Genomic Surveillance of SARS-COV-2 reveals diverse circulating variant lineages in Nairobi and Kiambu County, Kenya

Josiah O. Kuja (✉ josiah.kuja@bio.ku.dk)
University of Copenhagen https://orcid.org/0000-0002-3202-9540
Bernard N. Kanoi
Mount Kenya University
Clement Shiluli
Mount Kenya University
Michael Maina
Mount Kenya University
Harrison Waweru
Mount Kenya University
Kimita Gethii
United States Army Medical Research Unit, Kenya
Mary Mungai
Kenya Medical Research Institute, Nairobi, Kenya
Moses Masika
University of Nairobi, Nairobi, Kenya
Omu Anzala
University of Nairobi, Nairobi, Kenya
Matilu Mwau
Kenya Medical Research Institute, Nairobi, Kenya
Taane G. Clark
London School of Hygiene & Tropical Medicine, London, United Kingdom
John Waitumbi
United States Army Medical Research Unit, Kenya
Jesse Gitaka (✉ jgitaka@mku.ac.ke)
Mount Kenya University https://orcid.org/0000-0001-7516-522X

Research Article

Keywords: SARS-CoV-2, variant, surveillance, vaccine, genomic

Posted Date: May 13th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1636866/v1

Abstract

Genomic surveillance and identification of SARS-CoV-2 outbreaks are important in understanding the genetic diversity, phylogeny, and lineages of SAR-CoV-2. Genomic surveillance provides insights into circulating infections, and insights into the robustness and design of vaccines and other infection control approaches. We sequenced 56 SARS-CoV-2 isolates from a Kenyan clinical population, of which 52 passed the Ultrafast sample Placement on the existing tRE for the phylo-genome-temporal analyses across two regions in Kenya (Nairobi and Kiambu County). B.1.1.7 (Alpha; n = 32, 61.5%) and B.1 (n = 9, 17.3%) lineages were the most predominant variant with a wide-range of Ct values (5–31) and variant mutations across the two regions. Lineages B.1.617.2, B.1.1, A.23.1, A.2.5.1, B.1.596, A, and B.1.405 were also detected across the sampling sites within the target population. The lineages and genetic isolates were traced back to China (A), Costa Rica (A.2.5.1), Europe (B.1, B.1.1, A.23.1), USA (B.1.405, B.1.596), South Africa (B.1.617.2), and United Kingdom (B.1.1.7), indicating multiple introduction events. There were, however, no genetic isolates associated with the omicron (B.1.1.529) variant of concern that is less severe than the previous variants.

Introduction

The COVID-19 pandemic, caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) virus, has spread globally, with the arising of new variants (e.g., alpha, delta, and omicron) affecting infection control, leading to policy changes in social restrictions, and impacts on the efficacy of vaccines. An understanding of the genetic diversity, phylogeny and lineages of SAR-CoV-2 provides insights into circulating infections, and insights into the robustness and design of vaccines and other infection control tools. To date, there have been more than 11 million reported infections and 239,000 reported deaths caused by the novel coronavirus in Africa [1]. In the early months of the COVID-19 pandemic, Africa's rapid and coordinated response, informed by emerging data, led to infection control which mitigated effects of a first-wave and to a lesser degree a second wave [2]. This included rapid response through genomic surveillance to curb the spread of SARS-CoV-2. Nigeria took three days to sequence SARS-CoV-2 genome after the identification of the novel virus in the country [3]. Within the same period, the Network for Genomic Surveillance in South Africa (NGS-SA) was established to facilitate case confirmation and sequencing of the positive cases for phylogenetic and lineage updates [4]. Public and health officers in Uganda also set a sequencing program to facilitate genomic sequencing of the confirmed positive samples from the rapid contact tracing and international arrivals [5]. However, now in 2022, as vast vaccination campaigns have enabled the global north to gain some control over the pandemic, the vaccine roll-out in Africa lags because of inequities in accessing them. Kenya has vaccinated 12,652,991 people at a rate of 23.89% doses per 100 people [6].

