



Coding-Complete Genome Sequence and Phylogenetic Relatedness of a SARS-CoV-2 Strain Detected in March 2020 in Cameroon

Richard Njouom,^a Serge Alain Sadeuh-Mba,^a Jules Tchatchueng,^a Moussa Moïse Diagne,^b Ndongo Dia,^b Paul Alain Ngoupo Tagnouokam,^a Yap Boum,^{c,d} Achta Hamadou,^{a,d} Linda Ezzo,^d Ousmane Faye,^b Mathurin Cyrille Tejiokem,^a Marie Claire Okomo,^e Alain Etoundi,^d Elisabeth Carniel,^a Sara Eyangoh^{a,d}

^aCentre Pasteur du Cameroun, Yaounde, Cameroon

^bInstitut Pasteur de Dakar, Dakar, Senegal

^cEpicentre, Yaounde, Cameroon

^dPublic Health Emergency Operation Center, Ministry of Health, Yaounde, Cameroon

^eNational Public Health Laboratory, Yaounde, Cameroon

ABSTRACT We describe the coding-complete genome sequence of a severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) strain obtained in Cameroon from a 58-year-old French patient who arrived from France on 24 February 2020. Phylogenetic analysis showed that this virus, named hCoV-19/Cameroon/1958-CMR-YAO/2020, belongs to lineage B.1.5 and is closely related to an isolate from France.

A novel coronavirus, recently named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), that is related to the genus *Betacoronavirus* and the family *Coronaviridae* was identified in the city of Wuhan, China, in December 2019 (1). As of 30 December 2020, Cameroon has recorded 26,848 confirmed cases, 448 deaths (case fatality rate, 1.7%), and 25,468 recoveries (95%) (2). As of 30 December 2020, 294,185 complete genome sequences of SARS-CoV-2 have been obtained and deposited in the GISAID database (3). Phylogenetic analysis showed that these viruses diversified during the duration of the pandemic into two major lineages, A and B, with several sublineage diversifications (4). A recent study indicated that the African SARS-CoV-2 sequences also diversified into two lineages, A and B, with B being more diverse with multiple sublineages (5). Here, we report the coding-complete genome sequence of a SARS-CoV-2 strain obtained from a sample taken on 5 March 2020 in Cameroon from a 58-year-old French patient who arrived from France on 24 February 2020. Ethical clearance to conduct this study was obtained from the Cameroon National Ethics Committee (approval number 2020/05/1224/CE/CNERSH/SP).

Diagnosis of coronavirus disease 2019 (COVID-19) was conducted at the Centre Pasteur du Cameroun (Yaounde, Cameroon) by real-time reverse transcription (RT)-PCR (6). The nasopharyngeal swab sample obtained from the patient had cycle threshold (C_T) values of 31.9 for the E gene and 33.4 for the RdRP gene and was further sent to the Institut Pasteur de Dakar (Dakar, Senegal) for the first confirmatory diagnosis and sequencing. Briefly, viral RNA was extracted using the QIAamp viral RNA minikit and amplified by RT-PCR as described previously (7). The PCR products were purified and the DNA concentrations were measured with a Qubit 3 fluorometer (Invitrogen). DNA products (multiplex PCR pools A and B) were pooled in equal concentrations. DNA libraries were generated from the pooled amplicons using the Illumina DNA preparation kit according to the manufacturer's specifications. Whole-genome sequencing was performed with paired-end reads using the Illumina MiSeq reagent kit v3 (150 cycles) on an Illumina MiSeq instrument. The 187,596 reads were

Citation Njouom R, Sadeuh-Mba SA, Tchatchueng J, Diagne MM, Dia N, Tagnouokam PAN, Boum Y, Hamadou A, Ezzo L, Faye O, Tejiokem MC, Okomo MC, Etoundi A, Carniel E, Eyangoh S. 2021. Coding-complete genome sequence and phylogenetic relatedness of a SARS-CoV-2 strain detected in March 2020 in Cameroon. *Microbiol Resour Announc* 10:e00093-21. <https://doi.org/10.1128/MRA.00093-21>.

