JIDC | THE JOURNAL OF INFECTION IN DEVELOPING COUNTRIES

Coronavirus Pandemic

Partial spike gene sequencing for the identification of SARS-CoV-2 variants circulating in Cameroon in 2021

Chavely Gwladys Monamele[1,2], Serge Alain Sadeuh-Mba[1], Pauliana Vanessa Ilouga[1], Moïse Henri Y Moumbeket[1], Loique Landry E Messanga[1], Aristide Mounchili Njifon[1], Estelle L Madaha[1], Ripa M Njankouo[1], Paul Alain Tagnouokam-Ngoupo[1], Jules Brice M Tchatchueng[3], Mathurin Cyrille Tejiokem[3], Ronald Perraut[4], Sara Eyangoh[5], Richard Njouom[1]

[1] *Virology Service, Centre Pasteur of Cameroon, PO Box 1274, Yaounde, Cameroon*
[2] *Faculty of Health Sciences, University of Buea, PO Box 63, Buea, Cameroon*
[3] *Epidemiology and Public Health Service, Centre Pasteur of Cameroon, PO Box 1274, Yaounde, Cameroon*
[4] *Garoua Annex, Centre Pasteur of Cameroon, PO Box Garoua, Cameroon*
[5] *Scientific Department, Centre Pasteur of Cameroon, PO Box 1274, Yaounde, Cameroon*

Abstract

Introduction: Global monitoring of severe acute respiratory syndrome related coronavirus 2 (SARS-CoV-2) genetic sequences and associated metadata is essential for coronavirus disease 2019 (COVID-19) response. Therefore, Sanger's partial genome sequencing technique was used to monitor the circulating variants of SARS-CoV-2 in Cameroon.

Methodology: Nasopharyngeal specimen was collected from persons suspected of SARS-CoV-2 following the national guidelines between January and December 2021. All specimens with cycle threshold (Ct) below 30 after amplification were eligible for sequencing of the partial spike (*S*) gene of SARS-CoV-2 using the Sanger sequencing method.

Results: During the year 2021, 1481 real time reverse transcriptase polymerase chain reaction (RT-PCR) SARS-CoV-2 positive samples were selected for partial sequencing of the *S* gene of SARS-CoV-2. Amongst these, 878 yielded good sequencing products. A total of 231 probable variants (26.3%) were identified. The variants were mainly represented by Delta (70.6%), Alpha (15.6%), Omicron (7.4%), Beta (3.5%), Mu (1.7%) and Gamma (0.4%). Phylogenetic analysis of the probable variants from Cameroon with reference strains confirmed that all prior and current variants of concern (VOC) clustered with their respective reference sequences.

Conclusions: The surveillance strategy implemented in Cameroon, based on partial sequencing of the *S* gene enabled identification of the major circulating variants and provided information on the distribution of these variants, which contributed to implementing public health measures to control disease spread in the country.

Key words: SARS-CoV-2; genome; Sanger; sequencing; variants.

J Infect Dev Ctries 2024; 18(5):701-709. doi:10.3855/jidc.18155

(Received 28 February 2023 – Accepted 08 August 2023)

Introduction

In December 2019, the first cases of a novel coronavirus were reported in Wuhan, China; and eventually led to critical threats to global public health and economy [1]. On 11 March 2020, the World Health Organization (WHO) characterized coronavirus disease 2019 (COVID-19) as a pandemic. Initially, this virus was named 2019-nCoV and later designated as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) by the International Committee on Taxonomy of Viruses due to its taxonomic and genomic relationships with coronavirus species that are associated with severe acute respiratory syndrome (SARS) [2]. SARS-CoV-2 primarily attacks the respiratory system; and typical symptoms include fever, cough, fatigue, and loss of taste or smell. Chest discomfort and respiratory distress syndrome often occur in severe cases [3]. According to WHO, as of February 2023, there have been more than 754 million cases of COVID-19, including over 6 million deaths, globally [4]. The first case of COVID-19 in Cameroon was reported on 6 March 2020, and as of October 2022, there have been more than 123,000 confirmed cases with 1,965 deaths [5,6]. Four significant waves have been identified in the country, and each was driven by a variant. These waves occurred in May–June 2020, March–April 2021, September–October 2021, and December 2021 [5,6].

Since its first detection in Wuhan, SARS-CoV-2 has followed various evolutionary pathways. WHO has classified variants identified with fitness-enhancing

mutations as variants of concern (VOC), variants of interest (VOI) or variants under monitoring (VUM), based on their phenotypic characteristics, including transmissibility, disease severity, risk of reinfection, and impact on diagnostics and vaccine performance [7]. The classification of the SARS-CoV-2 evolves continuously. The currently circulating VOCs include the Omicron variant that has displaced previously circulating VOCs (Alpha (B.1.1.7), Beta (B.1.351), Gamma (P.1), and Delta (B.1.617.2)). There are currently no VOIs, however previously circulating VOIs included Epsilon (B.1.427 and B.1.429), Zeta (P.2), Eta (B.1.525), Theta (P.3), Iota (B.1.526), Kappa (B.1.617.1), Lambda (C.37), and Mu (B.1.621) [7].

Five VOCs that have been identified so far. Alpha was the first VOC described in the United Kingdom (UK) by end of December 2020. Beta and Gamma were detected in South Africa and Brazil respectively between December 2020 and January 2021. Delta was first described in India in October 2020. Omicron is the most recent VOC to be identified and was first reported in South Africa in November 2021 [8]. The common characteristic of these variants is that they share the N501Y mutation in the spike (S) protein, which is precisely the target of most COVID-19 vaccines. In addition, mutations such as N501Y, E484K, and K417N in the S protein can affect viral fitness and transmissibility [9]. WHO suggested that the region to be sequenced should cover at least the entire N-terminal and receptor binding domain (RBD) (amino acid 1–541, 1623 bp) to reliably differentiate between the circulating variants [10].

