MethodsX

Development of a Sequencing Protocol for SARS-CoV-2 Genotyping: An alternative Approach for Genomic Surveillance in low and middle-income settings

--Manuscript Draft--

Manuscript Number:	MEX-D-23-00489
Article Type:	Method Article - Direct Submission
Section/Category:	Biochemistry, Genetics and Molecular Biology
Keywords:	SARS-CoV-2; Variants of concern; Sanger sequencing; Resource-limited settings.
Corresponding Author:	Davy-Hyacinthe GOUISSI ANGUECHIA CAMEROON
First Author:	Davy-Hyacinthe GOUISSI ANGUECHIA, MsC
Order of Authors:	Davy-Hyacinthe GOUISSI ANGUECHIA, MsC
	Desire Takou Takoua
	Ezechiel Ngoufack Jagni Semengue, PhD
	Grace Beloumou
	Sandrine Djupsa
	Collins Chenwi, MD
	Alex Durand Nka, PhD
	Aissatou Abba
	Aude Christelle Ka'e
	Willy Le Roi Togna Pabo
	Aurelie Kengni
	Naomi Karell Etame
	Larissa Gaelle Moko
	Evariste Molimbou
	Derrick Tambe Ayuk Ngwese
	Rachel Audrey Nayang Mundoa
	Michel Tommo
	Nadine Fainguem Nadine Fainguema, PhD
	Vittorio Colizzi
	Carlo-Federico Perno
	Joseph Fokam, PhD
Abstract:	Timely detection of SARS-CoV-2 variants of concerns (VOCs) is key for an efficient response against COVID-19 and any future pandemic. However, access to Next Generation Sequencing (reference tool) remains challenging for resource-limited settings. Following experience with the HIV-1 Sanger-sequencing, we developed an in-house SARS-CoV-2 sequencing protocol using specific primers targeting a fragment of SARS-CoV-2 Spike protein region. Using the developed Sanger-sequencing protocol on 163 nasopharyngeal specimens, 130 Spike sequences were successfully generated, giving an overall performance of 79.75% (130/163) [72.76-85.64] (95% CI), effective on SARS-COv-2 lineages of origin and all variants (Alpha, Beta, Gamma, Delta, Omicron) referenced under accession numbers OQ248255 – OQ248384. Thus, in the frame of limited access to NGS, adapting this targeted sequencing protocol would contribute timeously in the genomic surveillance of pathogens with

	epidemic/pandemic potential.
Suggested Reviewers:	Kathleen Subramoney National Health Laboratory,Charlotte Maxeke Johannesburg Academic south Africa kathleen.subramoney@nhls.ac.za He is a South African virologist with articles on molecular biology techniques.
	Lavanya Singh University of KwaZulu-Natal Singhl@ukzn.ac.za He has expertise in the domain
	James E. San E. San University of KwaZulu-Natal sanemmanueljames@gmail.com has expertise in the field
	Anmol M. Kiran University of Liverpool Anmol.Kiran@liverpool.ac.uk has expertise in the subject
	Tulio de Oliveira University of KwaZulu-Natal deoliveira@ukzn.ac.za He has expertise in the subject
	Sonia Etenna Lekana-Douki Centre interdisciplinaire de recherche medical de Franceville(CIRMF),Franceville,Gabon s_lekana@yahoo.fr he has expertise in the field
	Rodney S. Daniels Worlwide Influenza Centre rod.daniels@crick.ac.uk he has expertise in the field
	Chika Kingsley Onwuamah Centre for Human virology and Genomics ck.onwuamah@nimr.gov.ng he has relevant expertise
Opposed Reviewers:	

Development of a Sequencing Protocol for SARS-CoV-2 Genotyping: An alternative Approach for Genomic Surveillance in low and middle-income settings