Editor Simon Roux, DOE Joint Genome Institute

Copyright © 2021 Njouom et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Richard Njouom, njouom@pasteur-yaounde.org.

Received 26 January 2021

Accepted 14 February 2021

Published 11 March 2021

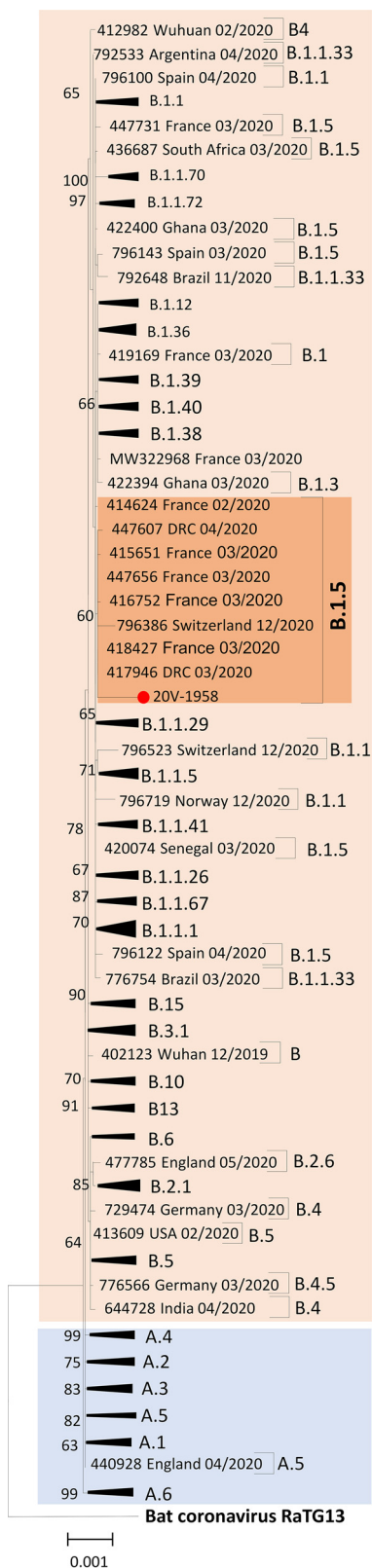


FIG 1 Phylogenetic analysis of the Cameroon SARS-CoV-2 complete genome sequence generated in this study (indicated with a red circle), with complete sequences from different countries available in the GISAID database as of 30 December 2020. Multiple sequence alignment and pairwise sequence comparisons were carried out with CLC Main Workbench v5.7.2 software. Inference of the phylogenetic tree was conducted by the maximum likelihood approach using PhyML v3.0 with SPR branch swapping. The reliability of the tree topology was estimated with 1,000 bootstrap replicates.

trimmed for quality and length and assembled by mapping to the reference genome from Wuhan (GenBank accession number [NC_045512.2](#)) using the already published USAMRIID-WRAIR pipeline (8). All tools were run with default parameters unless otherwise specified. Multiple sequence alignment and pairwise sequence comparisons with the hCoV-19/Cameroon/1958-CMR-YAO/2020 sequence obtained were carried out with CLC Main Workbench v5.7.2 software. Estimation of the best-fitting substitution model (GTR+I+Γ4) was conducted with Smart Model Selection (9) based on the Bayesian information criterion. Inference of the phylogenetic tree was conducted with the maximum likelihood approach using PhyML v3.0 (10) with subtree pruning and regrafting (SPR) branch swapping. The reliability of the tree topology was estimated with 1,000 bootstrap replicates.

The consensus sequence length was 29,686 bp, corresponding to 99.25% of the genome, with a GC content of 33.02%. The amino acid substitutions found in the hCoV-19/Cameroon/1958-CMR-YAO/2020 virus, compared with the reference, were M protein L206V, M protein N207D, nonstructural protein 6 (NSP6) L227F, and NSP12 P323L. The maximum likelihood phylogenetic tree (Fig. 1) shows that hCoV-19/Cameroon/1958-CMR-YAO/2020 clusters with B.1.1.5 SARS-CoV-2 lineage sequences isolated in France in March 2020.

