


SHORT REPORT

Open Access



# Whole-genome sequence analysis reveals the circulation of multiple SARS-CoV-2 variants of concern in Nairobi and neighboring counties, Kenya between March and July 2021

Samoel Ashimosi Khamadi<sup>1</sup>, Silvanos Opananda<sup>1</sup>, Samwel Lifumo Symekher<sup>1</sup>, Samson Konongoi Limbaso<sup>1</sup>, Solomon Langat<sup>1</sup>, Josyline Kaburi Cirindi<sup>3</sup>, Milkah Mwangi<sup>3</sup>, Nicholas Mwikwabe<sup>2</sup>, Seth Okeyo<sup>1</sup>, Edith Koskei<sup>1</sup>, James Mutisya<sup>1</sup>, Samwel Owaka<sup>1</sup>, Albert Nyunja<sup>1</sup>, Hellen Koka<sup>1</sup>, Meshack Wadegu<sup>1</sup>, Esther Chitech<sup>1</sup>, Rachel Achilla<sup>1</sup>, Janet Masitsa Majanja<sup>1</sup>, Lucy Kanyara<sup>1</sup>, Evans Amukoye<sup>4</sup> and Wallace Bulimo<sup>1\*</sup> 

## Abstract

The emergence and rapid spread of SARS-CoV-2 variants of concern (VOC) have been linked to new waves of COVID-19 epidemics occurring in different regions of the world. The VOC have acquired adaptive mutations that have enhanced virus transmissibility, increased virulence, and reduced response to neutralizing antibodies. Kenya has experienced six waves of COVID-19 epidemics. In this study, we analyzed 64 genome sequences of SARS-CoV-2 strains that circulated in Nairobi and neighboring counties, Kenya between March 2021 and July 2021. Viral RNA was extracted from RT-PCR confirmed COVID-19 cases, followed by sequencing using the ARTIC network protocol and Oxford Nanopore Technologies. Analysis of the sequence data was performed using different bioinformatics methods. Our analyses revealed that during the study period, three SARS-CoV-2 variants of concern (VOC) circulated in Nairobi and nearby counties in Kenya. The Alpha (B.1.1.7) lineage predominated (62.7%), followed by Delta (B.1.617.2, 35.8%) and Beta (B.1.351, 1.5%). Notably, the Alpha (B.1.1.7) VOC were most frequent from March 2021 to May 2021, while the Delta (B.1.617.2) dominated beginning June 2021 through July 2021. Sequence comparisons revealed that all the Kenyan viruses were genetically similar to those that circulated in other regions. Although the majority of Kenyan viruses clustered together in their respective phylogenetic lineages/clades, a significant number were interspersed among foreign strains. Between March and July 2021, our study's findings indicate the prevalence of multiple lineages of SARS-CoV-2 VOC in Nairobi and nearby counties in Kenya. The data suggest that the recent increase in SARS-CoV-2 infection, particularly in Nairobi and Kenya as a whole, is attributable to the introduction and community transmission of SARS-CoV-2 VOC among the populace. In conclusion, the findings provide a snapshot of the SARS-CoV-2 variants that circulated in Kenya during the study period.

**Keywords:** SARS-CoV-2, Coronaviruses, Lineage, Kenya

## Introduction

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) first emerged in Wuhan, China on 8th December 2019 [1] and rapidly spread across the world, causing coronavirus disease of 2019 (COVID-19) pandemic [1–3]. Until now, the SARS-CoV-2 continues to circulate in

\*Correspondence: Wbulimo@kemri.org

<sup>1</sup> Centre for Virus Research, Kenya Medical Research Institute (KEMRI), P.O. Box 54628-00200, Nairobi, Kenya  
Full list of author information is available at the end of the article



© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

many countries, causing multiple waves of COVID-19 epidemics [4]. Indeed, as of 21st July 2022, >564 million people have been infected worldwide, with >6 million deaths reported, with Kenya registering 336,904 cases and 5668 deaths [5]. Since its first appearance, the SARS-CoV-2 has undergone significant numbers of mutations over time, giving rise to multiple genetic variants, some of which have beneficial host adaptive fitness [4, 6–11]. The emerging SARS-CoV-2 variants have been described as variants of interest (VOI) and variants of concern (VOC) [10, 12]. Certainly, considerable attention has been focused on the latter as they harbor genetic attributes associated with increased transmissibility, increased virulence, and reduced effectiveness of public health countermeasures [4, 8–10].

Five major SARS-CoV-2 lineages, Alpha (B.1.1.7), first reported in the United Kingdom; Beta (B.1.351), first reported in South Africa; Delta (B.1.617.2) first seen in India; Gamma (P.1) first seen in Japan/Brazil; and Omicron (B.1.1.529) first reported in South Africa have been identified as variants of concern [4, 8–10, 13–16]. They possess point mutations in the receptor-binding domain (RBD) of the spike protein, which enhances binding affinity to the human angiotensin-converting enzyme 2 (ACE2) receptor. The Alpha variants contain asparagine to tyrosine mutation (N501Y) in the spike protein RBD; Beta (N501Y, K417N, & E484K); Delta (L452R) while the Gamma P1 variant has acquired 17 mutations of which K417N, E484K, and N501Y are contained in the spike protein [9, 10, 17, 18]. The Omicron variant is heavily mutated, with thirty amino acid variations within the surface spike protein, mostly affecting the receptor-binding domain (RBD) [19]. However, some of the concerning mutations including the N501Y, D614G, K417N, and T478K are present among the previous variants of concern [14].

