

Emergence and spread of two SARS-CoV-2 variants of interest in Nigeria

Received: 24 September 2022

Accepted: 26 January 2023

Published online: 13 February 2023

 Check for updates

A list of authors and their affiliations appears at the end of the paper

Identifying the dissemination patterns and impacts of a virus of economic or health importance during a pandemic is crucial, as it informs the public on policies for containment in order to reduce the spread of the virus. In this study, we integrated genomic and travel data to investigate the emergence and spread of the SARS-CoV-2 B.1.1.318 and B.1.525 (Eta) variants of interest in Nigeria and the wider Africa region. By integrating travel data and phylogeographic reconstructions, we find that these two variants that arose during the second wave in Nigeria emerged from within Africa, with the B.1.525 from Nigeria, and then spread to other parts of the world. Data from this study show how regional connectivity of Nigeria drove the spread of these variants of interest to surrounding countries and those connected by air-traffic. Our findings demonstrate the power of genomic analysis when combined with mobility and epidemiological data to identify the drivers of transmission, as bidirectional transmission within and between African nations are grossly underestimated as seen in our import risk index estimates.

With a population of over 200 million, Nigeria is the most populous country in Africa¹. Since the first report of SARS-CoV-2 in Nigeria on the 27th of February 2020, the cumulative number of confirmed COVID-19 cases in Nigeria has risen to more than 265 000 as of mid-September 2022². However, this burden is extremely low relative to SARS-CoV-2 infections in the rest of the world. Adjusting for population size, there have been 129 cases per 100,000 people in Nigeria as of mid-September, compared to more than 34,700 and 28,700 in the UK and USA, and 2,340 and 711 in Indonesia and Pakistan, respectively. Concurrently, Nigeria has also only had a little over 3,000 reported deaths (2 per 100,000), compared to the UK and USA, which had 303 and 315 cumulative deaths per 100,000 people². This relatively low but heterogeneous incidence in Nigeria and across the wider African region has been the cause of persistent speculation, including on the putative central role of case underascertainment^{3–5}. The underlying cause remains understudied and is likely multifactorial, with many other factors speculated to contribute, including skewed age structures of populations, more restricted human mobility in certain regions, host genetics, environmental factors, and potential pre-existing population immunity to related viruses^{6,7}. Before the emergence of globally sweeping variants such as Delta and Omicron, it was also unclear if the specific

genetic diversity circulating in the African region may have contributed to the heterogeneous incidence and mortality observed^{8–10}. It is therefore important to characterise the genomic epidemiology of SARS-CoV-2 in Nigeria over the course of the pandemic, to improve our understanding of what variants emerged to dominate the different epidemic waves. It is also important to improve our understanding of the drivers of transmission in the region, as this remains understudied in Africa and findings from other regions may not be generalizable^{9,11}. Nigeria is highly connected to its neighbouring countries and the wider African region, potentially acting as a dominant source of transmission via the high volume of movement across land as well as air borders¹².

Several SARS-CoV-2 variants, which have a constellation of mutations that are biologically significant to the virus, emerged during the pandemic^{9,13,14}. The B.1.525 (Eta) and B.1.1.318 variants of interest (VOIs) dominantly co-circulated with the Alpha variant during the second wave in Nigeria from December 2020 to March 2021². B.1.525 and B.1.1.318 were suggested to have emerged in Nigeria based on epidemiological reports from travellers in countries such as the UK, India, Mauritius, Canada, and Brazil^{15–19}. Notably, B.1.525 and B.1.1.318 both showed a significant increase in the infectivity of human and monkey cell lines experimentally, raising concerns of intrinsic

 e-mail: happic@run.edu.ng

increased transmissibility²⁰. Both of these lineages also share the well-characterised E484K substitution in the Spike protein receptor-binding domain (RBD), which effectively reduces antibody neutralisation²¹.

In this study, we examine the genomic epidemiology of SARS-CoV-2 in Nigeria from March 2020 to September 2021, across the first three epidemic waves. In particular, we investigate the timing and origin of the emergence of the B.1.525 and B.1.1.318 VOIs. In phylogeographic reconstructions, we characterized the source-sink profiles of these best-sampled VOIs to better understand the bidirectional transmission dynamics of SARS-CoV-2 in Nigeria and the wider African region. We compare our findings from our genomic data to integrated travel and epidemiological data to explore the role of Nigeria's regional and intercontinental connectivity in the estimated import and export dynamics.

Results

Lineage dynamics of SARS-CoV-2 in Nigeria over the first three epidemic waves

We generated a total of 1577 genomes from samples obtained between March 2020 and September 2021, across the first three waves of the pandemic (Fig. 1A, B). We collected samples from 25 of the 36 states and the Federal Capital Territory of Nigeria (Fig. 1C). The northern states were highly undersampled relative to the southern states, with sampling highly uneven across time, especially during the third wave (Fig. 1A, B).

To investigate the lineage dynamics across the different outbreak waves, we used the pangolin nomenclature tool to assign all sequences to corresponding lineages. We detected more than 35 lineages across the first three waves (Fig. 1D). We found that a number of ancestral lineages circulated during the first wave, including A, B, B.1, and B.1.1