Several strategies are employed for the identification and characterization of the SARS-CoV-2 variants to correlate amino acid mutation profiles with virus transmissibility, pathogenicity, clinical significance, and possible vaccine evasion [6]. Identification of variants involves sequencing the complete genome of the virus to establish the possible mutations in the variant by comparison with the reference strain first isolated in Wuhan. The whole genome surveillance technology currently used for SARS-CoV-2 is not always successful, particularly when there is not enough viral load in the specimen [11]. In addition, the next generation sequencing (NGS) technology is also known to be associated with computational errors and biases in base-calling [12]. In such cases, target specific mutation assays are needed to identify or verify variants of concern [11]. A second strategy is to sequence only the regions of the gene coding for the S protein in which the mutations that define each of the variants are found [6]. This can be performed with Sanger sequencing and is a "first-generation" DNA sequencing method that uses oligonucleotide primers to seek out specific regions of the DNA [13,14]. Although Sanger sequencing is too laborious and expensive for whole genome sequencing, it is routinely used when sequencing specific genes or fragments of genes [15].

Global monitoring of SARS-CoV-2 genetic sequences and associated metadata is a crucial component of the COVID-19 response, including tracking the geographic spread of SARS-CoV-2 over time, and rapidly detecting and assessing mutations that may affect pathogenicity and transmission of the virus, to determine what counter measures should be adopted. Countries must adopt a genomic surveillance strategy based on available resources for a sustainable monitoring system. Therefore, Sanger's partial genome sequencing technique was used in Cameroon to monitor circulating variants of SARS-CoV-2.

Methodology

Collection and preparation of samples

Nasopharyngeal specimens were collected from persons suspected to be infected with SARS-CoV-2 between January and December 2021. Sample collection was done following the national guidelines from recognized recruitment sites in Cameroon. Centre Pasteur Cameroon (CPC) was the first laboratory to initiate the diagnosis of SARS-CoV-2 and received specimens collected from specified sites in Yaounde and some other sites in regions where polymerase chain reaction (PCR) diagnosis was not established. Specimens were collected in viral transport medium and transported, in coolers containing ice packs, on the same day to the virology laboratory of CPC, where analyses were performed.

Detection of SARS-CoV-2 by molecular assays

RNA was extracted from the samples using the QIAamp Viral RNA Mini Kit (Qiagen, Hilden, Germany). Diagnosis of SARS-CoV-2 was performed using a real-time reverse transcription polymerase chain reaction (rtRT-PCR) test kit from DaAn Gene Co., Ltd. of Sun Yatsen University (DaAn Gene Co., Ltd, Guangzhou, China). This assay targets the nucleocapsid (*N*) and the open reading frames 1a and 1b (*ORF1ab*) genes of SARS-CoV-2. A cut-off value of positivity was fixed at 37 cycles for both genes. Samples with cycle threshold (Ct) below the cut-off value for at least one gene were considered positive. Each specimen was run with an internal control (*RNase P*) which was

positive for all samples irrespective of whether they were positive or negative for the target genes.

Sampling strategy

Specimens with Ct value below 30 cycles for at least one gene (*N* and/or *ORF1ab*) were eligible for amplifying the partial *S* gene of SARS-CoV-2 and subsequently for sequencing using Sanger's method. Randomized samples representative of the geographic and demographic distribution of SARS-CoV-2 infections in Cameroon according to WHO recommendation were selected for sequencing [16]. If there were less than 30 specimens with appropriate Ct values in any particular week, all samples from that week were considered for sequencing.

Amplification of partial S gene in samples positive for SARS-CoV-2

Amplification of a fragment of the *S* gene was performed using the Superscript III One Step Reverse Transcription-Polymerase Chain Reaction (RT-PCR) System (Invitrogen, Carlsbad, California, USA) enzyme. Two PCR reactions were required to obtain a PCR product of 934 bp. The primer set for the first PCR consisted of SPfwd1 (AATTAC-CTGTATAGATTGTTTAGG) and SPrev1 (TTTGTAAATTTGTTTGACTTGTGC). Meanwhile, the primer set for the second PCR consisted of SPfwd2 (GAGAGAGATATTTCAACTGAAATC) and SPrev2 (TGTCTTGTTCAACAGCTATTCCAG).

The 25 µL reaction mixture for the first PCR reaction was composed of 12.5 µL of 2x PCR buffer, 0.2 µM forward and reverse primers, 0.75 µL enzyme and 5 µL RNA. The first PCR product was a 1113 bp fragment of the *S* gene and the following program was used for the amplification: 45 °C for 30 min, 94 °C for 2 min, 35 cycles at (94 °C for 15 sec, 45 °C for 30 sec, 72 °C for 90 sec) and 72 °C for 10 min.

The second PCR reaction was performed using the *Taq* DNA polymerase enzyme 5 U/µL (Promega, Madison, USA), and the first PCR product was used as a template for nested amplification. A total reaction volume of 50 µL contained 5 µL of 10x PCR buffer, 2 µL of 10 mM dNTP mix, 2 µL of 50 mM $MgCl_2$, 1 µL of each primer (at a working concentration of 10 µM), 0.5 µL of *Taq* DNA polymerase 5U/µL, and 2 µL of the first PCR product. The following thermal cycling conditions were used for amplification of the 934 bp fragment of the *S* gene: 94 °C for 2 min, 35 cycles at (94 °C for 15 sec, 45 °C for 30 sec, 72 °C for 90 sec), and 72 °C for 5 min. Amplicons of the appropriate size after gel electrophoresis were sent for sequencing at GENEWIZ UK, LTD (Hope End, UK) with the sense and anti-sense primer set used for its amplification.

