

## 1 **Genomic epidemiology of SARS-CoV-2 in Sudan: A retrospective analysis 2020-2022**

2 **Authors:** Maureen W. Mburu<sup>1#</sup>, Arnold W. Lambisia<sup>1#</sup>, Musab Elnegoumi<sup>2#</sup>, Eman O. M.  
3 Nour<sup>2</sup>, Osama M. M. Khair<sup>3</sup>, Elamin Abualas<sup>2</sup>, Rehab M Elhassan<sup>2</sup>, Muatsim A. M. Adam  
4 <sup>2</sup>, Khadija Said Mohammed<sup>1,12,13</sup>, John M. Morobe<sup>1</sup>, Leonard Ndwiga<sup>1</sup>, Edidah O. Moraa<sup>1</sup>,  
5 Timothy O. Makori<sup>1</sup>, Nickson Murunga<sup>1</sup>, Esther Nyadzua Katama<sup>1</sup>, Dorcas W. Waruguru<sup>4</sup>,  
6 Sarah Mwangi<sup>4</sup>, Sofonias K. Tessema<sup>4</sup>, Mawahib H. E. Adam<sup>2</sup>, Iman Mahmoud<sup>2</sup>, Hanan  
7 SalmanBabikir Mohamed<sup>2</sup>, Yousif A. A. Mustafa<sup>2</sup>, Mohamedtaha M. A. Mohamed<sup>2</sup>,  
8 Rasheeda H.A. Ahmed<sup>2</sup>, Nuha Y.I. Mohamed<sup>2</sup>, Rahma H. Ali<sup>2</sup>, Rasha S.M. Ebraheem<sup>2</sup>, Huda  
9 H.H. Ahmed<sup>2</sup>, Hamadelniel E. Abdalla<sup>2</sup>, Izdihar Mukhtar<sup>2</sup>, Abdalkhalig E. M. W. Alla<sup>2</sup>, Omer  
10 H. A. Elsheikh<sup>2</sup>, Abdualmoniem O. Musa<sup>5,7,8</sup>, Mahmoud Taha Faki<sup>5</sup>, Gada Mustfa Ahmed<sup>5</sup>,  
11 Thoyba E. R. Ahmed<sup>2</sup>, Mohammed A. O. Abdalla<sup>2</sup>, Mujtaba O. K. Mohamed<sup>2</sup>, Rayan A. A.  
12 Ahmed<sup>2</sup>, Samah M. M. AwadElkarim<sup>2</sup>, Basamat A. E. M. Ahmed<sup>2</sup>, Zainab A. I. Shumo<sup>2</sup>,  
13 Eiman N. A. Ahmed<sup>2</sup>, Samia S. A. Rushwan<sup>6</sup>, Phillip Bejon<sup>1</sup>, D. James Nokes<sup>1,9</sup>, Lynette  
14 Isabella Ochola-Oyier<sup>1</sup>, Charles N. Agoti<sup>1,10</sup>, George Githinji<sup>1,11\*</sup>, Shahinaz A. Bedri<sup>2</sup>

### 15 16 **AFFILIATIONS:**

17 <sup>1</sup> Kenya Medical Research Institute-Wellcome Trust Research Programme (KWTRP), Kilifi,  
18 Kenya.

19 <sup>2</sup> National Public Health Laboratory, Khartoum, Sudan.

20 <sup>3</sup> Department of Molecular Biology, National University Biomedical Research Institute.

21 <sup>4</sup> Africa Centre for Disease Control and Prevention (Africa CDC), Addis Ababa, Ethiopia.

22 <sup>5</sup> General Administration of Laboratories and Blood Banks, Ministry of Health, Kassala State,  
23 Sudan.

24 <sup>6</sup> Port Sudan Public Health Laboratory, Red Sea State, Sudan.

25 <sup>7</sup> Microbiology Department, Faculty of Medicine and Health Sciences, Kassala University,  
26 Sudan.

27 <sup>8</sup> Microbiology Department, Faculty of Medical Laboratory Sciences, Gezira University,  
28 Sudan.

29 <sup>9</sup> School of Life Sciences and Zeeman Institute for Systems Biology and Infectious Disease  
30 Epidemiology Research (SBIDER), University of Warwick, Coventry, UK.

31 <sup>10</sup> Department of Public Health, Pwani University, Kilifi, Kenya.

32 <sup>11</sup> Department of Biochemistry and Biotechnology, Pwani University, Kilifi, Kenya.

33 <sup>12</sup>The Francis Crick Institute, London, UK.

**NOTE:** This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

34 <sup>13</sup> Division of Biosciences, Faculty of Life Sciences, University College London, London, UK.