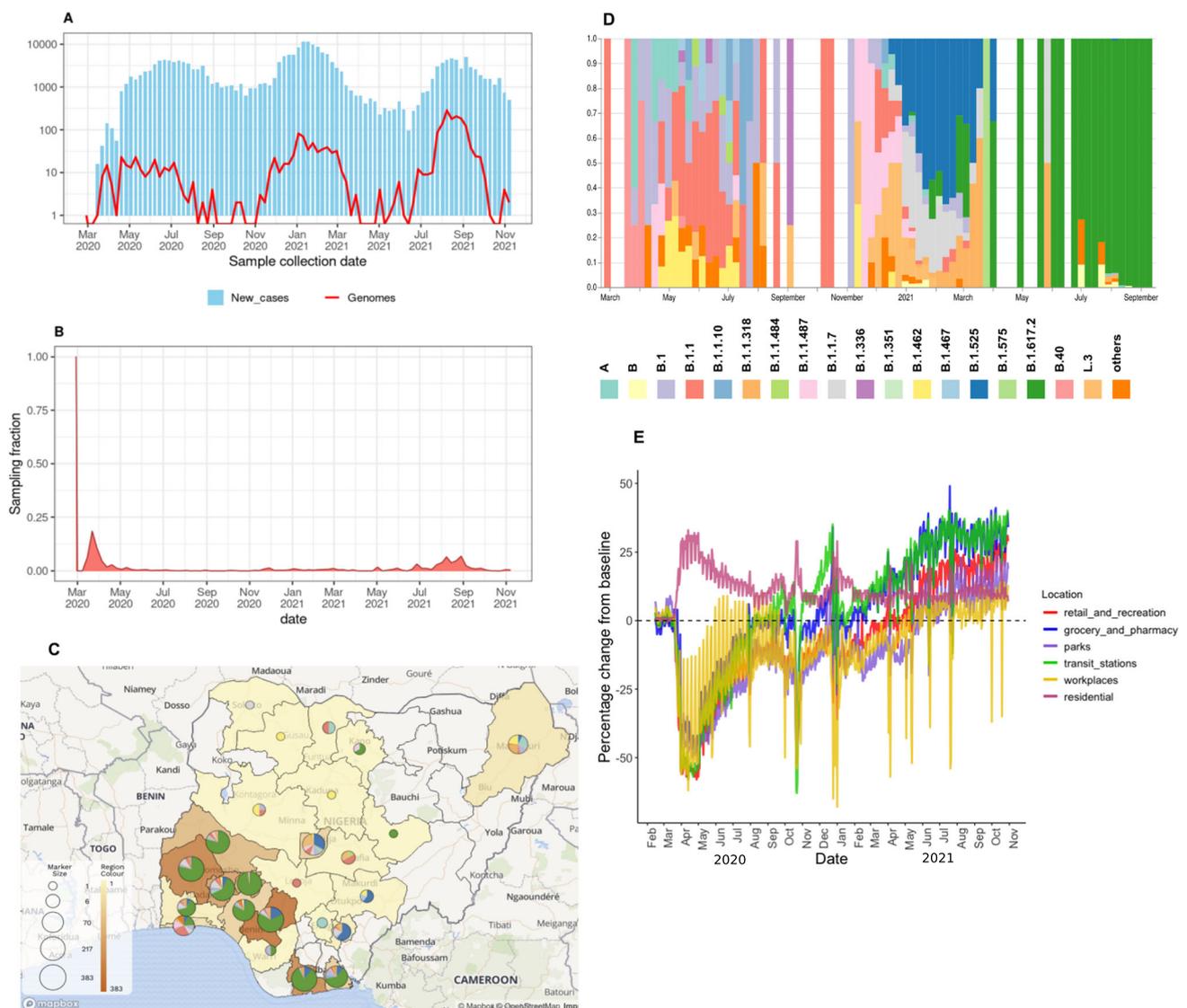


Fig. 1 | Genomic epidemiology of SARS-CoV-2 in Nigeria. **A** Epicurves of SARS-CoV-2 cases in Nigeria and genomes produced assembled during the first three waves with the y-axis transformed to a log-scale. **B** Time-varying sampling fraction of genomes produced in this study per new cases. **C** Geographic distribution of sequences generated in the current study. The map of Nigeria shows the number of genomes from states across the country as region colour and marker size display lineages per state. Maps © Mapbox (www.mapbox.com/about/maps) and © OpenStreetMap (www.openstreetmap.org/about). **D** Lineage frequency

profile for the study period. The frequency of PANGO lineages across the country over a period of 80 weeks. Lineages that are not VOCs or VOIs, or that do not appear more than three times over the course of the pandemic, are grouped as “others”. **E** Human mobility data from Google for retail shops, pharmacies, parks, public transport, workplaces, and residential activity places from January 2020 to November 2021, as compared to the pre-pandemic baseline (January to February 2021).

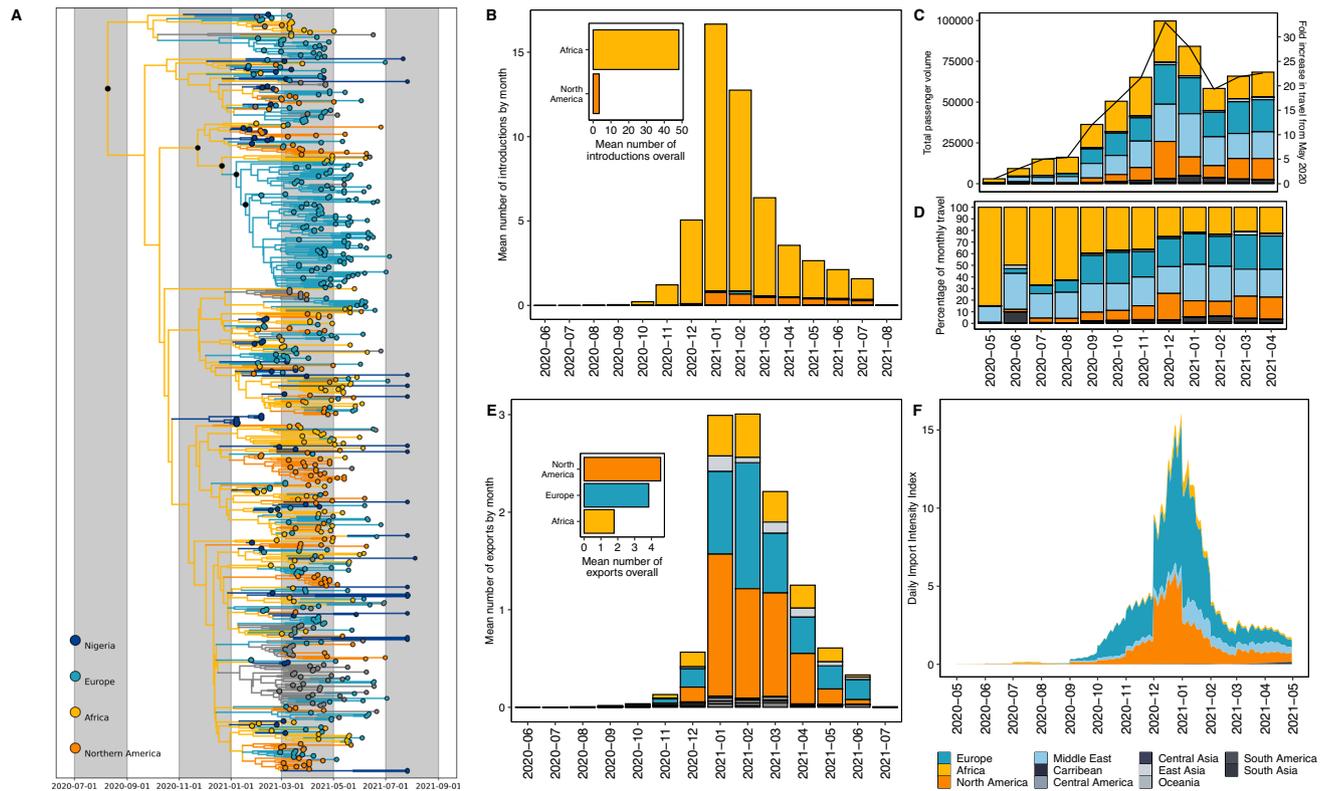


Fig. 2 | Bayesian phylogeographic reconstruction of the B.1.1.318 variant.

A Time-resolved B.1.1.318 phylogeny. Branches are coloured by region-level geographic state reconstruction. Locations with negligible contributions have been grouped and annotated in grey. Internal nodes annotated with black points represent posterior support > 0.75. **B** Mean number of introductions (Markov jumps) of B.1.1.318 into Nigeria from all source regions, binned by month. The mean number of introductions overall (inset). **C** Volume of monthly air passengers

inbound to Nigeria across 2020–2021 by source region (bar, left axis) and fold increase in travel volume compared to May 2020 baseline. **D** Estimated Introduction Intensity Index for Nigeria from 2020 to 2021. **E** Percentage of inbound air travel to Nigeria by source region across 2020–2021. See Supplementary Fig. 2 for country-level data. **F** Estimated Importation Intensity Index (III) for Nigeria from 2020 to 2021.

(Fig. 1D). The onset of community transmission in Nigeria was initially delayed, with the first wave initiated two months after the first case detection in March 2020. A nationwide lockdown and ban on local and international air travel were enforced, with the first wave ending in August 2020 after four months (Fig. 1B). During the lockdown, we found that there was up to a 50% decline in activity at workplaces, retail, supermarkets, and public transport stations, and up to a 25% increase in residential activity (Fig. 1E).

After this strictly adhered-to lockdown, Nigeria opened its air-space to international traffic in October 2020. This was shortly followed by an increase in new cases in November 2020, marking the start of the second wave (Fig. 1B), which was characterised by B.1.1.7, B.1.525, B.1.1.318 (VOC and VOI), and other variants such as L.3 and B.1.1.487 (Fig. 1D).