Determination of VOCs and VOIs based on identified mutations

The sequences were assembled following Sanger's sequencing using the reference sequence, SARS-CoV-2-Wuhan-Hu-1_NC045512, in CLC Main Workbench

Table 1. Centers for Disease Control and Prevention (CDC) classification of variants during the study period.

Variant	Mutations of interest	Pango lineage	Date of designation
Alpha	H69-V70 del, Y144 del, **N501Y, A570D, P681H, T716I,** S982A, D1118H	B.1.1.7 and Q lineages	**VOC: December 2020** **VUM:** September 2021
Beta	L18F, D80A, D215G, R246I, K417N, **E484K, N501Y, A701V**, 242-244 del	B.1.351	**VOC: December 2020** **VUM:** September 2021
Gamma	L18F, T20N, P26S, D138Y, R190S, K417T, **E484K, N501Y, H655Y**, T1027I	P.1	**VOC: December 2020** **VUM:** September 2021
Delta	T19R, Y144 del, 157-158 del, K417N, L452R, **T478K, D614G, P681R**, S950N	B.1.617.2	**VOC: June 2021** **VUM:** April 2022
Omicron	A67V, H69-V70 del, T95I, G142D, 143-145 del, N211 del, L212I, R214_insEPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, **S477N, S478K, E484A, Q493R, Q496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K**, D796Y, N856K, Q954H, N969K, L981F	B.1.1.529, BA.1, BA.1.1, BA.2, BA.3, BA.4 and BA.5 lineages	November 2021
Epsilon	S13I, W152C, L452R, D614G	B.1.427, B.1.429	**VOC: March 2021** **VOI: February 2021** **VUM:** September 2021
Zeta	**E484K, D614G**, V1176F	P.2	**VOI: February 2021** **VUM:** September 2021
Eta	Q52R, A67V, H69-V70 del, 144 del, **T478K, D614G, Q677H**, F888L	B.1.525	**VOI: February 2021** **VUM:** September 2021
Iota	L5F, T95I, D235G, **E484K, D614G, A701V**	B.1.526	**VOI: February 2021** **VUM:** September 2021
Kappa	L452R, **E484Q, D614G, P681R**	B.1.617	**VOI: May 2021** **VUM:** September 2021
Mu	T95I, Y144S, Y145N, R346K, **E484K, N501Y, D614G, P681H**, D950N	B.1.621	**VUM:** September 2021

Mutations in the ACE2 RBD that could be identified from the partial *S* gene sequencing are bolded. Del: deletion; VOC: variants of concern; VUM: variants under monitoring; VOI: variants of interest.

version 5.5. The sequences obtained in FASTA format were uploaded to the GISAID CoV mutation application [17] to identify mutations in the different specimens in comparison to the reference strain, hCoV-19/Wuhan/WIV04/2019. Identified mutations were compared to that of recognized VOCs and VOIs to assign a probable classification of the SARS-CoV-2 viruses from Cameroon. The sequences were further uploaded into Nextclade [18] to classify them into Pango lineages and WHO clades. When the results obtained using the two methods varied, judgement on the molecular epidemiology of SARS-CoV-2 at the time was made on a case-by-case basis. Sequences ranged between nucleotide position 1400 and 2327 of the *S* gene and thus could enable the identification of amino acid mutations ranging between position 454-794 of ACE2 receptor binding domain (RBD). Table 1 lists the amino acid mutations and the variants to which they correspond according to Centers for Disease Control and Prevention (CDC) classification [19].

Phylogenetic analysis of Cameroon sequences

Phylogenetic analyses of sequences generated in Cameroon was performed with respect to the reference sequences. The reference sequences were downloaded from National Center for Biotechnology Information (NCBI) and uploaded in UShER (Ultrafast Sample placement on Existing Tree) [20] to confirm the clade designation and the sequence's placement in the UCSC/UShER phylogenetic tree. Alignments were performed in Seaview version 4.6.1 and the phylogenetic tree of partial fragment of the *S* gene was generated by Molecular Evolutionary Genetics Analysis (MEGA) software version 6.0 [21]. Evolutionary history was inferred using the maximum likelihood method and distances were computed using the Tamura-Nei model. The number of boot-strap replications was set to 1000 and bootstrap values above 50% were labelled on major tree branches for reference. All sequences generated during this period were submitted to NCBI database with accession numbers ON514632 to ON515457.

Ethical considerations

This study was approved by the Cameroon National Ethics Committee (Reference N°: 2020/05/1224/CE/CNERSH/SP).

Results

A total of 59,785 nasopharyngeal specimens were screened within the one year of genomic surveillance activity in Cameroon, and 7,679 (12.8%) were confirmed positive for SARS-CoV-2 by rtRT-PCR. Out of these, 1,481 positive samples were selected for partial sequencing of gene *S* of the SARS-CoV-2 based on the Ct value, and geographic and demographic distribution. Amongst these, 929 samples were successfully amplified using the partial *S* gene PCR and 878 samples yielded high coverage sequencing products. Specimens which were successfully amplified had lower Ct values (mean 21.0 ± 5.6; range: 7.0–36.4) compared to those in which amplification of the partial *S* gene failed (mean: 24.7 ± 5.4; range: 10.9–35.9). Additionally, specimens which failed to amplify predominantly originated from localities other than the central region (69.0%) in comparison to their counterparts that effectively amplified (44.1%) (Table 2).

