

1 Molecular Epidemiology and Diversity of SARS-CoV-2 in Ethiopia, 2020- 2 2022

3

4

5 Abay Sisay^{1, 2, ¶, *}, Derek Tshiabuila^{3, 4, ¶}, Abraham Tesfaye^{1, 5}, Gerald Mboowa⁶, Samuel O. Oyola⁷,
6 Sofonias Kifle Tesema⁶, Cheryl Baxter^{3, 4}, Darren Martin^{3, 4}, Richard Lessells^{3, 4}, Houriiyah Tegally^{3, 4},
7 Stephanie van Wyk^{3, 4}, Monika Moir^{3, 4}, Jennifer Giandhari^{3, 4}, Sureshnee Pillay^{3, 4}, Lavanya Singh^{3, 4},
8 Yajna Ramphal^{3, 4}, Arisha Maharaj^{3, 4}, Yusasha Pillay^{3, 4}, Akhil Maharaj^{3, 4}, Yeshnee Naidoo^{3, 4},
9 Upasana Ramphal^{3, 4}, Lucious Chabuka^{3, 4}, Eduan Wilkinson^{3, 4}, Tulio de Oliveira^{3, 4, 8}, Adey Feleke
10 Desta² and James E. San^{3, 4, *}

11

12 ¹Department of Medical Laboratory Sciences, College of Health Sciences, Addis Ababa University,
13 Addis Ababa, Ethiopia

14 ²Department of Microbial, Cellular, and Molecular Biology, College of Natural and Computational
15 Sciences, Addis Ababa University, Addis Ababa, Ethiopia.

16 ³KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), Nelson R Mandela School
17 of Medicine, University of KwaZulu-Natal, Durban, South Africa.

18 ⁴Centre for Epidemic Response and Innovation (CERI), School of Data Science and Computational
19 Thinking, Stellenbosch University, Stellenbosch, South Africa.

20 ⁵Diagnostic Unit, Center for Innovative Drug Development and Therapeutic Trials for Africa, CDT-
21 Africa, Addis Ababa, Ethiopia

22 ⁶Institute of Pathogen Genomics, Africa Center for Disease Control and Prevention (Africa CDC),
23 Addis Ababa, Ethiopia.

24 ⁷International Livestock Research Institute (ILRI), Nairobi, Kenya.

25 ⁸Department of Global Health, University of Washington; Seattle, USA.

26

27

28 *Corresponding author(s): Abay Sisay (abay.sisay@aau.edu.et), James E. San
29 (sanemanueljames@gmail.com)

30 ¶ These authors contributed equally to this work

31 **Abstract**

32 Ethiopia is the second most populous country in Africa and the sixth most affected by
33 COVID-19 on the continent. Despite having experienced five infection waves, >499 000
34 cases, and ~7 500 COVID-19-related deaths as of January 2023, there is still no
35 detailed genomic epidemiological report on the introduction and spread of SARS-CoV-2
36 in Ethiopia. In this study, we reconstructed and elucidated the COVID-19 epidemic
37 dynamics. Specifically, we investigated the introduction, local transmission, ongoing
38 evolution, and spread of SARS-CoV-2 during the first four infection waves using 353
39 high-quality near-whole genomes sampled in Ethiopia. Our results show that whereas
40 viral introductions seeded the first wave, subsequent waves were seeded by local
41 transmission. The B.1.480 lineage emerged in the first wave and notably remained in
42 circulation even after the emergence of the Alpha variant. The B.1.480 was out-
43 competed by the Delta variant. Notably, Ethiopia's lack of local sequencing capacity was
44 further limited by sporadic, uneven, and insufficient sampling that limited the
45 incorporation of genomic epidemiology in the epidemic public health response in
46 Ethiopia. These results highlight Ethiopia's role in SARS-CoV-2 dissemination and the
47 urgent need for balanced, near-real-time genomic sequencing.