35

36 # Authors with equal contributions

37 \* Corresponding author

38

39 Corresponding: [gginthinji@kemri-wellcome.org](mailto:gginthinji@kemri-wellcome.org)

40

41 **Keywords:**

42 Sudan, SARS-CoV-2, VOC, VOI and introduction

43

44 **Words: Abstract (202 words), Main text: (3932 words)**

45

46

47

48 **Abstract**

49 SARS-CoV-2 was first detected in Sudan on 13<sup>th</sup> March 2020. Here, we describe the genomic  
50 epidemiology of SARS-CoV-2 in Sudan between May 2020 and April 2022 to understand the  
51 introduction and transmission of SARS-CoV-2 variants in the country. A total of 667 SARS-  
52 CoV-2 positive samples were successfully sequenced using the nCoV-19 ARTIC protocol on the  
53 Oxford Nanopore Technology ( $\geq 70\%$  genome completeness). The genomes were compared  
54 with a select contemporaneous global dataset to determine genetic relatedness and estimate  
55 import/export events. The genomes were classified into 37 Pango lineages within the **ancestral**  
56 **strain** (107 isolates across 13 Pango lineages), **Eta variant of interest (VOI)** (78 isolates in 1  
57 lineage), **Alpha variant of concern (VOC)** (10 isolates in 2 lineages), **Beta VOC** (26 isolates  
58 in 1 lineage), **Delta VOC** (171 isolates across 8 lineages) and **Omicron VOC** (242 isolates  
59 across 12 lineages). We estimated a total of 144 introductions of the observed variants from  
60 different countries across the globe. Multiple introductions of the Eta VOI, Beta VOC and  
61 Omicron VOC were observed in Sudan mainly from Europe and Africa. These findings suggest  
62 a need for continuous genomic surveillance of SARS-CoV-2 to monitor their introduction and  
63 spread consequently inform public health measures to combat SARS-CoV-2 transmission.

## INTRODUCTION

64 The first confirmed SARS-CoV-2 case in Sudan was detected on 13<sup>th</sup> March 2020 and by end  
65 of April 2022 the country had confirmed a total of 62,131 cases and 4,933 COVID-19 related  
66 deaths [1]. A population-based study conducted between 1<sup>st</sup> March to 10<sup>th</sup> April 2021 in  
67 Omdurman city in Sudan estimated a seroprevalence of 54.6% with higher SARS-CoV-2  
68 seropositivity among older people (>50 years)[2].

69

70 The first major SARS-CoV-2 wave of infections was observed between 7<sup>th</sup> April and 22<sup>nd</sup>  
71 September 2020 [1]. The government introduced various countermeasures such as closure of  
72 schools, airports and borders, restriction of movement across localities, dusk to dawn curfews  
73 and banning of religious and social gatherings in order to curb the spread of the virus, but these  
74 measures were relaxed from September 2020 [3]. Due to challenges in contact tracing earlier  
75 in the pandemic in Sudan, the SARS-COV-2 cases persisted and resulted in the second surge  
76 of infections between November 2020 and January 2021 [2] [3]. Notably, political instability  
77 in the country affected tracking and containment of SARS-CoV-2 infections. The third major  
78 wave was observed between December 2021 and February 2022 [1].

79

80 The major waves of SARS-CoV-2 cases observed in Sudan coincided with the infection waves  
81 observed in other African countries and these have been attributed to import/export events  
82 between different countries globally [4]. Sudan is bordered by seven countries namely, Eritrea,  
83 Ethiopia, South Sudan, Central African Republic, Chad, Libya, and Egypt with high movement  
84 across the borders and major cities lying either along the Red Sea or Nile River. Previously,  
85 transmission of viral infectious diseases such yellow fever and recurrent arbovirus outbreaks  
86 have attributed to rural-urban migration within the country[3]. However, the transmission  
87 dynamics of SARS-CoV-2 variants in Sudan and across its borders during the pandemic to date  
88 has not been previously described.

89

90 Genomic surveillance has been instrumental in monitoring and understanding SARS-CoV-2  
91 variants evolution and spread in many countries globally [5]–[8]. This study aims to describe  
92 SARS-CoV-2 variants and transmission dynamics in Sudan using 667 whole genome  
93 sequences obtained from PCR-confirmed cases collected between May 2020 and April 2022.

## 94 **METHODS**

### 95 **Ethical approval**

96 The genomic surveillance protocol used at Kenya Medical Research Institute (KEMRI)  
97 Wellcome Trust (KWTRP) was reviewed and approved by the Scientific and Ethics Review  
98 Committee (SERU) in KEMRI (SERU #4035). The SARS-CoV-2 positive samples were part  
99 of the public health emergency response thus individual participant consent was waived.

100

### 101 **SARS-CoV-2 sampling.**

102 A total of 667 SARS-CoV-2 positive samples were sent to KWTRP-Kenya, ILRI-Kenya and  
103 South Africa for whole genome sequencing. The dates of collection spanned from 20<sup>th</sup> May  
104 2020 to 18<sup>th</sup> April 2022.