Specimens selected for sequencing originated from eight of the country's ten regions; the west and southwest regions were not represented. Representative numbers of specimens were received throughout the year 2021 only from the central region. Specimens from the other regions were mostly collected during the months of March through May 2021, except for the northwest region where samples were collected from May through August. Information on the collection date was unavailable for specimens originating in the east and the far-north regions. The majority of samples were from the central region (n = 696), followed by Littoral (n = 397) and Adamawa (n = 153). The predominance of the probable variants was in the following descending order: centre, Littoral, northwest and north regions. Table 3 summarizes the number of samples tested per region and the variants identified.

Table 2. Characteristics of specimen with respect to results of amplification of the partial S gene.

Characteristics	Results of amplification of the partial *S* gene		*p* value
	Negative N = 552	Positive N = 929	
Mean cycle threshold (± SD)			
All genes	24.7 (± 5.4)	21.0 (± 5.6)	< 0.001
N gene	24.0 (± 4.6)	20.5 (± 5.1)	
ORF1ab gene	26.7 (± 4.8)	22.7 (± 5.1)	
Region			
Centre	171 (31.0%)	519 (55.9)	< 0.001
Other regions	381 (69.0%)	410 (44.1)	

SD: Standard deviation.

Table 3. Distribution of samples by region and variants identified.

Region	SARS-CoV-2 positive specimens selected	Number sequenced	Probable variants	Type of variants						
				Alpha	Beta	Gamma	Delta	Mu	Omicron	Recombinant
Centre	696	493	199	26	5	0	145	4	17	2
Littoral	397	222	13	9	3	0	1	0	0	0
Adamawa	153	4	0	0	0	0	0	0	0	0
Far North	67	57	0	0	0	0	0	0	0	0
North	66	54	2	1	0	1	0	0	0	0
East	47	11	0	0	0	0	0	0	0	0
South	3	2	0	0	0	0	0	0	0	0
North West	39	31	13	0	0	0	13	0	0	0
Unknown	13	4	4	0	0	0	4	0	0	0
TOTAL	**1481**	**878**	**231**	**36**	**8**	**1**	**163**	**4**	**17**	**2**

The results obtained from analyses of these sequences showed that, there were 647 unassigned variants (73.7%) and 231 probable variants (26.3%). The term 'probable variants' is used here as the full *S* gene sequence was not obtained. The probable variants were mainly represented by Delta (n = 163, 70.6%), Alpha (n = 36, 15.6%), Omicron (n = 17, 7.4%), Beta (n = 8, 3.5%), Mu (n = 4, 1.7%), Gamma (n = 1, 0.4%) and recombinants (n = 2, 0.9%) (Table 4). The unassigned sequences fell into Nextstrain clade 19A (Pango lineage B), 19B (A.2.5, A.21, A.27, A.29 lineages), 20A (B.1, B.1.618, B.1.620, B.1.640 lineages), 20B (B.1.1, B.1.1.318) and 21D (B.1.525).

Most study specimens were collected during the months of March to May. There was a shift in variants identified throughout the year. During the first six months, most variants were unassigned and some were the Alpha variant. The Delta variant dominated from July through November, and Omicron in December 2021 (Figure 1).

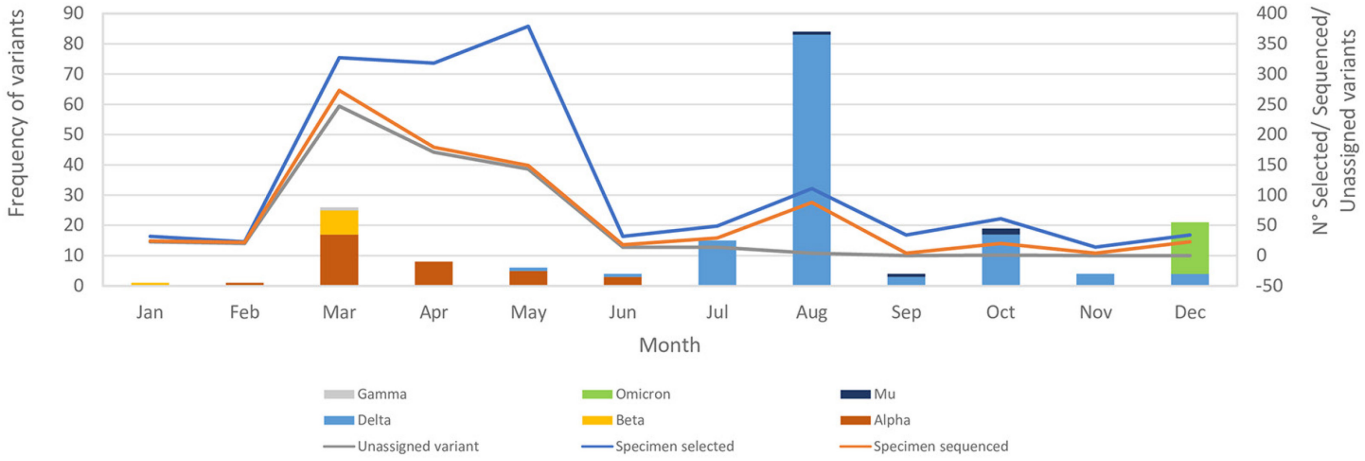
Phylogenetic analysis of the probable variants among the Cameroon strains was performed with reference strains from other regions of the world. The analyses confirmed that all current and prior VOCs clustered with their respective reference sequences, as indicated in Figure 2.

Protein alignment of the Omicron viruses was performed, as it is the unique VOC currently in circulation. Comparison of Cameroon sequences with the Wuhan-Hu-1 strain showed the presence of 14 of the mutations that have been described in the *S* gene of the Omicron variant: S477N, S478K, E484A, Q493R, Q496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K (Figure 3).

Table 4. Nextclade classification of variants.