After the emergence and rapid transmission, accompanied by devastating disease outcomes in the source countries, the SARS-CoV-2 VOC have spread to other regions, causing multiple waves of COVID-19 outbreaks [16, 18, 20]. Kenya in particular has experienced multiple waves of SARS-CoV-2 infection since the detection of the first case in the country [21]. The first peak in SARS-CoV-2 infection occurred in August 2020 and was attributed to the relaxation of public health restriction measures while the second and third waves occurred in November 2020 and March 2021 respectively, seemingly due to the emergence of genetic variants of concern [21]. However, after a short lull and beginning in June 2021, a surge in SARS-CoV-2 infection was observed in Kenya [21], especially in the western city of Kisumu and the capital city of Nairobi, yet genomic characteristics of these devastating variants have not been determined.

Whole-genome sequence analysis is recognized for its importance in providing insights into virus transmission patterns and evolutionary dynamics [22]. However, aside from the study by Githinji et al. [23] describing the importation and local transmissions of SARS-CoV-2 strains, full genome studies involving SARS-CoV-2 variants circulating in Kenya remain scanty. The current study sought to characterize the full genome sequences of SARS-CoV-2 that circulated in Nairobi and neighboring counties in Kenya between March and July 2021, during the third COVID-19 wave in the country, to describe the variants and classify them into VOC lineages. Because Nairobi is Kenya's travel hub, the data provided here provide a glimpse of the introductions and temporal dynamics of SARS-CoV-2 VOC circulation in Kenya.

## Methods

### Sampling

Respiratory samples were collected from persons who presented for voluntary COVID-19 testing at the Kenya Medical Research Institute (KEMRI). The testing was carried out as part of the Kenya Ministry of Health's (MoH) national response to control COVID-19. Persons presenting for a COVID-19 test filled out a questionnaire capturing their demographic and residence information.

### RNA extraction, genome amplification, and sequencing

Genomic surveillance for the SARS-CoV-2 was conducted between March 2021 and July 2021 to investigate the presence and spread of the SARS-CoV-2 variants of concern (VOC) in Nairobi and neighboring counties, Kenya. Samples received at KEMRI's Sample Management and Repository Facility (SMRF) were tested for COVID-19 using the real-time RT-PCR assay. Viral RNA was extracted from nasopharyngeal/oropharyngeal swab specimens using the QIAamp Viral RNA Mini Kit (QIAGEN, Germany) following the manufacturer's protocol. SARS-CoV-2 real-time RT-PCR assay was performed using different commercial kits viz. DaAn Gene Detection kit for 2019-nCoV (DaAnGene, China), RADI COVID-19 Detection kit (KHMedical, South Korea) and SD Biosensor Covid 19 RT PCR Test Kit (SD BIOSENSOR Republic of Korea), according to the World Health Organization's (WHO) guidelines [24]. Selected samples that tested positive with threshold (Ct) values < 30, for the E & RdRp or S & RdRp or ORF1ab & N target genes were selected and RT-PCR and SARS-CoV-2 genome sequencing performed following ARTIC's network protocol and Oxford Nanopore Technologies, as previously described [25, 26]. Viral genome assembly and variant assignment were carried out based on ARTIC Network bioinformatics protocol [27, 28]. Mutational analysis was performed

using various web tools including CoVsurver [29], CoV-GLUE [30], and Nexclade [31].