Davy-Hyacinthe Anguechia Gouissi[a,b,],Desire Takou[a], Ezechiel Ngoufack Jagni Semengue[a,c], Grace Beloumou[a], Sandrine Djupsa[a] , Collins Chenwi[a,d], Alex Durand Nka[a,c], Aissatou Abba[a], Aude Christelle Ka'e[a,c], Willy Le Roi Togna Pabo[a], Aurelie Kengni[a], Naomi Karell Etame[a], Larissa Gaelle Moko[a,b], Evariste Molimbou[a], Derrick Tambe Ayuk Ngwese[a,b], Rachel Audrey Nayang Mundo[a], Michel Tommo[a], Nadine Fainguem[a,c], Vittorio Colizzi[a,c,e], Carlo-Federico Perno[a,c,f], Joseph Fokam[a,b,g*]

a. *Chantal BIYA International Reference Centre for research on HIV/AIDS prevention and management, Yaounde, Cameroon;*
b. *Faculty of Medicine and Biomedical Sciences, University of Yaounde I, Yaounde, Cameroon;*
c. *University of Rome "Tor Vergata", Rome, Italy;*
d. *Mvangan District Hospital, Mvangan, Cameroon*
e. *Faculty of Science and Technology, Evangelic University of Cameroon, Bandjoun, Cameroon;*
f. *Bambino Gesu Pediatric Hospital, Rome, Italy;*
g. *Faculty of Health Sciences, University of Buea, Buea, Cameroon.*

Corresponding author:
*Contact: +237-222315450 (office); E-mail: josephfokam@gmail.com.

Keywords
SARS-CoV-2, Variants of concern, Sanger sequencing; Resource-limited settings.

Abstract
Timely detection of SARS-CoV-2 variants of concerns (VOCs) is key for an efficient response against COVID-19 and any future pandemic. However, access to Next Generation Sequencing (reference tool) remains challenging for resource-limited settings. Following experience with the HIV-1 Sanger-sequencing, we developed an in-house SARS-CoV-2 sequencing protocol using specific primers targeting a fragment of SARS-CoV-2 Spike protein region. Using the developed Sanger-sequencing protocol on 163 nasopharyngeal specimens, 130 Spike sequences were successfully generated, giving an overall performance of 79.75% (130/163) [72.76-85.64] (95% CI), effective on SARS-COv-2 lineages of origin and all variants (Alpha, Beta, Gamma, Delta, Omicron) referenced under accession numbers OQ248255 – OQ248384. Thus, in the frame of limited access to NGS, adapting this targeted sequencing protocol would contribute timeously in the genomic surveillance of pathogens with epidemic/pandemic potential.

- The protocol for sanger sequencing of SARS-CoV-2 was designed using new primers to amplify and sequence a fragment of the part of the Spike protein gene.

- This sequencing assay were successfully generated, giving an overall performance of 79.75% (130/163).
- Sanger-sequencing could represent a good alternative to population-based genotyping.

Specifications table

Subject area	Biochemistry, Genetics and Molecular Biology
More specific subject area	Medical virology, Biological Sciences
Name of your method	In-house Sanger sequencing Assay: Alternative method for SARS-CoV-2 Genotyping in low and middle-income settings
Name and reference of original method	Not applicable
Resource availability	Not applicable

Method details

Clinical specimens: A laboratory-based study was conducted from April 2020 throughout August 2022 on nasopharyngeal specimens of individuals confirmed positive to COVID-19 by real-time PCR at the Virology Laboratory of the "Chantal BIYA" International Reference Centre in Yaoundé, Cameroon. Briefly, CIRCB a national reference laboratory for SARS-CoV-2 molecular testing and a reference centre for genomic surveillance at country-level. Samples were selected based on cycle threshold (Ct) value under 35 (ORF1ab, N genes) for *DaAn*Gene RT-PCR assay (https://en.daangene.com/).

Detection of SARS-CoV-2 VOC with Sanger sequencing (Home made protocol)

Following experience in Sanger sequencing for HIV genotyping [5,6] , the protocol for sequencing of SARS-CoV-2 was designed using new primers to amplify and sequence a fragment of the part of the Spike protein gene.