48

49 **Keywords:** COVID-19, Molecular epidemiology, SARS-CoV-2, Whole Genome Sequence, Ethiopia

50 **Introduction**

51 Ethiopia is a landlocked country in Eastern Africa [1, 2]. It is further one of the most
52 populous countries on the continent, and following the global trend, has accordingly
53 endured severe economic and healthcare burdens due to the ongoing COVID-19
54 pandemic [3,4]. On 13 March 2020, the Federal Ministry of Health in Ethiopia reported
55 its first COVID-19 case in Addis Ababa. By February 2022, Ethiopia had recorded
56 approximately 494,760 confirmed cases and 7,572 deaths [5, 6]. Despite various
57 pandemic mitigation strategies implemented by the government (Table S1), the
58 number of COVID-19 cases continued to rise. Ethiopia currently constitutes one of the
59 most severely affected countries on the African continent. On the 8th of April 2020, the
60 Ethiopian government declared a 5-month national state of emergency to control and
61 contain the growth rate of the local epidemic and support mitigation mechanisms [7,8].
62 These emergency measures included the closing of schools and other academic
63 institutes, restrictions on large gatherings, and implementation of mask mandates.
64 These restrictions were complemented by encouraging social distancing and hand
65 washing using multimedia campaigns [7, 9, 10]. While genetic diversity is a major
66 driver of epidemic dynamics, until now, no detailed molecular epidemiology and
67 genetic diversity analyses of SARS-CoV-2 have been performed in Ethiopia. To
68 address this problem, we analyzed 353 high-quality near complete SARS-CoV-2
69 genomes sampled in Ethiopia against a globally representative set to characterize the
70 introduction, transmission, and spread of SARS-CoV-2 in Ethiopia.

71 **Materials and Methods**

72 **Ethical Consideration**

73 This study was reviewed and got approved by Addis Ababa University, College of
74 Health Science Institutional Review Board (IRB) (IRB # 029/20/Lab), IRB of the
75 Department of Medical Laboratory Sciences, Addis Ababa University (reference #-
76 MLS/174/21), IRB office of Addis Ababa City Administration Health Bureau,
77 AAPHREML (Reference #- AAHB/4039/227) and also from Addis Ababa University,
78 College of Natural and Computational Science IRB (IRB #- CNCSDO/604/13/2021),
79 by Yekatit 12 Hospital Medical College (IRB protocol # 75/20) and also reviewed and

80 approved by the Federal Democratic Republic of Ethiopia National Ethics Committee
81 at the Ministry of Science and Higher Education (MoSHE) (IRB #021/246/820/21).

82 **Study Design, Period, and Setting**

83 We leveraged samples collected from confirmed Ethiopian SARS-CoV-2 cases
84 recorded between July 2020 and February 2022 to produce whole and near-whole
85 genomes. Specifically, whole genome sequencing (WGS) was carried out on a subset
86 of RT-PCR positive samples (N=353) collected during the first four waves of the
87 pandemic in Ethiopia. The genomes can be further classified according to the
88 sampling strategy, i.e., community surveillance (n=170), contacts of confirmed cases
89 (n=78), and suspected cases (symptomatic but yet to be confirmed by laboratory
90 results) (n=105), (Figure S2). Nasopharyngeal (NP) swab specimens were collected
91 using 2ml of Viral Transport Medium (VTM), (China, Miraclean Technology Co., Ltd.,
92 www.mantacc.com) from patients in Addis Ababa, Amhara, Oromia, and the Southern
93 Nations, Nationalities, and Peoples Regional States (provinces) of the country. The
94 number of samples selected from each region was based on the epidemic distribution
95 of the diseases across the country (Figure S3). These areas are known to have high
96 population densities and cover almost three-quarters of the country's inhabitants [11,
97 12]. Only samples with cycle threshold (Ct) values under 30 were considered viable
98 and submitted for sequencing

99 **Quality Assurance**

100 We implemented a comprehensive total quality management system to ensure the
101 quality of our sequence data throughout the process (from pre-analytical to post-
102 analytical). The specimens with RT-PCR Ct values <30 were stored in a -80°C freezer
103 at Arsho Advance Medical Laboratories Plc: an ISO 15189 accredited laboratory [13].