105

### 106 **Viral RNA extraction and cDNA synthesis**

107 SARS-CoV-2 samples were shipped from the National Public Health Laboratory – Sudan to  
108 KWTRP-Kenya (n=405). The samples shipped to KWTRP-Kenya were extracted using  
109 QIAamp Viral RNA Mini kit (52906, Qiagen, Hilden, Germany) following the manufacturer's  
110 instructions. SARS-CoV-2 positivity screening was done using various quantitative PCR kits.  
111 Samples with a cycle threshold (Ct)  $\leq$  34 were selected for cDNA synthesis. 8 $\mu$ l of RNA was  
112 reverse transcribed using 2  $\mu$ l LunaScript® RT SuperMix Kit.

113

### 114 **Whole genome amplification, library preparation and sequencing at KWTRP-Kenya**

115 The SARS-CoV-2 genome was amplified using Q5® Hot Start High Fidelity 2X Master Mix  
116 (NEB, US) and the modified V4 ARTIC primers [9]. The amplicons were cleaned using 1x  
117 AMPure XP beads (Beckman Coulter, US) as per the manufacturer's protocol, eluted in  
118 nuclease free water and the eluent quantified using Qubit dsDNA HS Assay Kit (ThermoFisher  
119 Scientific, US). Library preparation was performed as previously described [9], [10].

120

121 End-prep was done using 50ng of PCR amplicons according to the ARTIC nCoV-2019  
122 sequencing protocol v3 (LoCost) [11] using the NEBNext Ultra II End repair/dA-tailing Kit  
123 (NEB, US). Barcode ligation was done using the native barcoding expansion (EXP-NBD196)  
124 kit (Oxford Nanopore Technology, UK) and NEBNext Blunt/TA Ligase Master Mix (NEB,  
125 US). Thereafter, adaptor ligation was done using NEBNext Quick Ligation Module reagents  
126 (NEB, US), Adaptor Mix II (ONT, UK) and 100ng of the barcoded sample following the

127 manufacturers instruction. The final loading concentration was normalised to 15ng and loaded  
128 on a SpotON R9 flow cell and sequenced on a GridION device [12].

129

### 130 **SARS-CoV-2 genome assembly**

131 Basecalling and demultiplexing were executed simultaneously during sequencing using Guppy  
132 version 4.0.5 (<https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html>) with a  
133 minimum Q score of 1. Consensus SARS-CoV-2 genomes were generated using the ARTIC  
134 bioinformatics protocol (<https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html>).  
135 Briefly, the reads were filtered based on length (300-700bp) and mapped to the reference  
136 genome (GenBank accession MN908947.3) using Minimap2  
137 (<https://github.com/lh3/minimap2>). Fast5 reads were used to polish the consensus genomes  
138 using Nanopolish (<https://github.com/jts/nanopolish>) and consensus genomes generated with a  
139 read depth of 20X.

140

### 141 **Additional Sudan data**

142 Additional sequences from Sudan generated at ILRI-Kenya (n= 199) and South Africa (n=204)  
143 available in the Global Initiative on Sharing All Influenza Data (GISAID) were also included  
144 in the subsequent analysis [13].

145

### 146 **Lineage assignment, phylogenetics and migration analysis**

147 The consensus genomes were assigned lineages using Pangolin version 4.0.6 [13]. Clade  
148 assignment and mutation profiling was done using NextClade v1.13.2 [14].

149

150 All sequences originating from Sudan were analysed against a global dataset of 13,122 SARS-  
151 CoV-2 genomes to ensure that global representation of sequences. The selected global subset  
152 had  $\geq 95\%$  genome coverage and were selected evenly from across the Pango lineage, time  
153 (month of collection) and continents. The genomes were collected from 1<sup>st</sup> September 2020 to  
154 19<sup>th</sup> January 2022. All global reference genome sequences and associated metadata in this  
155 dataset are published in GISAID's EpiCoV database <https://doi.org/10.55876/gis8.220815ga>.

156

157 The sequences from Sudan and the global dataset were aligned using Nextalign version 1.4.1  
158 (<https://github.com/neherlab/nextalign>) against the SARS-CoV-2 reference genome. A  
159 maximum likelihood (ML) tree was generated using IQTREE version 2.1.3  
160 (<http://www.iqtree.org/>) and a branch support was assessed using 1000 bootstrap iterations.

161 The ML trees were converted to time resolved phylogenetic trees using TreeTime  
162 (<https://github.com/neherlab/treetime>).