Kenya joined the genomic surveillance of the SARS-CoV-2 pandemic after reporting the first case on 13th March, 2020 (Health, 2020). Most of the earlier cases were dominated by the lineage B.1, which was introduced into African countries from the international arrivals, particularly from European origin. Early public health measures in Kenya included restricted movements through limitation of social interactions

and gatherings, but these failed to prevent transmission [7]. By the end of July 2020, the Kenyan Ministry of Health had reported 20,636 PCR confirmed cases and 341 SARS-CoV-2 associated deaths. Most of the cases were from Nairobi and Mombasa, which were exposed to cross-border interactions and international arrivals, with individuals who did not undergo rapid testing procedures at the points of border control. At the time, seroprevalence surveillance of the national blood bank revealed the existence of SARS-CoV-2 in the population before the published 13th March 2020 [8]. The growing prevalence was confirmed by the community-based modelling teams [9] who relied solely on the seroprevalence, PCR confirmed cases, and genomic data that were able to present different variants for each wave in the country.

Genomic surveillance is an essential approach to characterise the transmission dynamics and the prevalence of SARS-CoV-2 within a population. Most of the sequences published in Kenya have been closely related to the Wuhan reference sequences clustered between 4 and 16 nucleotide substitutions [10]. The predominant nucleotide substitutions were associated with mutations at positions A23403G (S:D614G), P970L, P314L, R203K and G204R [11]. However, genomic surveillance revealed D614G to dominate the prevalence across the country and across the borders despite its initial appearance in the earlier stages of the pandemic [12].

Here, we sequenced RT-PCR confirmed SARS-CoV-2 positive samples, sourced from Nairobi and Kiambu County, collected between September 2020 and March 2021, spanning in between the severe alpha and delta variants of concern within the country and across borders. This work led to 52 SARS-CoV-2 isolate sequences for the phylo-genomic analyses across Nairobi and Kiambu County, one of the largest genomic epidemiology studies in the Nairobi metropolitan.

Methods

Sample Collection

Samples collection and testing were conducted according to the Kenya Ministry of Health (MoH, Kenya) COVID-19 pandemic surveillance protocols and guidelines [13]. Sampling and whole genome sequencing protocol was reviewed and approved by the Ethics Review Committee (ERC-MKU/ERC/1613) of Mount Kenya University. The study was conducted between September 2020 and March 2021 when the nasopharyngeal samples were collected using nasal swabs. The collected swabs were stored in viral transport media tubes until use. 150µl of each sample was processed for RNA extraction for sequencing.

SARS-CoV-2 diagnosis and RNA Extraction

RNA extraction was performed using the Sacace Biotechnologies Ribo Virus kit protocol (Sacace™ SARS-CoV-2 Variants Typing Real-TM; Srl-Via Scalabrini, Como Italy) according to manufacturer's instructions. Subsequently, positive infections were quickly identified through RNA purification followed by real time reverse-transcription Polymerase Chain Reaction (RT-PCR) using a Ribo virus column (Srl-Via Scalabrini, Como Italy).

SARS-CoV-2 genome amplification, library preparation and sequencing

The purified RNA was used to synthesise complementary DNA (cDNA) using random primers with the Superscript IV one step reverse transcriptase kit (Thermo Fisher Scientific, CA, USA). The cDNA was then amplified using the multiplex ARTIC primer-pools A and B version 3 (Sevinsky et al., 2020) using the NEBNext Q5 High-Fidelity 2X Master Mix (New England Biolabs, MA, USA). The resulting PCR products were pooled together and cleaned using 1× AMPure XP beads (Beckman Coulter), and then used for library preparation with the NexteraXT library preparation kit (Illumina, CA, USA), following the manufacturer's instructions. The final library was normalised to 12 pM, spiked with 10% Phix genome and sequenced on Illumina MiSeq platform (Illumina, CA, USA) using 600 V3 paired end chemistry.

SARS-CoV-2 lineage and clade assignment

Amplicon sequences were assembled against SARS CoV-2 reference genomes and assigned lineages using the IDseq platform. Resulting phylogenetic clades were quality checked in NextClade (version 0.13.0) and phylogenetic assignments performed using Ultrafast sample Placement on the existing tRE (Usher) through maximally parsimonious placement (https://genome.ucsc.edu/cgi-bin/hgPhyloPlace) with a total of 52 sequences.