Variant name	Frequency n (%)	Overall Proportion (%)
Unassigned sequences	647	73.7
19A	01 (0.2)	
19B	23 (3.6)	
20A	398 (61.5)	
20B	91 (14.1)	
21D	134 (20.7)	
Probable variants	231	26.3
Alpha	36 (15.6)	
Beta	08 (3.5)	
Gamma	01 (0.4)	
Delta	163 (70.6)	
Mu	04 (1.7)	
Omicron	17 (7.4)	
Recombinant	02 (0.9)	
TOTAL	878	100%

Figure 1. Distribution of variants identified by months in the year 2021.



Discussion

Documenting the circulation dynamics of SARS-CoV-2 variants in different regions of the world is crucial for monitoring virus transmission and contributing to global efforts towards combating the pandemic. In this study, we describe partial genome characterization of the SARS-CoV-2 virus in Cameroon and demonstrate the reliability of Sanger sequencing in identifying VOCs and VOIs. The *S* gene of SARS-CoV-2 has already been reported as sufficient to discriminate circulating VOCs as it harbours the major mutations [10]. With the knowledge of the epidemiology of COVID-19 in each population, it is possible to identify with high probability the variant in circulation using the partial fragment of the S gene [6]. Sanger's sequencing has been used in several settings to implement a molecular surveillance strategy for SARS-CoV-2 in association with whole genome sequencing or rtRT-PCR method [22–24].

In this study, we performed partial genome sequencing of the 934 bp long segment of the *S* gene of SARS-CoV-2, including signature mutations, to identify the variants. The partial *S* gene of SARS-CoV-2 was successfully amplified from about 60% of our specimens. Low viral loads and specimens originating from localities other than the central region were factors that were significantly associated with poorer outcomes. The specimen from other localities were transported through regular transport agencies to the CPC with occasional delays in reception due to poor road infrastructure and logistic issues. As some specimens with Ct value below 30 failed to amplify, other factors may have altered the success of the partial *S* gene PCR technique. Such factors include mutations in the *S* gene resulting in *S* gene target failure (SGTF) and consequently diagnosis failure [25] or mismatch between primers and target gene due to variability in the SARS-CoV-2 genes as reported by other studies [26].

Cameroon has experienced four major waves of COVID-19, three of which were noted in 2021 and occurred in the months of March-April, September-October, and December. Based on the identified variants in this dataset, the two latter waves were driven by Delta, and Omicron, although the subvariant could not be determined. These peaks can be attributed to the re-opening of the academic year and the end-of-year festivities, respectively, emphasizing the impact of crowded areas and close contacts in the transmission of the SARS-CoV-2 virus [27]. Similar trends have been observed in several other countries worldwide.

Figure 2. Phylogenetic analyses of Cameroon sequences (probable variants).

The sequences ranged between nucleotide position 1400 and 2327 of the S gene and the phylogenetic analysis was generated with the program MEGA version 6.0 using the maximum likelihood method. The distances were computed using the Tamura-Nei model. The number of bootstrap replications was set to 1000 and bootstrap values above 50% were labelled on the major tree branches for reference. The designation of clades was performed with respect to reference sequences downloaded from NCBI. The branches grouping the Alpha, Beta Delta and Omicron variants are colored blue, green, purple, and red, respectively.

By July 2021, Tunisia experienced four waves of COVID-19 in the months of March 2020, September 2020–January 2021, April 2021, and June–July 2021 [17]. The peak observed between September–January coincided with the start of school and the celebration of the new year.

