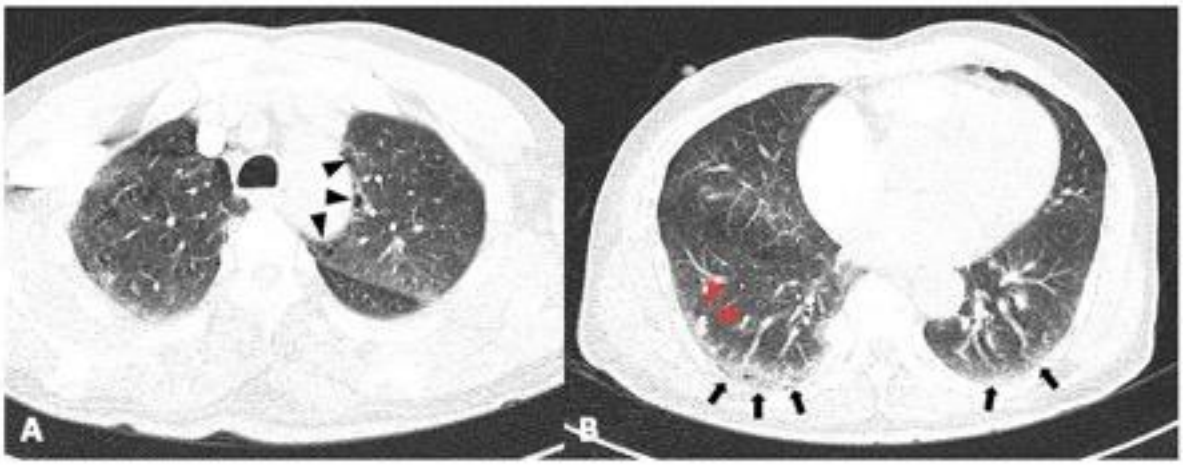

Figure 2: Serial axial computed tomography (CT) of the upper and lower thorax images in the lung window showing: (A) - Several pulmonary bullae (black arrowheads) at the medial apicoposterior segment of the left upper lobe. (B) - Peripherally distributed ground-glass densities at both posterobasal segments (black arrows), typically seen in COVID-19 pneumonia. Also, note few nodular densities (red arrowheads) at the right laterobasal segment.

He showed marked clinical improvement subsequently and was discharged on day 12 of admission. His genome surveillance showed he contracted the B.1.351 (beta) variant. He lived at Taman Tanah Putih Baru and visited Tanah Putih Mosque, where one of the attendees was positive for COVID-19.

Case 2. A 17-year-old man who worked at Pasar Kemunting market presented with headache and fever for 1 week associated with anosmia and loose. He was treated as COVID-19 Category 2 and quarantined at PKRC

Gambang for nearly 11 days. He was discharged well. No blood or radiological investigation was done. His genome surveillance showed that he contracted B.1.617 (delta variant).

SARS-CoV-2 Identification

Table 1. SARS-CoV-2 identification of genome surveillance of the clusters

	IIUM 6472	IIUM 9104
Virus name	hCoV-19/Malaysia/IIUM6472/2021	hCoV-19/Malaysia/IIUM9104/2021
Accession ID	EPI_ISL_2622089	EPI_ISL_3221324
Type	betacoronavirus	betacoronavirus
Clade	GH	GK
Pango Lineage	B.1.351 (Pango v.3.1.20 2022-02-28), Beta (B.1.351-like) (Scorpio)	AY.25.3 (Pango v.3.1.20 2022-02-28), Delta (B.1.617.2-like) (Scorpio)
AA Substitutions	Spike A243del, Spike A701V, Spike D80A, Spike D215G, Spike D614G, Spike E484K, Spike G1223C, Spike K417N, Spike L5F, Spike L18F, Spike L242del, Spike L244del, Spike N501Y, E P71L, M H155Y, N Q9H, N T205I, NS3 A72S, NS3 Q57H, NS3 S171L, NS8 I121L, NSP2 T85I, NSP3 K837N, NSP3 S1699F, NSP5 K90R, NSP5 K100R, NSP6 F108del, NSP6 G107del, NSP6 M164V, NSP6 S106del, NSP12 P323L, NSP13 T255S, NSP15 T48I	Spike D950N, Spike L452R, Spike P681R, Spike T478K, M I82T, N D63G, NS3 S26L, NS7b T40I, NSP3 A488S, NSP3 P1228L, NSP3 P1469S, NSP4 T492I, NSP4 V167L, NSP12 G671S, NSP12 P323L, NSP13 P77L, NSP14 A394V

