SARS-CoV-2 Circulation, Guinea, March 2020–July 2021

Solène Grayo,[1] Cécile Troupin,[1,2] Moussa Moïse Diagne, Houlou Sagno, Isabelle Ellis, Bakary Doukouré, Amadou Diallo, Jean-Mathieu Bart, Mohamed Lamine Kaba, Benoit Henry, Billy Sivahera Muyisa, Mamadou Saliou Sow, Ndongo Dia, Ousmane Faye, Sakoba Keita, Noël Tordo

Author affiliations: Institut Pasteur de Guinée, Conakry, Guinea (S. Grayo, C. Troupin, H. Sagno, I. Ellis, B. Doukouré, N. Tordo); Institut Pasteur de Dakar, Dakar, Senegal (M.M. Diagne, A. Diallo, N. Dia, O. Faye); Institut de Recherche pour le Développement-Programme National de Lutte contre la Trypanosomiase Humaine Africaine, Montpellier, France (J.-M. Bart); Camp Militaire Alpha Yaya, Conakry (M.L. Kaba); Centre Médico-Social de l'Ambassade de France, Conakry (B. Henry); Centre Hospitalier de Donka (M.S. Sow); Alliance for International Medical Action, Guinea (B.S. Muyisa), Agence Nationale de Sécurité Sanitaire, Conakry (S. Keita)

DOI: https://doi.org/10.3201/eid2802.212182

This overview of severe acute respiratory syndrome coronavirus 2 circulation over 1.5 years in Guinea demonstrates that virus clades and variants of interest and concern were progressively introduced, mostly by travellers through Conakry, before spreading through the country. Sequencing is key to following virus evolution and establishing efficient control strategies.

In Guinea, the index coronavirus disease (COVID-19) case-patient identified on March 12, 2020, was an expatriate traveling back from Europe. Immediately, a COVID-19 task force was established by the Agence Nationale de Sécurité Sanitaire; 6 national laboratories were involved in the diagnosis of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections. As of July 16, 2021, a total of 24,668 confirmed cases (23,571 recovered persons and 188 deaths) have been reported (https://www.anss-guinee.org). The Institut Pasteur de Guinée has contributed to the testing of >25,000 human nasopharyngeal swab samples. Most samples originated in the Conakry area from the Donka University Hospital and the Alpha Yaya Military Hospital, which serve the general population, and from the Health Center of the French Embassy, which serves mostly expatriates or travelers. We selected a panel of 252 (12.26%) SARS-CoV-2–positive samples taken during March 12, 2020–July 16, 2021, for whole-genome sequencing, which was performed at the World Health Organization Collaborative Centre of the Institut Pasteur de Dakar, to examine the evolution of SARS-CoV-2 in Guinea.

From these 252 samples, 226 sequences were generated; we excluded 90 sequences showing >10% missing nucleotides. We analyzed the remaining 136 (54%) sequences by using Nextclade (https://clades.nextstrain.org) and Pangolin software (https://cov-lineages.org). The Guinea sequences are distributed into 7 clades (Appendix Figure, https://wwwnc.cdc.gov/EID/article/28/2/21-2182-App1.pdf): 20A clade (n = 55, 40.44%), 20B clade (n = 31, 22.80%), 20C clade (n = 1, 0.74%), 20D clade (n = 8, 5.88%), 20I clade (20I/B.1.1.7/Alpha; n = 19, 13.97%), 21A clade (21A/B.1.617.2/Delta; n = 16, 11.76%), and 21D clade (21D/B.1.525/Eta; n = 6, 4.41%) (Figure, panel A). The 7 clades are subdivided into subclades. None of these subclades gather sequences from specific prefectures in Guinea, suggesting that SARS-CoV-2 viruses circulating inside the country are related to Conakry cases. At the time of this writing, ≥21 sublineages of SARS-CoV-2 viruses were circulating in Guinea (Table).