104 **RNA Extraction and Viral Detection**

105 Extraction, purification, and viral detection using RT-PCR were done as per our
106 previous work and respective manufacturer recommendations [14-17]. Viral detection
107 was done at the Ethiopian Public Health Institute (EPHI), AAPHREM center
108 laboratory, Yekatit 12 Hospital, Arsho medical laboratory, and Addis Ababa University.

109 These were amongst the established RT-PCR laboratories where the majority of
110 COVID-19 testing was performed. SARS-CoV-2 positive samples from the first three
111 waves of the pandemic were shipped to CERI/KRISP (South Africa), and from the
112 fourth wave (103) to ILRI (Kenya) for whole genome sequencing. All biological
113 specimens were transported following a triple packaging biosafety and biosecurity
114 protocol. Samples at CERI/ KRISP were sequenced between 22 September 2021 and
115 18 October 2022, while samples sent to ILRI were sequenced on March 2022.

116 **Next-Generation Sequencing of SARS-CoV-2**

117 WGS was performed using either the Illumina or the Oxford Nanopore Technology
118 (ONT) sequencing platforms. Briefly, RNA was extracted on the automated Chemagic
119 360 system (Perkin Elmer) as per the manufacturer's instructions. Samples were lysed
120 using proteinase K and bound using silica magnetic beads. Beads were then washed
121 and the RNA was eluted and stored at -80 °C until use. SuperScript IV reverse
122 transcriptase (Life Technologies) was then used, in combination with random hexamer
123 primers, to convert extracted RNA to complementary DNA. Gene-specific multiplex
124 PCR using the ARCTIC protocol was then performed. PCR was used to generate 400
125 base pair (bp) amplicons, with an overlap of 70 bp to cover the 30 kilobases (kb)
126 SARS-CoV-2 genome. PCR products were purified using AmpureXP purification
127 beads (Beckman Coulter, High Wycombe, UK), and quantification was performed
128 using the Qubit double-strand DNA (dsDNA) High Sensitivity Assay Kit on a Qubit 4.0
129 instrument (Life Technologies). These libraries were then used for Illumina MiSeq and
130 ONT GridION sequencing.

131 Illumina MiSeq indexed paired-end libraries were prepared using the Nextera DNA
132 Flex Library Prep Kits (Illumina) as per the manufacturer's instructions. Tagmented
133 amplicons were barcoded with a unique barcode using the Nextera CD Indexes
134 (Illumina), which allows for library pooling. Prior to pooling, libraries were normalized
135 to 4 nM and pooled, then denatured using sodium acetate. Pooled libraries were then
136 diluted and spiked with 1 % PhiX Control v3 and sequenced using a 500-cycle v2
137 MiSeq Reagent Kit on the Illumina MiSeq instrument (Illumina, San Diego, CA, USA)
138 [18]. ONT Native Barcoding Expansion kits were used for ONT GridION sequencing

139 according to the manufacturer's protocol. Samples were multiplexed on the FLO-
140 MIN106 flowcells and ran on the GridION X5. The MinKNOW software was used to
141 monitor sequencing performance in real-time [19].

142

143 **Metadata Management**

144 Clinical and demographic metadata were collected from the confirmed cases using a
145 structured assessment tool by trained healthcare professionals at the triage of the
146 health facilities and from the community. A medical record number (MRN) was
147 assigned to each patient as a unique anonymous identifier as their names were not
148 linked to the MRNs submitted to sequence repositories such as GISAID,
149 EPI_SET_221214bg doi: [10.55876/gis8.221214bg](https://doi.org/10.55876/gis8.221214bg).