Results

SARS-CoV-2 testing and Clinical Characteristics

During the study period, SARS-CoV-2 RT-PCR testing identified 65 positive samples from the five COVID-19 testing sites in Nairobi and Kiambu county. Specifically, the samples obtained from Nairobi were 33 (St. Francis Community Hospital, n = 24; Uhai Neema Hospital, n = 3; and Mbagathi Hospital, n = 6) and Kiambu 19 (Kiambu Level 5 hospital n = 7, Gatundu Level 5 hospital, n = 19) counties (Table S1). The observed Ct values (representing relative abundance of virus material in a sample) from the positive samples ranged between 3.96–31.59 with a total of 25 asymptomatic and 31 symptomatic patients. Out of the 25 asymptomatic patients, 12 (46.2%) were males while 13 (53.8%) were female. For the 31 symptomatic patients, 20 (64.5%) were males and 11 (35.5%) were females. The precise mean of the Ct values of the samples varied across Nairobi and Kiambu county (Fig. 1). However, the variation was associated with the lineage variants (Fig. 1).

Inferred lineages

The Ultrafast sample Placement on the existing tRE (Usher) through maximally parsimonious placement https://genome.ucsc.edu/cgi-bin/hgPhyloPlace skipped four isolates due to too many N-bases (> 0.50). Sequence analyses of the 52 isolates (n < 0.50) revealed 8 major SARS-CoV-2 lineage variants (A, A.2.5.1, B.1, B.1.1, B.1.1.7, B.1.405, B.1.596, B.1.617.2), originating from major hotspots across the globe. The predominant variant was associated with the B.1.1.7 (Alpha; n = 32, 61.5%) lineage (Fig. 2). B.1 was the

second most dominant variant (n = 9, 17.3%). In St. Francis Hospital (n = 24), lineages B.1.1.7 (13; 54.2%) and B.1 (5 20.8%) were the most dominant, followed by A.2.5.1 (8.3%), A (4.2%), B.1.1 (4.2%), B.1.405 (4.2%), and B.1.596 (4.2%). B.1.1 (n = 1) and B.1.1.7 (n = 5) variants were associated only with Mbagathi. B.1 (n = 3), B.1.617.2 (n = 3) and B.1.1 (n = 1) lineages dominated Kiambu Level 5 hospital. Gatundu Level 5 hospital was dominated by B.1.1.7 (n = 11; 91.7%), with B.1 (n = 1; 8.3%) present. B.1.1.7 was the only variant (n = 3) detected at Uhai Neema hospital in Nairobi County (Fig. 2).

SARS-CoV-2 Sequence Diversity

Sequence variants across the 52 isolates included nucleotide mutations (range: 5 to 48) leading to 4 amino acid changes (range 4 to 23) and deletions (range: 2 to 8) (https://clades.nextstrain.org/). The variants were found across twelve genes, where the greatest number of mutations were in isolates associated to variant B.1.1.7. Twelve genes were identified with a size range of 5–4000 base-pairs (Fig. 3). Gene ORF1a had the highest number (500–4000) of base-pairs, followed by ORF1b (500–2500), gene S (200–1200), gene N (50–400). The other genes, ORF3a (50–250), gene M (50–200), ORF8(20–120), ORF9b (10–90), gene E (10–70), ORF6 (10–60), ORF7a (20–120), and ORF7b (5–40) had almost a similar and overlapping base-pairs indicated by the clade assignments, mutation calls and sequence quality check in Nextclade v1.13.1 (https://clades.nextstrain.org/). Genes associated with the ORF1a were the most diverse with highly conserved regions across the variants followed by the ORF1b, and gene S. (Fig. 3). Phylogenetic analyses revealed the prevalence and introduction of lineages in Nairobi and Kiambu counties. Lineages were clustered based on their specific variants indicating B.1.1.7 as the most predominant lineage in the phylogeny (Fig. 4) despite the mutational variation and temporal differences within isolates.