During March–August 2020, the sequences were exclusively distributed into 2 clades, 20A and 20B, globally circulating in West and Central Africa (Table; Figure, panel B) (*1*–*3*). Their ancestral position in the maximum-likelihood tree outlines their introduction in Guinea, most likely from Europe as illustrated by the index case. Their circulation has persisted in a nonexclusive manner up to May–July 2021. The 20D clade, sparsely detected in Africa (Table), was observed in Guinea through ≥2 introductions in September and October 2020, according to the topology of the maximum-likelihood tree (Figure, panel B). Moreover, a single case of 20C clade originating from North America was detected in January 2021 in a person traveling from Haiti (Table; Figure, panel B).

In 2021, new SARS-CoV-2 variants of concern (VOC) and variants of interest, reputed to be more transmissible, emerged in Guinea (*4*). The VOC 20I/B.1.1.7/Alpha variant, which originally emerged in the United Kingdom, was first identified in Guinea in January 2021, increased in incidence up to March 2021, and then decreased from April to June 2021, corresponding to the dynamic described in Africa (Figure, panel B) (*1*–*3*,*5*; E.A. Ozer et al., unpub. data, https://www.medrxiv.org/content/10.1101/2021.04.09.21255206v3). The variant of interest 21D/B.1.525/Eta was identified in Guinea and other countries in Central and West Africa in February–May 2021 (Table) (*5*; E.A.

[1]These authors contributed equally to this article.

[2]Current affiliation: Institut Pasteur du Cambodge, Phnom Penh, Cambodia.

Ozer et al., unpub. data). The topology of the Guinea maximum-likelihood tree with only one subclade of this variant suggests a unique introduction in this study. Finally, the 21A/B.1.617.2/Delta VOC was first detected in May 2021 in Guinea (Figure, panel B). By July, it had become dominant; >90% of the sequenced

