Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Insights from SARS-CoV-2 nanopore genomic surveillance in Edo State, Nigeria

Cyril Erameh
Irrua Specialist Teaching Hospital
Ehizojie Ehiremen Emua
Irrua Specialist Teaching Hospital
Emanuele Gustani-Buss
KU Leuven
Joseph Okoeguale
Irrua Specialist Teaching Hospital
Alexandru Tomazatos
tomazatos@bnitm.de
Bernhard Nocht Institute for Tropical Medicine
Racheal Omiunu
Irrua Specialist Teaching Hospital
Sarah Ryter
Bernhard Nocht Institute for Tropical Medicine
Precious Olachukwu Njigha
Irrua Specialist Teaching Hospital
Ganiyu Igenegbale
Irrua Specialist Teaching Hospital
Obehiaghe Benedicta Okogbenin
Irrua Specialist Teaching Hospital
Paulson Ehichioya Ebhodaghe
Irrua Specialist Teaching Hospital
Kékoura Ifono
Bernhard Nocht Institute for Tropical Medicine
Emily Victoria Nelson
Bernhard Nocht Institute for Tropical Medicine
Mette Hinrichs
Bernhard Nocht Institute for Tropical Medicine
Julia Hinzmann
Bernhard Nocht Institute for Tropical Medicine
Christine Jacobsen

Bernhard Nocht Institute for Tropical Medicine

Beate Becker-Ziaja
Bernhard Nocht Institute for Tropical Medicine

Annick Renevey
Bernhard Nocht Institute for Tropical Medicine

N'Faly Magassouba
Centre for Virology Research - Laboratory for Viral Haemorrhagic Fevers in Guinea (CRV-LFHVG)

Sanaba Boumbaly
Centre for Virology Research - Laboratory for Viral Haemorrhagic Fevers in Guinea (CRV-LFHVG)

Jacob Camara
Centre for Virology Research - Laboratory for Viral Haemorrhagic Fevers in Guinea (CRV-LFHVG)

Fara Raymond Koundouno
Bernhard Nocht Institute for Tropical Medicine

Meike Pahlmann
Bernhard Nocht Institute for Tropical Medicine

Lisa Oestereich
Bernhard Nocht Institute for Tropical Medicine

Ikponmwosa Odia
Irrua Specialist Teaching Hospital

Donatus Itsoghiaonode Adomeh
Irrua Specialist Teaching Hospital

Ephraim Ogbaini-Emovon
Irrua Specialist Teaching Hospital

Stephan Günther
Bernhard Nocht Institute for Tropical Medicine

Reuben Agbons Eifediyi
Irrua Specialist Teaching Hospital

Liana Eleni Kafetzopoulou
KU Leuven

Giuditta Annibaldis
Bernhard Nocht Institute for Tropical Medicine

Sylvanus Okogbenin
Irrua Specialist Teaching Hospital

Sophie Duraffour
Bernhard Nocht Institute for Tropical Medicine

Research Article

Keywords: SARS-CoV-2, nanopore sequencing, genomic surveillance, capacity building, Nigeria, virus variant

Posted Date: November 13th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11465448/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

In response to the COVID-19 pandemic, local SARS-CoV-2 nanopore sequencing capacity was established at Irrua Specialist Teaching Hospital (ISTH), a Lassa fever reference centre based in Edo State, Nigeria. This was enabled through training of local staff in Nigeria and Guinea in both laboratory and bioinformatic SARS-CoV-2 sequencing protocols. Between March 2020 and December 2023, the laboratory at ISTH conducted 29,643 diagnostic tests for SARS-CoV-2, of which 4,575 (15.4%) were positive. Between April and December 2022, 410 SARS-CoV-2 positive samples were selected and retrospectively sequenced in 19 runs. From these, 213 near-complete (> 93%) viral genomes were generated and submitted to GISAID. Phylogenetic and phylogeographic analyses supported that Nigeria served simultaneously as a source and recipient of SARS-CoV-2 lineages, driving global spread of the AY.36 Delta variant and receiving Omicron sublineages (BA.1.1 and BQ.1) from North America and Europe. Here we show that a pre-existing long-term laboratory capacity development program for Lassa virus, a high-risk pathogen, facilitated the successful integration of novel workflows for a new emerging pathogen.

Introduction

Since the emergence of SARS-CoV-2 in late 2019, almost 780 million infections have been reported, resulting in more than 7.1 million deaths and severe socio-economic disruption across the world [1]. The evolution of SARS-CoV-2 during the pandemic gave rise to successively dominant lineages such as Alpha, Delta, and Omicron, which displaced earlier strains through increased transmissibility and/or immune escape[2]. Variants of concern (VOC) and variants of interest (VOI) were defined based on their potential for spread, lineage replacement, and disease severity[3]. Between March 2020 and March 2022, four SARS-CoV-2 infection waves have swept across West Africa, each of them driven by distinct virus variants[2]. The successive lineage waves driven by emerging variants placed major strain on healthcare systems worldwide, particularly in resource-limited settings, where diagnostic and genomic surveillance capacity was limited and unevenly distributed[4].

Despite the considerable disruptions of the pandemic, COVID-19 catalysed the establishment of diagnostic and sequencing facilities across the continent[5]. Efforts into surveillance and diagnostic upscaling were supported by targeted training in sequencing and bioinformatics, aimed at building local expertise and strengthening real-time genomic surveillance across the continent [6,7]. However, a major obstacle in disease control remained local access to diagnostic and surveillance capabilities. Retrospective assessment of global publicly-accessible SARS-CoV-2 genomic data revealed a stark contrast in sequencing intensity and timely variant detection between lower- or medium-income countries (LMICs) and high-income countries (HICs), with the latter producing twice as more sequencing data than the former[8]. Effective genomic surveillance depends on both the spatial and temporal coverage of sampling and on sequencing turnaround time, therefore limited diagnostic and sequencing

infrastructure had substantial consequences for COVID-19 surveillance, particularly in low-resource settings.