Discussion

The study revealed B.1.1.7 lineage as the predominant variant in Nairobi and Kiambu county (two neighbouring counties with high population). The cumulative prevalence of B.1.1.7 lineage is 19% globally [14] and this correlates to its prevalence in Kenya, which is estimated to be between 10–20% with a total of 947 positive samples out of 5,175 sequences [14] up to the time of sampling. The prevalence of B.1.1.7 lineage in Kenya, particularly within the capital city, is an indication of local transmission since the Delta variant (lineage B1.617.2) was emerging at the time of sample collection. Consequently, B.1.1.7 lineage represented wide-range of Ct values (5–31), which implies higher risk of infectivity of the variant in the population. Ct values of less than 30 indicate high infectivity rates, and as such patients would require isolation to contain the transmission of the virus within the population. In contrast, Ct values above 30 are an indication of viral particles in circulation, with patients that would not require isolation and quarantine, especially if the symptoms appeared ten days before the RT-PCR test [15]. In this study the mean Ct values were relatively lower than 25 (Fig. 1) and this indicates how the patients were highly infectious. They would be referred to as super spreaders if the viral preparations could be calibrated to one million copies of SARS-CoV-2 RNA per ml [16, 17].

B.1 lineage was the second most abundant (17.3%) variant in Nairobi and Kiambu counties. Its abundance was skewed to Kiambu clinical sites. In Nairobi County, it was detected in St. Francis Community Hospital with a relative abundance of 20.8%. The variation in lineage distribution across the two counties is an indication of local transmission through a geographical timescale. Globally 102,145 sequences of B.1 lineage have been identified with a cumulative frequency of 25 [14]. Since its first identification in the United States of America, the variant has spread across countries including Kenya with a cumulative prevalence of 20–35%. B.1 was reported to be the predominant lineage at the Kenyan Coast, Mombasa County [18]. Most of the cases at the Kenyan coast were from international arrivals and travellers. By then Mombasa County was experiencing high death rates with exponential positivity rates unlike any other county in Kenya. Lineages identified in our study population must have been localized transmissions after a wave of the B.1 at the coastal region. The rest of the lineages were homogenously distributed across the clinical units. Their occurrence was, however, specific to healthcare facilities indicating the success of local transmissions across the counties. Apart from the B.1 and A that were detected at the Kenyan coastal region [19], the rest of the lineages in our study were specific to Nairobi and Kiambu county. Even the abundant B.1.1.7 alpha lineage was never detected at the Kenyan coast and its borders at the time of the B.1 outbreak in the region.

Dynamics of the identified lineages in the population across the two counties is potentially due to many factors, including socio-economic parameters. The genomic epidemiology of the variants, however, underlies the epidemic waves across Africa and Kenya. Nairobi is a cosmopolitan city with diverse interactions from international travellers. Kiambu is equally a busy county in which populations interact through trade and travels. These two counties are also characterized by the high number of student populations, hence the variation within sub-populations between these two counties are significant and are likely to determine SARS-CoV-2 variants prevalence outcome. Ostensibly, sub-populations belonging to the lower socioeconomic status are likely to encounter SARS-CoV-2 variants as opposed to those from the higher socioeconomic status [9]. Though isolates in this study do not show any relations to omicron (B.1.1.529) variant of concern, the variant is likely to spread in Nairobi and Kiambu counties due to high rates of mutation from the previous variants (Delta) and super spreading between populations. It would occur with symptoms similar to those of the previous variants with less severe infections. The severity of omicron and any other emerging COVID-19 variant infection can, however, be prevented by vaccination as the best public health measure protecting people from severe illnesses.

In conclusion, the SARS-CoV-2 lineages and genetic isolates identified in this study could be traced back to multiple countries including China (A), Costa Rica (A.2.5.1), Europe (B.1, B.1.1, A.23.1), USA (B.1.405, B.1.596), and South Africa for B.1.617.2 and Alpha B1.1.7, indicating a significant rate of transmission across borders into Kenya. However, there were no genetic isolates associated with the omicron (B.1.1.529) that has been associated with increased infectivity and less severe infection outcome. Using the established platforms, continued surveillance will be required to give a deeper understanding of the spread of SARS-CoV-2 and detect any emerging variants that may be of interest to support pandemic control.