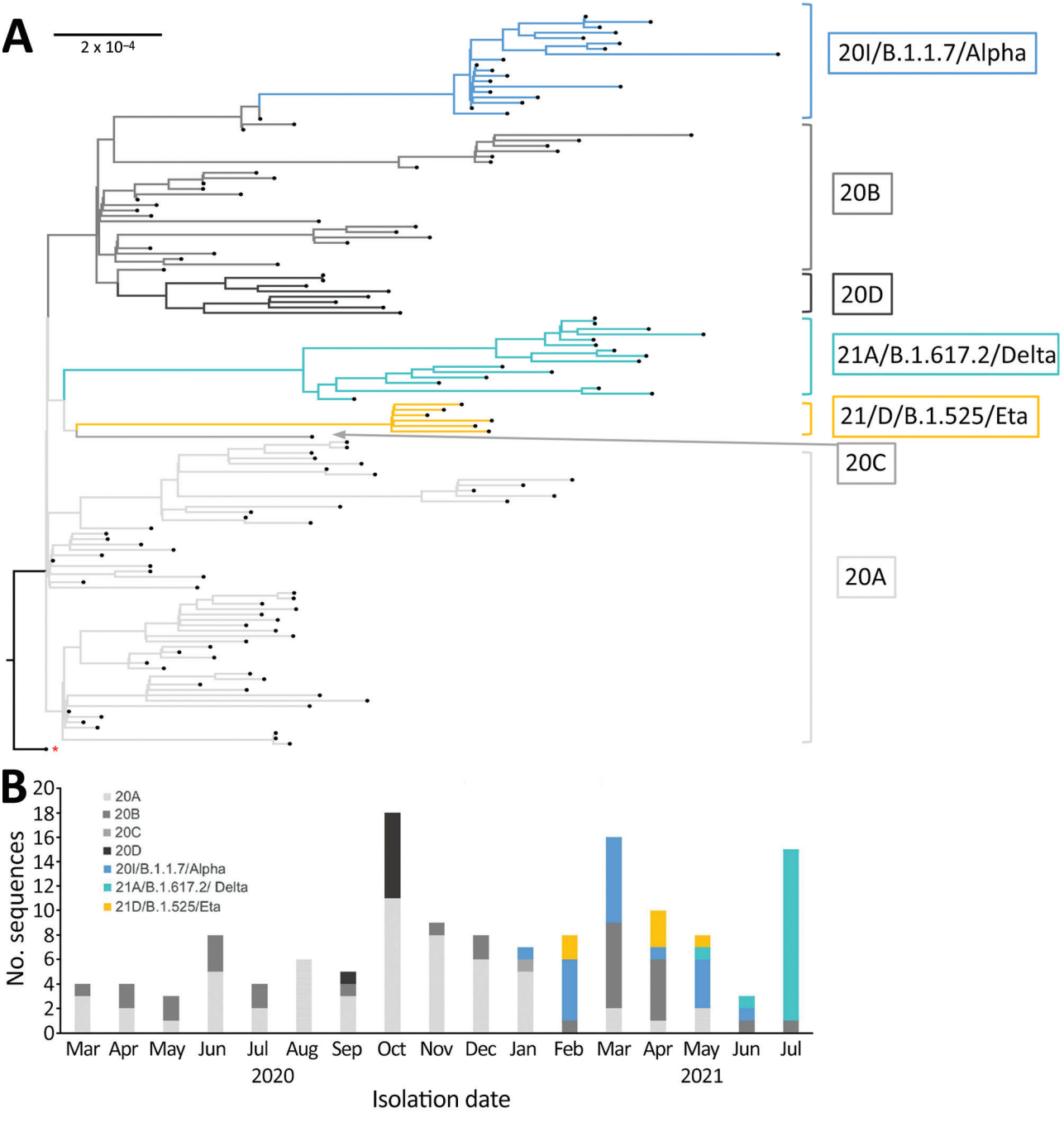


Figure. Phylogenetic and temporal descriptions of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) sequences from Institut Pasteur de Guinée from samples collected in Guinea during March 12, 2020–July 16, 2021. A) Maximum-likelihood phylogenetic tree of 136 SARS-CoV-2 genomic sequences. The tree was constructed with IQ-tree software by using multiple-genome sequence alignment and Wuhan-Hu-1 strain (GenBank accession no. NC 045512) as outgroup reference sequence, indicated by the red asterisk. Branches and the sequence names are colored according to Nextclade assigned clades: 20A, light gray; 20B, medium gray; 20C, dark gray; 20D, black; 20I/B.1.1.7/Alpha, blue; 21A/B.1.617.2/Delta, azure; 21D/B.1.525/Eta, yellow. Each sequence is highlighted by a black tip. Scale bar indicates the distance corresponding to substitution per site. B) Chronologic distribution of SARS-CoV-2 genomic variants over 17 months in Guinea. The 136 selected sequences are assigned by Nextclade and classified according to sampling date from March 31, 2020, to July 16, 2021. Clades are colored as in panel A.

Table. Characteristics of clades and lineages identified among the Institut Pasteur de Guinée SARS-CoV-2 sequences from samples taken in Guinea during March 12, 2020–July 16, 2021*

	Worldwide			Africa			Guinea†	
Clade and lineage	1st described	Location	No. sequences	1st described	Location	No. sequences	1st described	No. sequences
20A								
B.1	2020 Jan	UK	83,632	2020 Mar	RDC	2,816	2020 Mar	43
B.1.36.10	2020 Mar	United States	824	2020 Apr	South Africa	17	2021 Jan	1
B.1.210	2020 Mar	India	403	No	No	0	2020 Oct	1
B.1.243	2020 Mar	United States	13,091	2020 Jun	Kenya	6	2020 Jun	1
B.1.298	2020 Mar	United States	397	No	No	0	2020 Oct	1
B.1.540	2020 Feb	India	2,186	2020 Mar	Gambia, Kenya	134	2020 Jun	2
B.1.622	2021 Jan	Réunion	76	No	No	0	2020 Sep	1
B.1.629	2021 Jan	Belgium	84	Unknown	Guinea	14	2021 Mar	5
20B								
B.1.1	2020 Jan	UK	48,119	2020 Feb	Nigeria	1,361	2020 Mar	16
B.1.1.39	2020 Mar	Switzerland	1,861	No	No	0	2021 Jan	1
B.1.1.142	2020 Mar	Australia	51	No	No	0	2021 Apr	1
B.1.1.236	2020 Feb	UK	1,404	2020 Mar	South Africa	36	2020 Mar	1
B.1.1.316.1‡	2020 Jan	Sierra Leone	10,444	2020 Jan	Sierra Leone	35	2020 Dec	4
B.1.1.317	2020 Feb	Russia	2,435	2020 Jun	Zimbabwe	4	2020 Aug	1
B.1.1.318	2021 Jan	UK	3,350	2021 Jan	Nigeria	360	2021 Feb	6
B.1.1.372	2020 Mar	UK	1,381	2020 May	South Africa	16	2020 Jul	1
20C								
B.1.575	2020 Oct	United States	3,026	2020 Dec	Senegal	12	2021 Jan	1
20D								
B.1.1.1	2020 Mar	UK	3,078	2020 Mar	RDC	169	2020 Sep	8
20I								
B.1.1.7 (Alpha)	2020 Sep	UK	1,045,206	2020 Dec	Ghana	2,047	2021 Jan	19
21A								
B.1.617.2 (Delta)	2020 Nov	India	261,339	2021 Mar	South Africa	1,662	2021 May	16
21D								
B.1.525 (Eta)	2020 Dec	UK, Nigeria	7,752	2020 Dec	Nigeria	581	2021 Jan	6