150 **Sequence Alignment, Assembly, and Phylogenetic Analysis**

151 Fastq sequences were assembled using Genome Detective 1.133 ([http://www.](http://www.genomedetective.com)
152 [genomedetective.com](http://www.genomedetective.com)) [20] and the Coronavirus tool [21]. The initial assemblies
153 obtained from Genome Detective were visualized and manually edited using Geneious
154 Prime version 2022.2.2 [22,23] to remove low-quality mutations. Nextclade was used
155 to determine the quality control profiles of the sequences. The quality metrics used for
156 the sequences in this study included genome coverage greater than 80%, ensured the
157 removal of unknown frameshifts, stop codon, and clustered mutations. Additionally,
158 submissions with exceedingly high and missing data (N) <3000Ns) were excluded
159 from downstream analyses. Nextclade was further employed to assign clades and
160 lineage classifications to the sequences [24,25]. Sequences that met these quality
161 control thresholds were deposited in the GISAID (<https://www.gisaid.org/>) database,
162 EPI_SET_221214bg doi: [10.55876/gis8.221214bg](https://doi.org/10.55876/gis8.221214bg)].

163 To present a comprehensive analysis of the genomic epidemiology of SARS-CoV-2 in
164 Ethiopia, the genomes generated in this study were combined with a globally
165 representative set of genomes (n= 13589) that were sampled during the same period.
166 All samples from the bordering countries to Ethiopia were included in the analysis to
167 better model the relationship between Ethiopia and its Neighbors and infer cross-

168 border transmission. This did not cause a sampling bias as these countries were
169 already characterized by a low number of sequences. Additionally, all global
170 sequences belonging to the B.1.480 lineage which was predominant in Ethiopia and
171 remained in circulation across the first three waves were included. Together, this
172 resulted in a dataset of 13,942 genomes. Acknowledgment was given to the
173 sequencing laboratories that contributed to the analysis set (EPI_SET_221214bg,
174 DOI: [10.55876/gis8.221214bg](https://doi.org/10.55876/gis8.221214bg)).

175 Sequences were aligned using NextAlign [26] to obtain a good codon alignment of the
176 sequences. A maximum likelihood tree topology was inferred from the resulting
177 alignment in IQTREE 2 [27] using the General Time Reversible model of nucleotide
178 substitution [28] The inferred phylogeny was used along with the associated metadata
179 that was obtained from GISAID, to map discrete geographical locations to each of the
180 tips and infer locations for the internal nodes. This was done using the Muration
181 package extension of TreeTime [29]. A custom Python script was then used to count
182 the number of discrete changes occurring as we transcended the topology from the
183 root towards the tips. In essence, this provided a crude estimate of the number and
184 timing of viral exchanges (import-export) between Ethiopia and the rest of the world.

185 **Data Analysis and Visualization**

186 The data were analyzed and visualized using R version 4.2.0 using available
187 packages and custom scripts. The tree was visualized using the R package ggtree
188 version 3.0.4[30].

189

190

191

192

193 **Results**

194

195 **Socio-demographic Characteristics of the Study Participants along with SARS-** 196 **CoV-2 Variants of Concerns (VOCs)**

197 A total of 1300 nasopharyngeal swabs were collected from four provinces in the
198 country. These samples included 181 (51.3%) specimens obtained from female
199 patients while 172 (48.7%) were sampled from male patients. The patient ages ranged
200 between 2 and 86 years old, although most of the cases were between 21- and 30-
201 year-old patients, with an overall estimated median age of 31.45 (IQR=26-46) years.
202 More than 77 % of the patients presented with mild cases ([Table 1](#)). Of these, we
203 discarded 300 due to their high CT values (i.e., > 30). The remaining 1000 samples
204 were subjected to the library preparation and 808 yielded sufficient RNA for WGS
205 following RT-PCR amplification. The WGS resulted in 353 SARS-CoV-2 sequences
206 that passed bioinformatics quality controls. (i.e., with genome coverage greater than
207 80%, a limited number of private mutations, no clustered mutations, and no misplaced
208 stop codons). [Figure S2](#) illustrates a detailed breakdown of the sample processing.
209 During the study period, a combination of genetically distinct SARS-CoV-2 lineages
210 were identified. The majority of these assemblies (91.7%) were classified as variants
211 of concern (VOC).