The following analyses focused on the lineages (B.1.525 and B.1.1.318) that contributed to the surge during the second wave and have also been shown to possess genotypic traits for increased transmissibility and also increased virulence in human and animal cell lines, such as the E484K spike protein RBD mutation²¹. We excluded B.1.1.7 from our analysis as it has already been well studied^{22,23}, and it was not the dominant variant in Nigeria during the second wave, unlike in other places where it was reported, such as the UK and the USA^{24–26}. Moreso, the second wave began after travel restrictions were lifted and human mobility was back to normal, thus the need to investigate the emergence of B.1.1.318 and B.1.525 in Nigeria. The third wave was initiated in June 2021, with the Delta variant (B.1.617.2) and its sub-lineages sweeping to dominance in sampling (Fig. 1B, D). In the other analysis described in this work, we

also sought to identify the timings of emergence and number of introductions of these VOI.

Regional connectivity drove the introduction of B.1.1.318 into Nigeria

Using Bayesian phylogeographic reconstructions, we investigated the timing and origin of the B.1.1.318 emergence to better understand Nigeria's connectivity in the global SARS-CoV-2 transmission networks (see Methods). B.1.1.318 was first detected in Nigeria in Lagos State in December 2020. The lineage was detected in multiple countries within and outside Africa, notably resulting in a large outbreak in Mauritius¹⁹.

In our phylogeographic reconstruction, we found that B.1.1.318 emerged in Africa (root state posterior support = 0.95–6) in early August 2020 [mean tMRCA = 5 August, 95% HPD 25 June to 20 September] across two replicates (Fig. 2A). We estimated that B.1.1.318 was introduced to Nigeria on at least 53 independent occasions [mean introductions, 95% HPD 50–59] beginning in November 2020, after travel restrictions were lifted in October 2020 (Fig. 2B). We found that the majority of introductions originated from other African nations (Fig. 2B), indicating Nigeria's strong connectivity to the region. However, the number and origin of introductions and exports estimated from genomic data are sample dependent, with bidirectional transmission to and from undersampled regions obscured by uneven sampling globally. We quantified air travel patterns to and from Nigeria over time to better understand which countries were most likely to act as transmission sources or sinks based on their connectivity to Nigeria.

Before travel restrictions were lifted in October 2020, we found that the low levels of incoming air travel volume predominantly

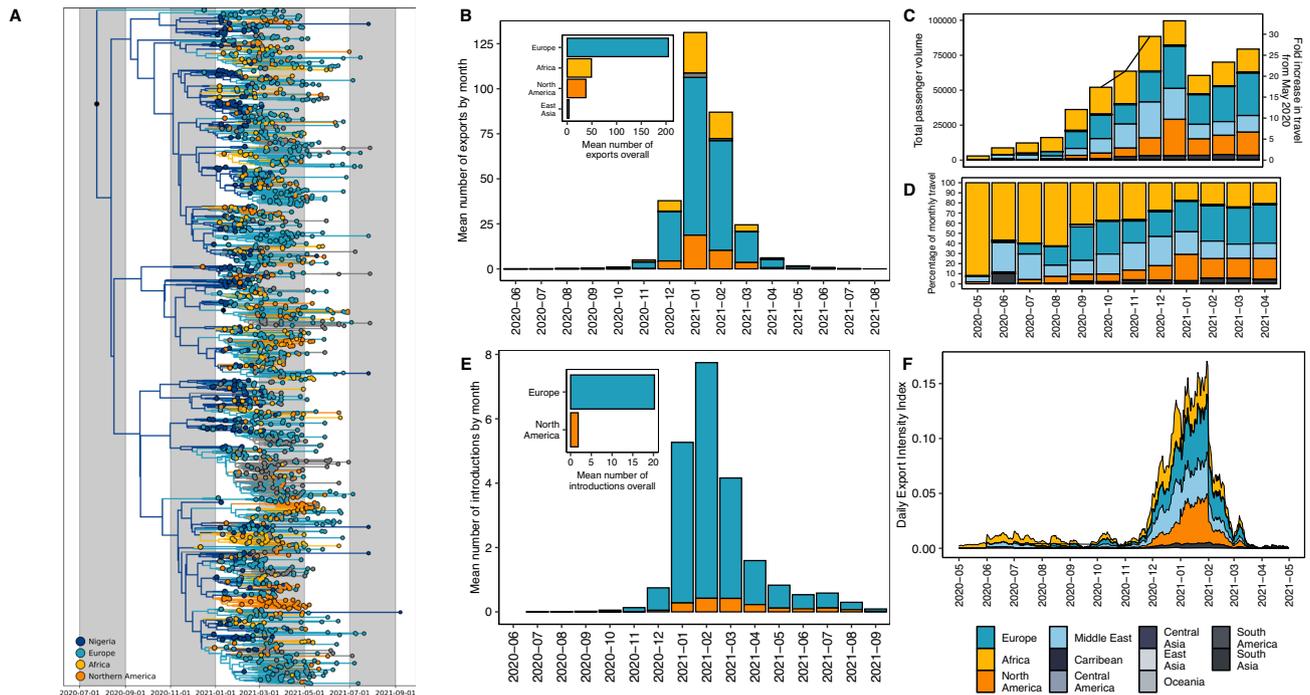


Fig. 3 | Bayesian phylogeographic reconstruction of the B.1.525 variant. A Time-resolved B.1.525 phylogeny. Branches are coloured by region-level geographic state reconstruction. Locations with negligible contributions have been grouped and annotated in grey. Internal nodes annotated with black points represent posterior support > 0.75. **B** Mean number of exports (Markov jumps) of B.1.525 from Nigeria to destination regions, binned by month. Mean number of exports overall (inset). **C** Volume of monthly air passengers outbound from Nigeria across 2020–2021 by

destination region (bar, left axis) and fold increase in travel volume compared to baseline (May 2020). **D** Percentage of outbound air travel from Nigeria by destination region across 2020–2021. See Supplementary Fig. 1 for country-level data. **E** Mean number of introductions (Markov jumps) of B.1.525 into Nigeria from all origin regions, binned by month. The mean number of introductions overall (inset). **F** Estimated Exportation Intensity Index (EII) for Nigeria from 2020 to 2021.

originated from other African nations, followed by the Middle East (driven by the UAE) and Europe (driven by the UK) (Fig. 2C,D, Supplementary Fig. 2). This suggests that the introduction risk was predominantly driven by regional connectivity during the period when travel restrictions were in place. We sought to investigate this further by quantifying the introduction risk for all countries connected to Nigeria by air travel. We estimated the introduction intensity index (III), which accounts for the number of cases in the source countries as well as the travel volume to Nigeria (see Methods). Overall, we found that the introduction risk from air travel was very low during the period of travel restrictions (May 2020–October 2020), as incoming travellers originated largely from other African nations with low incidence (Fig. 2F).

We found that the number of incoming air passengers peaked in December 2020–January 2021 (Fig. 2C, D). We also observed that the number of inbound passengers from Europe, the Middle East, and North America increased relative to other African nations after restrictions were lifted, resulting in comparable volumes from Africa, Europe, and the Middle East (Fig. 2C, D). The 30-fold increase in incoming passengers by December 2020 was reflected in the corresponding peak in the III, attributable to increased introduction risk from Europe and North America from November 2020 into early 2021 (Fig. 2F). The surge was predominantly driven by the high-incidence second waves of the UK and USA, respectively (Supplementary Fig. 1).

We found that the III peak but not the estimated sources of risk were consistent with the number and origin of introductions of B.1.1.318 in our phylogeographic reconstructions (Fig. 2B). We estimated the highest introduction risk for Europe after travel restrictions were lifted, whereas the majority of introductions estimated from genomic data originated in other African countries. The III from other African nations was relatively low, despite comparable incoming air

travel volume. This is likely both a factor of the relatively low incidence in most African countries during this period (Fig. 2F, Supplementary Fig. 3) as well as the fact that our travel data does not account for land-based travel, which will expectedly severely underestimate the introduction risk from surrounding countries. The III from the African region was largely driven by South Africa, which was experiencing a peaked incidence from its second wave (Supplementary Fig. 3). In our III analyses, we found negligible introduction risk from outside of Europe and North America.