Phylogenetic Tree Analysis

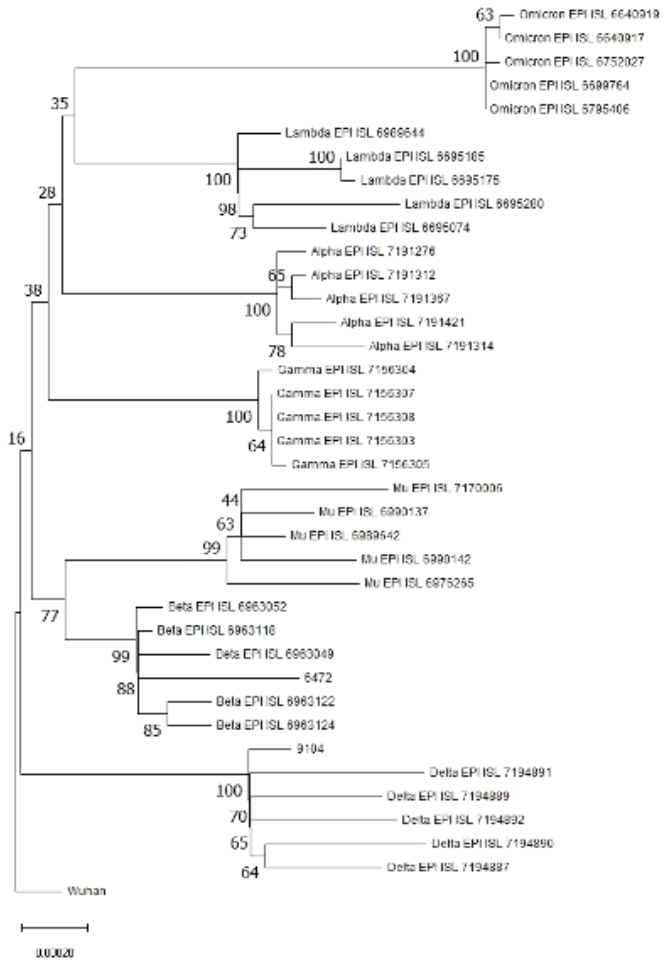

Figure 3: Phylogenetic tree of SARS CoV-2 VOCs and VOIs. Phylogenetic reconstruction of five whole genomes of each SARS CoV-2 VOC and VOIs as designated by WHO. Samples 9104 and 6472 lie in Delta and Beta Clade, respectively. The tree is reconstructed by the Neighbor-Joining method [10] with 1000 bootstrap replicates. Bootstrap values are indicated at nodes. Rectangular orientation are shown.

Phylogenetic Tree Analysis

The pandemic of COVID-19 is devastating, and Malaysia is currently affected by the 5th wave. Taman Tanah Putih Baru and Kemunting clusters are the two most significant clusters of COVID-19 reported in Pahang, Malaysia, during the 4th wave.

Reverse transcriptase PCR (RT-PCR) remains the gold standard for diagnosing COVID-19, despite its inability to distinguish variant types. On the other hand, it is anticipated that the combination of real-time whole-genome sequence analysis and epidemiology will aid in controlling SARS-CoV-2 transmission in the population and aiding early clinical decision-making. Taman Tanah Putih Baru and Kemunting clusters are the two most significant clusters of COVID-19 reported in Pahang, Malaysia. Using the local population of Pahang, Malaysia, as an example, we describe a substantial chain of transmission that began with an initial local infection with SARS-CoV-2 Beta and Delta variant.