Nigeria recorded nearly 270,000 confirmed COVID-19 cases, with the majority reported in Lagos state. Over 3,000 of these were fatal, representing approximately 2 deaths per 100,000 individuals [9]. The Irrua Specialist Teaching Hospital (ISTH), a federal tertiary hospital in Edo State, is a national reference centre for Lassa fever (LF) and other infectious diseases. ISTH is located in a LF endemic region and has developed laboratory and clinical expertise in its surveillance, diagnosis and case management, in addition to other emerging infections, in a resource-limited setting [10–14]. Capitalising on its expertise on LF, ISTH adapted to the new epidemiological situation and in February 2020 became one of the first laboratory for SARS-CoV-2 molecular testing in the country [15]. Nanopore genomic surveillance was first initiated in-country in 2018 for Lassa virus and efforts to build and establish local genomic surveillance capacity were ongoing until 2020, at which point they had to be put on hold due to the start of COVID-19 [10]. During this period (2018–2020), staff were primed to nanopore sequencing protocols through training events in Nigeria and Germany. Capitalizing on this pre-existing local competence and building on the successful SARS-CoV-2 establishment of a genomic surveillance unit in Guinea [16], a genomic surveillance unit for SARS-CoV-2 was initiated in 2022 at ISTH, Nigeria.

Materials and Methods

Samples, viral RNA extraction and SARS-CoV-2 diagnostic at ISTH

Nasopharyngeal swabs were collected as part of the in-country COVID-19 surveillance and processed at ISTH for real-time reverse transcription PCR (RT-PCR) as part of routine diagnostic. For viral RNA extraction, QIAamp Viral RNA Mini Kit (Qiagen, Germany) and ZYMO Quick-RNA™ kit (Zymo Research, USA), was used according to manufacturer's instructions. Different RT-PCR kits were used for COVID-19 diagnostics, depending on availability; all kits were used according to manufacturers' instructions and included: (i) Novel Coronavirus (2019-nCoV) Nucleic Acid Diagnostic Kit (Sansure Biotech Inc, China), (ii) COVID-19 test kit GeneFinder™ (OSANG Healthcare Co. Ltd, Korea), (iii) RealStar® SARS-CoV-2 RT-PCR Kit 1.0 (altona Diagnostics, Germany) and (iv) Liferiver® Novel Coronavirus (2019-nCoV) Real Time Multiplex RT-PCR Kit (Shanghai ZJ Bio-Tech, China). QuantStudio™ 5 (Thermo Fisher, U.S.) and the Rotor Gene Q® (Qiagen) were used as real-time thermocycling platforms. Nucleic acid of selected SARS-CoV-2-positive samples was re-extracted with the Qiagen kit before sequencing and cycle threshold (Ct) values were re-evaluated with the RealStar® SARS-CoV-2 RT-PCR Kit 1.0 (altona Diagnostics, Germany). All SARS-CoV-2-positive samples used for identification of virus variants covered a period spanning from March 2020 to October 2022.

Genome sequencing and variant identification

Amplicon-based sequencing was carried out using the Midnight RT-PCR Expansion kit (EXP-MRT001, Oxford Nanopore Technologies (ONT), United Kingdom) as described previously [16]. Each sequencing run included maximum of 24 samples per flow cell. Samples were selected based on the Ct value range (< 35) and the availability of material in storage. For each sequencing run, samples were batched within a maximum Ct range of 5 (e.g., samples with Ct from 20 to 25 were pooled together in one sequencing run). Libraries were prepared from amplified cDNA and sequenced for approximately 24 hours on a MinION MK1C device (ONT). Raw sequencing data was demultiplexed and basecalled with Guppy (v.5.0.16) using high accuracy settings followed by convertion to fastq files. Consensus genomes were generated using different available versions of the epi2me™ ARTIC SARS-CoV-2 pipeline (*wf-artic* workflow) (https://epi2me.nanoporetech.com/epi2me-docs/workflows/wf-artic/). SARS-CoV-2 lineages were assigned using the Pangolin tool [17] and clades were identified using Nextclade [18]. The study dataset for variant identification comprised sequences generated at ISTH with sufficient genome recovery (> 93%).

Data curation, phylogenetic and phylogeographic analyses

Publicly available genomes and associated metadata available for Nigeria, covering an interval from January 2020 to September 2023, were downloaded from GISAID[19] (5 February 2026). To provide global phylogenetic context, sequences from six global regions were included and subsampled per each region per epidemiological week to ensure a more balanced spatio-temporal representation and reduce sampling bias. Only sequences assigned to the lineages of interest, Delta (AY.36), Omicron BA.1 (BA.1.1), and Omicron BA.5 (BQ.1.1 and BE.1.1) were retained for downstream analyses.

Sequences were aligned with Nextclade CLI (v3.7.1), after which the alignment was inspected and polished in AliView [20]. Maximum likelihood phylogenetic trees were inferred using IQ-TREE2 (v.3.0.1) under the best-fit substitution model selected by ModelFinder. Temporal signal was assessed by root-to-tip regression in TempEst (v1.5.3)[21] and TreeTime [22]. Sequences identified as outliers (i.e. significantly deviating from the molecular clock expectation identified by TreeTime) were excluded prior to Bayesian phylogeographic analysis. For each lineage-specific clade dataset, the number of sequences contributing to the downstream analysis per continental region was as follows: Delta (AY.36): Nigeria (n = 67), Asia (n = 2), E, Europe (n = 363), North America (n = 60), South America (n = 20), Oceania (n = 0); Omicron BA.1 Omicron BA.1 (BA.1.1): Africa (n = 2), Asia (n = 34), Europe (n = 179), North America (n = 160), South America (n = 20), Oceania (n = 3); Omicron BA.5 (BQ.1.1 + BE.1.1): Africa (n = 7), Asia (n = 5), Europe (n = 199), North America (n = 46), South America (n = 24), Oceania (n = 1).