*Clades and lineages are respectively assigned according to Nextclade definition (https://github.com/nextstrain/ncov/blob/master/docs/src/reference/naming_clades.md) and PANGO lineages list (https://github.com/cov-lineages/pangolin) at the same assignment date (August 14, 2021). The Guinea sequences are distributed in 21 lineages clustered into 7 clades: 20A clade (n = 55, 40.44%) with 8 lineages (B.1, B.1.36.10, B.1.210, B.1.243, B.1.298, B.1.540, B.1.622, and B.1.629), 20B clade (n = 31, 22.80%) with 8 lineages (B.1.1, B.1.1.39, B.1.1.142, B.1.1.236, B.1.1.316.1, B.1.1.317, B.1.1.318, and B.1.1.372), 20C clade (n = 1, 0.74%) with 1 lineage (B.1.575), 20D clade (n = 8, 5.88%) with 1 lineage (B.1.1.1), 20I clade (n = 19, 13.97%) with 1 lineage (B.1.1.7 [Alpha]), 21A clade (n = 16, 11.76%) with 1 lineage (B.1.617.2 [Delta]) and 21D clade (n = 6, 4.41%) with 1 lineage (B.1.525 [Eta]). For each lineage, the first worldwide and African descriptions are provided (date and location), as well as the number of deposited sequences in GISAID (August 16, 2021). RDC, Democratic Republic of the Congo; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2; UK, United Kingdom.
†First description and number of sequences in this study.
‡B.1.1.316.1 lineage alias R.1.

viruses by Institut Pasteur de Guinée demonstrated the same dynamics observed during May–August 2021 in Africa (*6*). The maximum-likelihood tree suggests ≥2 main introductions of this variant in Guinea.

In summary, although only 20A and 20B clades circulated in Guinea for the first 6 months of the pandemic (March–August 2020), the reopening of borders and commercial flights have progressively enabled the introduction of variants from surrounding parts of Africa (21D/B.1.525/Eta) and globally (20I/B.1.1.7/Alpha and 21A/B.1.617.2/Delta) several months after their original detection (Table). Although the 20I/B.1.1.7/Alpha and 21A/B.1.617.2/Delta variants have spread successfully in the population, the 21D/B.1.525/Eta variant has only occasionally been detected. We did not detect other variants previously found in Africa, such as the 20H/B.1.351/Beta variant (which populated 50% of sequences in Africa during January–May 2021) and variants from the sublineage A, including the A.23.1 lineage from East Africa and the A.27 lineage of uncertain origin, in this study (*1*–*3*,*5*; E.A. Anoh et al., unpub. data, https://www.medrxiv.org/content/10.1101/2021.05.06.21256282v1).

This overview of the circulation of SARS-CoV-2 viruses in Guinea furthers the examination of infectious diseases control strategies in Africa, which faces vaccination implementation delay (*7*). Beside classical quantitative reverse transcription PCR diagnostic testing, strengthening of the sequencing capacity is the cornerstone of tracking and fighting the emergence of SARS-CoV-2 variants in real time (*8*). Making countries autonomous in sequencing is the next challenge in fighting COVID-19, as well as other emerging zoonoses, in Africa.