The Eta variant of interest (B.1.525) likely emerged in Nigeria

The B.1.525 (Eta) lineage drove the second wave from January to March 2021. B.1.525 was initially reported to have originated in Nigeria or the UK^{16–18}. We investigate the timing, origin, and transmission dynamics of B.1.525 with Bayesian phylogeographic reconstructions (see Methods). We estimated that the B.1.525 lineage emerged in Nigeria (root state posterior support = 0.998) in late July 2020 [mean tMRCA 23 July, 95% HPD 7 June to 4 September], during the final weeks of Nigeria's first wave (Fig. 3A). This suggests that this lineage circulated cryptically for several months before its first detection in Nigeria on 12 December 2020, likely owing to low-incidence-associated sparse sampling (Fig. 1A).

We sought to better understand Nigeria's connectivity in global SARS-CoV-2 transmission networks by identifying the most likely destinations of B.1.525's exportation, as the lineage (1) emerged in Nigeria and (2) is Nigeria's most sampled lineage, excluding the Delta variant sublineages. To investigate the spread of B.1.525 from Nigeria, we reconstructed the timing and pattern of geographic transitions out of and into Nigeria across the full posterior of our Bayesian phylogeographic reconstructions. We estimated that B.1.525 was exported from Nigeria a lower bound of 295 times [mean Markov jumps, 95%

HPD 259-335] (Fig. 3B). The majority of sampled exports were destined for Europe, followed by the rest of Africa and Northern America from December 2020 to March 2021 (encompassing Nigeria's second wave) (Fig. 3B). During this period, we also found that B.1.525 was re-introduced from Europe a lower bound of 20 [95% HPD 6-37] times (Fig. 3E). These estimated source-sink dynamics support Nigeria's strong connectivity, particularly to Europe, but will underestimate exports to undersampled regions such as neighbouring African countries. To mitigate the effect of global sampling biases on genomic estimates of transmission dynamics, we again analysed changes in air travel to and from Nigeria over time to understand the sources and seeds of Nigeria's bidirectional transmission dynamics.

African nations dominated the reduced level of outbound travel during the period of travel restrictions (May to October 2020) (Fig. 3C, D), suggesting that regional connectivity drove Nigeria's export risk during this period. We integrated the air travel data with Nigeria's epidemic incidence data to quantify an exportation intensity index (EII)²⁷. The EII quantifies the temporal trend in the daily estimated number of viral exports from Nigeria to sink countries (see Methods). Given the data available, Nigeria's estimated export intensity was low overall compared to countries with high air traffic data to Nigeria, as the country's SARS-CoV-2 incidence remained comparatively low on a global-scale for the entire period under investigation (Fig. 3F). Nigeria's first wave had relatively low incidence, peaking at about 4000 cases in a week in late June (week 18) (Supplementary Fig. 4B). Combined with reduced travel volumes, the period from May 2020 to September 2020 was characterised by negligible export intensity (Fig. 3F).

Outbound air travel recovered gradually over 2020, peaking in December 2020–January 2021 (Fig. 3C, D). We found that this peak corresponded with the peak of the larger second wave in Nigeria and therefore also the EII, as well as the distribution of exports of B.1.1.525 from Nigeria (Fig. 3B, F). The proportion of outbound air travel destined for countries in Europe, the Middle East, and North America increased relative to travel destined for African countries after travel restrictions were lifted (Fig. 3C, D). We estimated the highest EII for Europe (driven predominantly by travel to the UK) and Northern America (driven by travel to the USA) from December 2020 to February 2021 (Supplementary Fig. 4). The export intensity to Europe is consistent with the high number of estimated exports in the genomic data, though North America was disproportionately underestimated in the genomic data (Fig. 3B). Overall, air travel volume destined for Europe was only moderately higher (10%) than Africa, North America, and the Middle East from December (after first detection and at the start of the second wave) to April 2021 (end of the second wave) (Fig. 3C, D). However, Europe had 4-fold more sampled introductions of B.1.525 than African nations or the USA. Notably, the comparatively high EII for the Middle East region (representing 20% of outbound volume overall) was not represented in our phylogeographic reconstructions (Fig. 3B). This highlights how unevenly distributed surveillance capacity can lead to an underestimation of transmission events (Supplementary Fig. 3A).

Discussion

In this study, we combined genomic, travel and epidemiological data to characterise the emergence and bidirectional transmission of two focal variants of interest in Nigeria. The focal VOIs, B.1.525 and B.1.1.318, were suggested to have emerged in Nigeria before spreading globally, resulting in large-scale outbreaks in Brazil and Mauritius^{15,19}. In our phylogeographic reconstructions, we found that the B.1.1.318 lineage most likely emerged in the African region in early August 2020, with its dominance in Nigeria driven by multiple introductions during the second wave from November 2020 onwards. We also found that the Eta variant (B.1.525) emerged in Nigeria in late July, with high levels of export to Europe reflected in the genomic and estimated export index.

The true extent of bidirectional transmission with other African nations will be severely underestimated in phylogeographic reconstructions, as Africa remains disproportionately undersampled despite heroic efforts²⁸. Our findings should be interpreted in the light of our own sampling fraction, which was 0.026% of reported cases. We attempted to mitigate these global surveillance biases by supplementing genomic estimates with sampling-independent metrics like the introduction and export intensity indices. The III and EII indices should be interpreted to quantify the temporal trend in the daily estimated number of introductions or exports rather than absolute values. The indices are based on the back-extrapolated time series of deaths and are therefore restricted by the associated reporting delays and biases. Notably, they will be underestimated if there is large-scale underascertainment in source countries, as previously shown for African nations, including Nigeria, based on excess mortality data²⁹⁻³¹. Most notably, our import risk index underestimates the importance of regional connectivity in driving introduction estimates, as they likely result from shorter distance connectivity to surrounding countries, which we could not collect data on. One of the biases recorded in our study was that we were not able to distinguish samples based on travel and community testing due to insufficient metadata. Hence, a sensitivity analysis to exclude travel-based testing was not possible. Another limitation is the inability to sequence samples from all the states in Nigeria at the same proportion across time, as the northern states were undersampled during the third wave. Hence, the within-country dynamics could not be determined accurately due to complications that arose relating to sample transfer and data sharing from other partnering labs during the pandemic. This also calls for better data-sharing agreements and guidelines for genomic surveillance of pathogens in Nigeria where trust is a big issue.

At an unprecedented speed (72 h from receiving the sample), ACEGID generated the first whole-genome sequence of the virus in Africa in March 2020. This catalysed and built confidence in SARS-CoV-2 genomic sequencing on the African continent. This immense collaborative effort, involving seventeen partner laboratories in Nigeria, helped to characterise the pandemic in near real-time during its first three waves in the country. As we have seen in recent years, novel variants of interest or concern can emerge from anywhere around the world, notably including undersampled regions such as Africa^{13,14}. In a highly connected world, there is a proactive need to adopt early warning systems for pandemic preemption and response in Africa, such as SENTINEL³². This would enable equality in global genomic surveillance, especially in countries with fragile public health systems, in order to effectively detect and curb emerging outbreaks before they spread.

Methods

Study population/sampling

The study was approved by the National Health Research Ethics Committee of Nigeria (NHREC), with protocol number NHREC/01/01/20017-08/08/2020 and approval number NHREC/01/01/2007-30/11/2021B. Informed written consent was obtained directly from the patient as part of the routine surveillance program in Nigeria, therefore there was no participant compensation. Samples were collected from people who reported to community testing centres for COVID-19 tests (travellers included) and hospitalised individuals from February 2020 to October 2021 across the country. Sampling fraction was deduced by the number of SARS-CoV-2 genome(s) assembled against the number of confirmed cases in a given day.