Discrete spatiotemporal diffusion was reconstructed for supported subtrees corresponding to the predominant lineages of the Delta (AY.36), Omicron BA.1 (BA.1.1), and Omicron BA.5 (BE.1.1 and BQ.1.1) variants of concern using BEAST X (v10.5)[23] coupled with BEAGLE v4.0.0 [24] to accelerate likelihood computation. Nucleotide substitution was modelled under the GTR model with gamma-distributed among-site rate variation. Evolutionary rates were estimated under a strict molecular clock with a fixed

substitution rate of 8×10^{-3} substitutions per site per year [25]. Demographic history was modelled using a non-parametric Bayesian Skygrid coalescent prior informed by a root height prior distribution. Viral dispersal among geographic locations was inferred through an asymmetric discrete-state diffusion model implemented as a continuous-time Markov chain (CTMC), combined with Bayesian stochastic search variable selection (BSSVS) to identify statistically supported migration routes. The significance of transition rates was evaluated using Bayes factor tests (BF > 10), with posterior inclusion probabilities reported as state posterior probabilities (SPP). Markov jump histories between locations were extracted using the TreeMarkovJumpHistoryAnalyzer [26] utility within BEAST X. Each Markov chain Monte Carlo (MCMC) analysis was run for 2×10^8 steps, sampling every 10,000 steps. Convergence and mixing were assessed by inspecting effective sample sizes (ESS > 200) for all parameters in Tracer (v1.7.2) [27], and independent runs were combined using LogCombiner. Maximum clade credibility (MCC) trees were summarised from the posterior tree distribution using TreeAnnotator, with a 10% burn-in. All phylogenetic trees were visualised and annotated using the *ggtree* package (v3.x) in R environment [28].

Results

SARS-CoV-2 molecular diagnostics at ISTH

Following the emergence of COVID-19, SARS-CoV-2 diagnostic testing was quickly implemented at ISTH. Between March 2020 and December 2023, the laboratory conducted a total of 29,643 tests for SARS-CoV-2, of which 4,575 (15.43%) were confirmed as positive (Fig. 1, **Table S1**). Overall, ISTH diagnostic testing efforts represented approximately 0.5% of the 5.9 million SARS-CoV-2 tests conducted nationally in Nigeria during the pandemic (data extracted from the Nigeria Centre for Disease Control (NCDC) report dated 26 February 2023). The positive cases detected at ISTH (n = 4,575) accounted for approximately 1.7% of all confirmed cases nationwide (n = 266,313) during the reporting period[29]. Diagnostic testing was more frequent in Q2 and Q3 of 2020 (May-August), Q1 of 2021 (January-March) and Q3 and Q4 of 2021 (July-December), while from Q4 of 2022 (October-December) it decreased steadily to minimal testing activity (Fig. 1, **Table S1**). The majority of the tested samples originated from Edo State, with fewer samples originating from other states, mainly Delta State (Fig. 1, **Table S1**). The tested population was skewed toward males (62.8% versus 37.2% tested females) (**Figure S1**). Positivity rates varied across the testing period, reaching a peak of 38.0% of positive tested samples in December 2021 (Fig. 1**).**

Establishment of SARS-CoV-2 genomic sequencing at ISTH

The development and feasibility of nanopore sequencing capacity at ISTH was initiated as a pilot study during the 2018 Lassa fever outbreak in Nigeria [10], and further developed in 2019–2020 with improved laboratory and bioinformatic workflows to advance Lassa virus molecular epidemiology using metagenomic nanopore sequencing. During this period, laboratory staff underwent hands-on trainings

locally in Nigeria and in Hamburg (Germany), focusing on metagenomic nanopore sequencing and basic command line practices. These foundational efforts established relevant infrastructure and basic local expertise that could be redirected for pandemic response, hence providing the groundwork for the setup of amplicon-based SARS-CoV-2 genomic surveillance. To develop this capacity, a total of six ISTH laboratory staff, including one previously trained on metagenomic sequencing, underwent a two-week training in both laboratory and bioinformatics workflows at the *Centre de Recherche en Virologie – Laboratoire des Fièvres Hémorragiques Virales de Guinée* (CRV-LFHVG) in Conakry, Guinea, in November 2021 and March 2022 [16]. This collaboration facilitated knowledge transfer, strengthened regional laboratory networks in West Africa, and provided hands-on experience in sequencing and bioinformatic analysis. Based on these joint trainings, the ISTH sequencing laboratory became fully operational on 21 April 2022. Between April and December 2022, a total of 19 sequencing runs were performed on 410 SARS-CoV-2 positive samples. Of these runs, 13 generated 213 genomes with > 93% recovery (median Ct value: 21.3, range: 12.5–32), which were further deposited on GISAID (Fig. 2, **Figure S2, Table S1**). As the establishment of the in-country SARS-CoV-2 genomic surveillance capacity coincided with the decline in positivity rate in 2022, the majority of the sequencing runs were conducted on retrospective samples. The 213 genomes represented 4.66% of the total confirmed cases at ISTH (**Figure S2**). All 213 near-complete SARS-CoV-2 genomes were suitable for inclusion in downstream phylogenetic analysis. From April to December 2022, four reports integrating results of SARS-CoV-2 sequencing were compiled and shared with the Ministry of Health to support national genomic surveillance of COVID-19 (Fig. 1).

Variants of concern and lineage dynamics

Among the SARS-CoV-2 sequences obtained from Nigeria, the most frequently detected lineages were BA.1.1 (21K Omicron, 21.13%, n = 45), followed by BE.1.1.1 (22B Omicron, 12.21%, n = 26), AY.36 (21J Delta, 10.8%, n = 23), and B.1.1 (20B, 7.98%, n = 17) (Fig. 2, **Table S3**). The remaining lineages, including BQ.1.12 (22E Omicron, 6.57%, n = 14), BE.1.1 (22B Omicron, 6.1%, n = 13), BQ.1 (22E Omicron, 5,16%, n = 11), and BA.1.15 (21K Omicron, 5.16%, n = 11), were detected at lower frequencies. Given the predominance of specific lineage groups, phylogeographic analyses focused on the three most epidemiologically relevant clusters with sufficient sample size for robust spatial diffusion estimates: AY.36 (dominant Delta variant), BA.1.1 (largest Omicron BA.1 sub-lineage), and BQ.1 combined with BE.1.1 (most prevalent BA.5-derived sub-lineages from later Omicron waves). These lineages clustered with high statistical support (ultrafast bootstrap > 95%) in our dataset of 7,511 genomes.