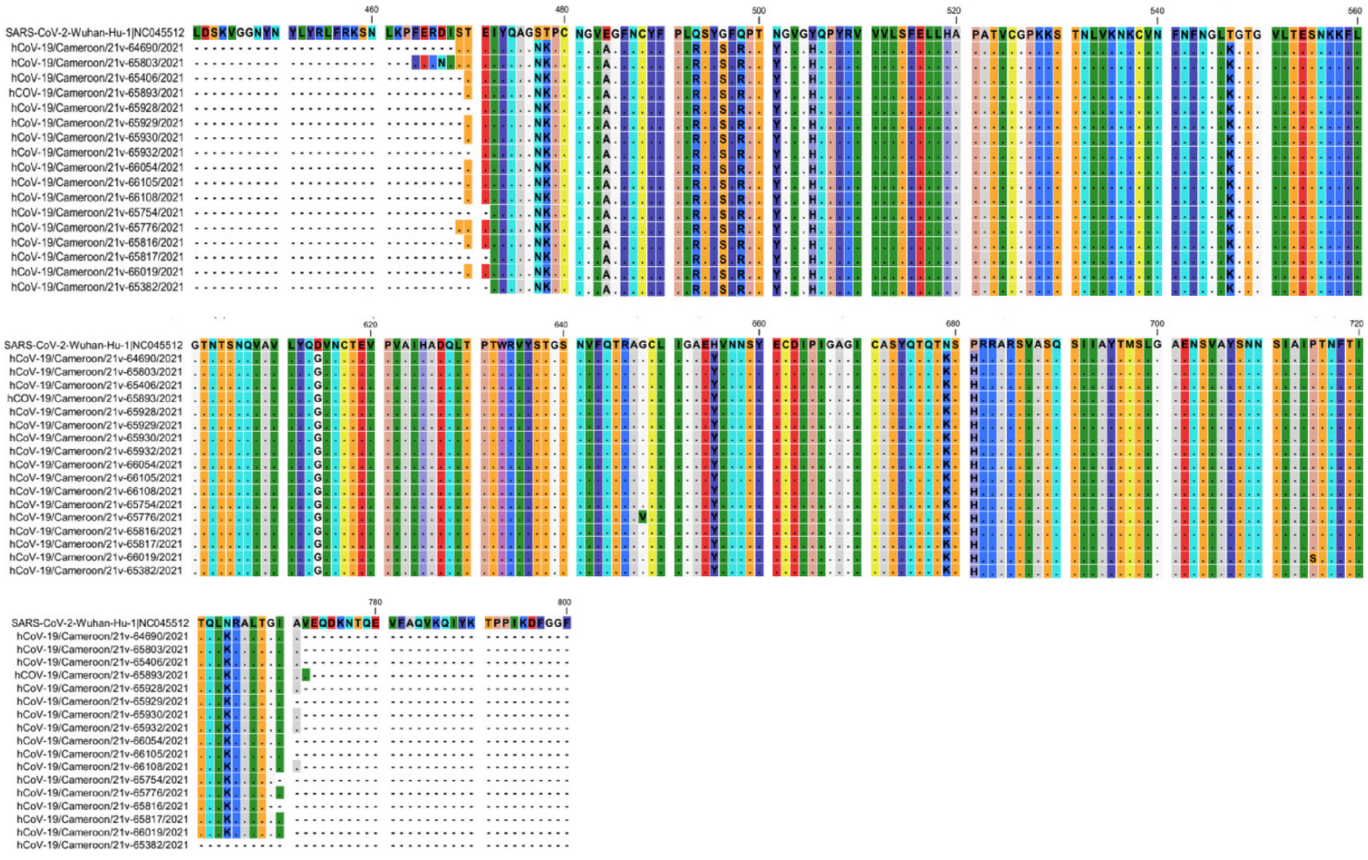
Regarding data from Central African countries submitted in GISAID in 2021, Alpha, Beta, and Eta variants were the most identified in the region between January and May. Delta was the dominant variant observed from May–September 2021. The last trimester of 2021 was initially driven by Delta and later, in December, by Omicron [28]. No specific variant could be identified in the first months of 2021 because most sequences fell in the unassigned group as no distinct mutation profile enabled their classification in any of the existing VOCs, VOIs or VUMs. The unassigned sequences fell into Nextstrain clade 19A, 19B, 20A, 20B and 21D corresponding to varied Pango lineages (A.2.5, A.21, A.27, A.29, B, B.1, B.1.1.318, B.1.525, B.1.618, B.1.620, and B.1.640 lineages). These lineages represent the earlier identified lineages as SARS-CoV-2 was initially grouped into lineage A or B due to likeliness to the early samples from Wuhan [25] as they lack specific mutations to identify them by sequencing. In addition, we noted a mixture of signature mutations among the sequences which fell in the unassigned group rendering their differentiation into any lineages impossible.

Apart from the three variants that dominated during the peak periods of activity in Cameroon, other variants being monitored were also detected in our samples in the year 2021, including Beta, Mu, and Gamma. Gamma and Mu have been observed at relatively lower frequencies in the African region; similar proportions were observed in this study. These variants have now been de-escalated from VOIs because they have significantly decreased, have had little impact on the overall epidemiological situation, and have not been associated with any concerning properties based on scientific evidence [29].

Two specimens had mutations in the *S* gene which were consistent with recombinant strains. However, whole genome sequencing is required to confirm the presence of recombinant viruses. This highlights the limitations of partial sequencing of the SARS-CoV-2 *S* gene and the need for further confirmation of the lineages identified especially in case of doubt.

The occurrence of mutations in the genome of the SARS-CoV-2 virus is an expected event since the virus

Figure 3. Protein alignment of the probable omicron viruses identified in Cameroon.



is an RNA virus; the low or no correction activity of the viral polymerase protein responsible for the nucleotide synthesis results in high mutation rates [30]. In addition to this natural predisposition, geographical factors, intrinsic viral factors (mutagenic properties), host factors (selective pressure of the host immune system, immune escape), and environmental factors could contribute to the appearance of SARS-CoV-2 variants [31].

Possible factors contributing to the propagation of COVID-19 variants across the country include importation by travelers, non-adherence to public health measures (social distancing, wearing of face masks), mass gathering events including the African Nation Football Championship in January 2021, and vaccine hesitancy [32,33].

A significant limitation of this study is the lack of metadata to map the presence of variants to sociodemographic and clinical characteristics such as travel history, uptake of COVID-19 vaccine, severity of COVID-19 disease (symptomatic/ asymptomatic), hospitalization status and presence of other comorbidities [16]. Additionally, comparisons of the regional distribution of variants could not be made because specimens collected throughout the year for all regions were not available.

Conclusions

In conclusion, the surveillance strategy implemented in Cameroon, based on partial sequencing of the *S* gene, enabled identification of the major circulating variants, and contributed to implementing public health measures to control the spread of the disease. The main variants found in 2021 in Cameroon were Delta, Alpha, Omicron, and to a lesser extent Beta, Mu and Gamma. As CPC is among the reference laboratories contributing to COVID-19 laboratory data in Cameroon, the observations made in this study can be extrapolated to the country.

Acknowledgements

We acknowledge all the staff involved in the laboratory detection of SARS-CoV-2 at the Centre Pasteur of Cameroon and the various laboratories that contributed the SARS-CoV-2 positives used in this study for their help with the processing and analysis of the samples. We are thankful to Mrs Seraphine Mojoko Eko for proofreading this manuscript.