**Phylogenetic analysis**

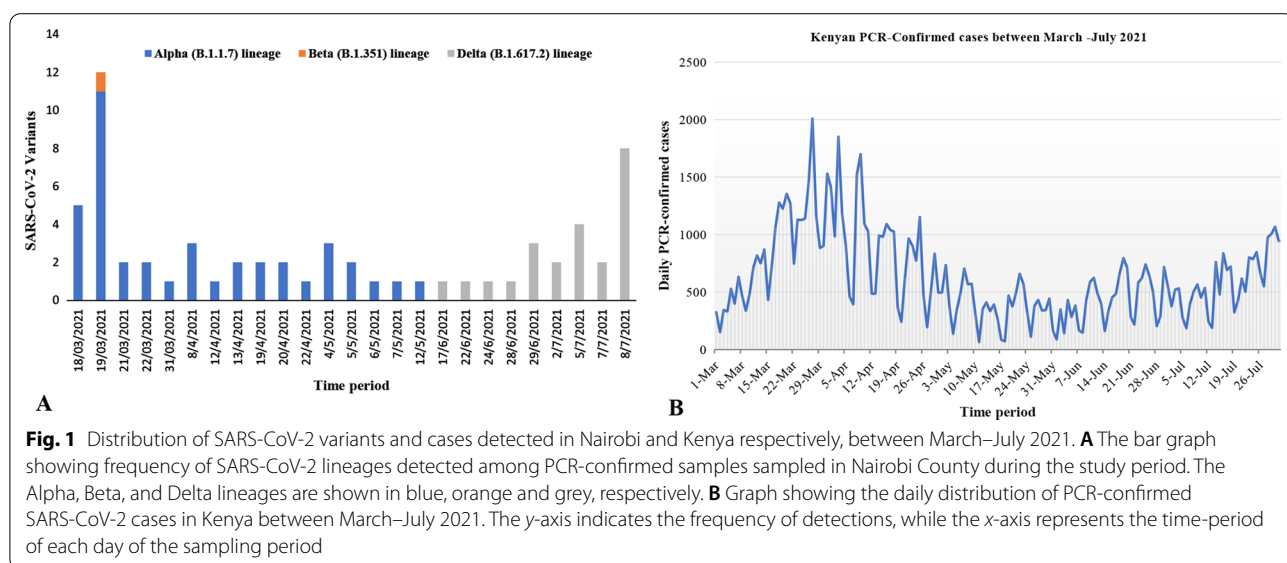
The SARS-CoV-2 genome sequences were grouped into genetic lineages based on the Phylogenetic Assignment of Named Global Outbreak LiNeages (PANGOLIN) [32] and NextStrain [33] classification methods. The genome sequences were analyzed relative to NC\_045512.2-Wuhan-Hu-1 and those obtained from other regions between March 2021 and July 2021. The latter were sampled from the coronavirus data collection of the Global initiative on sharing all influenza data (GISAID) [34] based on phylogenetic placement and GISAID’s BLAST tool within the EpiCoV™ browser [35]. Sequences with a match quality of  $\geq 0.990$  were selected. Low-quality, divergent, and incomplete genomes (<29,000 bp) were filtered out, leaving 135 reference genomes. Multiple sequence alignment of the local SARS-CoV-2 genomes alongside those of foreign strains sampled from the GISAID [34], resulting in a final dataset of 199 sequences was performed using MAFFT software v7 [36]. The 3’ and 5’ untranslated regions were trimmed using Unipro UGENE v39.0 software [37], followed by masking of the three known sequencing error-prone regions (13,402, 24,389, and 24,390). The best-fit substitution model for the dataset was determined using ModelFinder [38]. A maximum-likelihood tree was re-constructed with IQ-TREE v1.6.12 [39] utilizing the GTR+F+R2 model. Branch support was established by ultrafast bootstrap and SH-like approximate likelihood ratio test (SH-aLRT) with 1000 replicates. The consensus tree was annotated

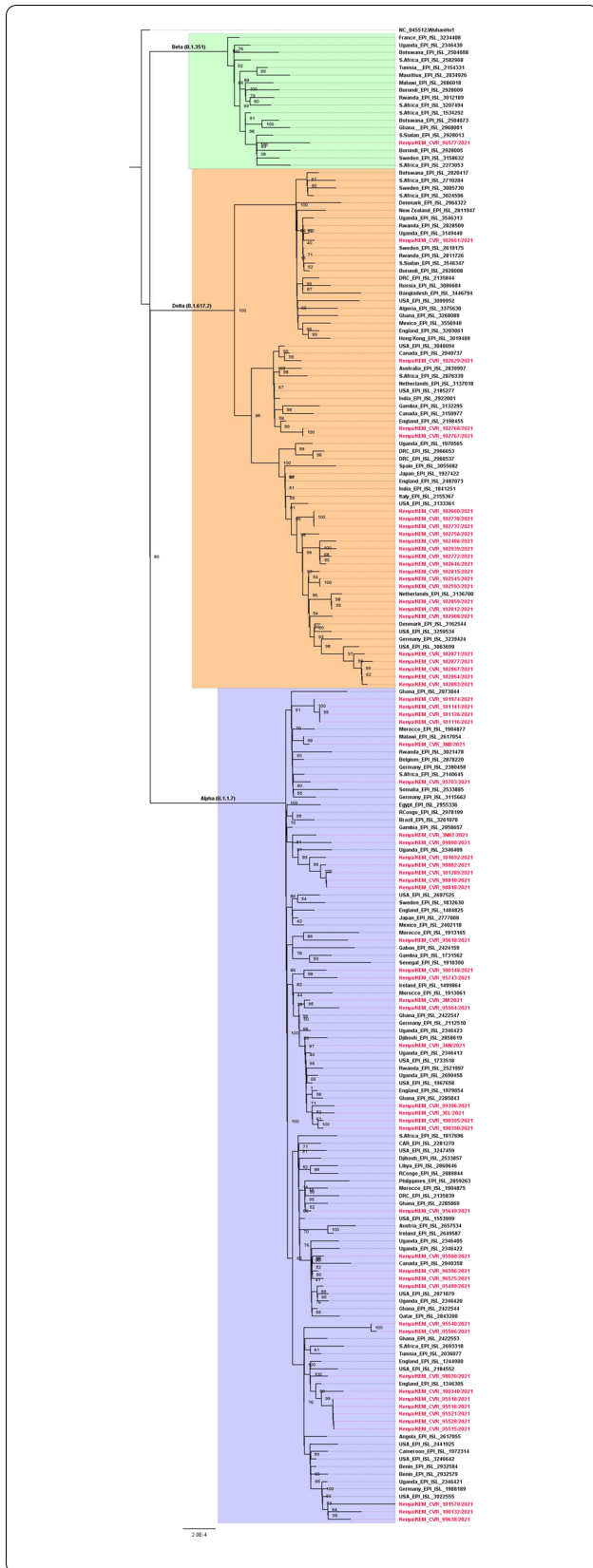
and displayed with the interactive tree of life v4 (iTOL) [40].