The analysis of Delta sub-lineage AY.36 offered robust evidence for Nigeria as a major source population (posterior probability support [PPS] = 1), corroborating previous research indicating the predominance of this lineage within Nigeria [30]. Bayesian phylogeographic analysis also supported Nigeria as the major source of AY.36 global dissemination, with a total of 76 Markov jump-inferred transition events to Europe [95% HPD: 64– 86], 30 to North America [95% HPD: 22–36] and 15 within the study site [95% HPD: 12–17]. Fewer transitions were inferred to Asia (median = 2) and South America (median = 1). The earliest estimated export from Nigeria to Europe occurred around late March 2021 (mean = 29 March 2021 [95%

HPD: 28 February–24 April 2021]), followed by exports to North America (mean = 19 May 2021 [95% HPD: 16 April–14 June 2021]) and Asia (mean = 28 June 2021 [95% HPD: 23 May–8 November 2021]) (Fig. 3, **Figure S3**).

In contrast, the BA.1.1 Omicron sub-lineage exhibited a distinct dispersal pattern, with North America identified as the most probable ancestral location (PPS = 0.73). The earliest estimated introduction from North America to the study site occurred around late April 2021 (mean = 27 April 2021 [95% HPD: 10 February–26 August 2021]), followed by subsequent introductions from Europe (median = 15 transitions [95% HPD: 8–22]). From the study site, the BA.1.1 lineage underwent substantial onward dissemination, with the majority of transition events (Markov jumps) directed toward Europe (median = 69 [95% HPD: 51–86]), North America (median = 27 [95% HPD: 14–37]), and Asia (median = 15 [95% HPD: 8–20]), with limited spread to Oceania (median = 2 [95% HPD: 1–3]). These findings suggest that following its introduction from North America, the BA.1.1 lineage expanded rapidly within Nigeria before being re-exported to multiple global regions, particularly Europe (Fig. 3, **Figure S3**).

For the BE.1 and BQ.1 Omicron sub-lineages, Europe was inferred as a potential source of the earliest clade (PPS = 0.78). The earliest estimated introduction from Europe to the study site occurred around early March 2022 (mean = 7 March 2022 [95% HPD: 1 January–5 May 2022]), with a total of 12 transitions from Europe to the study site [95% HPD: 7–15]. Additional introductions from Asia were estimated around late January 2022 (mean = 30 January 2022 [95% HPD: 7 August 2021–18 August 2022]). Together, these results suggest that the BE.1 and BQ.1 lineages were likely introduced into Nigeria from European countries, with Europe serving as the dominant source of this Omicron sub-lineage (Fig. 3, **Figure S3**).

Discussion

The SARS-CoV-2 pandemic was a turning point for the development of disease control capacity in LMICs[6]. Aside from increasing diagnostic capabilities, genomic surveillance was boosted substantially, thus changing the landscape of laboratory capacity in Africa and across the word. Substantial disparities between the volumes of pathogen genomic data generated in LMICs and Europe and other high-income settings regions existed before the COVID-19 pandemic. Still present today, they largely reflect inequitable access to laboratory infrastructure and surveillance capabilities [31]. Nigeria was one of the few African countries to surpass the threshold of 0.5% confirmed COVID-19 sequenced cases (nearly 8,000 genomes as of April 2023[32]), highlighting the substantial expansion of the national genomic surveillance capacity during the pandemic [6,8]. During this time, the newly established ISTH nanopore sequencing unit contributed with 2.7% to the total of Nigerian institutional submissions of SARS-CoV-2 genomes to GISAID (as of April 2023[32]), while representing 4.66% of the cases confirmed locally by the ISTH diagnostic unit.

Here we provide an example of how existing capabilities initially developed for field nanopore genomic surveillance of Lassa virus were tailored to respond to a new public health crisis. Development of

sequencing capacity at ISTH was possible through local and international (Guinea) trainings in sequencing and bioinformatic protocols, the latter also facilitated by publicly available tools (e.g. epi2me, Pangolin[17] and Nextclade[18]). Pre-existing robust diagnostic capacity at ISTH rapidly expanded towards SARS-CoV-2 testing enabling ISTH to conduct nearly 30,000 SARS-CoV-2 tests, concomitantly to the upkeep of all other routine diagnostic activities (e.g., Lassa virus and other viral emerging infections). The cause of relatively low incidence in Nigeria and across Africa has been debated and one of the main causes is considered to be underreporting [33–35].

Our study also offers additional information about the SARS-CoV-2 genomic epidemiology in West Africa, highlighting Nigeria's dual role as a source and recipient of SARS-CoV-2 lineage diversity. The country was the primary hub of AY.36 Delta dissemination globally, while acting as a sink for the Omicron sub-lineages (BA.1.1 and BQ.1) introduced later from North America and Europe, respectively. These results are consistent with previous studies indicating that SARS-CoV-2 introductions during transitional and post-pandemic periods shifted from initial emergence locations, potentially due to restriction measures. The shift away from the presumed location of origin is even more pronounced for Omicron variants[36].

Study limitations

These findings should be interpreted in light of several limitations. The down sampling strategies employed to construct the genomic background for the downstream analysis, including the heterogeneous variable sequencing and epidemiological data availability in public databases across the countries, may influence the inferred patterns of viral dispersal and the estimated timing of introduction events [36]. Sample selection for sequencing was made based on the Ct values reported from the diagnostic tests (< 35); this was necessary to increase the genome recovery and phylogenetic signal, but was likely to bias the representation of variant diversity and their dynamic across time. Significant investments were made to establish and expand local sequencing infrastructure, enabling African countries to monitor the genetic evolution of SARS-CoV-2 directly rather than rely on international sequencing [6]. This has the potential to substantially reduce turnaround times between sampling and data dissemination. However, because much of this capacity at ISTH was established while the pandemic was already unfolding, sequencing was performed retrospectively, limiting its applicability to real-time public health decision-making.

Implications for genomic surveillance in LMICs

Alongside the establishment of the SARS-CoV-2 sequencing capacity at ISTH, another program was conducted in parallel in Guinea, where the *Centre de Recherche en Virologie – Laboratoire des Fièvres Hémorragiques Virales de Guinée* (CRV-LFHVG) served as the training hub for both the Guinean and the Nigerian teams [16]. This approach proved effective under local conditions and fostered a productive regional collaboration, as also shown in other contexts in the African region [7]. Despite the rapid expansion of sequencing capacity and pathogen genomic data generation across the continent, the

African Centre for Disease Control has highlighted persistent shortages of skilled bioinformaticians, data scientists, molecular epidemiologists, and infectious disease modelers, as well as challenges in data management and the timely, interoperable, harmonized analysis of genomic data as key barriers to translating data into public health action [31]. Strengthening these analytical capacities and data systems therefore represents priorities for building the capacity of genomic surveillance across the continent. Establishing an autonomous sequencing unit in resource-limited settings revealed key challenges such as unstable power grid, limited internet access, shortages of laboratory reagents and travel constraints for training during the pandemic. Despite these hurdles, the pandemic response enabled the creation of sequencing capacity able to support future genomics surveillance of emerging pathogens in the region beyond the COVID-19 pandemic.