163

164 A Python script was used to estimate SARS-CoV-2 import and export events in Sudan based  
165 on the origin of sequence and time resolved tree's tips and internal nodes. The import/export  
166 events were visualized using R packages alluvial and tidyverse [15]

167

## 168 **RESULTS**

### 169 **Baseline demographic characteristics**

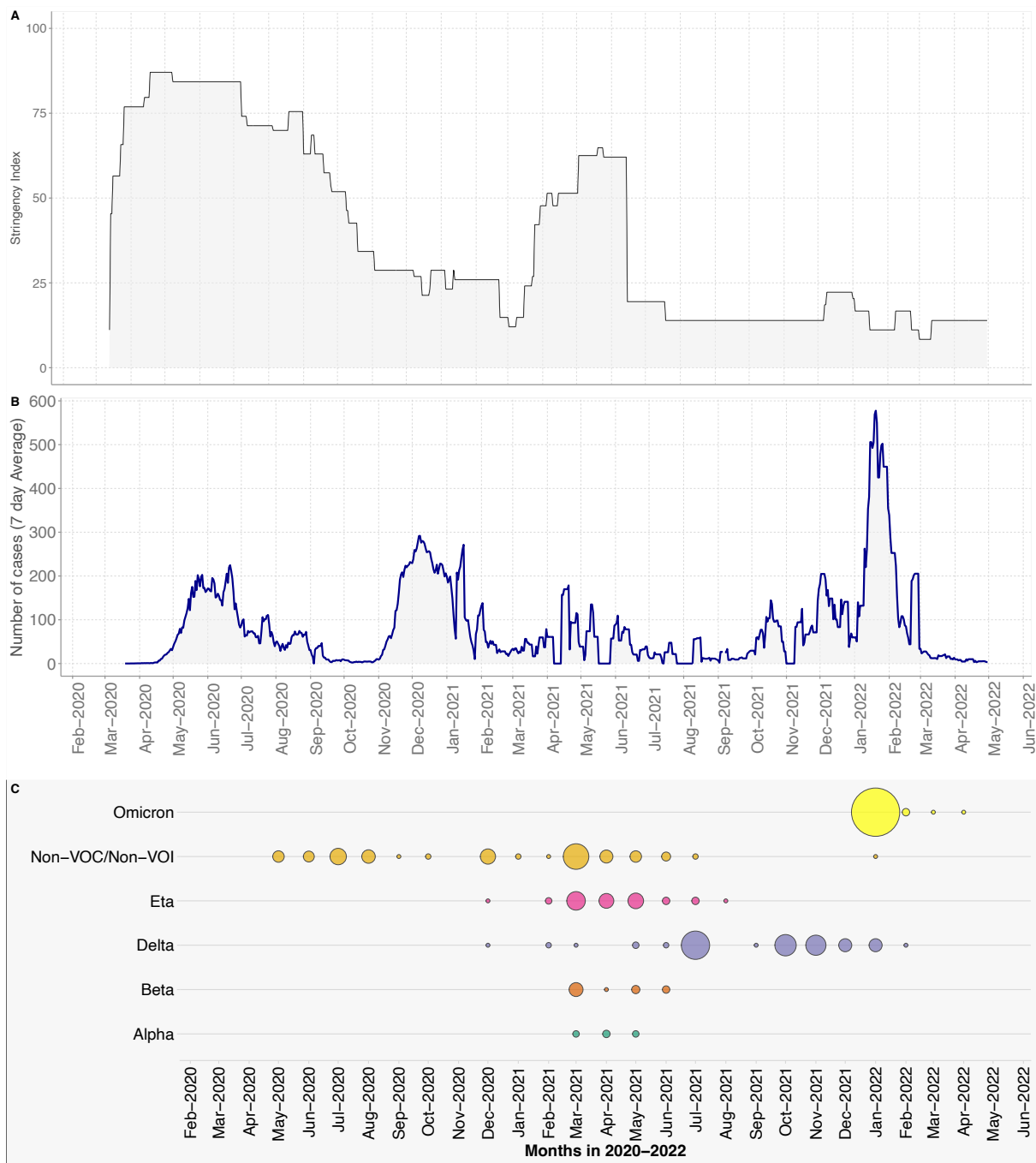
170 A total of 667 SARS-CoV-2 samples were collected from 11 different locations in Sudan and  
171 a majority of these samples were from Khartoum (n=390, 58.5%), Kassala (n=180, 27%) and  
172 Port Sudan (n=66, 9.9%). Among this population, the median age was 40 years. Out of the 667  
173 cases, 237 (35.5%) were presented with symptoms and only 4 (0.6%) had known international  
174 travel history. Metadata for variables such as vaccination status, travel history, sex, age, and  
175 reason for testing was lacking for >10% of the cases (**Supplementary table 1**).

176

### 177 **SARS-CoV-2 variants surveillance in Sudan**

178 Genomic surveillance of SARS-CoV-2 started in May 2020 in Sudan and the 667 genomes  
179 sequenced were classified into multiple variants namely, non-variant of concern/ non-variant  
180 of internet (non-VOC/non-VOI (n=140), Eta (n=78), Alpha (n=10), Beta (n=26), Delta (n=171)  
181 and Omicron (n=242) (**Figure1C**).