Declarations

Data availability

The datasets generated and/or analysed during the current study are available from the corresponding author on reasonable request. The data has also been deposited to https://submit.ncbi.nlm.nih.gov/subs/sra/SUB11087556/files

Conflicts of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Authors' contributions

JOK, BNK, CS, MM, HW, JG conceived and designed the experiment. CS, MM, HW performed sample preparation and molecular experiments. KG, JW library preparation and sequencing. JOK, KG, JW compiled results and analysed the data. JOK, BNK, CS, MM, HW, TGC, JW, JG write up and final edit of the manuscript. All authors JOK, BNK, CS, MM, HW, KG, MM, MM, OA, MM, TGC, JW, JG read and approved the final manuscript.

Funding

This work was funded in part by The African Academy of Sciences grant number SARSCoV2-4-20-010 to JG. TGC is funded by a UKRI Global Effort on COVID-19 (GECO) Health Research grant (Ref. GEC2211; MR/V036890/1). The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We acknowledge the Mount Kenya University Rapid Response Team for sample collection and SARS-CoV-2 diagnosis, we also acknowledge United States Army Medical Research Unit, Kenya for sample sequencing.

References

1. Khan M, Adil SF, Alkhathlan HZ, Tahir MN, Saif S, Khan M, Khan ST (2020) COVID-19: A Global Challenge with Old History, Epidemiology and Progress So Far. Molecules 26:39
2. Salyer SJ, Maeda J, Sembuche S, et al (2021) The first and second waves of the COVID-19 pandemic in Africa: a cross-sectional study. The Lancet 397:1265–1275
3. Oluwagbemi OO, Oladipo EK, Dairo EO, Ayeni AE, Irewolede BA, Jimah EM, Oyewole MP, Olawale BM, Adegoke HM, Ogunleye AJ (2022) Computational construction of a glycoprotein multi-epitope

subunit vaccine candidate for old and new South-African SARS-CoV-2 virus strains. Inform Med Unlocked 28:100845

4. Giandhari J, Pillay S, Wilkinson E, et al (2021) Early transmission of SARS-CoV-2 in South Africa: An epidemiological and phylogenetic report. Int J Infect Dis 103:234–241
5. Wilkinson E, Giovanetti M, Tegally H, et al (2021) A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science 374:423–431
6. [No title]. https://www.health.go.ke/wp-content/uploads/2022/02/MINISTRY-OF-HEALTH-KENYA-COVID-19-IMMUNIZATION-STATUS-REPORT-1ST-FEBRUARY-2022.pdf. Accessed 1 May 2022
7. Brand SPC, Aziza R, Kombe IK, Agoti CN, Hilton J, Rock KS, Parisi A, James Nokes D, Keeling MJ, Barasa EW Forecasting the scale of the COVID-19 epidemic in Kenya. https://doi.org/10.1101/2020.04.09.20059865
8. Adetifa IMO, Uyoga S, Gitonga JN, et al (2021) Temporal trends of SARS-CoV-2 seroprevalence during the first wave of the COVID-19 epidemic in Kenya. Nat Commun 12:3966
9. Brand SPC, Ojal J, Aziza R, et al (2021) COVID-19 transmission dynamics underlying epidemic waves in Kenya. Science 374:989–994
10. Githinji G, de Laurent ZR, Mohammed KS, et al (2021) Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. Nat Commun 12:4809
11. Githinji G, de Laurent ZR, Mohammed KS, et al (2021) Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. Nat Commun 12:4809
12. Korber B, Fischer WM, Gnanakaran S, et al (2020) Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell 182:812–827.e19
13. Barasa E, Kazungu J, Orangi S, Kabia E, Ogero M, Kasera K (2021) Indirect health effects of the COVID-19 pandemic in Kenya: a mixed methods assessment. BMC Health Services Research. https://doi.org/10.1186/s12913-021-06726-4
14. Gangavarapu K, Latif AA, Mullen JL, et al Outbreak.info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. https://doi.org/10.1101/2022.01.27.22269965
15. Ade C, Pum J, Abele I, Raggub L, Bockmühl D, Zöllner B (2021) Analysis of cycle threshold values in SARS-CoV-2-PCR in a long-term study. Journal of Clinical Virology 138:104791
16. (2004) Emerging Infectious Diseases.
17. Website. Robert-Koch-Institut, Covid-19 Entlassungskriterien aus der Isolierung. htt ps://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/ Entlassmanagement-Infografik.pdf?_blob = publicationFile; doi.org/10.25646/ 6957.3.
18. Githinji G, de Laurent ZR, Mohammed KS, et al (2021) Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. Nat Commun 12:4809
19. Githinji G, de Laurent ZR, Mohammed KS, et al (2021) Tracking the introduction and spread of SARS-CoV-2 in coastal Kenya. Nat Commun 12:4809