212 **Local Epidemic Dynamics**

213 To illuminate the local epidemic dynamics in Ethiopia, we performed an in-depth study
214 of the 353 high-quality SARS-CoV-2 genomes ([Figure 1](#)). The study period
215 corresponds with the first four COVID-19 waves experienced in the country. The first
216 wave occurred between August and September 2020, the second between February
217 and March 2021, and the third occurred between July and October 2021. Each wave
218 was characterized by higher rates of transmission and fatality than previous waves
219 ([Figure 1A](#)). Notably, the third wave was more severe and lasted longer than the two
220 previous waves. The fourth wave began with an uptick in cases in December 2021
221 and continued to February 2022. The wave was characterized by the highest rates of

222 community transmission yet observed (Figure 1A). However, compared to the global
223 trend at the time, it resulted in fewer deaths.

224 Across the four waves, the peak number of new cases remained relatively consistent
225 with the first and third waves reflecting a 7-day rolling average of approximately 1500
226 individuals. The second wave had a slightly higher 7-day rolling average of 2000
227 individuals whilst the fourth had the highest at over 4000 individuals (Figure 1A). The
228 number of new deaths across each wave had a similar trend to the average number of
229 new cases with the second and third waves having the highest number of new deaths
230 at approximately 50 individuals per day (Figure 1B). Vaccine rollout commenced in
231 April 2021 in Ethiopia and by February 2022, approximately 25% of the population had
232 received at least one dose of Oxford-AstraZeneca or Sinopharm COVID-19 vaccines
233 (Figure 1B).

234 **Phylogenetic reconstruction and Variant Detection in Ethiopia**

235 Ancestral state reconstruction showed that viral imports into Ethiopia occurred
236 between August 2020 and January 2021, with the majority of importation events
237 occurring in August 2020 [n=123], December 2020 [n=130], and January 2021 [n=169]
238 (Figure 1C). In comparison, September and November 2020 corresponded with the
239 highest numbers of inferred SARS-CoV-2 exports from Ethiopia (n=109 and n =137,
240 respectively; Figure 1D, Figure S4). The majority of inferred importation events were
241 from Eastern Africa [n=324] origins, whereas most of the inferred exportation events
242 were generally attributed to Western European and Asian origins (n=20 and n=141
243 respectively). Overall, we identified 570 importation and 446 exportation events during
244 the study period.

245 Our early sequences sampled between November and December 2020 represented
246 two lineages, B.1/B.1.1 and B.1.480, with most of the sequences classified as
247 B.1/B.1.1. These two lineages continued to circulate throughout the first quarter of
248 2021 (January to March). The B.1.480 lineage was thought to have originated in
249 Ethiopia. In our study, 21 genome assemblies were assigned to this emerging lineage.
250 (EPI_SET_230126dy, doi: 10.55876/gis8.230126dy). Phylogenetic inference

251 supported that this lineage was a monophyletic clade. Despite its persistence in
252 Ethiopian populations for an extended period, this variant did not disseminate globally
253 as was observed for other VOC and ultimately only constituted only 473 GISAID
254 submissions during the progression of the pandemic, EPI_SET_230126vf, doi:
255 10.55876/gis8.230126vf. Nevertheless, according to GISAID submissions, this lineage
256 was detected in neighboring countries such as Kenya, Somalia, Djibouti, and Guinea-
257 Bissau. This lineage was further disseminated internationally where it was detected in
258 Australia, the United States of America, Canada, Japan, Israel, Hong Kong, and South
259 Korea. This period was followed by the detection and rise to dominance of VOCs in
260 the country.