RNA extraction

Viral RNA was manually extracted from 200 µL of nasopharyngeal clinical swabs using the *DaAn*Gene viral RNA Mini kit according to the manufacturer's protocol (www.daangene.com). SARS-CoV-2 RNA extract was then processed directly for conventional reverse transcription and amplification.

Preparation of the primers

The primers arrived in lyophilised form and were resuspended with RNase-DNase free molecular grade water to a stock solution concentrated at 100µM.

Reverse transcription and PCR amplification

Viral RNA was retro-transcribed and amplified using the kit One-Step Invitrogen (SuperScript® One-Step for long templates RT/ PCR; Foster City, CA) and 2 different sequence-specific primers (Table2): **38F** as the forward, and **1191R** as the reverse (given position refers to the Wuhan strain of SARS-CoV-2).

RT-PCR Reagents and Conditions

The RT-PCR reaction contained for each sample, 25 µl reaction mix which includes:

- 12µl 2Xreaction Mix (Invitrogen),
- 4 µl $MgSO_4$ (5 mM),
- 1.5 µl DNAse-RNAse-free water,
- 0.375 µl sense primer 38F (10 mM stock),
- 0.375 µl antisense primer 1191R (10 mM stock),
- 0.75 µl RNAseOUT (5 U/ µl Invitrogen),
- 0.75 µl RT-Taq (Superscript III RT/Platinum high fidelity), and
- 05 µl of extracted RNA.
- Run the RTPCR-reaction in the thermal cycler with the appropriate program (Table1), using sequence-specific primers (Table 2).

Table1 : PCR Cycler program

Temperature (°C)	Time	Cycles
50	30 min	1
94	2 min	1
95	30 s	40x
52	30s	
72	90s	
72	10 min	1
4	Forever	1

Table2: Primers used for Reverse Transcriptase Amplification and Sequencing reactions.

Name	Primer sequence (5′–3′)	Length (bp)	Direction	Use
38F	GTCAGTGTGTTAATCTTACAACCAG	25	Forward	Amplification and Sequencing reaction
514F	TCTCAGCCTTTTCTTATGGACCT	23	Forward	RT-PCR reaction
655R	CCTGAGGGAGATCACGCACTA	21	Reverse.	RT-PCR reaction
1191R	TGCATAGACATTAGTAAAGCAGAGA	25	Reverse.	Amplification and Sequencing reaction

Legend. Bp : base pair ; F : Forward ; R : Reverse ; RT-PCR : Reverse-Transcriptase Polymerase Chain Reaction

Gel Electrophoresis

PCR products were analysed on 1% agarose gel electrophoresis by visual comparison of the amount of PCR product (1200bp expected size) with a low-molecular-weight DNA ladder (Figure1).

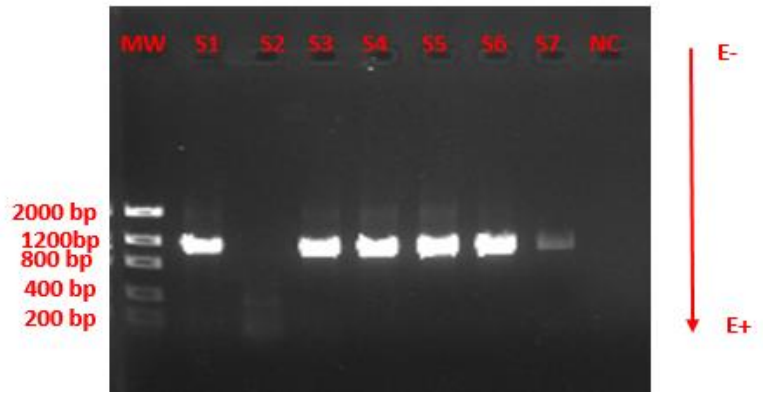

Figure 1: Image of gel electrophoresis

Legend. NC: Negative Control; MW: Molecular Weight low DNA mass ladder; bp: base pair; S: Samples

Purification

Following revelation on agarose-gel electrophoresis, samples with a positive result (PCR products) were purified by using ExoSAP-IT™ kit as per the manufacturer's instructions (Applied Biosystems™, Lithuania). Purified PCR products were then referred for sequencing.