Figures

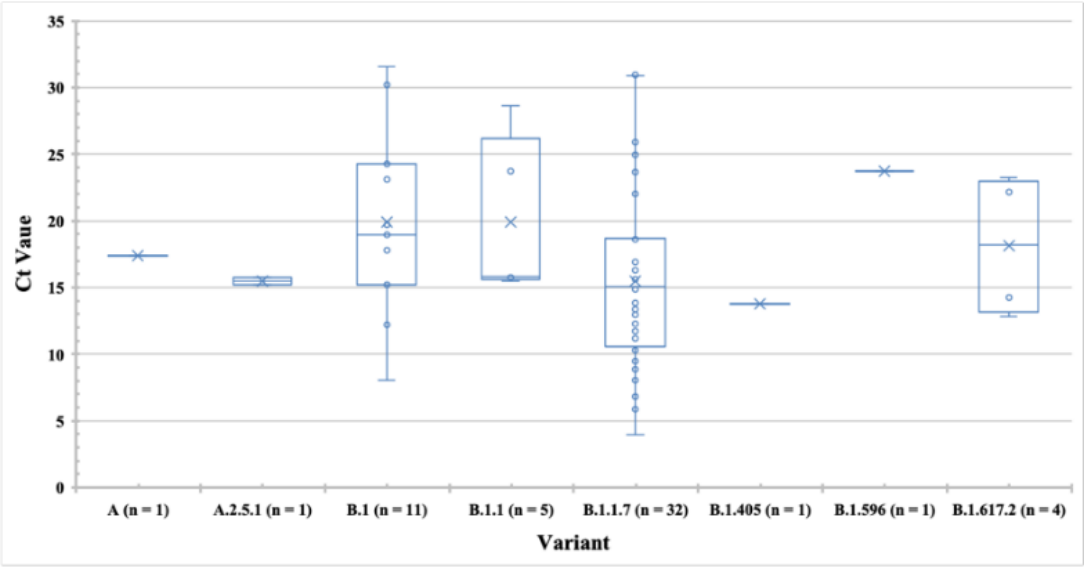

Figure 1. Box-Whisker-Plots of cycle threshold (Ct) values in different variants. n: number of samples related to the variant in the Pangolin lineage.

Figure 1

See image above for figure legend.