References

1. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY, Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jon L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao Ge, Li H, Tao Z, Yang Y, Deng Z, Liu B, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT (2020) Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med 382: 1199-1207. doi: 10.1056/NEJMoa2001316.
2. Coronaviridae Study Group of the International Committee on Taxonomy of Viruses (2020) The species severe acute respiratory syndrome-related coronavirus: classifying 2019-nCoV and naming it SARS-CoV-2. Nat Microbiol 5: 536-544. doi: 10.1038/s41564-020-0695-z.
3. Centers for Disease Control and Prevention (2022) Symptoms of COVID-19. Available: https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html. Accessed: 12 April 2022.
4. World Health Organization (2022) WHO coronavirus (COVID-19) dashboard. Available: https://covid19.who.int/. Accessed: 4 February 2023.
5. World Health Organization (2023) Cameroon: WHO coronavirus disease (COVID-19) dashboard with vaccination data. Available: https://covid19.who.int. Accessed: 4 February 2023.
6. CCOUSP (nd) Cameroon COVID-19 situation report N°151. Available: https://www.ccousp.cm/download/cameroon-covid-19-situation-report-n134/. Accessed: 12 January 2023.
7. World Health Organization (2022) Tracking SARS-CoV-2 variants. Available: https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/. Accessed: 12 April 2022.
8. Cantón R, De Lucas Ramos P, García-Botella A, García-Lledó A, Gómez-Pavón J, González Del Castillo J, Hernández-Sampelayo T, Martín-Delgado MC, Martín Sánchez FJ, Martínez-Sellés M, Molero García JM, Moreno Guillén S, Rodríguez-Artalejo FJ, Ruiz-Galiana J, Bouza E (2021) New variants of SARS-CoV-2. Rev Esp Quimioter 34: 419-428. doi: 10.37201/req/071.2021.
9. Jia Z, Gong W (2021) Will mutations in the spike protein of SARS-CoV-2 lead to the failure of COVID-19 vaccines? J Korean Med Sci 36: e124. doi: 10.3346/jkms.2021.36.e124.
10. World Health Organization Regional Office for Europe (2021) Methods for the detection and identification of SARS-CoV-2 variants. Available: https://apps.who.int/iris/handle/10665/340067. Accessed: 7 September 2021.
11. Buchan BW, Wolk DM, Yao JD (2021) Rapid communication. SARS-CoV-2 variant testing. The AMP, IDSA, and PASCV. Available: https://www.amp.org/AMP/assets/File/clinical-practice/COVID/AMP_RCVariantTestingforSARSCOV2_4_28_21.pdf?pass=91. Accessed: 7 September 2021.
12. Abnizova I, Te Boekhorst R, Orlov Y (2017) Computational errors and biases in short read next generation sequencing. J Proteomics Bioinform 10: 1-17. doi: 10.4172/jpb.1000420.
13. Hagemann IS (2015) Overview of technical aspects and chemistries of next-generation sequencing. In Kulkarni S, Pfeifer J, editors. Clinical Genomics. Boston: Academic Press: 3-19. doi: 10.1016/B978-0-12-404748-8.00001-0.
14. Gomes A, Korf B (2018) Chapter 5 - Genetic Testing Techniques. Robin NH, Farmer MB, eds. Pediatric Cancer Genetics. Elsevier: 47-64. doi: 10.1016/B978-0-323-48555-5.00005-3.
15. Vernet G (2017) Genomics of infectious diseases and private industry. Tibayrenc M, ed. Genetics and evolution of infectious