**Results**

Sixty-four SARS-CoV-2 genome sequences were obtained over the study period. Of these, forty-one [41] were found to belong to the Alpha (B.1.1.7) lineage, twenty-two [22] to Delta (B.1.617.2), and one (1) to Beta (B.1.351). The majority of the individuals from which viral sequences were obtained were male (n=43; 68.3%). The patients’ ages ranged from 16 to 69 years. Markedly, the Alpha and Beta SARS-CoV-2 genomes were recovered from the samples collected between March and May 2021, while viruses of the Delta lineages originated from samples collected during June and July 2021. From these data sets it can be concluded that the sequenced samples from Kenya during this period represent only 0.1% or less of the co-circulating SARS-CoV-2 causing confirmed infections. Figure 1 summarizes the distribution of SARS-CoV-2 variants detected in the study, as well as the PCR-confirmed cases in the country from March – July 2021.

Genetic analyses of the SARS-CoV-2 genomes relative to the NC\_045512.2-Wuhan-Hu-1 strain revealed a varying degree of amino acid polymorphisms. All the Kenyan sequences contained a D614G substitution in the spike protein-encoding gene. Those belonging to the Alpha (B.1.1.7) lineage displayed 69 -70del and Y144del mutations, while a small proportion (30%) of the Delta (B.1.351) lineage sequences had N354X change. However, these mutations were not peculiar to the Kenyan viruses as they were noted among sequences of global strains available in the GISAID database. Phylogenetic analysis revealed





**Fig. 2** Maximum-likelihood tree of Kenyan and global strains. Phylogenetic analysis depicting assignment of the genome sequences analyzed in the study to SARS-CoV-2 Alpha, Beta, and Delta genetic clades/lineages, signified by different colours. The Kenyan strains are identified in red. The bootstrap numbers at the nodes denote statistical branch support values. The bar indicates the number of nucleotide substitutions per site

that the local strains of the Alpha and Delta lineages were separated into about 15 and 6 clades of varying sizes, respectively, interspersed among the foreign strains. This interspersed of Kenyan clades among the foreign strains indicates multiple introductions of the viruses from outside Kenya, resulting in local transmission (Fig. 2).

**Discussion**

Surveillance of the SARS-CoV-2 by whole genome sequencing is crucial in providing insights into its transmission and evolutionary dynamics [4, 22]. It's equally valuable in identifying potential variants and virulence determinants [4, 41, 42]. This type of information is critical in evaluating vaccine effectiveness and guiding the design of drugs or diagnostics [41]. In this study, we investigated sixty-four (N=64) genome sequences of SARS-CoV-2 VOCs in Nairobi and neighboring counties in Kenya between March 2021 and July 2021, utilizing various bioinformatics approaches. Our analyses revealed the circulation of multiple SARS-CoV-2 variants of concern in Nairobi and neighboring counties during the studied period. The Alpha (B.1.1.7) lineage was predominant (65%) from March to May 2021, with Beta (B.1.351) being the minority (<2%), while Delta (B.1.617) lineage dominated beginning June 2021. This result mirrors findings of related studies conducted elsewhere [43–45] and suggests that the surge in SARS-CoV-2 infection observed in Nairobi and its environs during these periods were largely caused by the introduction and subsequent transmission into the population of the highly transmissible SARS-CoV-2 Alpha (B.1.1.7) and Delta (B.1.617) variants into the population. SARS-CoV-2 community transmissions and the occurrence of different waves of SARS-CoV-2 infection have been documented in Kenya [21, 23]. Moreover, the marked decline in the frequency of virus infection associated with the Alpha (B.1.1.7) variant around May 2021 could possibly be attributed to the public health countermeasures that were instituted to mitigate further seeding of newer virus clusters and the spread of the virus in the population. Overall, our findings indicate that there were multiple introductions of the viruses from outside Kenya, resulting in community transmission that contributed to the large number of new infections observed during the study period. This

observation emphasizes the importance of strengthening quarantine and border controls to combat virus imports into Kenya.