The representative sample of Taman Tanah Putih Baru cluster patient genome surveillance showed that he contracted with B.1.351 (beta) variant, and the representative sample of Pasar Kemunting cluster patient genome surveillance showed that he contracted with B.1.617 (delta variant). Both variants had been classified as variants of concern (VOC). This mutation was reported previously and will not impact diagnostic testing [11], and our patient is shown to have a positive PCR test.

We highlighted this case to show that the potential transmission of this VOC is greater, and a large population was affected. It has been proven previously that this variant had caused a rapid spread of disease due to multiple spike mutations. It was predicted that B.1.351.5 resulted in significant selection pressure in a population with high seroprevalence and immunity [12]. Multiple data is still needed to ascertain the disease severity, and early reports show that there is no considerable difference in the severity of infection [13].

The D614G variation is a precursor that accounts for 90.30 percent of all COVID-19 infections in Malaysia, and this mutation is found in all new variants under development [6]. Because of positive natural selection, D614G improves the infectivity, viral fitness, transmission rate, and efficiency of cellular entry for the SARS-CoV-2 virus across a wide range of human cell types [5]. Despite this, the D614G mutation hasn't been proven to cause an increase in COVID-19 mortality or clinical severity, nor has it been shown to alter the efficacy of existing laboratory diagnoses, medicines, immunizations, or public health preventative efforts [14]. The Delta variation is thought to have a greater viral replication rate, causing viral loads in Delta infected individuals to be 1000 times higher than other strain infections on the day the testing is positive [15]. This indicates that the Delta variation is more infectious during the early stages of infection, and the frequency of population screening should be maximized [15]. Surveillance data suggest that the Delta variant swiftly overtook the previously successful Alpha (B.1.1.7) strain in other countries. The Delta variant has the fastest growth rate of all the variants found, reflecting its biological features and the setting in which it is transmitted [16]. The Delta variant has also been linked to a higher risk of hospitalization than the Alpha variant [17, 18], however, research suggests that the current COVID-19 vaccinations provide almost similar protection against the Delta variant after two doses as they do against the Alpha variant [19].

Because of the increased prevalence of the Beta and Delta variants globally, disease control strategies such as population testing, quarantine during the pre-symptomatic period, and genetic surveillance should be updated. Because of the higher infectivity rate of Delta variant infections in the pre-symptomatic phase, suspected individuals or close contacts should be quarantined immediately. Despite a single case, it can eventually become dominant in the population if the pandemic is not successfully contained.

Rapid identification of variants is required to implement public health actions promptly. It has been demonstrated that public health expertise and Oxford Nanopore Technologies (ONT) can aid in comprehending genomic epidemiology. Furthermore, combining short and long reads dramatically enhanced the assembly of the SARS-CoV-2 genome and established a novel method for correcting erroneous frame shifts from a single sequencing effort [20].

CONCLUSION

Oxford Nanopore Technologies (ONT) long-read sequencing devices offer a viable alternative with many advantages. ONT devices are tiny, affordable, and need minimal laboratory infrastructure and technical skills for sample preparation. In addition, they can be used to undertake flexible, rapid sequencing analyses that aid in the comprehension of genomic epidemiology.

ACKNOWLEDGEMENT

This research was funded by the Kurita Water and Environment Foundation Overseas Research Grant (and the SASMEC Research Grant (SRG21-040-0040). The funders played no role in the study's design, data collection and analysis, publication decision, or manuscript preparation.

REFERENCES

[1] F. Wu *et al.*, “A new coronavirus associated with human respiratory disease in China,” *Nat. 2020 5797798*, vol. 579, no. 7798, pp. 265–269, Feb. 2020.

[2] “WHO Coronavirus (COVID-19) Dashboard | WHO Coronavirus (COVID-19) Dashboard With Vaccination Data.” [Online]. Available: https://covid19.who.int/. [Accessed: 30-Jul-2021].

[3] H.-L. Ser, L. T.-H. Tan, J. W.-F. Law, V. Letchumanan, N.-S. A. Mutalib, and L.-H. Lee, “Genomic analysis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strains isolated in Malaysia,” *Prog. Microbes Mol. Biol.*, vol. 3, no. 1, Jun. 2020.

[4] L. Rampal and L. Boon Seng, “Malaysia’s third COVID-19 wave – a paradigm shift required.”