182



183

184 **Figure 1: (A)** Stringency index showing the government intervention levels. **(B)** A weekly  
 185 rolling average of the daily confirmed SARS-CoV-2 cases in Sudan as populated in  
 186 <https://ourworldindata.org/coronavirus/country/sudan>. **(C)** Temporal distribution of detected  
 187 SARS-CoV-2 variants.

188

189 The predominant lineages within the non-VOC/non-VOI group were A.29 (n=43, 30.7%), A  
 190 (n=42, 30.0%) and B.1 (n=29, 20.7%). These lineages persisted through the first two major  
 191 waves of SARS-CoV-2 infections in Sudan (**Figure 2**). The Eta VOI (B.1.525) was first  
 192 observed in December 2020 and was sustained until August 2021 with the peak cases seen in



193 March 2021 (**Figure 2**). Alpha and Beta VOC cases were seen to cocirculate between March  
 194 2021 and May 2021.

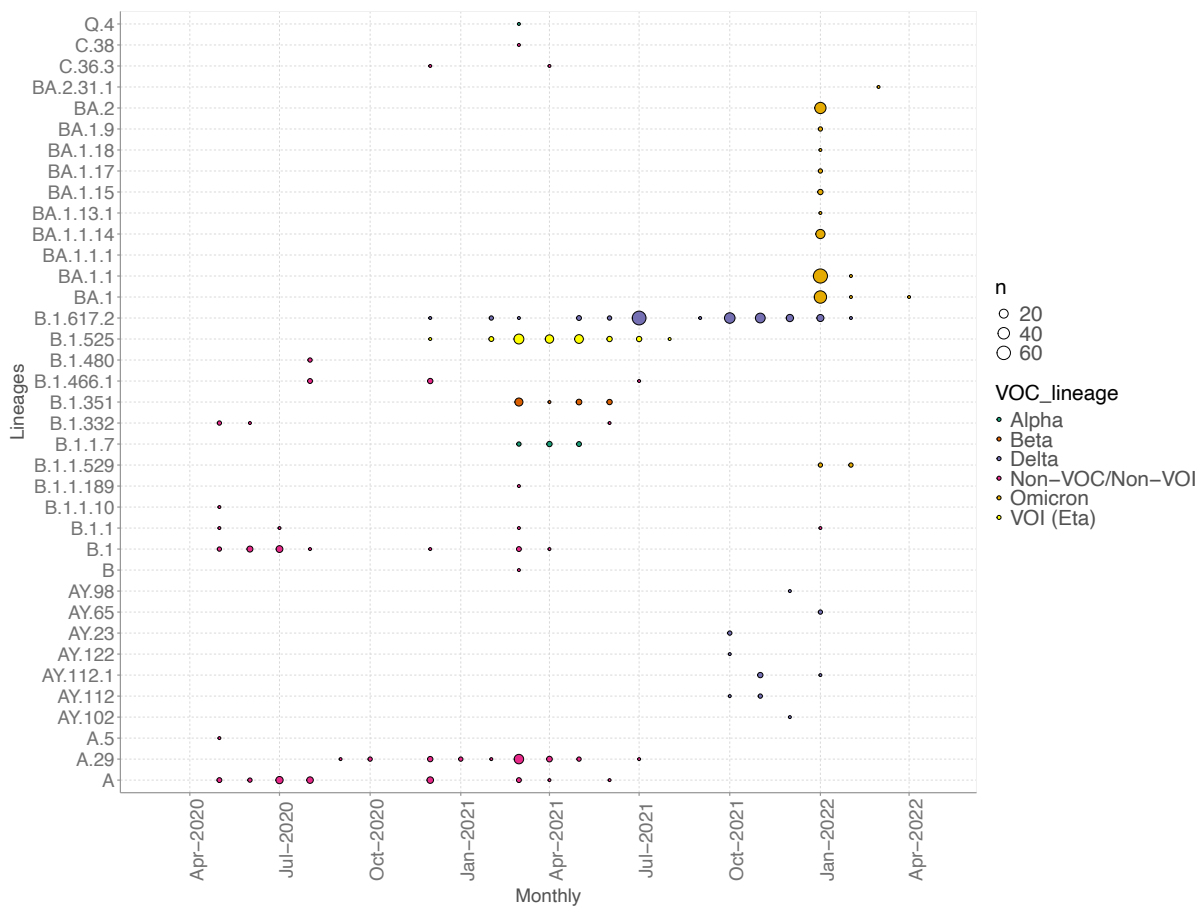
195

196 Nine Delta VOC lineages were observed between December 2020 and February 2022. The  
 197 most predominant lineage across this period was B.1.617.2 (n=156, 91.2%). The other Delta  
 198 VOC lineages were observed between October 2021 and February 2022 (**Figure 2**). By July  
 199 2021, the Delta VOC was the predominant SARS-CoV-2 variant in circulation in Sudan and  
 200 become dominant by August 2021.