Sample processing

RNA was extracted from nasopharyngeal swabs in viral transport media and saliva/sputum in PBS using the QiAmp viral RNA mini kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and MagMax pathogen RNA/DNA kit (Applied Biosystems,

Massachusetts, USA) using a Kingfisher Flex purification system (ThermoFisher Scientific, Massachusetts, USA) according to the manufacturer's instructions. RT-qPCR screening of suspected samples was carried out targeting N, ORF1ab and RdRP genes of the virus using commercially available kits: genesig® (Primerdesign Ltd, UK), DaAn-Gene (Daan Gene Co., Ltd, China), Liferiver (Shanghai ZJ Bio-Tech Co., Ltd, China), GeneFinder (OSANG Healthcare Co., Ltd, Korea), and Sansure (Sansure Biotech Inc., China).

SARS-CoV-2 whole-genome sequencing

Random samples with moderate to high viral load detected with RT-qPCR (Ct value <30) were selected for sequencing, which were samples collected from routine sentinel surveillance and hospitalised individuals spatially and temporally. These are people who: (i) had COVID-19 symptoms and reported to testing centres, (ii) were hospitalised or under quarantine due to COVID-19 related complications, (iii) are in-bound or out-bound travellers who had to take mandatory COVID-19 tests, or (iv) were tested based on reported outbreaks within the community. Samples were collected from both private and public COVID-19 testing labs across the country in twenty-six states (Adamawa, Akwa Ibom, Benue, Borno, Delta, Ebonyi, Edo, Ekiti, Enugu, FCT Abuja, Kaduna, Kano, Katsina, Kogi, Kwara, Lagos, Nasawara, Niger, Ogun, Ondo, Osun, Oyo, Plateau, Rivers, Sokoto, and Zamfara states). The sampling gap was the inability to sequence samples from every state during all different waves of the pandemic in Nigeria. The Illumina COVIDseq protocol was used for sequencing preparation (<https://www.illumina.com/products/by-type/clinical-research-products/covidseq-assay.html>). RNA was converted to cDNA, followed by tiling amplification (400 bp) with ARTIC V3 primer pools, which covers the whole genome of the virus. Nextera DNA flex libraries were made from the amplicons. Libraries were pooled and sequenced on the Illumina MiSeq, NextSeq 2000, and NovaSeq 6000 platforms at the African Center of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Ede, Nigeria.

Genome assembly

We used the viral-ngs pipeline v2.1.19 (<https://github.com/broadinstitute/viral-ngs>) for demultiplexing, quality control, and genome assembly, consistent with prior work^{33–35}. The demux_plus applet of viral-ngs was used for demultiplexing basecalls to BAM files. Genome assembly was carried out using the assemble_refbased workflow on the viral-ngs pipeline which maps each unmapped BAM file to the SARS-CoV-2 reference genome (NC_045512.2) to generate coverage plots and FASTA files of successfully assembled genomes.

Lineage/clade assignment

We used Phylogenetic Assignment of Named Global Outbreak Lineages (PANGOLIN)³⁶ v3.1.12 to identify the lineages of SARS-CoV-2 circulating within Nigeria and to identify the lineages circulating in specific states of Nigeria. Nextclade v1.3.0 (<https://clades.nextstrain.org/>) was used to assign the sequences to globally circulating viral clades and to investigate mutations in the genome that could affect primer amplification during PCR. We visualised the lineage dynamics across Nigeria using Microreact³⁷ and a GeoJSON file of the map of Nigeria.

Maximum likelihood phylogenetic analysis

1577 genomes with sequence length $\geq 20,930$ (covering 70% or more of the reference–NC_045512.2) were assembled and aligned using MAFFT³⁸ v7.490. We used IQTREE³⁹ v2.1.2 and the GTR (generalised time-reversible) model with a bootstrap value of 1000 to construct the phylogenetic tree. TreeTime⁴⁰ v0.92 was used for quantifying the molecular clock and identifying ancestral phylogeny in the augur pipeline of Nextstrain⁴¹ v3.0.3.

Bayesian phylogenetic analysis

All B.1.1.318 ($n = 3858$ with major locations: Europe, USA, and Africa accounting for 94%) and B.1.525 ($n = 8278$ with major locations: Europe, USA, and Africa accounting for 86%) sequences were downloaded from GISAID⁴² on 18 August 2021. Sequences with >5% ambiguous nucleotides, a length <95%, or incomplete dates were discarded. Sequences were aligned to the reference (NC_045512.2) using minimap2, with the 5' and 3' UTRs and known problematic sites (GitHub–W-L/ProblematicSites_SARS-CoV2 3) masked⁴³. Both lineage-specific datasets were downsampled for representativeness and computational tractability with a phylogenetic-informed downsampling scheme. A maximum-likelihood phylogeny was reconstructed for each lineage with automatic model selection in IQTREE⁴⁴. The phylogeny was downsampled by root-to-tip tree-traversal, with all internal nodes subject to two rules: (1) if 95% of the leaves subtended by the internal node represented a single country, the earliest representative was retained alongside a random sequence; (2) if leaves from the same location were separated by a zero-branch length, the earliest representative was retained alongside a random sequence. All countries with more than the median number of sequences across countries were then downsampled randomly to the median. This yielded a total number of 1118 and 1746 genomes for B.1.1.318 and B.1.525 respectively. All Nigerian sequences were retained ($n = 73$ for B.1.1.318; $n = 256$ for B.1.525).

We reconstructed Bayesian time-scaled phylogenies for each lineage using BEAST v1.10.5⁴⁵. For both lineages, the time-scaled phylogenies were reconstructed under an HKY substitution model with a gamma-distributed rate for variation among sites⁴⁶, a relaxed molecular clock with a log-normal prior⁴⁷, and an exponential growth coalescent tree prior⁴⁸. For each dataset, we combined two independent Markov Chain Monte Carlo (MCMC) chains of 200 million states, run with the BEAGLE package⁴⁹ to improve run time. Parameters and trees were sampled every 20,000 steps, with the first 20% of steps discarded as burn-in. Convergence and mixing of the MCMC chains were assessed in Tracer v1.7, to ensure the effective sample size (ESS) of all estimated parameters was > 200⁵⁰.

We performed an asymmetric discrete trait analysis using BEAST version 1.10.5 to reconstruct the location-transition history across an empirical distribution of 4000 time-calibrated trees (sampled from each of the posterior tree distributions estimated above). We aggregated the country of sampling on a regional level for computational tractability. We used Bayesian stochastic search variable selection (BSSVS) to infer non-zero migration rates and identify the statistically supported transition routes into and out of Nigeria by a Bayes factor test⁵¹. In addition to the discrete trait analysis, we used a Markov jump counting approach to estimate the timing and origin of geographic transitions into Nigeria to account for uncertainty in phylogeographic reconstruction associated with sparse sampling and low sequence variability⁵². We used the TreeMarkovJumpHistoryAnalyzer from the pre-release version of BEAST v1.10.5 to extract the Markov jumps from posterior tree distributions⁵³. We used TreeAnnotator v1.10 to construct Maximum clade credibility (MCC) trees for all datasets. Trees were visualised using baltic (<https://github.com/evogytis/baltic>).