[5] U. A. Zainulabid *et al.*, "Whole genome sequence analysis showing unique SARS-CoV-2 lineages of B.1.524 and AU.2 in Malaysia," *PLoS One*, vol. 17, no. 2, p. e0263678, 2022.
[6] A. S. Mat Yassim, M. F. F. Asras, A. M. Gazali, M. S. Marcial-Coba, U. A. Zainulabid, and H. F. Ahmad, "COVID-19 outbreak in Malaysia: Decoding D614G mutation of SARS-CoV-2 virus isolated from an asymptomatic case in Pahang.," *Mater. today. Proc.*, vol. 48, pp. 828–836, 2022.
[7] "Clustal W and Clustal X Multiple Sequence Alignment." [Online]. Available: http://www.clustal.org/clustal2/. [Accessed: 29-Mar-2022].
[8] "BioEdit 7.2 Download (Free) - BioEdit.exe." [Online]. Available: https://bioedit.software.informer.com/7.2/. [Accessed: 29-Mar-2022].
[9] K. Tamura, G. Stecher, and S. Kumar, "MEGA11: Molecular Evolutionary Genetics Analysis Version 11," *Mol. Biol. Evol.*, vol. 38, no. 7, pp. 3022–3027, Jul. 2021.
[10] N. Saitou and M. Nei, "The neighbor-joining method: a new method for reconstructing phylogenetic trees.," *Mol. Biol. Evol.*, vol. 4, no. 4, pp. 406–425, Jul. 1987.
[11] L. Peñarrubia *et al.*, "Multiple assays in a real-time RT-PCR SARS-CoV-2 panel can mitigate the risk of loss of sensitivity by new genomic variants during the COVID-19 outbreak," *Int. J. Infect. Dis.*, vol. 97, pp. 225–229, Aug. 2020.
[12] H. Tegally *et al.*, "Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa," *medRxiv*, p. 2020.12.21.20248640, Jan. 2020.
[13] "Risk Assessment: Risk related to the spread of new SARS-CoV-2 variants of concern in the EU/EEA – first update." [Online]. Available: https://www.ecdc.europa.eu/en/publications-data/covid-19-risk-assessment-spread-new-variants-concern-eueea-first-update. [Accessed: 30-Jun-2021].
[14] U. A. Zainulabid *et al.*, "Near-Complete Genome Sequences of Nine SARS-CoV-2 Strains Harboring the D614G Mutation in Malaysia," *Microbiol. Resour. Announc.*, vol. 10, no. 31, Aug. 2021.
[15] "Viral infection and transmission in a large well-traced outbreak caused by the Delta SARS-CoV-2 variant - SARS-CoV-2 coronavirus / nCoV-2019 Genomic Epidemiology - Virological." [Online]. Available: https://virological.org/t/viral-infection-and-transmission-in-a-large-well-traced-outbreak-caused-by-the-delta-sars-cov-2-variant/724. [Accessed: 04-Dec-2021].
[16] "Investigation of SARS-CoV-2 variants of concern: technical briefings - GOV.UK." [Online]. Available: https://www.gov.uk/government/publications/investigation-of-novel-sars-cov-2-variant-variant-of-concern-20201201. [Accessed: 04-Dec-2021].
[17] A. Sheikh, J. McMenamin, B. Taylor, and C. Robertson, "SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness," *Lancet (London, England)*, vol. 397, no. 10293, pp. 2461–2462, Jun. 2021.
[18] H. Allen *et al.*, "Household transmission of COVID-19 cases associated with SARS-CoV-2 delta variant (B.1.617.2): national case-control study," *Lancet Reg. Heal. - Eur.*, p. 100252, Oct. 2021.
[19] J. Lopez Bernal *et al.*, "Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant," *N. Engl. J. Med.*, vol. 385, no. 7, pp. 585–594, Aug. 2021.
[20] U. A. Zainulabid *et al.*, "Hybrid de novo Whole-Genome Assembly and Annotation of SARS-CoV-2 Virus from Nosocomial Infection in Pahang, Malaysia," *Int. J. Infect. Dis.*, vol. 116, p. S19, Mar. 2022.