Sanger-based Sequencing Reaction

The amplified products from the Orf-Spike region were completely sequenced in the sense and antisense directions using on a Applied Biosystem thermal cycler ProFlex PCR system, using overlapping sequence-specific primers (Table 2): **38F, 514F, 665R, 1191R.**

Sequencing Reaction Reagents and Conditions

The reaction mixture for the sequencing reaction contained:

- 1.5 µl Big Dye Terminator V3.1,
- 6.5 µl big dye Buffer (from the kit),
- 4.8 µl DNAse/ RNAse-free water,
- 3.2 µl primer (1 µM) and
- 2 µl of purified cDNA.

Sequencing Reaction Procedure

1- Prepare 4 microtubes for each sample, each microtube corresponding to a specific primer;
2- Dispense 13 µl of master mix, with each primer in a corresponding microtube;
3- Dispenser 2 µl of purified DNA in 4 microtubes each corresponding to a specific primer;
4- Run the sequence reaction in the thermal cycler with the appropriate program.

Sequencing conditions were as follows: 35 cycles (96°C, 10 s; 55°C, 10 s; 60°C, 4 min); 1 cycle of 4°C for 30 min. The sequencing product with four overlapping primers (Figure 2), was then purified by gel filtration chromatography using Sephadex G-50 resin (Sigma-Aldrich) to eliminate excess primers, unincorporated dideoxynucleotides (ddNTPs) and salts.