Declarations

Ethical approval and consent to participate

This descriptive research using anonymized diagnostics and genomic surveillance data was under exempt status by the Health Research Ethics Committee (HREC) of ISTH, Nigeria, under the number ISTH/HREC/20250906/728. The ISTH-HREC approved the exemption from obtaining individual consent in accordance with paragraph 32 of the Declaration of Helsinki.

Consent for publication

All authors have given their consent for publication of present work.

Competing Interests

The authors declare that they have no competing interests.

Contributions

Conceived and designed the study: C.E., E.E.E., L.E.K., G.A., S.O., S.D.

Performed sequencing and/or sequence validation: E.E.E., R.O., S.R., P.O.N., G.I., O.B.O., P.E.E.

Data collection and/or analysis: C.E., E.E.E., J.O., A.T., R.O., S.R., I.O., D.I.A., E.O.E., G.A.

Phylogenetic analysis: E.E.E., E.G.B., A.T., S.R.

Phylogeographic analysis: E.G.B

Project implementation: C.E., E.E.E., J.O., S.R., K.I., E.V.N., M.H., C.J., B.B.Z., A.R., N.F.M., S.B., J.C., F.R.K., M.P., L.O., I.O., D.I.A., E.O.E., S.G., R.A.E., L.E.K., G.A., S.O., S.D.

Funding acquisition: M.P., L.O., S.G., S.D.

Wrote the manuscript: C.E., E.E.E., E.G.B., A.T., L.E.K., G.A., S.O., S.D.

Edited the manuscript: all authors.

All authors read and approved the contents of the manuscript.

Use generative AI and AI-assisted technologies

During the preparation of this work the author(s) used the ChatGPT / free version in order to edit some sentences. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Funding

The work was supported by the German Federal Ministry of Health through support of the WHO Collaborating Centre for Arboviruses and Haemorrhagic Fever Viruses at the Bernhard-Nocht-Institute for Tropical Medicine (agreement ZMV I1-2517WHO005 to S.G.); the Global Health Protection Program (GHPP, agreements ZMV I1-2517GHP-704, ZMVI1-2519GHP704, and ZMI1-2521GHP921 until end of 2022 to S.G., and 2023–2025 agreements ZMI5-2523GHP006 to S.G. and ZMI5-2523GHP008 to S.D., and 2026-ongoing ZMBII2-2525GHP008 and ZMBII2-2525GHP006 to S.D.), the COVID-19 surge fund (BMG ZMVI1-2520COR001) to S.G., the Research and Innovation Programme of the European Union under H2020 grant agreement n°871029-EVA-GLOBAL to S.G. L.E.K. acknowledges the Research Foundation - Flanders (Fonds voor Wetenschappelijk Onderzoek – Vlaanderen, G005323N, 1SH2V24N, 12X9222N). E.G-B. acknowledges funding from EU Horizon 2020 grants MOOD (H2020–347 874850). The funders had no role in the study design, collection, analyses or interpretation of data, in the writing of the manuscript or in the decision to publish the results.

Author Contribution

Conceived and designed the study: C.E., E.E.E., L.E.K., G.A., S.O., S.D.Performed sequencing and/or sequence validation: E.E.E., R.O., S.R., P.O.N., G.I., O.B.O., P.E.E.Data collection and/or analysis: C.E., E.E.E., J.O., A.T., R.O., S.R., I.O., D.I.A., E.O.E., G.A.Phylogenetic analysis: E.E.E., E.G.B., A.T., S.R. Phylogeographic analysis: E.G.BProject implementation: C.E., E.E.E., J.O., S.R., K.I., E.V.N., M.H., C.J., B.B.Z., A.R., N.F.M., S.B., J.C., F.R.K., M.P., L.O., I.O., D.I.A., E.O.E., S.G., R.A.E., L.E.K., G.A., S.O., S.D. Funding acquisition: M.P.,

L.O., S.G., S.D.Wrote the manuscript: C.E., E.E.E., E.G.B., A.T., L.E.K., G.A., S.O., S.D. Edited the manuscript: all authors.All authors read and approved the contents of the manuscript.

Acknowledgement

We thank the staff of the Irrua Specialist Teaching Hospital who were involved in the diagnostics, the care and the management of patients with SARS-CoV-2 during the pandemic period.

Data Availability

The genome sequences generated in the present study can be accessed through the persistent digital object identifier https://doi.org/10.55876/gis8.261103wf and the GISAID identifier EPI_SET_ 261103wf. The genome sequences available publicly and used for phylogenetic contextualization of newly obtained sequences can be accessed through the persistent digital object identifier https://doi.org/10.55876/gis8.261103yf and the GISAID identifier EPI_SET_261103yf. The complete sequence dataset used in the analysis can be accessed through the persistent digital object identifier https://doi.org/10.55876/gis8.261103xv and the GISAID identifier EPI_SET_ 261103xv.

References

1. Smit J, Nacer E, Sikorski A, Goddard C, Magdziarz W. Social and Economic Consequences of COVID-19. *Publication for the special committee on COVID-19 pandemic: lessons learnedand recommendations for the future (COVI), Policy Department for Economic, Scientific and Quality of Life Policies, European Parliament, Luxembourg* (2023).
2. Uraki R, Korber B, Diamond MS, Kawaoka Y. SARS-CoV-2 variants: biology, pathogenicity, immunity and control. Nat Rev Microbiol. 2026;24:8–28.
3. World Health Organisation. Tracking SARS-CoV-2 variants. *www.who.int* https://www.who.int/activities/tracking-SARS-CoV-2-variants (2023). Accessed 30 September 2026.
4. Silva SJR. da. SARS-CoV-2 detection and genomic surveillance: moving to ensure accessible tools in low-income and middle-income countries. Lancet Microbe 5, (2024).
5. Wilkinson E, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science. 2021;374:423–31.
6. de Oliveira T, Baxter C. Investing in Africa's scientific future. Science. 2024;383:eadn4168.
7. Agboli E, et al. Building Pathogen Genomic Sequencing Capacity in Africa: Centre for Epidemic Response and Innovation Fellowship. Trop Med Infect Disease. 2025;10:90.
8. Brito AF, et al. Global disparities in SARS-CoV-2 genomic surveillance. Nat Commun. 2022;13:7003.
9. Tsueng G, et al. Outbreak.info Research Library: a standardized, searchable platform to discover and explore COVID-19 resources. Nat Methods. 2023;20:536–40.