As reported elsewhere and as a common feature of the SARS-CoV-2 Alpha, Beta, and Delta lineages [22, 41], all the local strains harbored the D614G mutation in the spike glycoprotein, associated with increased infectivity [9, 46–48]. Interestingly, all of the Alpha (B.1.1.7) viruses in this study had the 69-70del and Y144del deletions in the spike protein, which are linked to increased infectivity and decreased monoclonal antibody binding [49–52]. Furthermore, the H69del and V70del mutations have been associated with a false-negative result for the target S gene in the standard RT-PCR SARS-CoV-2 assay [53, 54]. Similar to the findings of Parvin et al. [4], 30% of the study Delta (B.1.351) viruses had N354X spike mutation, associated with increased virus infectivity [2]. Meanwhile, phylogenetic analysis showed the majority of the Kenyan viruses clustered together in the respective SARS-CoV-2 lineages, except a few that appeared interspersed among the foreign strains. The clustering and interspersed of the Kenyan SARS-CoV-2 genome sequences among those of foreign strains included in the analysis plausibly suggest multiple routes of introduction and transmission into the community.

A limitation of this study was that it did not include samples from all regions of Kenya in order to present a nationwide picture. The other limitation is the relatively small sequence dataset, which may not be representative enough and may have introduced some degree of bias in our findings. However, despite the limitations, we have demonstrated that multiple lineages of SARS-CoV-2 variants of concern (VOC) viz B.1.1.7, B.1.351, and B.1.617 circulated in Nairobi and neighboring counties in Kenya between March and July 2021. These viruses were genetically similar to those that circulated in other regions of the world and are attributable to the different waves of SARS-CoV-2 infection experienced particularly in Nairobi, and Kenya at large. In conclusion, our findings provide a snapshot of information about the SARS-CoV-2 variants co-circulating in Kenya, and more data are needed to construct a more comprehensive picture of the epidemic in Kenya.

#### Abbreviations

ACE2: Angiotensin-converting enzyme 2; COVID-19: Coronavirus disease 2019; E gene: Envelope protein gene; GISAID: Global initiative on sharing all influenza data; iTOL: Interactive tree of life; KEMRI: Kenya Medical Research Institute; MoH: Ministry of Health; N gene: Nucleocapsid protein gene; ORF1ab: Open Reading Frame 1ab; PANGOLIN: Phylogenetic Assignment of Named Global Outbreak Lineages; RBD: Receptor-binding domain; RNA: Ribonucleic acid; RT-PCR: Reverse Transcription Polymerase Chain Reaction; S gene: Spike protein gene; SARS-CoV-2: Severe acute respiratory syndrome coronavirus-2; SERU: Scientific and Ethical Review Unit; SMRF: Sample Management and Repository Facility; VOI: Variants of interest; VOC: Variants of concern; WHO: World Health Organization.

#### Acknowledgements

We are grateful to GISAID for sharing the genomic sequences with the public and all contributors of the SARS-CoV-2 genomic sequence data included in our analysis.

#### Author contributions

Conceptualization: SAK, WB, SO, EA, SLS. Supervision: SAK, WB, SLS, SKL, RA. Whole-genome sequencing: SO, SL, JMM. Formal analysis: SO, SL, WB. Methodology: JMM, JKC, MM, NM, SO, EK, JM, SO, AN, MW, HK, EC, LK, RA. Writing—Original draft preparation: SO, SL, SLS. Writing – Review and editing: WB, SAK, EA. All authors read and approved the final manuscript.

#### Funding

This study was made possible by financial support from the government of Kenya through the Kenya Medical Research Institute.

#### Availability of data and materials

All the SARS-CoV-2 genomes reported in this work have been deposited in the GISAID database under the accession numbers: EPI\_ISL\_4472133; EPI\_ISL\_4472540; EPI\_ISL\_4472541; EPI\_ISL\_4472846; EPI\_ISL\_4473129; EPI\_ISL\_4473132; EPI\_ISL\_4473149; EPI\_ISL\_4473150; EPI\_ISL\_4473153; EPI\_ISL\_4473226; EPI\_ISL\_4473242; EPI\_ISL\_4473270; EPI\_ISL\_4473302; EPI\_ISL\_4473307 to EPI\_ISL\_4473309; EPI\_ISL\_4473504; EPI\_ISL\_4473482; EPI\_ISL\_4473483; EPI\_ISL\_4474043; EPI\_ISL\_4474044; EPI\_ISL\_4474467; EPI\_ISL\_4474512; EPI\_ISL\_4474518; EPI\_ISL\_4474539; EPI\_ISL\_4474581; EPI\_ISL\_4474771; EPI\_ISL\_4505350; EPI\_ISL\_4505457; EPI\_ISL\_4505540 to EPI\_ISL\_4505546; EPI\_ISL\_4506003; EPI\_ISL\_4506016; EPI\_ISL\_4506017; EPI\_ISL\_4506812; EPI\_ISL\_4507490; EPI\_ISL\_4507883; EPI\_ISL\_4508206; EPI\_ISL\_4508258; EPI\_ISL\_4508426; EPI\_ISL\_4515253; EPI\_ISL\_4515254; EPI\_ISL\_4515337 to EPI\_ISL\_4515340; EPI\_ISL\_4515415; EPI\_ISL\_4515456; EPI\_ISL\_45155537; EPI\_ISL\_4515844; EPI\_ISL\_4515845; EPI\_ISL\_4515707; EPI\_ISL\_4515900 to EPI\_ISL\_4515905; EPI\_ISL\_455616.