201

202 Twelve lineages within the Omicron VOC were observed in between January and April 2022  
 203 during the third major wave in Sudan (**Figure 2**). The most common Omicron lineages were  
 204 BA.1.1 (n=87, 36.0%), BA.1 (n=60, 24.8%), BA.2 (n=49, 20.2%), and BA.1.1.14 (n=28,  
 205 11.6%).

206



207

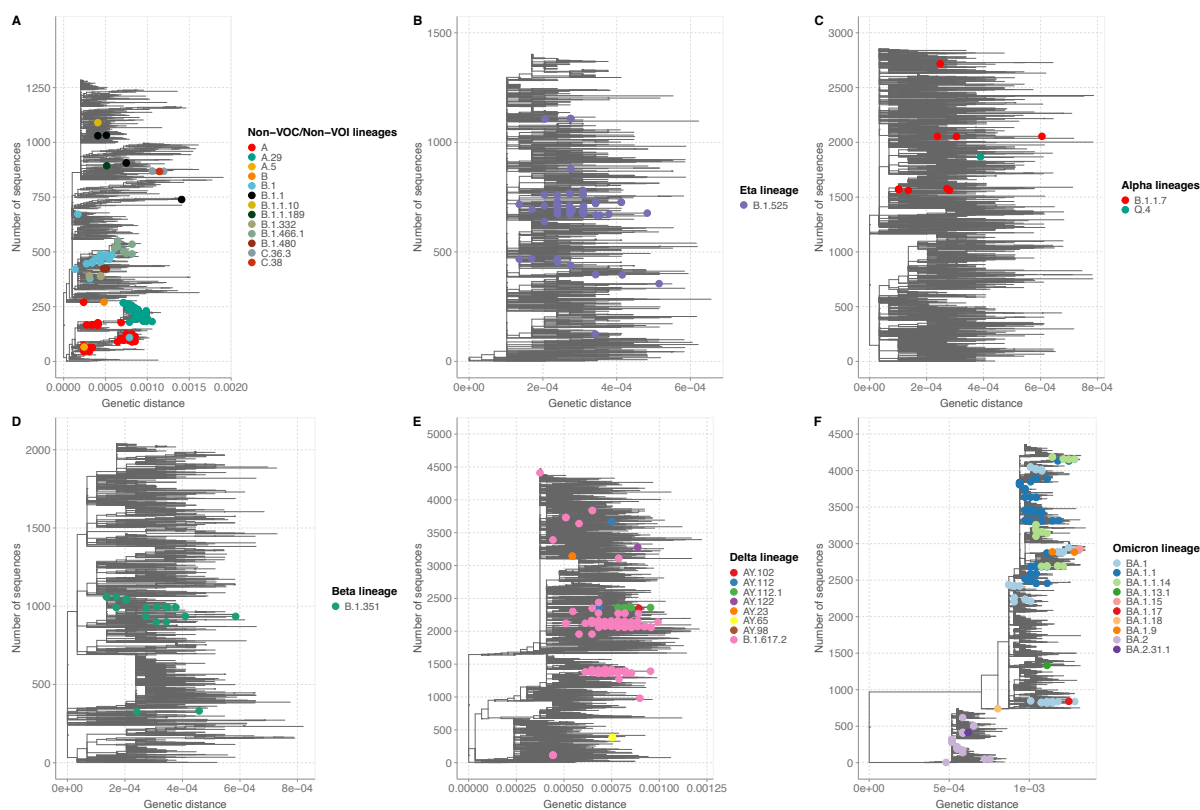
208

209 **Figure 2:** Temporal distribution of detected SARS-CoV-2 lineages in Sudan between May  
 210 2020 to April 2022. The colours show the different VOCs/VOI, and the circle sizes indicate  
 211 the frequency.

212

### 213 Genetic diversity of SARS-CoV-2 in Sudan

214 Multiple clusters and singletons from the Sudan genomes were observed with each of the  
 215 detected variants. The Sudan genomes were interspersed within the global datasets among all  
 216 the variants (**Figure 3**).



217

218 **Figure 3:** Maximum likelihood phylogenetic trees of SARS-CoV-2 sequences from different  
 219 lineages (A) non-VOC/non-VOI (B) Eta VOI (C) Alpha VOC (D) Beta VOC (E) Delta VOC,  
 220 and (F) Omicron VOC from Sudan (coloured dots) and sub-sampled global sequences.

221

### 222 Importations of SARS-CoV-2 variants in Sudan

223 Virus importation events into Sudan were estimated from the time-resolved phylogenetic trees.  
 224 A total of 144 virus introduction events were estimated for all the variants from different parts  
 225 of the globe. The introduction events for each variant were as follows: non-VOC/non-VOI  
 226 (n=13), Alpha (n=7), Beta (n=11), Delta (n=24), Eta (n=13) and Omicron (n=76)  
 227 (**Supplementary figure 1**).