Air traffic and human movement data

In order to associate international air travel and local human movement with variant movement across borders, we used air traffic data from the International Air Transportation Association (IATA) obtained from BlueDot (<https://bluedot.global/>) to quantify the volume of international travellers to and from Nigeria. The dataset included the number of travellers by origin and destination country, aggregated by month, across all international airports in Nigeria. Data was obtained for the period May 2020 to April 2021, encompassing the emergence and spread of B.1.525 and B.1.1.318. We also curated human movement

data within Nigeria from Google Mobility (<https://www.google.com/covid19/mobility/>) from January 2020 to November 2021, which reflects a baseline before lockdown (January–March 2020) and the study's period of interest. We used R to display movement patterns after grouping human movement into six divisions: retail & recreation, grocery & pharmacy, parks, transport stations, workplaces, and residential areas.

Estimated introduction and exportation intensity index

From genomic data, estimates of the number and origin/destination of bidirectional transmissions with Nigeria are dependent on the sample's representativeness of the population and are therefore limited by global and local sampling biases. To limit these sampling biases, we supplemented our phylogeographic analyses by estimating the introduction (III) and export intensity index (EII) according to Du Plessis et al.²⁷. The III estimates the daily risk of introductions into Nigeria from each country as a product of the number of asymptotically infected individuals in each source country on that day (estimated from the time series of deaths) and their likelihood of travelling to Nigeria (based on the volume of inbound air travel from the source country). The EII is calculated similarly, based on the number of asymptotically infected individuals in Nigeria and their likelihood of travel to each destination by air. Notably, we could not obtain data for land-based travel. Connectivity to regional and neighbouring countries and the associated introduction and export intensity are therefore likely severely underestimated. A time series of reported deaths for each country were collected from the outbreak.info R package⁵⁴. As air travel data was aggregated by month, we conservatively assumed that travel was uniform across all days of the month. We quantified the III and EII for the period May 2020 to April 2021, encompassing the emergence and spread of B.1.525 and B.1.1.318.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The datasets and associated metadata used in this study are available in GISAID's EpiCoV database under the EPI_SET_ID accession numbers EPI_SET_221227vp (<https://doi.org/10.55876/gis8.221227vp>) and EPI_SET_221227pc (<https://doi.org/10.55876/gis8.221227pc>). XML files used for the BEAST analysis are deposited in GitHub: https://github.com/acegid/SARS-CoV-2_Manuscript_Supplemental_Data. The SARS-CoV-2 consensus genome assembly data generated in this study have been deposited in NCBI GenBank database under the accession numbers: OQ050230 to OQ052977 ([https://www.ncbi.nlm.nih.gov/nucleotide/?term=OQ050230:OQ052977\[accn\]](https://www.ncbi.nlm.nih.gov/nucleotide/?term=OQ050230:OQ052977[accn])). DNA sequences have been deposited in NCBI SRA under the BioProject PRJNA916503. Source data are provided as a Source Data file. Reference genome Wuhan-Hu-1 available in GenBank under accession MN908947.3. Air travel data can be requested for release from Bluedot (info@bluedot.global), with use pending approval by Bluedot. Source data are provided with this paper.

Code availability

The scripts used for analysis reported in this study are publicly available at https://github.com/acegid/SARS-CoV-2_Manuscript_Supplemental_Data/tree/main/Figures⁵⁵.

References

- World Bank, World Development Indicators. <http://wdi.worldbank.org/tables> (2022).
- Tsueng, G. et al. Outbreak. info Research Library: a standardized, searchable platform to discover and explore COVID-19 resources and data. Preprint at *BioRxiv* <https://doi.org/10.1101/2022.01.20.477133> (2022).
- Nordling, L. The pandemic appears to have spared Africa so far. Scientists are struggling to explain why. <https://www.science.org/content/article/pandemic-appears-have-spared-africa-so-far-scientists-are-struggling-explain-why>. (2020).
- Kreier F. Morgue data hint at COVID's true toll in Africa. *Nature* **603**, 778–779 (2022).
- Tessema, S. K. & Nkengasong, J. N. Understanding COVID-19 in Africa. *Nat. Rev. Immunol.* **21**, 469–470 (2021).
- Gilbert, M. et al. Preparedness and vulnerability of African countries against importations of COVID-19: a modelling study. *Lancet* **395**, 871–877 (2020).
- Borrega, R. et al. Cross-reactive antibodies to SARS-CoV-2 and MERS-CoV in pre-COVID-19 blood samples from Sierra Leoneans. *Viruses* **13**, 2325 (2021).
- Bugembe, D. L. et al. Emergence and spread of a SARS-CoV-2 lineage A variant (A. 23.1) with altered spike protein in Uganda. *Nat. Microbiol.* **6**, 1094–1101 (2021).
- Butera, Y. et al. Genomic sequencing of SARS-CoV-2 in Rwanda reveals the importance of incoming travelers on lineage diversity. *Nat. Commun.* **12**, 1–1 (2021).
- Wilkinson, E. et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. *Science* **374**, 423–31. (2021).
- Parker, E. et al. Regional connectivity drove bidirectional transmission of SARS-CoV-2 in the Middle East during travel restrictions. *Nat. Commun.* **13**, 1–4 (2022).
- Bogoch, I. I. et al. Assessment of the potential for international dissemination of Ebola virus via commercial air travel during the 2014 west African outbreak. *Lancet* **385**, 29–35 (2015).
- Tegally, H. et al. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature* **592**, 438–43. (2021).
- Viana, R. et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* **603**, 679–686 (2022).
- Pereira, F. et al. Genomic surveillance activities unveil the introduction of the SARS-CoV-2 B.1.525 variant of interest in Brazil: Case report. *J. Med. Virol.* **93**, 5523–5526 (2021).
- Dresch, M. Nigerian Covid-19 variant under investigation with 38 new cases found. <https://www.mirror.co.uk/news/uk-news/breaking-nigerian-covid-19-variant-23511168> (2021).
- Vadlapatla, S. B.1.525 Variant Linked to Nigeria now Found in Telangana - <https://timesofindia.indiatimes.com/city/hyderabad/b-1-525-variant-linked-to-nigeria-now-found-in-telangana/articleshow/83133727.cms>. (2021).
- Takeuchi C. COVID-19 in B.C.: New variant from Nigeria detected, ski awareness campaign launched, and more. <https://www.straight.com/covid-19-pandemic/february-12-coronavirus-update-bc-vancouver-new-variant-from-nigeria-detected-ski-awareness-campaign-and-more>Takeuchi. (2021).
- Tegally, H. et al. A novel and expanding SARS-CoV-2 Variant, B. 1.1. 318, dominates infections in Mauritius. Preprint at *medRxiv* <https://doi.org/10.1101/2021.06.16.21259017> (2021).
- Zhang, L. et al. Ten emerging SARS-CoV-2 spike variants exhibit variable infectivity, animal tropism, and antibody neutralization. *Commun. Biol.* **4**, 1–0 (2021).
- Weisblum, Y. et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *Elife* **9**, e61312 (2020).
- Galloway, S. E. et al. Emergence of SARS-CoV-2 b. 1.1. 7 lineage—united states, december 29, 2020–january 12, 2021. *Morbidity Mortal. Wkly. Report.* **70**, 95 (2021).
- Emary, K. R. et al. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of concern 202012/01 (B. 1.1. 7): an exploratory analysis of a randomised controlled trial. *Lancet* **397**, 1351–62. (2021).