#### Declarations

##### Ethics approval and consent to participate

Written informed consent was obtained from all study participants following the Kenya Ministry of Health's (MoH) national response to COVID-19 pandemic protocols and guidelines. The study was carried out under the whole genome sequencing study protocol, approved by the Scientific and Ethical Review Unit (SERU) at KEMRI (SERU # 4035).

##### Consent for publication

This manuscript does not contain any individual personal data in any form (including any individual details, images, or videos). Thus, consent for publication was not required.

##### Competing interests

The authors declare no competing interests.

##### Author details

<sup>1</sup>Centre for Virus Research, Kenya Medical Research Institute (KEMRI), P.O. Box 54628-00200, Nairobi, Kenya. <sup>2</sup>Centre for Traditional Medicine and Drug Research, Kenya Medical Research Institute (KEMRI), Nairobi, Kenya. <sup>3</sup>Centre for Biotechnology Research and Development, Kenya Medical Research Institute (KEMRI), Nairobi, Kenya. <sup>4</sup>Centre for Respiratory Diseases Research, Kenya Medical Research Institute (KEMRI), Nairobi, Kenya.

Received: 27 October 2021 Accepted: 5 October 2022

Published online: 08 November 2022

#### References

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382:727–33. <https://doi.org/10.1056/nejmoa2001017>.
- Ou X, Liu Y, Lei X, Li P, Mi D, Ren L, et al. Characterization of spike glycoprotein of SARS-CoV-2 on virus entry and its immune cross-reactivity