261 In January 2021, the first case of the Alpha variant was detected. Within a span of
262 one month, it had reached equal prevalence to B.1/B.1.1. The B.1.1 continued to
263 circulate in low prevalence with the cases detected until May 2021. Alpha was
264 however short-lived as the Delta variant emerged. The Delta variant was first detected
265 in February 2021 and continued circulating in low prevalence until May 2021.
266 Thereafter it outcompeted all other lineages and variants in circulation to become the
267 most prevalent variant (Figure S5). The predominant Delta variant sub lineages were
268 AY.120 (116 sequences, 46.4%) and B.1.617.2 (50 sequences, 20%). The Delta
269 variant seeded the 3rd wave in Ethiopia. Although no sequence data was available
270 from September to December 2021, the majority of the cases sampled and sequenced
271 from January 2022 were attributed to the Omicron lineage (Fig 1A & 2A). Over the
272 course of the study period, we identified 37 distinct lineages in circulation (Fig 2 A, Fig
273 2B, Figure S6).

274 The phylogenetic reconstruction revealed close relatedness (not substantially
275 observed levels of divergence) among sequences assigned to the same lineages.
276 However, this observation did not hold for samples assigned to the Delta variant
277 (Figure 2C). Sequences attributed to the Delta variant showed greater variation in
278 sequence composition. Moreover, higher levels of phylogenetic clusters, associated
279 with community transmission, were observed during the study period. Phylogenetic

280 clusters and associated community transmission events were also observed for the
281 Omicron lineage (Figure 2C).

282

283 **Discussion**

284 In this study, we investigated the introduction of SARS-CoV-2 viral lineages, local
285 transmission, and its evolution within Ethiopia. This study represents the first of its
286 kind as no other molecular epidemiology investigation has been conducted for this
287 country and provided an overview of the pandemic progression relative to a global
288 context. Our results show that SARS-CoV-2 was introduced into Ethiopia on multiple
289 occasions, with most introductions originating from Eastern African countries. These
290 introductions included 37 distinct SARS-CoV-2 lineages and four VOCs, namely
291 Alpha, Beta, Delta, and Omicron lineages. Of the introduced lineages the B.1.480
292 persisted and remained in circulation from December 2020 to August of 2021.

293 Our findings show that as of February 2022, Ethiopia had experienced four distinct
294 COVID-19 waves. Although various preventive strategies such as the mandatory
295 wearing of masks, and media awareness campaigns were adopted together with a
296 mandatory 14-day self-quarantine regulation for all international travelers arriving in
297 the country [5], Ethiopia remained vulnerable to SARS-CoV-2 introductions. Our
298 results showed that despite preventative measures viral introduction events continued
299 to occur and were followed by community transmission. This in turn facilitated a
300 continued increase in case numbers. This could be attributed to the partial restriction
301 of public transport which is an import mode of SARS-CoV-2 transmission [5, 31].
302 During lockdowns, the operation of economic activities did not cease as they were
303 deemed necessary to support the already fragile economic and socioeconomic
304 situation of the country. This made lockdowns ineffective in preventing local
305 transmission and epidemic growth. By 30 April 2021, the total number of cases and
306 deaths had surpassed 257442 and 3688, respectively [1, 3], and has currently
307 increased to >499000 and 7500 deaths.

308 Our findings show that Ethiopia experienced poor sampling mostly limited to Addis
309 Ababa. Temporal sampling was also hampered by periods of limited sampling and,
310 lack of local sequencing capacity resulting in substantial delays that in turn may cause
311 the degradation of samples during transit. Yet, an in-depth understanding of the
312 COVID-19 pandemic progression in Ethiopia remains of paramount importance to aid
313 the navigation of ongoing and future pandemics within these regions [32]. Indeed,
314 Ethiopia has had several viral outbreaks in recent history, in addition to the ongoing
315 COVID-19 pandemic. Some of the notable outbreaks include Measles, and mosquito-
316 borne illnesses such as Yellow fever, Chikungunya, and Malaria [32, 33]. Thus,
317 scaling up of local and regional pathogen genomic sequencing capacity and a robust
318 pathogen genomic surveillance infrastructure is thus required to shorten the
319 turnaround time and to allow for early public health response [34]. Additionally, during
320 the onset of the pandemic, a survey on outbreak readiness highlighted a dire need for
321 not only pathogen surveillance in Ethiopia but Ethiopia further scored only 52% on the
322 Ready Score index, which shows that the nation still has progress to be made to
323 increase case tracing and identification, increase healthcare and testing facilities,
324 provide clear operational guidelines on preventive measures across various
325 organizations, businesses, and community settings, as well as proactive steps to
326 maintain life during social and economic lockdowns[3, 35]. Ethiopia, as with many
327 other African countries, has limited resources to deal with these outbreaks, and often
328 depends on international aid to help control and prevent these diseases [10, 36, 37].