Acknowledgments

We are thankful to the members of the SARS-CoV-2 surveillance system network in Guinea for their fruitful contributions in field activities.

This study was funded by the French Ministry for Europe and Foreign Affairs (MEAE) through the project "REPAIR Covid-19-Africa," coordinated by the Pasteur Network association, and by the European Union through the EBO-SURSY project (EU-FOOD/2016/379-660). We are also grateful to the Africa Pathogen Genomics Initiative (Africa PGI) at the Africa Centres for Disease Control and Prevention supported by the Bill and Melinda Gates Foundation.

About the Author

Dr. Grayo is a virologist at Institute Pasteur de Guinée. Her research interests are emerging viral infectious diseases and zoonoses, with a current focus on viral hemorrhagic fevers in West Africa.

References

1. GISAID/Institut Pasteur de Dakar. Phylodynamics of pandemic coronavirus in west Africa. 2021 [cited 2021 Apr 5]. https://www.gisaid.org/phylodynamics/west-africa
2. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, et al. A year of genomic surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science. 2021;374:423–31. https://doi.org/10.1126/science.abj4336
3. Wruck W, Adjaye J. Detailed phylogenetic analysis tracks transmission of distinct SARS-COV-2 variants from China and Europe to West Africa. Sci Rep. 2021;11:21108. https://doi.org/10.1038/s41598-021-00267-w
4. Harvey WT, Carabelli AM, Jackson B, Gupta RK, Thomson EC, Harrison EM, et al.; COVID-19 Genomics UK (COG-UK) Consortium. SARS-CoV-2 variants, spike mutations and immune escape. Nat Rev Microbiol. 2021;19:409–24. https://doi.org/10.1038/s41579-021-00573-0
5. Sander AL, Yadouleton A, De Oliveira Filho EF, Tchibozo C, Hounkanrin G, Badou Y, et al. Mutations associated with SARS-CoV-2 variants of concern, Benin, early 2021. Emerg Infect Dis. 2021;27:2889–903. https://doi.org/10.3201/eid2711.211353
6. World Health Organization. COVID-19 situation update for the WHO African Region, 26 August 2020 [cited 2020 Dec 26]. https://apps.who.int/iris/bitstream/handle/10665/334003/SITREP_COVID-19_WHOAFRO_20200826-eng.pdf
7. Ariyo OE, Oladipo EK, Osasona OG, Obe O, Olomojobi F. COVID-19 vaccines and vaccination: how prepared is Africa? Pan Afr Med J. 2021;39:107. https://doi.org/10.11604/pamj.2021.39.107.27912
8. Tsanni A. Covid-19: Africa scrambles to increase genomic testing capacity as variants spread. BMJ. 2021;373:n1122. https://doi.org/10.1136/bmj.n1122

Address for correspondance: Solène Grayo, Virology Unit, Institut Pasteur de Guinée, Route de Donka, BP 4416, Conakry, Guinée; email: solene.grayo.ext@pasteur.fr

Probable Transmission of SARS-CoV-2 Omicron Variant in Quarantine Hotel, Hong Kong, China, November 2021

Haogao Gu, Pavithra Krishnan, Daisy Y.M. Ng, Lydia D.J Chang, Gigi Y.Z. Liu, Samuel S.M. Cheng, Mani M.Y. Hui, Mathew C.Y. Fan, Jacob H.L. Wan, Leo H.K. Lau, Benjamin J. Cowling, Malik Peiris, Leo L.M. Poon

Author affiliation: The University of Hong Kong, Hong Kong, China

DOI: https://doi.org/10.3201/eid2802.212422

We report detection of severe acute respiratory syndrome coronavirus 2 Omicron variant (B.1.1.529) in an asymptomatic, fully vaccinated traveler in a quarantine hotel in Hong Kong, China. The Omicron variant was also detected in a fully vaccinated traveler staying in a room across the corridor from the index patient, suggesting transmission despite strict quarantine precautions.