- with SARS-CoV. *Nat Commun.* 2020;11:1620. <https://doi.org/10.1038/s41467-020-15562-9>.
3. Otu A, Agogo E, Ebenso B. Africa needs more genome sequencing to tackle new variants of SARS-CoV-2. *Nat Med.* 2021;27:744–5. <https://doi.org/10.1038/s41591-021-01327-4>.
  4. Parvin R, Afrin SZ, Begum JA, Ahmed S, Nooruzzaman M, Chowdhury EH, et al. Molecular analysis of SARS-CoV-2 circulating in Bangladesh during 2020 revealed lineage diversity and potential mutations. *Microorganisms.* 2021;9:1035. <https://doi.org/10.3390/microorganisms9051035>.
  5. WHO Health Emergency Dashboard. 2022. <https://covid19.who.int/region/afro/country/ke>. Accessed 21 July 2022.
  6. Srivastava S, Banu S, Singh P, Sowpati DT, Mishra RK. SARS-CoV-2 genomics: an Indian perspective on sequencing viral variants. *J Biosci.* 2021;46:22. <https://doi.org/10.1007/s12038-021-00145-7>.
  7. Koyama T, Platt D, Parida L. Variant analysis of SARS-CoV-2 genomes. *Bull World Health Organ.* 2020;98:495–504. <https://doi.org/10.2471/blt.20.253591>.
  8. Eurosurveillance Editorial Team. Updated rapid risk assessment from ECDC on the risk related to the spread of new SARS-CoV-2 variants of concern in the EU/EEA—first update. *Euro Surveill.* 2021;26:2101211. <https://doi.org/10.2807/1560-7917.ES.2021.26.3.2101211>.
  9. Sanyaolu A, Okorie C, Marinkovic A, Haider N, Abbasi AF, Jaferi U, et al. The emerging SARS-CoV-2 variants of concern. *Ther Adv Infect Dis.* 2021;8:20499361211024372. <https://doi.org/10.1177/20499361211024372>.
  10. WHO. COVID-19 weekly epidemiological update: proposed working definitions of SARS-CoV-2 variants of interest and variants of concern 2021. <https://www.who.int/publications/m/item/covid-19-weekly-epidemiological-update>. Accessed 21 July 2022.
  11. Parums DV. Revised World Health Organization (WHO) terminology for variants of concern and variants of interest of SARS-CoV-2. *Med Sci Monit Int Med J Exp Clin Res.* 2021;27:e933622. <https://doi.org/10.12659/MSM.933622>.
  12. Konings F, Perkins MD, Kuhn JH, Pallen MJ, Alm EJ, Archer BN, et al. SARS-CoV-2 variants of interest and concern naming scheme conducive for global discourse. *Nat Microbiol.* 2021;6:821–3. <https://doi.org/10.1038/s41564-021-00932-w>.
  13. Wolter N, Jassat W, Walaza S, Welch R, Moultrie H, Groome M, et al. Early assessment of the clinical severity of the SARS-CoV-2 omicron variant in South Africa: a data linkage study. *The Lancet.* 2022;399:437–46. [https://doi.org/10.1016/s0140-6736\(22\)00017-4](https://doi.org/10.1016/s0140-6736(22)00017-4).
  14. Araf Y, Akter F, Tang Yd, Fatemi R, Parvez MSA, Zheng C, et al. Omicron variant of SARS-CoV-2: genomics, transmissibility, and responses to current COVID-19 vaccines. *J Med Virol.* 2022;94:1825–1832. <https://doi.org/10.1002/jmv.27588>.
  15. Cedro-Tanda A, Gómez-Romero L, de Anda-Jauregui G, Garnica-López D, Alfaro-Mora Y, Sánchez-Xochipa S, et al. Early genomic, epidemiological, and clinical description of the SARS-CoV-2 omicron variant in Mexico City. *Viruses.* 2022;14:545. <https://doi.org/10.3390/v14030545>.
  16. Viana R, Moyo S, Amoako DG, Tegally H, Scheepers C, Althaus CL, et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature.* 2022;603:679–86. <https://doi.org/10.1038/s41586-022-04411-y>.
  17. Petra M, Steven K, Mahesh Shanker D, Guido P, Bo M, Swapnil M, et al. SARS-CoV-2 B. 1.617. 2 Delta variant replication and immune evasion. *Nature.* 2021;599:114–119. <https://doi.org/10.1038/s41586-021-03944-y>.
  18. Faria NR, Mellan TA, Whittaker C, Claro IM, Candido DdS, Mishra S, et al. Genomics and epidemiology of the P. 1 SARS-CoV-2 lineage in Manaus, Brazil. *Science.* 2021;372:815–821. <https://doi.org/10.1126/science.abh2644>.
  19. Garcia-Beltran WF, Denis KJS, Hoelzemer A, Lam EC, Nitido AD, Sheehan ML, et al. mRNA-based COVID-19 vaccine boosters induce neutralizing immunity against SARS-CoV-2 Omicron variant. *Cell.* 2022;185:457–466.e4. <https://doi.org/10.1016/j.cell.2021.12.033>.
  20. Bekker LG, Ntusi NAB. Lessons from two SARS-CoV-2 waves in South Africa. *Lancet Global Health.* 2021;9:e1177–8. [https://doi.org/10.1016/s2214-109x\(21\)00313-2](https://doi.org/10.1016/s2214-109x(21)00313-2).
  21. Brand SPC, Ojal J, Aziza R, Were V, Okiro EA, Kombe IK, et al. COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science.* 2021;374:989–94. <https://doi.org/10.1126/science.abk0414>.
  22. Ntoumi F, Mapanguly CCM, Tomazatos A, Pallerla SR, Casadei N, Angelov A, et al. Genomic surveillance of SARS-CoV-2 in the Republic of Congo. *Int J Infect Dis.* 2021;105:735–8. <https://doi.org/10.1016/j.ijid.2021.03.036>.
  23. Githinji G, de Laurent ZR, Mohammed KS, Omuoyo DO, Macharia PM, Morobe JM, et al. Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. *Nat Commun.* 2021;12:4809. <https://doi.org/10.1038/s41467-021-25137-x>.
  24. WHO. Diagnostic testing for SARS-CoV-2 2020. <https://www.who.int/publications/item/diagnostic-testing-for-sars-cov-2>. Accessed 21 July 2022.
  25. Freed NE, Vlková M, Faisal MB, Silander OK. Rapid and inexpensive whole-genome sequencing of SARS-CoV-2 using 1200 bp tiled amplicons and Oxford Nanopore Rapid Barcoding. *Biol Methods Protoc.* 2020;5:bpaa014. <https://doi.org/10.1093/biomethods/bpaa014>.
  26. Tyson JR, James P, Stoddart D, Sparks N, Wickenhagen A, Hall G, et al. Improvements to the ARTIC multiplex PCR method for SARS-CoV-2 genome sequencing using nanopore. *BioRxiv.* 2020:2020.09.04.283077. <https://doi.org/10.1101/2020.09.04.283077>.
  27. Loman N, Rowe W, Rambaut A. nCoV-2019 novel coronavirus bioinformatics protocol. *Artic Network.* 2020. <https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html>. Accessed 21 July 2022.
  28. Quick J. nCoV-2019 sequencing protocol v3 (LoCost) V. 3. 2020. 2021. <https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8bye>. Accessed 21 July 2022.
  29. GISAID.CoVsurver: mutation analysis of hCoV-19. <https://www.gisaid.org/epiflu-applications/covsurver-mutations-app/>. Accessed 21 July 2022.
  30. Singer J, Gifford R, Cotten M, Robertson D. CoV-GLUE: a web application for tracking SARS-CoV-2 genomic variation. 2020. <https://doi.org/10.20944/preprints202006.0225.v1>.
  31. Aksamentov I, Roemer C, Hodcroft EB, Neher RA. Nextclade: clade assignment, mutation calling and quality control for viral genomes. *J Open Source Softw.* 2021;6:3773. <https://doi.org/10.21105/joss.03773>.
  32. Rambaut A, Holmes EC, O'Toole Á, Hill V, McCrone JT, Ruis C, et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol.* 2020;5:1403–7. <https://doi.org/10.1038/s41564-020-0770-5>.
  33. Hadfield J, Megill C, Bell SM, Huddleston J, Potter B, Callender C, et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics.* 2018;34:4121–3. <https://doi.org/10.1093/bioinformatics/bty407>.
  34. Shu Y, McCauley J. GISAID: global initiative on sharing all influenza data—from vision to reality. *Eurosurveillance.* 2017;22:30494. <https://doi.org/10.2807/1560-7917.es.2017.22.13.30494>.
  35. GISAID.EpiCoV browser. <https://www.epicov.org/epi3/frontend#3c1bd>. Accessed 21 July 2022.
  36. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol.* 2013;30:772–80. <https://doi.org/10.1093/molbev/mst010>.
  37. Okonechnikov K, Golosova O, Fursov M, Team U. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics.* 2012;28:1166–7. <https://doi.org/10.1093/bioinformatics/bts091>.
  38. Kalyanamoothy S, Minh BQ, Wong TK, Von Haeseler A, Jermin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods.* 2017;14:587–9. <https://doi.org/10.1038/nmeth.4285>.
  39. Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res.* 2016;44:W232–5. <https://doi.org/10.1093/nar/gkw256>.
  40. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 2019;47:W256–9. <https://doi.org/10.1093/nar/gkz239>.
  41. Tokgun O, Caliskan A, Coskun C, Tokgun PE, Akca H. Whole genome sequencing and phylogenetic analysis of SARS-CoV-2 strains in Turkey. *J Infect Dev Countr.* 2021;15:470–7. <https://doi.org/10.3855/jidc.14560>.
  42. Cyranoski D. Alarming COVID variants show vital role of genomic surveillance. *Nature.* 2021;589:337–8. <https://doi.org/10.1038/d41586-021-00065-4>.
  43. Pascall DJ, Mollett G, Blacow R, Bulteel N, Campbell R, Campbell A, et al. The SARS-CoV-2 Alpha variant is associated with increased clinical severity of COVID-19 in Scotland: a genomics-based retrospective cohort analysis. *MedRxiv.* 2022; 2021-08. <https://doi.org/10.1101/2021.08.17.21260128>.