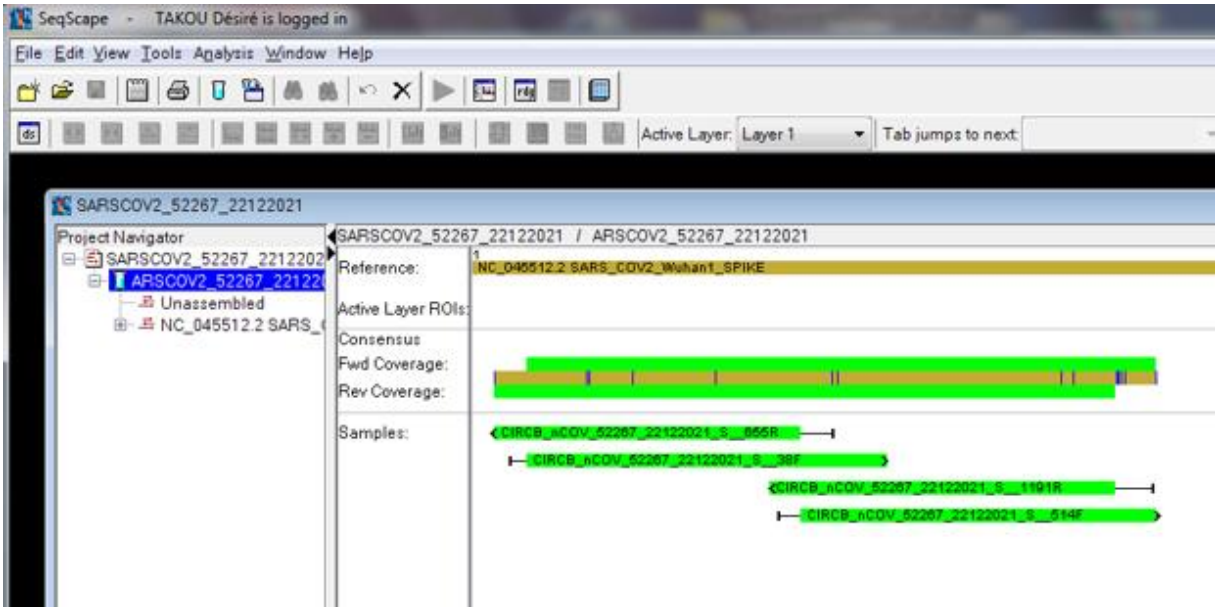

Figure 2: Mapping of sequencing primers coverage

Legend. 38F and 514F are forward primers; 655R and 1191R are reverse primers

SARS-CoV-2 sequence analysis and interpretation

Capillary electrophoresis was performed using an automated genetic analyser (ABI 3500, Applied Biosystems Tokyo, Japan). Sequences were aligned, assembled and edited with reference to the Wuhan strain of SARS-CoV-2 using SeqScape Version 2.7. Spike Sequences were interpreted using the COV19 database Stanford algorithm (https://covdb.stanford.edu) and NCBI (National Center for biotechnology information). SARS-CoV-2 Spike gene nucleotide sequences were submitted to GenBank using Bankit (https://www.ncbi.nlm.nih.gov/WebSub/).

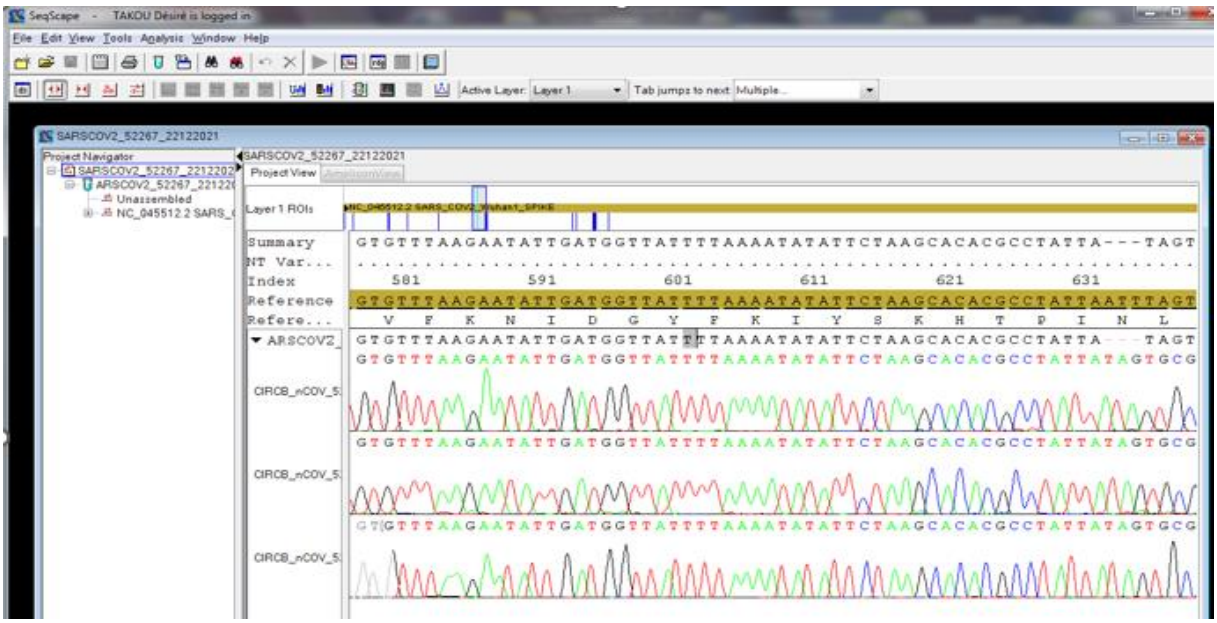

Figure 3: Electrophoregram of SARS-CoV-2 sequences

Sanger sequencing assay performance

Of the 163 positive nasopharyngeal specimens processed, 88.96 % (145/163) were successfully amplified and 130 Spike-sequences were successfully generated out of the 145 amplicons (89.65 %), giving an overall sequencing performance rate of 79.75% (130/163) [72.76-85.64] (95%CI). This sequencing assay detected all strains of SARS-CoV-2 (viruses of the original lineage, Alpha, Beta, Gamma, Delta and Omicron variants).

Data availability:

SARS-CoV-2 sequences generated in this study are available in Genbank under the following accession numbers: OQ248255 – OQ248384.

Conclusion

In the frame of SARS-CoV-2 molecular epidemiology, Sanger sequencing contributes in detecting all viral lineages and VOCs circulating in Cameroon. This in-house sequencing protocol is suitable alternative to NGS that can be adapted for a rapid identification of any other pathogen of pandemic or epidemic potentials in resource-limited countries.