10. Kafetzopoulou LE, et al. Metagenomic sequencing at the epicenter of the Nigeria 2018 Lassa fever outbreak. Science. 2019;363:74–7.
11. Ehichioya DU, et al. Phylogeography of Lassa Virus in Nigeria. J Virol. 2019;93:e00929–19.
12. Ajogbasile FV, et al. Real-time Metagenomic Analysis of Undiagnosed Fever Cases Unveils a Yellow Fever Outbreak in Edo State, Nigeria. Sci Rep. 2020;10:3180.
13. Ogbaini-Emovon E, et al. Virus Load Kinetics in Lassa Fever Patients Treated With Ribavirin: A Retrospective Cohort Study From Southern Nigeria. Open Forum Infect Dis. 2024;11:ofae575.
14. Duvignaud A, et al. Lassa fever outcomes and prognostic factors in Nigeria (LASCOPE): a prospective cohort study. Lancet Global Health. 2021;9:e469–78.
15. Osaigbovo II, Igbarumah IO, Muoebonam EB, Obaseki DE. Setting up a molecular diagnostic laboratory for SARS-CoV-2 testing: Experience of a single centre in a resource-constrained setting. Afr J Lab Med. 2021;10:7.
16. Magassouba N, et al. Two years of SARS-CoV-2 genomic surveillance capacity development in Guinea. Sci Rep. 2026;16:11225.
17. O'Toole Á, et al. Assignment of epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol. 2021;7:veab064.
18. Aksamentov I, Roemer C, Hodcroft E, Neher R, Nextclade. Clade Assignment, Mutation Calling and Quality Control for Viral Genomes. J Open Source Softw 6, (2021).
19. Khare S, et al. GISAID's Role in Pandemic Response. China CDC Wkly. 2021;3:1049–51.
20. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics. 2014;30:3276–8.
21. Rambaut A, Lam TT, Carvalho M, L., Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2016;2:vew007.
22. Sagulenko P, Puller V, Neher RA, TreeTime. Maximum-likelihood phylodynamic analysis. Virus Evol. 2018;4:vex042.
23. Baele G, et al. BEAST X for Bayesian phylogenetic, phylogeographic and phylodynamic inference. Nat Methods. 2025;22:1653–6.
24. Ayres DL, et al. BEAGLE 3: Improved Performance, Scaling, and Usability for a High-Performance Computing Library for Statistical Phylogenetics. Syst Biol. 2019;68:1052–61.
25. Bedford T, et al. Cryptic transmission of SARS-CoV-2 in Washington state. Science. 2020;370:571–5.
26. Lemey P, et al. Untangling introductions and persistence in COVID-19 resurgence in Europe. Nature. 2021;595:713–7.
27. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. Syst Biol. 2018;67:901–4.
28. Yu G, Smith DK, Zhu H, Guan Y, Lam T. T.-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods Ecol Evol. 2017;8:28–36.

29. Nigerian Centre for Disease Control and Prevention. *COVID-19 Situation Report: Bi-Weekly Epidemiological Report 22 - Epi Week 7–8*. covid19.ncdc.gov.ng (2023). Accessed 20 September 2026.
30. Ozer EA, et al. Multiple expansions of globally uncommon SARS-CoV-2 lineages in Nigeria. Nat Commun. 2022;13:688.
31. Mboowa G, et al. Africa in the era of pathogen genomics: Unlocking data barriers. Cell. 2024;187:5146–50.
32. Kono TJY, et al. Genomic diversity and surveillance of SARS-CoV-2 in Nigeria. BMC Genomics. 2025;26:917.
33. Nordling L. Africa's pandemic puzzle: why so few cases and deaths? Science. 2020;369:756–7.
34. Tessema SK, Nkengasong JN. Understanding COVID-19 in Africa. Nat Rev Immunol. 2021;21:469–70.
35. Kreier F. Morgue data hint at COVID's true toll in Africa. Nature. 2022;603:778–9.
36. Tegally H, et al. Dispersal patterns and influence of air travel during the global expansion of SARS-CoV-2 variants of concern. Cell. 2023;186:3277–e329016.