## 228 **DISCUSSION**

229 This study describes the genomic epidemiology of SARS-CoV-2 in Sudan between September  
230 2020 and January 2022. Over this period, 37 lineages which were either non-VOCs/non-VOIs  
231 Alpha VOC, Eta VOI, Beta VOC, Delta VOC or Omicron VOC were in circulation in Sudan.  
232 These SARS-CoV-2 variants sustained the cases throughout 2021 until the third major wave  
233 driven by Omicron VOC in January 2022. The third wave coincided with both the relaxation  
234 of the stringency measures in Sudan and the introduction of Omicron, with its increased  
235 transmissibility and infectivity compared to other VOCs such as Delta and Beta [16], [17].

236  
237 Lineage A.29 was first detected in Gambia in August 2020 and then spread with minimal  
238 infectivity and severity globally [18]. The lineage was first detected in Sudan in September  
239 2020 and was sustained until February 2021 based on additional sequences deposited on  
240 GISAID [19]. Alpha VOC, Eta VOI, Beta VOC, and Delta VOC were detected and were in  
241 circulation between February 2021 and July 2021. However, no major waves were observed in  
242 Sudan coincident with the introduction of these variants, which contrasts the association of  
243 these variants with surges of cases seen in other countries within Africa [6]. The common  
244 Omicron lineages detected in cocirculation in Sudan were BA.1.1 (30.7%), BA.1 (27.3%) and  
245 BA.2 (23.9%) and the frequency of these variants increased rapidly across the weeks in January  
246 2022. Similar findings in transition of lineages and these omicron subvariants causing surges  
247 in cases have been reported elsewhere [20]–[23].

248  
249 Our phylogenetic analysis showed the sequences from Sudan were interspersed within the  
250 global tree, suggesting the frequent global spread of the detected variants due to the movement  
251 of people across different countries. Multiple introductions of the Eta VOI, Beta VOC and  
252 Omicron VOC were observed into Sudan, mainly from Europe and Africa. Import/export  
253 events of the virus within Africa and across Europe have been previously described [6], [8].  
254 Interestingly, only one Delta lineage B.1.617.2 was identified in Sudan and the phylogeny  
255 suggests a single introduction event that led further spread of Delta in Sudan through local  
256 transmission. However, this conclusion may be influenced by the limited sampling, especially  
257 lack of sampling between August and December 2021 following the replacement of the Delta.  
258 Our study had some limitations. First, the samples used in this analysis are only from two  
259 major cities: Khartoum; and Port Sudan, and we do not have detailed epidemiological data.  
260 Additionally, there is a lack of continuous sampling over time, making it difficult to draw  
261 solid conclusions on the complete diversity of SARS-CoV-2 in Sudan during the pandemic.

262

263 In conclusion, this study reveals the genetic diversity of SARS-CoV-2 variants in circulation  
264 in after multiple introductions of these variants in the country. These findings show that  
265 continuous genomic surveillance is key in monitoring and understanding the introduction and  
266 spread of SARS-CoV-2 variants within countries as it helps in upholding measures public  
267 health measures that can mitigate the spread of the virus.

268

269

270 **ACKNOWLEDGMENTS**

271 We thank the Sudan Ministry of Health, Public Health Laboratory, Africa-CDC for facilitating  
272 sharing of sequenced SARS-CoV-2 samples. We would like to thank all the laboratories that  
273 have shared SARS-CoV-2 sequence data in GISAID that we included as comparison data in  
274 our analysis.

275

276 **Funding Information**

277 This work was supported by the National Institute for Health and Care Research (NIHR)  
278 (project references 17/63/82 and 16/136/33) using UK aid from the UK Government to support  
279 global health research, The UK Foreign, Commonwealth and Development Office and  
280 Wellcome Trust (grant# 220985/Z/20/Z) The views expressed in this publication are those of  
281 the author (s) and not necessarily those of NIHR or the Department of Health and Social Care,  
282 Foreign Commonwealth and Development Office. This work was also supported by the Sudan  
283 Public Health Laboratory (Public Health Authority), Africa-CDC, WHO-Afro, ASLM, and  
284 WHO-Kenya offices.

285

286 **Data Availability Statement:**

287 The Sudan daily case data for the study period was obtained and is freely available from the  
288 Our World in Data database. All genome sequences and associated metadata in this dataset  
289 are published in GISAID's EpiCoV database <https://doi.org/10.55876/gis8.231110pw>.

290 The epidemiological data is available on the Virus Epidemiology and Control (VEC)  
291 Dataverse (DOI: <https://doi.org/10.7910/DVN/782VRG>). Data are available under the terms  
292 of the [Creative Commons Attribution 4.0 International license](#) (CC-BY 4.0).

293

294 **Conflicts of Interest:** The authors have declared no competing interest.