A new variant of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), B.1.1.529, was identified in Botswana and South Africa in early November 2021 and was designated as variant of concern (VOC) Omicron by the World Health Organization on November 26, 2021 (*1*). As of December 1, 2021, ≈220 sequences were available on GISAID (https://www.gisaid.org), and this variant has been detected in countries in Africa and beyond since mid-November (*2*,*3*). This variant contains >30 spike protein amino acid mutations that might be associated with increased transmissibility, severity, and capacity for immune escape. With supporting evidence of epidemiologic and molecular epidemiologic findings, we report the probable transmission of Omicron in a quarantine hotel in Hong Kong, China. We also compare its mutational profile with other VOCs and variants of interest.

Two cases of infection with VOC Omicron (cases A and B) were detected in Hong Kong. Case-patient A arrived in Hong Kong from South Africa on November 11, 2021, and case-patient B arrived in Hong Kong from Canada on November 10, 2021. Both case-patients had previously received 2 vaccine doses (Pfizer-BioNTech, https://www.pfizer.com); the second dose was given on June 4, 2021, for case-patient A and on May 25, 2021, for case-patient B. Both case-patients tested negative by reverse transcription PCR (RT-PCR) for SARS-CoV-2 within 72 hours before arrival. On arrival at the Hong Kong

Article DOI: https://doi.org/10.3201/eid2802.212182

SARS-CoV-2 Circulation, Guinea, March 2020–July 2021

Appendix

Materials and Methods

Samples Origin

Institut Pasteur de Guinée (IPGui) is in charge of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) molecular diagnostics at 3 sites in Conakry, Guinea, including 2 Centre de Traitement des Épidémies (CTEPI): the CTEPI of the Donka University Hospital (serving the general public up to severe cases requiring oxygen supplementation), the CTEPI of the Alpha Yaya Military Hospital (for members of the military and their families), and the sociomedical center of the French Ambassy (for expatriates [suspected and control patients]).

Coronavirus Disease Diagnosis

Viral RNA isolation

Nucleic acid extraction was carried out by using ID Gene Mag Fast Extraction Kit and IDEAL 32 extraction robot (ID-Vet, https://www.id-vet.com) according to the manufacturer's instructions from 140 µL of nasopharyngeal swab specimens. The extracted RNA was eluted in 80 µL of elution buffer. Positive and negative controls are included in each extraction run.

Quantitative Reverse Transcription PCR

SARS-CoV-2 genome detection was performed on extracted RNA (no freezing) using the 2019-nCoV Nucleic Acid Diagnostic Kit (Sansure Biotech, https://www.sansureglobal.com) under US Food and Drug Administration Emergency Use Authorization. This quantitative reverse transcription PCR kit targets 2 SARS-CoV-2 genes (nucleocapsid N and polymerase ORF1ab) and the human RNase P gene, as Internal Control (IC).

Positive and negative controls are included in each amplification run performed in a LightCycler Roche (Roche, https://www.roche.com) machine. Left extracted RNA is immediately stored at −80°C in the IPGui Biobank.

Sample Selection for Sequencing

During March, 12 2020-July 16, 2021, 22,975 human nasopharyngeal swab samples were stored at −80°C at IPGui Biobank. Among the 2,055 positive patients, 252 (12.26%) were selected with the following characteristics: a) having a cycle threshold value <30 to improve sequence quality; b) following the dynamics of the epidemic with ≥10 samples per month with higher numbers from 2021 and the arrival of the variant of concern and variant of interest in Africa; c) including samples from outside Conakry (n = 47, 39%). Aliquots of 500 µL of nasopharyngeal swabs from the 252 selected samples were shipped at 4°C to Institut Pasteur de Dakar to perform Illumima sequencing.

Next-Generation Sequencing

Next-generation sequencing was performed at Institut Pasteur de Dakar. Briefly, viral RNA was extracted using the QIAamp viral RNA minikit (QIAGEN, https://www.qiagen.com) following manufacturer recommendations. SARS-CoV-2 genomes were generated by an amplicon-based approach using either the Illumina DNA Prep, (M) Tagmentation (96 Samples) kit or the US Food and Drug Administration–approved Illumina COVIDSeq kit (Illumina, https://www.illumina.com), depending on availability.