The report of a SARS-CoV-2 coding-complete genome from Cameroon will allow investigators to follow the spread of this strain in the country and to evaluate its potential accumulation of mutations over time, improving our understanding of the molecular epidemiology of SARS-CoV-2 in Cameroon.

Data availability. The viral sequence from the patient was deposited in the GISAID EpiCoV SARS-CoV-2 database with the following identifiers: virus name, hCoV-19/Cameroon/1958-CMR-YAO/2020; GISAID accession number, EPI_ISL_512873; NCBI GenBank accession number, [MW566800](#). The raw reads were submitted to the NCBI Sequence Read Archive (SRA) under BioProject accession number [PRJNA701254](#) and SRA accession number [SRR13677539](#).

ACKNOWLEDGMENTS

We are thankful to the members of the SARS-CoV-2 surveillance system network in Cameroon for their fruitful contributions in field activities. This study was funded by the French Ministry for Europe and Foreign Affairs via the project “REPAIR Covid-19-Africa,” coordinated by the Pasteur International Network association.

REFERENCES

- Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R, Niu P, Zhan F, Ma X, Wang D, Xu W, Wu G, Gao GF, Tan W, China Novel Coronavirus Investigating and Research Team. 2020. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med* 382:727–733. <https://doi.org/10.1056/NEJMoa2001017>.
- Cameroon Ministry of Public Health. 2020. Statistiques COVID-19. <https://covid19.minsante.cm/>.
- Global Initiative on Sharing All Influenza Data. 2021. GISAID initiative. <https://www.epicov.org/epi3/>. Accessed 8 January 2021.
- Rambaut A, Holmes EC, O’Toole Á, Hill V, McCrone JT, Ruis C, Du Plessis L, Pybus OG. 2020. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat Microbiol* 5:1403–1407. <https://doi.org/10.1038/s41564-020-0770-5>.
- Motayo BO, Oluwasemowo OO, Oluola BA, Akinduti PA, Arege OT, Obafemi YD, Faneye AO, Isibor PO, Aworunse OS, Oranusi SU. 2021. Evolution and genetic diversity of SARS-CoV-2 in Africa using whole genome sequences. *Int J Infect Dis* 103:282–287. <https://doi.org/10.1016/j.ijid.2020.11.190>.
- Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, Bleicker T, Brünink S, Schneider J, Schmidt ML, Mulders DG, Haagmans BL, van der Veer B, van den Brink S, Wijsman L, Goderski G, Romette J-L, Ellis J, Zambon M, Peiris M, Goossens H, Reusken C, Koopmans MP, Drosten C. 2020. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 25:2000045. <https://doi.org/10.2807/1560-7917.ES.2020.25.3.2000045>.
- Quick J. 2020. nCoV-2019 sequencing protocol v2 (Guntt) V.2. https://www.protocols.io/view/ncov-2019-sequencing-protocol-v2-bdp7i5rn?version_warning=no.
- Kilianski A, Carcel P, Yao S, Roth P, Schulte J, Donarum GB, Fochler ET, Hill JM, Liem AT, Wiley MR, Ladner JT, Pfeffer BP, Elliot O, Petrosov A, Jima DD, Vallard TG, Melendrez MC, Skowronski E, Quan PL, Lipkin WI, Gibbons HS, Hirschberg DL, Palacios GF, Rosenzweig CN. 2015. Pathosphere.org: pathogen detection and characterization through a Web-based, open source informatics platform. *BMC Bioinformatics* 16:416. <https://doi.org/10.1186/s12859-015-0840-5>.
- Lefort V, Longueville JE, Gascuel O. 2017. SMS: Smart Model Selection in PhyML. *Mol Biol Evol* 34:2422–2424. <https://doi.org/10.1093/molbev/msx149>.
- Guindon S, Gascuel O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 52:696–704. <https://doi.org/10.1080/10635150390235520>.