## 295 REFERENCES

- 296 [2] W. Moser *et al.*, “SARS-CoV-2 Antibody Prevalence and Population-Based Death  
297 Rates, Greater Omdurman, Sudan,” *Emerg. Infect. Dis.*, vol. 28, no. 5, pp. 1026–1030,  
298 2022, doi: 10.3201/eid2805.211951.
- 299 [3] M. M. Mukhtar and M. Khogali, “The accelerating COVID-19 epidemic in Sudan,”  
300 *Nat. Immunol.*, vol. 22, no. 7, pp. 2021, doi: 10.1038/s41590-021-00950-0.
- 301 [5] X. Deng *et al.*, “Genomic surveillance reveals multiple introductions of SARS-CoV-2  
302 into Northern California,” *Science (80-. )*, vol. 369, no. 6503, pp. 582–587, 2020, doi:  
303 10.1126/science.abb9263.
- 304 [6] H. Tegally *et al.*, “The evolving SARS-CoV-2 epidemic in Africa: Insights from  
305 rapidly expanding genomic surveillance
- 306 [7] Z. Chen *et al.*, “Global landscape of SARS-CoV-2 genomic surveillance and data  
307 sharing,” *Nat. Genet.*, vol. 54, no. 4, pp. 499–507, 2022, doi: 10.1038/s41588-022-  
308 01033-y.
- 309 [8] E. Wilkinson *et al.*, “A year of genomic surveillance reveals how the SARS-CoV-2  
310 pandemic unfolded in Africa,” *Science (80-. )*, vol. 374, no. 6566, pp. 2021, doi:  
311 10.1126/science.abj4336.
- 312 [9] A. W. Lambisia *et al.*, “Optimization of the SARS-CoV-2 ARTIC Network V4  
313 Primers and Whole Genome Sequencing Protocol,” *Front. Med.*, vol. 9, no. February,  
314 pp. 2022, doi: 10.3389/fmed.2022.836728.
- 315 [10] G. Githinji *et al.*, “Tracking the introduction and spread of SARS-CoV-2 in coastal  
316 Kenya,” *Nat. Commun.*, vol. 12, no. 1, p. 4809, Dec. 2021, doi: 10.1038/s41467-021-  
317 25137-x.
- 318 [11] J. Quick, “nCoV-2019 sequencing protocol v3 (LoCost) V.3,” 2020.
- 319 [12] J. R. Tyson *et al.*, “Improvements to the ARTIC multiplex PCR method for SARS-  
320 CoV-2 genome sequencing using nanopore.,” *bioRxiv Prepr. Serv. Biol.*, 2020, doi:  
321 10.1101/2020.09.04.283077.
- 322 [13] A. Rambaut *et al.*, “A dynamic nomenclature proposal for SARS-CoV-2 lineages to  
323 assist genomic epidemiology,” *Nat. Microbiol.*, vol. 5, no. 11, pp. 1403–1407, 2020,  
324 doi: 10.1038/s41564-020-0770-5.
- 325 [14] J. Hadfield *et al.*, “NextStrain: Real-time tracking of pathogen evolution,”  
326 *Bioinformatics*, vol. 34, no. 23, pp. 4121–4123, 2018, doi:  
327 10.1093/bioinformatics/bty407.
- 328 [15] R. C. R Core Team, “R: A Language and Environment for Statistical Computing,”

- 329 *Vienna, Austria*, 2018.
- 330 [16] D. Ao *et al.*, “SARS-CoV-2 Omicron variant: Immune escape and vaccine  
331 development,” *MedComm*, vol. 3, no. 1. pp. doi: 10.1002/mco2.126.
- 332 [17] J. Chen, R. Wang, N. B. Gilby, and G. W. Wei, “Omicron Variant (B.1.1.529):  
333 Infectivity, Vaccine Breakthrough, and Antibody Resistance,” *J. Chem. Inf. Model.*,  
334 vol. 62, no. 2, pp. 412–422, 2022, doi: 10.1021/acs.jcim.1c01451.
- 335 [18] and A. R. Áine O’Toole, Emily Scher, “Pango Lineages: Latest epidemiological  
336 lineages of SARS-CoV-2,” 2022.
- 337 [19] Y. Shu and J. McCauley, “GISAID: Global initiative on sharing all influenza data –  
338 from vision to reality,” *Eurosurveillance*, vol. 22, no. 13, pp. 2017, doi:  
339 10.2807/1560-7917.ES.2017.22.13.30494.
- 340 [20] P. Elliott *et al.*, “Twin peaks: The Omicron SARS-CoV-2 BA.1 and BA.2 epidemics in  
341 England,” *Science (80-. )*, vol. 376, no. 6600, 2022, doi: 10.1126/science.abq4411.
- 342 [21] O. Eales *et al.*, “The new normal? Dynamics and scale of the SARS-CoV-2 variant  
343 Omicron epidemic in England,” *medRxiv*, p. 2022.03.29.22273042, 2022, doi:  
344 10.1038/s41467-022-32096-4.
- 345 [22] NGS-SA, “Network for Genomic Surveillance in South Africa (NGS-SA) SARS-CoV-  
346 2 sequencing update.”
- 347
- 348