The first protocol consisted of a reverse complement step on the RNA extracts for cDNA synthesis before performing generation of tiled amplicons by reverse transcription PCR (RT-PCR) made with the ARTIC primers version 3 of the nCoV-2019 as previously described (https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye). The PCR products were purified and DNA was quantified with a Qubit 3 fluorometer (ThermoFisher Scientific, https://www.thermofisher.com). DNA products (multiplex PCR pools A and B) were pooled in equal concentrations and libraries were generated by using the Illumina DNA Prep, (M) Tagmentation (96 Samples) according to the manufacturer's specifications. Whole-genome sequencing was performed with paired-end reads by using the Illumina MiSeq reagent kit v3 (150 cycles) on an Illumina MiSeq instrument.

The Illumina COVIDSeq protocol is a modified version of the previous one, with 98 amplicons designed to amplify SARS-CoV-2–specific sequences, combined with proven Illumina sequencing technology (*1*). Final libraries were loaded in a NextSeq550 sequencer according to the manufacturer's recommendations.

Genome Assembly

Genome consensus was generated using the EDGE COVID-19 pipeline (C.-C. Lo, unpub. data, https://arxiv.org/abs/2006.08058), a tailored bioinformatics platform based on the fully open-source EDGE bioinformatics software (*2*). Preprocessing (data quality control) was performed using FastQCs. This analysis included trimming low-quality regions of reads and filtering reads that either failed a quality threshold or minimum length. We used the align_trim.py (from ARTIC) to soft clip primer regions from the alignment file (BAM) based on the position of primers in the reference genome (using primer BED file).

Reads were aligned to the original reference genome (NC_045512.2) after removing the PolyA-tail from the 3′ end (33 nt). BWA mem was used as the default aligner, which was then automatically followed by variant calling and generation of a consensus sequence.

Variant calling used bcftools mpileup command to convert the aligned BAM file into genomic positions and call genotypes, reduced the list of sites to those found to be variants by passing this file into bcftools call. The variant calls were further filtered by vcfutils.pl (from samtools). The consensus workflow used a maximum of 8000× depth coverage reads for computational efficiency. Various other parameters were defaulted including a minimum depth coverage (5×) of support of a variant site coverage per base (otherwise the consensus will be "N"), base quality (20), alternate base threshold (0.5) to support an alternative for the consensus to be changed, indels threshold (0.5) to support an INDEL for the consensus to be changed, and minimum mapping quality of 60. All the sequences generated in this study have been submitted to GISAID

Clade and Lineage Assignation

Nextclade tool (version 16.0) (https://clades.nextstrain.org) has been used to identify mutations (n = 136) compared with the SARS-CoV-2 reference sequence (Wuhan-Hu-1; NC_045512). The Nexclade tool uses these mutations to assign the sequences to specific clades and to place them on a reference phylogenetic tree with a subset of all sequences available in GISAID.

Phylogenetic Analysis

A multiple sequence alignment of the 136 new SARS-CoV-2 genomes from Guinea combined with the SARS-CoV-2 reference sequence NC_045512 from Genbank was constructed using MUSCLE (http://www.ebi.ac.uk/Tools/msa/muscle) (*3*) with default parameters implemented in Galaxy (*4–6*) (https://www.research.pasteur.fr/en/tool/pasteur-galaxy-platform). A phylogenetic tree was then estimated by using W-IQ-TREE, a web

interface and server for IQ-TREE that is a phylogenetic software for maximum likelihood (ML) analysis (*7*) (http://iqtree.cibiv.univie.ac.at/). The best-fit model for nucleotide substitution according to the Bayesian Information Criterion (BIC) determined by Model Finder (*8*) was the general time-reversible model with proportion of invariable sites plus gamma-distributed rate heterogeneity (GTR+I+G4). This project was approved by the Comité National d'Éthique pour la Recherche en Santé (CNERS) with reference number 109/CNERS/20 (28/08/2020) 152/CNERS/20 (05/011/2020)

References

1. Bhoyar RC, Jain A, Sehgal P, Divakar MK, Sharma D, Imran M, et al. High throughput detection and genetic epidemiology of SARS-CoV-2 using COVIDSeq next-generation sequencing. PLoS One. 2021;16:e0247115. PubMed https://doi.org/10.1371/journal.pone.0247115