24. Chaillon, A. & Smith, D. M. Phylogenetic analyses of SARS-CoV-2 B.1.1.7 lineage suggest a single origin followed by multiple exportation events versus convergent evolution. *Clin. Infect. Dis.* **73**, 2314–2317 (2021).
25. Washington, N. L. et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **184**, 2587–2594 (2021).
26. Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England. *Science* **372**, eabg3055 (2021).
27. du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**, 708–712 (2021).
28. Tegally, H. et al. The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding genomic surveillance. *Science* **378**, 6615 (2022).
29. Kleynhans, J. et al. SARS-CoV-2 Seroprevalence in a rural and urban household cohort during first and second waves of infections, South Africa, July 2020–March 2021. *Emerg. Infect. Dis.* **27**, 3020 (2021).
30. Wang, H. et al. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis of COVID-19-related mortality, 2020–21. *Lancet* **399**, 1513–1536 (2022).
31. Cohen, C. et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: results of the PHIRST-C cohort study, South Africa, 2020–21. *The Lancet Infectious Diseases*, (2022).
32. Botti-Lodovico, Y. et al. The origins and future of sentinel: an early-warning system for pandemic preemption and response. *Viruses* **13**, 1605 (2021).
33. Siddle, K. J. et al. Genomic analysis of Lassa virus during an increase in cases in Nigeria in 2018. *N. Engl. J. Med.* **379**, 1745–1753 (2018).
34. Ajogbasile, F. V. et al. Real-time metagenomic analysis of undiagnosed fever cases unveils a yellow fever outbreak in Edo State, Nigeria. *Sci. Rep.* **10**, 1–6 (2020).
35. Oluniyi, P. First African SARS-CoV-2 genome sequence from Nigerian COVID-19 case. *Virological* <https://virological.org/t/first-african-sars-cov-2-genome-sequence-from-nigerian-covid-19-case/421> (2020).
36. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* <https://doi.org/10.1038/s41564-020-0770-5> (2020).
37. Argimón, S. et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Micro. Genom.* **2**, e000093 (2016).
38. Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).
39. Minh, B. Q. et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).
40. Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018).
41. Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).
42. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID's innovative contribution to global health. *Glob. Chall.* **1**, 33–46 (2017).
43. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
44. Kalyanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jermini, L. S. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
45. Suchard, M. A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evol.* **4**, vey016 (2018). 1.
46. Hasegawa, M., Kishino, H. & Yano, T. A. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).
47. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).
48. Drummond, A. J., Rambaut, A., Shapiro, B. E. T. H. & Pybus, O. G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005).
49. Ayres, D. L. et al. BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).
50. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901 (2018).
51. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**, e1000520 (2009).
52. Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 564–570 (2020).
53. Lemey, P. et al. Untangling introductions and persistence in COVID-19 resurgence in Europe. *Nature* **595**, 713–717 (2021).
54. Gangavarapu, K. et al. Outbreak. info genomic reports: scalable and dynamic surveillance of SARS-CoV-2 variants and mutations. Research square. <https://doi.org/10.21203/rs.3.rs-1723829/v1>. (2022).
55. Olawoye I, Parker E, & acegid. acegid/SARS-CoV-2_Mancript_Supplemental_Data: Supplemental Data for 'Emergence and spread of two SARS-CoV-2 variants of interest in Nigeria' Nature Comms Manuscript (v1.0). Zenodo. <https://doi.org/10.5281/zenodo.7502930> (2023).

Acknowledgements

We gratefully acknowledge all data contributors, i.e., the Authors and their Originating laboratories responsible for obtaining the specimens, and their Submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which this research is based. This work is made possible by the support provided to ACEGID by a cohort of generous donors through TED's Audacious Project, including the ELMA Foundation, MacKenzie Scott, the Skoll Foundation, and Open Philanthropy. This work was also partly supported by grants from the National Institute of Allergy and Infectious Diseases (<https://www.niaid.nih.gov>), NIH-H3Africa (<https://h3africa.org>) (U01HG007480 and U54HG007480), the World Bank (projects ACE-019 and ACE-IMPACT), the Rockefeller Foundation (Grant #2021 HTH), the Africa CDC through the African Society of Laboratory Medicine [ASLM] (Grant #INV018978), the Wellcome Trust (Project 216619/Z/19/Z), the WARN-ID/CREID grant # U01AI151812, and the Science for Africa Foundation. We want to especially thank the Redeemer's University Management, the State Government of Osun and NCDC for their support during the course of the COVID-19 pandemic.

Author contributions

I.B. Olawoye, P.E.O., and E.P., conceived the study. C.T.H., P.C.S., K.G.A., obtained the funding. J.U.O., J.N.N., A.T.K., T.J.O., F.V.A., I.B. Olawoye, P.E.O., developed the methodology, sequencing, and wrote the original draft. P.E.E., O.A.F., A.N.H., O.T., B.L.S., C.I., I.M.A., M.M.B., F.O., S.F.S., N. Ndodo, I.N., R.A., coordinated the study. P.A., T.A.S., C.A.U., U.E.G., F.A., K. Akano, N.E.O., I.F., K. Adedotun-Sulaiman, F.B.B., B.B.A., C.P., R.A.A., G.C.C., M.I.A., O.O.O., S.G.O., O.A.O., M.F.S., A.E.S., G.O.E., O.G.J., J.O.A., O.O. Akinlo, O.O.F., T.O.I., D.C.N., A.E.O., I.B. Omwanghe, C.A.T., J.O., O. Ayo-Ale, O.I., E.B., G.O.N., A.E.P., O. Blessing, A.M., A.J., J.O.A., P.E., O.R., E.R., G.E.R., E.S., E.A., Y.E., A.O.C., A.I.D., E.O., M.Y.T., H.E.O., M.B., R.A.A., C.K.O., J.O.S., A.O., A.E.A., B.F.D., I.F.Y., A. Fajola, N. Ntia, J.J.E., A.E.M., B.W.M., O.E.F., M.A., I.M.K., A.B.S., Z.W.W., O.O. Adeyemi, O.A.A., A. Ahumibe, A. Akinpelu, O. Ayansola, O. Babatunde, A.A.O., C.C., N.G.M., E.C.O., O. Olisa, O.K.A., I.E.N., M.A.E., E.N., R.L.E., R.O.D., A.A., E.O., V.O., C.K.O., S.O., D.I., J.A.A., M.O.A., O.O., O.O., O.K.A., I.E.N.,

M.A.E., E.N., R.L.E., R.O.D., A.A., E.O., V.O., C.K.O., S.O., D.I., E.O.O., N.A.A., C.N.U., K.N.U., N.I.U., C.A., N.A., O. Ayodeji, A.A.L., R.O.I., G.G., A.F., contributed to detection of SARS-CoV-2 samples and selection of positive samples for sequencing. Sequencing analysis and figures were generated by I.B. Olawoye, P.E.O., E.P., D.P. Travel data were provided by K.K. B.A.P., B.L.M., K.J.S., A.Fowotade, S.O., P.O.O., G.A., C.O.E., P.C.S., C.T.H., critically reviewed and edited the manuscript. All authors read and approved the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-36449-5>.

Correspondence and requests for materials should be addressed to Christian T. Happi.