Figures

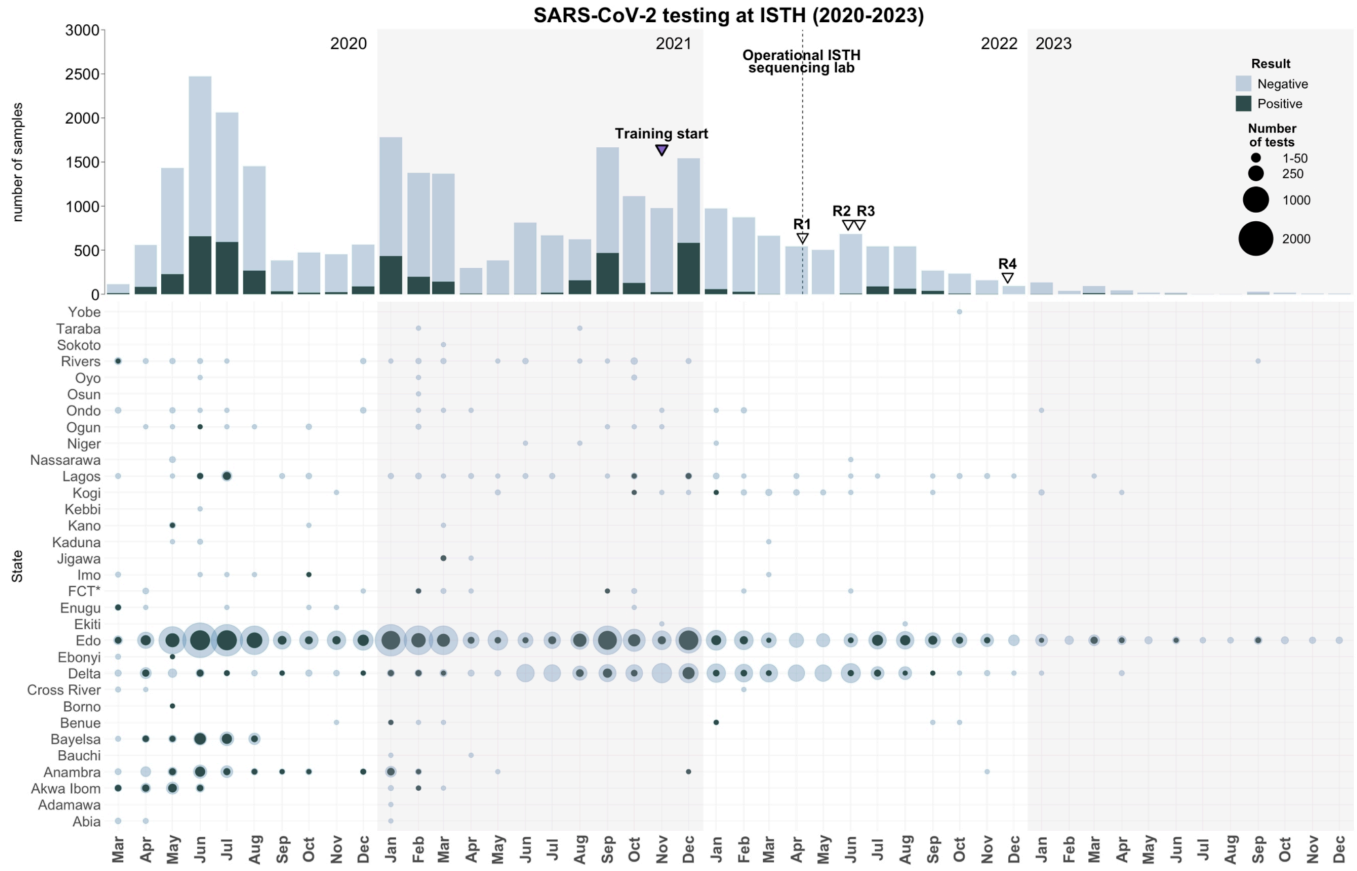


Figure 1

Overview of ISTH diagnostic activities and timeline of sequencing capacity development. The bar plot shows the data contained in Table S1, indicating the number of samples tested monthly for SARS-CoV-2 by real time RT-PCR. Main events in the sequencing capacity implementation are pointed by triangles: training (purple triangle), started in November 2021, inauguration of the ISTH sequencing lab (April 21st 2022) and submission of four variant reports (R1-R4) to Nigerian Ministry of Health (white triangles). The dot plot shows the same results as the bar plot according to the federal states of origin. *FCT: Federal Capital Territory.