2. Li PE, Lo CC, Anderson JJ, Davenport KW, Bishop-Lilly KA, Xu Y, et al. Enabling the democratization of the genomics revolution with a fully integrated web-based bioinformatics platform. Nucleic Acids Res. 2017;45:67–80. PubMed https://doi.org/10.1093/nar/gkw1027

3. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 2004;5:113. PubMed https://doi.org/10.1186/1471-2105-5-113

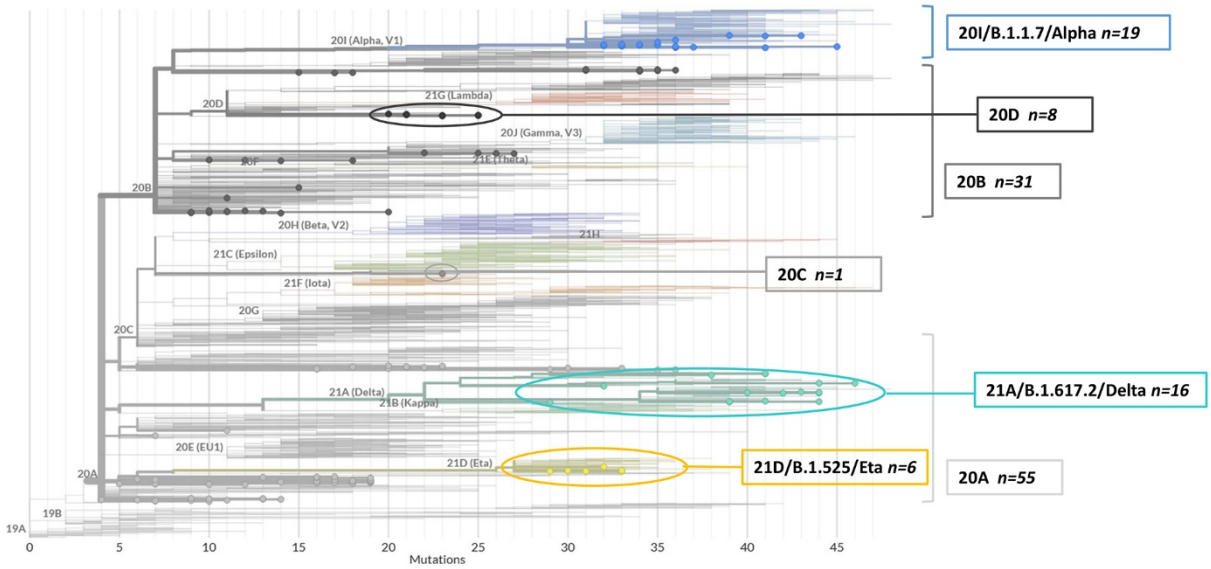
4. Goecks J, Nekrutenko A, Taylor J, Galaxy Team T; Galaxy Team. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. Genome Biol. 2010;11:R86. PubMed https://doi.org/10.1186/gb-2010-11-8-r86

5. Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, et al. Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol. 2010; Chapter 19:Unit 19.10.1–21.

6. Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, et al. Galaxy: a platform for interactive large-scale genome analysis. Genome Res. 2005;15:1451–5. PubMed https://doi.org/10.1101/gr.4086505

7. Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res. 2016;44:W232–5. PubMed https://doi.org/10.1093/nar/gkw256

8. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat Methods. 2017;14:587–9. PubMed https://doi.org/10.1038/nmeth.4285



Appendix Figure. Institut Pasteur de Guinée severe acute respiratory syndrome coronavirus 2 sequences' position on the global Nextclade phylogenetic tree. The tree is generated from Nextclade online software (https://clades.nextstrain.org/) and rooted with the Wuhan sequence reference (NC 045512 Wuhan-Hu-1, WIV04-reference). The x-axis labels the mutation number compared with the Wuhan sequence. Each Guinea sequence is highlighted by a bold circle to facilitate its position in the global tree. The 136 sequences from Institut Pasteur de Guinée Biobank are distributed into 7 clades: 20A, light gray; 20B, medium gray; 20C, dark gray; 20D, black; 20I/B.1.1.7/Alpha, blue; 21A/B.617.2/Delta, turquoise; 21D/B.1.525/Eta, yellow.