diseases (Second Edition). London: Elsevier: 421-434. doi: 10.1016/B978-0-12-799942-5.00018-4.
16. World Health Organization (2021) Guidance for surveillance of SARS-CoV-2 variants: Interim guidance 9 August 2021. Available: https://www.who.int/publications/i/item/WHO_2019-nCoV_surveillance_variants. Accessed: 6 February 2023.
17. Bernasconi A, Cilibrasi L, Al Khalaf R, Alfonsi T, Ceri S, Pinoli P, Canakoglu A (2021) EpiSurf: metadata-driven search server for analyzing amino acid changes within epitopes of SARS-CoV-2 and other viral species. Database 2021: baab059. doi: 10.1093/database/baab059.
18. Nextrain (nd) Nextclade as pango lineage classifier: methods and validation - Nextclade documentation. Available: https://docs.nextstrain.org/projects/nextclade/en/stable/user/algorithm/nextclade-pango.html. Accessed: 6 May 2024.
19. CDC (2020) Coronavirus Disease 2019 (COVID-19). Available: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-classifications.html. Accessed: 25 April 2024.
20. Hinrichs A, Ye C, Turakhia Y, Corbett-Detig R (2024) The ongoing evolution of UShER during the SARS-CoV-2 pandemic. Nat Genet 56: 4-7. doi: 10.1038/s41588-023-01622-5.
21. Tamura K, Stecher G, Peterson D, Filipski A, Kumar S (2013) MEGA6: Molecular evolutionary genetics analysis Version 6.0. Mol Biol Evol 30: 2725-2729. doi: 10.1093/molbev/mst197.
22. Chouikha A, Fares W, Laamari A, Haddad-Boubaker S, Belaiba Z, Ghedira K, Kammoun Rebai W, Ayouni K, Khedhiri M, Ben Halima S, Krichen H, Touzi H, Ben Dhifallah I, Guerfali FZ, Atri C, Azouz S, Khamessi O, Ardhaoui M, Safer M, Ben Alaya N, Guizani I, Kefi R, Gdoura M, Triki H (2022) Molecular epidemiology of SARS-CoV-2 in Tunisia (North Africa) through several successive waves of COVID-19. Viruses 14: 624. doi: 10.3390/v14030624.
23. Torres C, Mojsiejczuk L, Acuña D, Alexay S, Amadio A, Aulicino P, Debat H, Fay F, Fernández F, Giri AA, Goya S, König G, Lucero H, Nabaes Jodar M, Pianciola L, Sfalcin JA, Acevedo RM, Bengoa Luoni S, Bolatti EM, Brusés B, Cacciabue M, Casal PE, Cerri A, Chouhy D, Dus Santos MJ, Eberhardt MF, Fernandez A, Fernández PDC, Fernández Do Porto D, Formichelli L, Gismondi MI, Irazoqui M, Campos ML, Lusso S, Marquez N, Muñoz M, Mussin J, Natale M, Oria G, Pisano MB, Posner V, Puebla A, Re V, Sosa E, Villanova GV, Zaiat J, Zunino S, Acevedo ME, Acosta J, Alvarez Lopez C, Álvarez ML, Angeleri P, Angelletti A, Arca M, Ayala NA, Barbas G, Bertone A, Bonnet A, Bourlot I, Cabassi V, Castello A, Castro G, Cavatorta AL, Ceriani C, Cimmino C, Cipelli J, Colmeiro M, Cordero A, Cristina C, Di Bella S, Dolcini G, Ercole R, Espasandin Y, Espul C, Falaschi A, Fernandez Moll F, Foussal MD, Gatelli A, Goñi S, Jofré ME, Jaramillo J, Labarta N, Lacaze MA, Larreche R, Leiva V, Levin G, Luczak E, Mandile M, Marino G, Massone C, Mazzeo M, Medina C, Monaco B, Montoto L, Mugna V, Musto A, Nadalich V, Nieto MV, Ojeda G, Piedrabuena AC, Pintos C, Pozzati M, Rahhal M, Rechimont C, Remes Lenicov F, Rompato G, Seery V, Siri L, Spina J, Streitenberger C, Suárez A, Suárez J, Sujansky P, Talia JM, Theaux C, Thomas G, Ticeira M, Tittarelli E, Toro R, Uez O, Zaffanella MB, Ziehm C, Zubieta M, Mistchenko AS, Valinotto L, Viegas M (2021) Cost-effective method to perform SARS-CoV-2 variant surveillance: detection of Alpha, Gamma, Lambda, Delta, Epsilon, and Zeta in Argentina. Front Med (Lausanne) 8: 755463. doi: 10.3389/fmed.2021.755463.
24. Fares W, Ghedira K, Gdoura M, Chouikha A, Haddad-Boubaker S, Khedhiri M, Ayouni K, Lamari A, Touzi H, Hammemi W, Medeb Z, Sadraoui A, Hogga N, Ben Alaya N, Triki H (2021) Sequencing using a two-step strategy reveals high genetic diversity in the *S* Gene of SARS-CoV-2 after a high-transmission period in Tunis, Tunisia. Microbiology Spectrum 9: e00639-21. doi: 10.1128/Spectrum.00639-21.
25. Mahilkar S, Agrawal S, Chaudhary S, Parikh S, Sonkar SC, Verma DK, Chitalia V, Mehta D, Koner BC, Vijay N, Shastri J, Sunil S (2022) SARS-CoV-2 variants: impact on biological and clinical outcome. Front Med (Lausanne) 9: 995960. doi: 10.3389/fmed.2022.995960.
26. Alkhatib M, Carioti L, D'Anna S, Ceccherini-Silberstein F, Svicher V, Salpini R (2022) SARS-CoV-2 mutations and variants may muddle the sensitivity of COVID-19 diagnostic assays. Microorganisms 10: 1559. doi: 10.3390/microorganisms10081559.
27. Azuma K, Yanagi U, Kagi N, Kim H, Ogata M, Hayashi M (2020) Environmental factors involved in SARS-CoV-2 transmission: effect and role of indoor environmental quality in the strategy for COVID-19 infection control. Environ Health Prev Med 25: 66. doi: 10.1186/s12199-020-00904-2.
28. Shu Y, McCauley J (2017) GISAID: Global initiative on sharing all influenza data - from vision to reality. Eurosurveillance 22: 30494. doi: 10.2807/1560-7917.ES.2017.22.13.30494.
29. European Centre for Disease Prevention and Control (2022) SARS-CoV-2 variants of concern as of 25 May 2022. Available: https://www.ecdc.europa.eu/en/covid-19/variants-concern#:~:text=ECDC%20regularly%20assesses%20new%20evidence,screening%2C%20or%20other%20scientific%20sources Accessed: 25 September 2022.
30. Denison MR, Graham RL, Donaldson EF, Eckerle LD, Baric RS (2011) Coronaviruses. RNA Biol 8: 270-279. doi: 10.4161/rna.8.2.15013.
31. Parra-Lucares A, Segura P, Rojas V, Pumarino C, Saint-Pierre G, Toro L (2022) Emergence of SARS-CoV-2 variants in the world: how could this happen? Life (Basel) 12: 194. doi: 10.3390/life12020194.
32. Esso L, Epée E, Bilounga C, Abah A, Hamadou A, Dibongue E, Kamga Y, Belinga S, Eyangoh S, Okomo MC, Mounagué C, Tiwoda C, Mandeng N, Onana T, Mendjime P, Mahamat F, Mballa GAE, Boum Y (2021) Cameroon's bold response to the COVID-19 pandemic during the first and second waves. Lancet Infect Dis 21: 1474-4457.33. doi: 10.1016/S1473-3099(21)00388-1.
33. Dinga JN, Njoh AA, Gamua SD, Muki SE, Titanji VPK (2022) Factors driving COVID-19 vaccine hesitancy in Cameroon and their implications for Africa: a comparison of two cross-sectional studies conducted 19 months apart in 2020 and 2022. Vaccines 10: 1401. doi: 10.3390/vaccines10091401.

Corresponding author
Professor Richard Njouom, PhD.
Centre Pasteur of Cameroon
PO Box1274, Yaounde, Cameroon
Tel: + 237 699654767
Email: njouom@pasteur-yaounde.org

Conflict of interests: No conflict of interests is declared.