44. Saha S, Tanmoy AM, Tanni AA, Goswami S, Al Sium SM, Saha S, et al. New waves, new variants, old inequity: a continuing COVID-19 crisis. *BMJ Glob Health*. 2021;6: e007031. <https://doi.org/10.1136/bmjgh-2021-007031>.
45. Tuite A, Fisman D, Oduyayo A, Bobos P, Allen V, Bogoch I. COVID-19 hospitalizations, ICU admissions and deaths associated with the new variants of concern. *Sci Briefs Ont COVID-19 Sci Advis Table*. 2021; 1:1. <https://doi.org/10.47326/ocsaat.2021.02.18.1.0>.
46. 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;182:812-827.e19. <https://doi.org/10.1016/j.cell.2020.06.043>.
47. Zhang L, Jackson CB, Mou H, Ojha A, Rangarajan ES, Izard T, et al. SARS-CoV-2 spike-protein D614G mutation increases virion spike density and infectivity. *Nat Commun*. 2020;11:6013. <https://doi.org/10.1038/s41467-020-19808-4>.
48. Ogawa J, Zhu W, Tonnu N, Singer O, Hunter T, Ryan AL, et al. The D614G mutation in the SARS-CoV2 Spike protein increases infectivity in an ACE2 receptor dependent manner. *Biorxiv*. 2020. <https://doi.org/10.1101/2020.07.21.214932>.
49. Wang R, Zhang Q, Ge J, Ren W, Zhang R, Lan J, et al. Analysis of SARS-CoV-2 variant mutations reveals neutralization escape mechanisms and the ability to use ACE2 receptors from additional species. *Immunity*. 2021;54:1611-1621.e5. <https://doi.org/10.1016/j.immuni.2021.06.003>.
50. McCarthy KR, Rennick LJ, Nambulli S, Robinson-McCarthy LR, Bain WG, Haidar G, et al. Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science*. 2021;371:1139-42. <https://doi.org/10.1126/science.abf6950>.
51. Greaney AJ, Starr TN, Gilchuk P, Zost SJ, Binshtein E, Loes AN, et al. Complete mapping of mutations to the SARS-CoV-2 spike receptor-binding domain that escape antibody recognition. *Cell Host Microbe*. 2021;29:44-57.e9. <https://doi.org/10.1016/j.chom.2020.11.007>.
52. Kemp SA, Collier DA, Datir RP, Ferreira IA, Gayed S, Jahun A, et al. SARS-CoV-2 evolution during treatment of chronic infection. *Nature*. 2021;592:277-82. <https://doi.org/10.1038/s41586-021-03291-y>.
53. Bal A, Destras G, Gaymard A, Stefic K, Marlet J, Eymieux S, Regue H, Semanas Q, d'Aubarede C, Billaud G, Laurent F. Two-step strategy for the identification of SARS-CoV-2 variant of concern 202012/01 and other variants with spike deletion H69-V70, France, August to December 2020. *Euro Surveill*. 2021;26:2100008. <https://doi.org/10.2807/1560-7917.ES.2021.26.3.2100008>.
54. Washington NL, White S, Cirulli ET, Bolze A, Lu JT. S gene dropout patterns in SARS-CoV-2 tests suggest spread of the H69del/V70del mutation in the US. *MedRxiv*. 2020. <https://doi.org/10.1101/2020.12.24.20248814>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