Ethics statements

Ethical approval was obtained from the Cameroon National Ethics Committee for research on human health (2022/01/1430/CE/CNERSH/SP), within the frame of the EDCTP (*European and Developing Countries Clinical Trials Partnership) PERFECT-Study (RIA2020-EF3000*). A study information sheet was provided to each individual, and a written informed consent was obtained from all study participants.

CRediT author statement

Davy-Hyacinthe Anguechia Gouissi, Joseph Fokam, Desire Takou, Ezechiel Ngoufack Jagni Semengue, Vittorio Colizzi, Carlo-Federico Perno.**: Conceptualized and initiated the protocol.** Davy-Hyacinthe Anguechia gouissi, Ezechiel Ngoufack Jagni Semengue, Collins Chenwi, Grace Beloumou, Sandrine Djupsa, Alex Durand Nka, Willy Pabo, Aissatou Abba, Aude Christelle Ka'e, Aurelie Kengni, Naomi Karell Etame, Larissa Gaelle Moko, Evariste Molimbou, Rachel Audrey Nayang Mundo, Michel Tommo, Nadine Fainguem: **Collected and Analysed the samples.** Davy-Hyacinthe Anguechia gouissi, Ezechiel Ngoufack Jagni Semengue, Aude Christelle Ka'e, Vittorio Colizzi, Carlo-Federico Perno, Joseph Fokam, Desire Takou: **Interpreted the results.** All authors: **Revised and Approved the final version of the manuscript.**

Acknowledgments

We would like to thank the Chantal Biya International Reference Centre (CIRCB) for research on HIV/AIDS prevention and care, in Yaoundé-Cameroon for hosting the current study. We also acknowledge the national public health emergency operations coordination centre on COVID-19 for contributing to the project implementation.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

References

[1] Chiara M, D'Erchia AM, Gissi C, Manzari C, Parisi A, Resta N, et al. Next generation sequencing of SARS-CoV-2 genomes: challenges, applications and opportunities. Briefings in Bioinformatics 2021;22:616–30. https://doi.org/10.1093/bib/bbaa297.

[2] Pereira R, Oliveira J, Sousa M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. Journal of Clinical Medicine 2020;9:132. https://doi.org/10.3390/jcm9010132.

[3] Pennell NA, Mutebi A, Zhou Z-Y, Ricculli ML, Tang W, Wang H, et al. Economic Impact of Next-Generation Sequencing Versus Single-Gene Testing to Detect Genomic Alterations in Metastatic Non–Small-Cell Lung Cancer Using a Decision Analytic Model. JCO Precision Oncology 2019:1–9. https://doi.org/10.1200/PO.18.00356.

[4] Dorlass EG, Lourenço KL, Magalhães RDM, Sato H, Fiorini A, Peixoto R, et al. Survey of SARS-CoV-2 genetic diversity in two major Brazilian cities using a fast and affordable Sanger sequencing strategy. Genomics 2021;113:4109–15. https://doi.org/10.1016/j.ygeno.2021.10.015.

[5] Fokam J, Salpini R, Santoro MM, Cento V, D'Arrigo R, Gori C, et al. Performance evaluation of an in-house human immunodeficiency virus type-1 protease-reverse transcriptase genotyping assay in Cameroon. Arch Virol 2011;156:1235–43. https://doi.org/10.1007/s00705-011-0982-3.

[6] Takou D, Fokam J, Teto G, Santoro M-M, Ceccherini-Silberstein F, Nanfack AJ, et al. HIV-1 drug resistance testing is essential for heavily-treated patients switching from first- to second-line regimens in resource-limited settings: evidence from routine clinical practice in Cameroon. BMC Infectious Diseases 2019;19:246. https://doi.org/10.1186/s12879-019-3871-0.

Graphical Abstract

Graphical abstract: Detecting SARS-CoV-2 VOC with Sanger Spike-sequencing