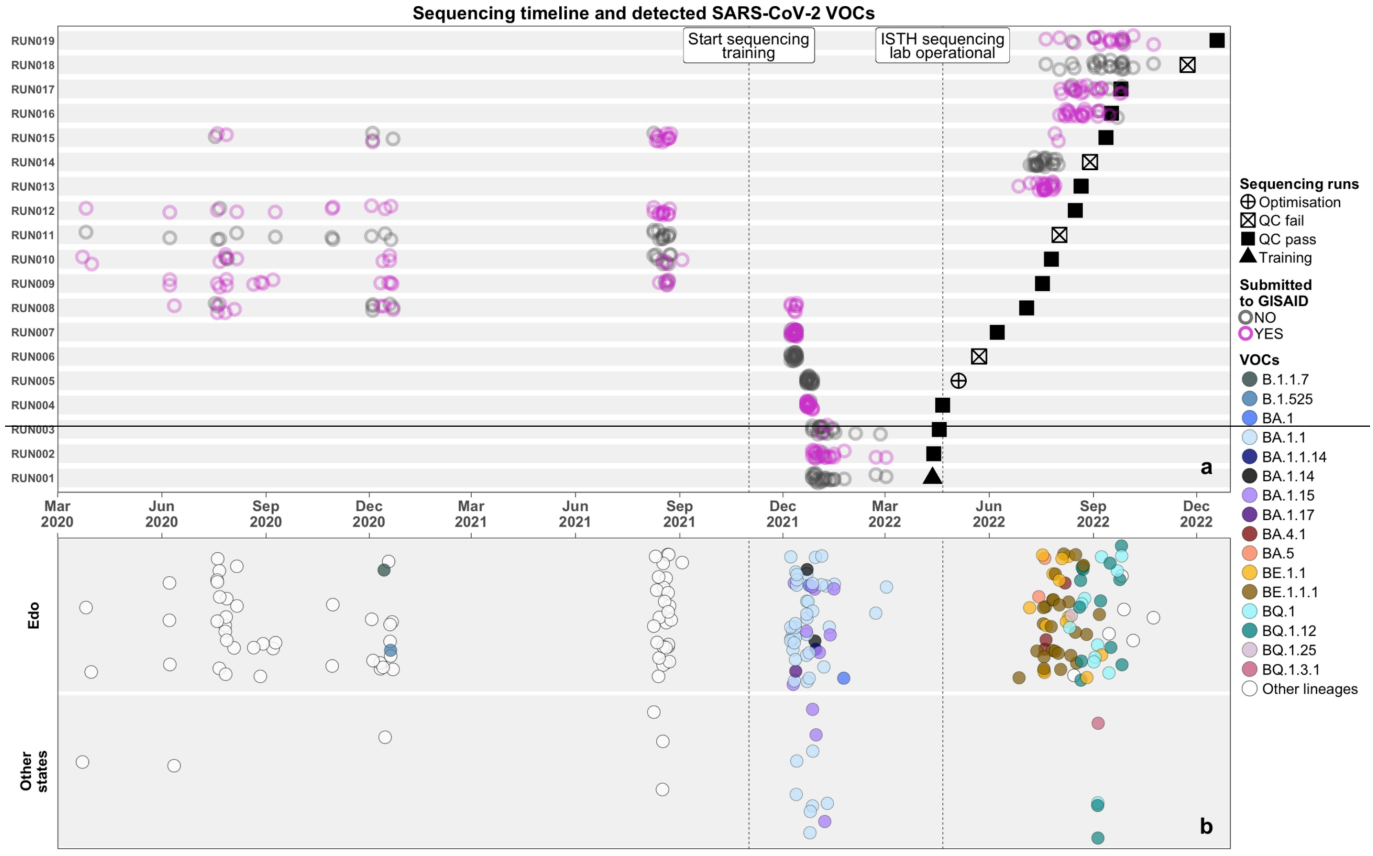


Figure 2

Overview of SARS-CoV-2 genomic surveillance conducted by the ISTH sequencing lab. a) Timeline of sequencing events. Sequencing activities were initiated in November 2021 with an initial training period until formal opening of the sequencing unit in April 2022 at ISTH. The black symbols the right side indicate each individual sequencing run (19 in total). The samples from the training run (solid triangle) were resequenced in the next run indicated above. Samples from runs with failed quality control (QC) were repeated in the subsequent run. The optimisation run was performed when the sequencing kit version was changed. b) Overview of SARS-CoV-2 variants of concern (VOC) identified by the ISTH

sequencing lab from March 2020 to December 2022 using high-coverage (>93%) genomes submitted to GISAID.

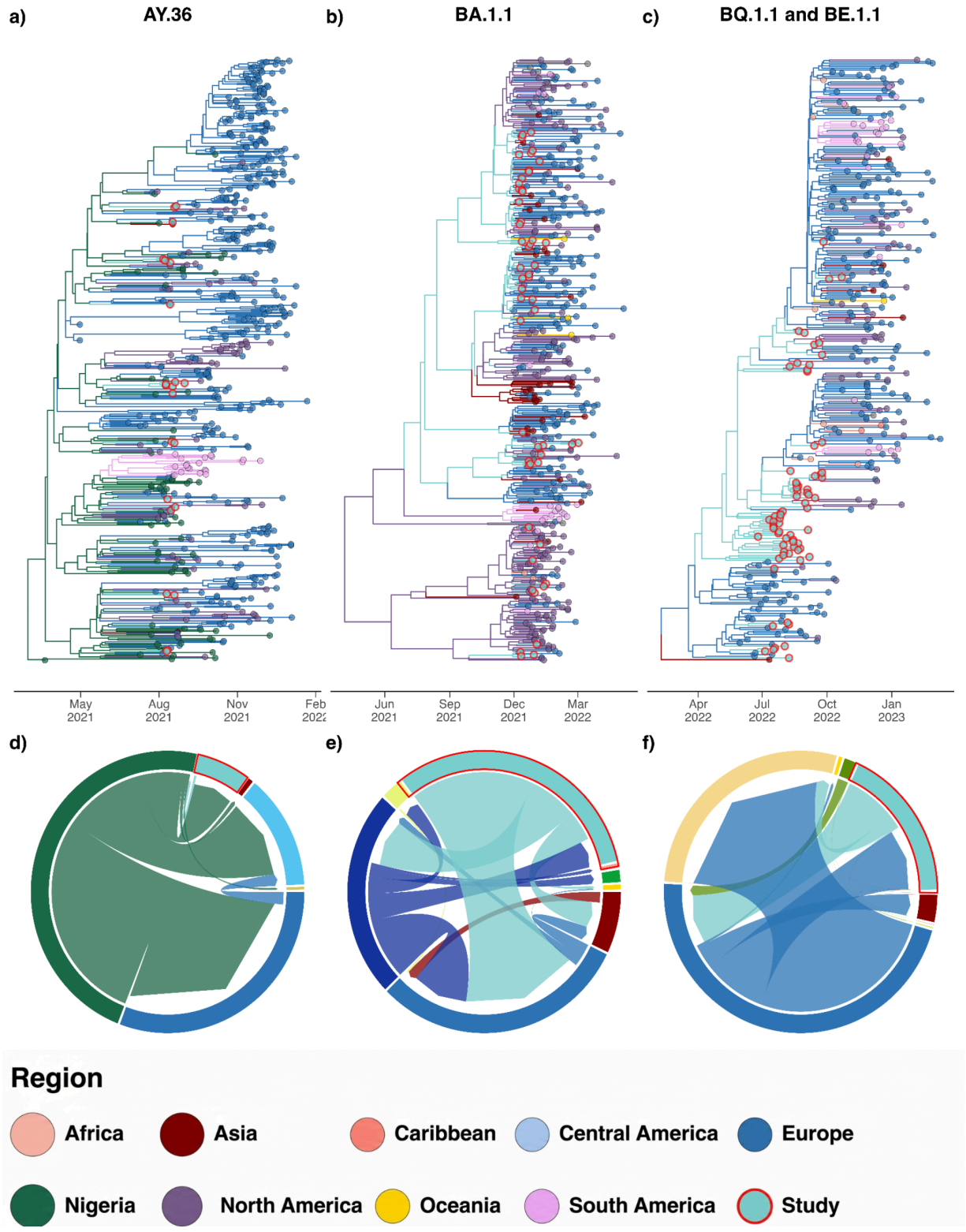


Figure 3

Phylogeographic analysis and migration patterns of SARS-CoV-2 variants in Nigeria. (A-C) Maximum clade credibility (MCC) phylogenetic trees showing evolutionary relationships and geographic

distribution of a) AY.36 (Delta sublineage), b) BA.1.1 (Omicron sublineage), and c) BQ.1.1 and BE.1.1 (Omicron sublineages) variants. Branch colors and tip points indicate geographic regions with Nigerian sequences highlighted in green and study samples in dark slate gray. Time scales display months and years with most recent sampling dates indicated. (d-f) Circle plots below represent virus migration fluxes between geographic regions of d) AY.36 (Delta sublineage), e) BA.1.1 (Omicron sublineage), and f) BQ.1.1 and BE.1.1 (Omicron sublineages) variants. Arrow thickness is proportional to migration strength and colours correspond to source regions.

Supplementary Files

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

- EramehetalNIGCOVIDSeqSupplementaryTables.xlsx
- EramehetalNIGCOVIDSeqSupplementaryFigures.docx