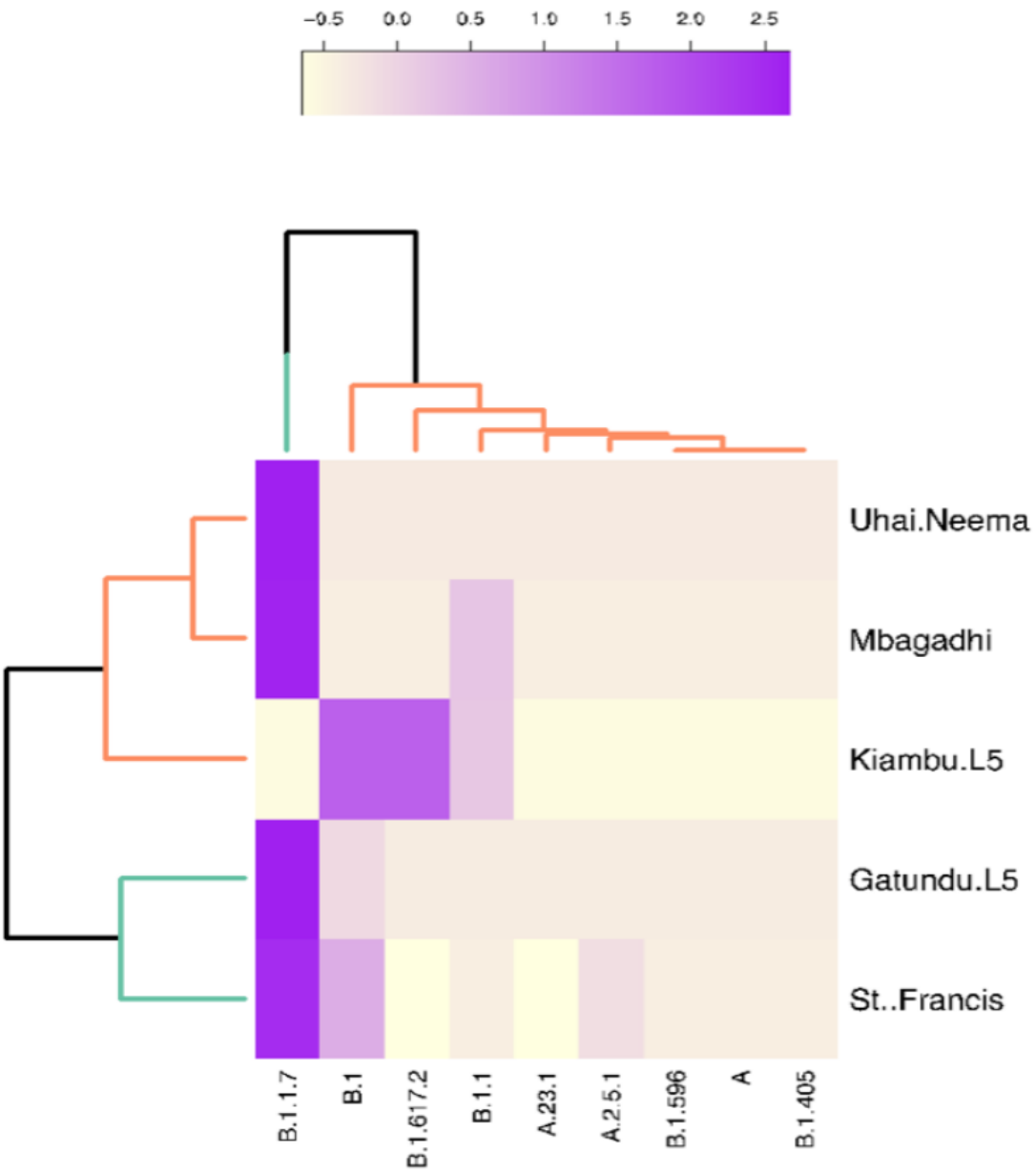

Figure 2. Hierarchical clustering of the identified lineages across Nairobi and Kiambu County. The colour scale represents the frequency of occurrence of specific lineages within a specific site in a scale of -0.5 – 2.5

Figure 2

See image above for figure legend.

Figure 3

See image above for figure legend.

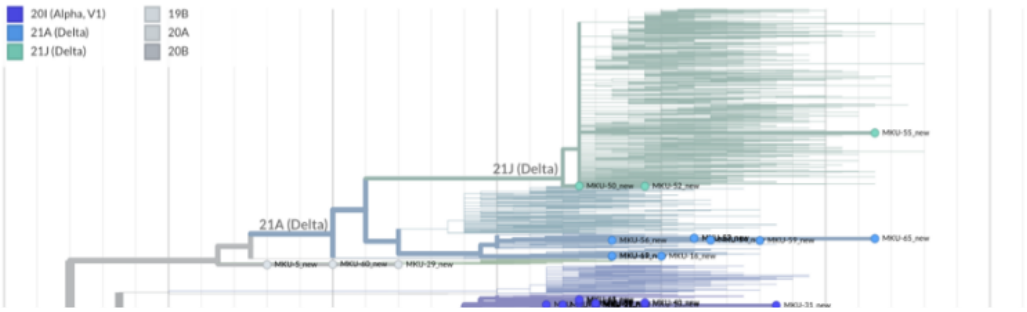

Figure 4

See image above for figure legend.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryfile.docx