Peer review information *Nature Communications* thanks the anonymous reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

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

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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023

Idowu B. Olawoye^{1,2,35}, Paul E. Oluniyi^{1,2,35}, Judith U. Oguzie^{1,2,35}, Jessica N. Uwanibe^{1,2,35}, Tolulope A. Kayode^{1,2,35}, Testimony J. Olumade^{1,2,35}, Fehintola V. Ajogbasile^{1,2,35}, Edyth Parker^{3,35}, Philomena E. Eromon², Priscilla Abechi^{1,2}, Tope A. Sobajo^{1,2}, Chinedu A. Ugwu^{1,2}, Uwem E. George^{1,2}, Femi Ayoade^{1,2}, Kazeem Akano^{1,2}, Nicholas E. Oyejide², Iguosadolo Nosamiefan², Iyanuoluwa Fred-Akintunwa², Kemi Adedotun-Sulaiman², Farida B. Brimmo², Babatunde B. Adegboyega², Courage Philip², Richard A. Adeleke², Grace C. Chukwu², Muhammad I. Ahmed², Oludayo O. Ope-Ewe², Shobi G. Otitoola², Olusola A. Ogunsanya², Mudasiru F. Saibu², Ayotunde E. Sijuwola², Grace O. Ezekiel², Oluwagboadurami G. John^{1,2}, Julie O. Akin-John^{1,2}, Oluwasemilogo O. Akinlo², Olanrewaju O. Fayemi², Testimony O. Ipaye², Deborah C. Nwodo², Abolade E. Omoniyi², Iyobosa B. Omwanghe², Christabel A. Terkuma², Johnson Okolie², Olubukola Ayo-Ale², Odia Ikponmwosa⁴, Ebo Benevolence⁴, Grace O. Naregose⁴, Akhilomen E. Patience⁴, Osiemi Blessing⁴, Airende Micheal⁴, Agbukor Jacqueline⁴, John O. Aiyepada⁴, Paulson Ebhodaghe⁴, Omiunu Racheal⁴, Esumeh Rita⁴, Giwa E. Rosemary⁴, Ehikhametator Solomon⁴, Ekanem Anieno⁴, Yerumoh Edna⁴, Aire O. Chris⁴, Adomeh I. Donatus⁴, Ephraim Ogbaini-Emovon⁴, Mirabeau Y. Tatfeng⁵, Hannah E. Omunakwe⁶, Mienye Bob-Manuel⁶, Rahaman A. Ahmed⁷, Chika K. Onwuamah⁷, Joseph O. Shaibu⁷, Azuka Okwuraiwe⁷, Anthony E. Ataga⁸, Andrew Bock-Oruma⁹, Funmi Daramola¹⁰, Ibrahim F. Yusuf¹⁰, Akinwumi Fajola¹¹, Nsikak-Abasi Ntia¹², Julie J. Ekpo¹³, Anietie E. Moses¹³, Beatrice W. Moore-Igwe¹⁴, Oluwatosin E. Fakayode¹⁵, Monilade Akinola¹⁶, Ibrahim M. Kida¹⁷, Bamidele S. Oderinde¹⁷, Zara W. Wudiri¹⁸, Oluwapelumi O. Adeyemi¹⁹, Olusola A. Akanbi²⁰, Anthony Ahumibe²⁰, Afolabi Akinpelu²⁰, Oyeronke Ayansola²⁰, Olajumoke Babatunde²⁰, Adesuyi A. Omoare²⁰, Chimaobi Chukwu²⁰, Nwando G. Mba²⁰, Ewean C. Omoruyi²¹, Olasunkanmi Olisa²², Olatunji K. Akande²², Ifeanyi E. Nwafor²³, Matthew A. Ekeh²³, Erim Ndoma²³, Richard L. Ewah²³, Rosemary O. Duruihuoma²³, Augustine Abu²³, Elizabeth Odeh²³, Venatius Onyia²³, Chiedozie K. Ojide²³, Sylvanus Okoro²⁴, Daniel Igwe²⁴, Emeka O. Ogah²⁴, Kamran Khan^{25,26}, Nnennaya A. Ajayi²⁷, Collins N. Ugwu²⁷, Kingsley N. Ukwaja²⁷, Ngozi I. Ugwu²⁸, Chukwuyem Abejegah²⁹, Nelson Adedosu²⁹, Olufemi Ayodeji²⁹, Ahmed A. Liasu²⁹, Rafiu O. Isamotu³⁰, Galadima Gadzama³¹, Brittany A. Petros³², Katherine J. Siddle³², Stephen F. Schaffner³², George Akpede⁴, Cyril Oshomah Erameh⁴, Marycelin M. Baba^{16,17}, Femi Oladiji³³, Rosemary Audu⁷, Nnaemeka Ndodo²⁰, Adeola Fowotade²³, Sylvanus Okogbenin⁴, Peter O. Okokhere⁴, Danny J. Park³², Bronwyn L. Mcannis³², Ifedayo M. Adetifa²⁰, Chikwe Ihekweazu²⁰, Babatunde L. Salako⁷, Oyewale Tomori², Anise N. Happi², Onikepe A. Folarin^{1,2}, Kristian G. Andersen³, Pardis C. Sabeti^{32,34} & Christian T. Happi^{1,2,34} ✉

¹Department of Biological Sciences, Faculty of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria. ²African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Ede, Osun State, Nigeria. ³Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA. ⁴Irrua Specialist Teaching Hospital, Irrua, Edo State, Nigeria. ⁵Department of Medical Laboratory Science, Niger Delta University, Amassoma, Bayelsa State, Nigeria. ⁶Satellite Molecular Laboratory, Rivers State University Teaching Hospital, Port Harcourt, Rivers State, Nigeria. ⁷The Nigerian Institute of Medical Research, Yaba, Lagos State, Nigeria. ⁸Molecular Laboratory, Regional Centre for Biotechnology and Bioresources Research,

University of Port Harcourt, Port Harcourt, Rivers State, Nigeria. ⁹Family Physician, SPDC, Port Harcourt, Rivers State, Nigeria. ¹⁰Clinical Health, SPDC, Port Harcourt, Rivers State, Nigeria. ¹¹Regional Community Health, SPDC, Port Harcourt, Rivers State, Nigeria. ¹²Occupational Health, SPDC, Port Harcourt, Rivers State, Nigeria. ¹³Department of Medical Microbiology and Parasitology, University of Uyo, Uyo, Akwa Ibom State, Nigeria. ¹⁴Rivers State University, Port Harcourt, Rivers State, Nigeria. ¹⁵Department of Public Health, Ministry of Health, Ilorin, Kwara State, Nigeria. ¹⁶WHO Polio Laboratory, University of Maiduguri Teaching Hospital, Maiduguri, Borno State, Nigeria. ¹⁷Department of Immunology, University of Maiduguri Teaching Hospital, Maiduguri, Nigeria. ¹⁸Department of Community Medicine, University of Maiduguri Teaching Hospital, Maiduguri, Borno State, Nigeria. ¹⁹Department of Medical Microbiology and Parasitology. Faculty of Basic Clinical Sciences. College of Health Sciences, University of Ilorin, Ilorin, Kwara State, Nigeria. ²⁰Nigeria Centre for Disease Control, Abuja, Nigeria. ²¹Medical Microbiology and Parasitology Department, College of Medicine, University of Ibadan, Ibadan, Nigeria. ²²Biorepository Clinical Virology Laboratory, University of Ibadan, Ibadan, Nigeria. ²³Virology Laboratory, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Ebonyi State, Nigeria. ²⁴Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ²⁵Department of Medicine, University of Toronto, Toronto, Canada. ²⁶BlueDot, Toronto, Canada. ²⁷Internal Medicine Department, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ²⁸Haematology Department, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria. ²⁹Federal Medical Center, Owo, Ondo State, Nigeria. ³⁰Ministry of Health, Osogbo, Osun State, Nigeria. ³¹Department of Medical Microbiology, University of Maiduguri Teaching Hospital, Maiduguri, Borno State, Nigeria. ³²Broad Institute of Harvard and MIT, Cambridge, MA, USA. ³³Department of Epidemiology and Community Health, Faculty of Clinical Sciences, College of Health Sciences, University of Ilorin, Ilorin, Nigeria. ³⁴Department of Immunology and Infectious Diseases, Harvard TH Chan School of Public Health, Boston, MA, USA. ³⁵These authors contributed equally: Idowu B. Olawoye, Paul E. Oluniyi, Judith U. Oguzie, Jessica N. Uwanibe, Tolulope A. Kayode, Testimony J. Olumade, Fehintola V. Ajogbasile, Edyth Parker. ✉ e-mail: happic@run.edu.ng