329 Over the course of four waves, we estimated 570 introductions and 446 exportation
330 events occurred. This is consistent with the relatively small number of lineages
331 identified during the study period, although could also be attributed to under and
332 sparse sampling that likely omitted lineages circulating at low prevalence. The short
333 generation time associated with RNA viruses could also imply that viruses circulating
334 during the period preceding the emergence of Omicron that became extinct would not
335 have been detected. The phylogenetic inference and that of the geographical
336 distribution of the Ethiopian-emerging B.1.480 lineage suggest that Ethiopia serves as
337 a source for viral introductions not only to neighboring countries but also further
338 abroad. Ethiopia serves as a travel hub for this African region connecting this region in

339 Africa to neighboring countries and further abroad. Possible disease dissemination
340 was most likely a result of travel to and from Ethiopia via air, land, and waterways [37,
341 38]. Ethiopia is strategically located on the Red Sea, making it an important transit
342 point for goods moving between Africa, the Middle East, and Asia. The main
343 international airport in Ethiopia is Addis Ababa Bole International Airport facilitates
344 high travel volumes which most likely contributed to viral amplification in the country
345 followed by Ethiopian-originating viral exportation events. Therefore, concerted,
346 collaborative action is imperative to achieve pandemic mitigation.

347 We observed a reduction in the number of new deaths reported during the fourth
348 wave. This is consistent with the immunity acquired protection from a large number of
349 infections from the preceding waves and increases in the total number of vaccinated
350 individuals. A sero-prevalence study among healthcare workers revealed a
351 seropositivity rate of 39.6% [39] while another study estimated the prevalence rate at
352 over 50% [40]. However, only about 0.4% of the population has reportedly been
353 infected with COVID-19: much fewer than the estimated sero-prevalence rates
354 estimated for the population. Therefore, taken together these statistics suggest that
355 case numbers may remain widely underreported. Reduced testing further hinders
356 effective genome surveillance as fewer samples were available to be subjected to
357 whole genome analyses and influences the accuracy of epidemiological conclusions
358 drawn.

359 In comparison to other countries, Ethiopia has very low vaccination coverage. As of
360 December 2021, at least 1 000 000 Ethiopians had been vaccinated. It is important to
361 note that COVID-19 vaccines in Ethiopia were prioritized based on criteria such as
362 occupation, age, medical condition, people that work in aggregate settings, the elderly,
363 and people with comorbidities known to increase the risk of adverse COVID-19
364 outcomes [41]. Moreover, priority was given to those deemed as frontline medical
365 staff, followed by non-frontline medical staff. This strategy was effective in slowing the
366 evolutionary rate of the virus as SARS-CoV-2 is prone to evolution in
367 immunosuppressed individuals experiencing chronic infection [42-44]. Vaccination has
368 also been shown to decrease the overall SARS-CoV-2 infection rate from 9.0% (95%

369 CrI: 8.4% – 9.4%) without vaccination to 4.6% (95% CrI: 4.3% – 5.0%) with the
370 highest relative reduction noted among individuals aged 65 and above. Adverse
371 outcomes, such as deaths, are reduced by approximately 69.3% (95% CrI: 65.5%-
372 73.1%) with lowered vaccination coverage [45].

373 Analysis of genomes and lineages across the study period highlighted the changes in
374 the population dynamics of the viral lineages circulating in Ethiopia. The B.1 and B.1.1
375 lineages dominated the first wave, persisted throughout the next two waves, and were
376 characterized by the well-known D614G mutation [46]. This mutation has been linked
377 to increased viral replication and transmissibility [38, 46]. Subsequent waves were
378 dominated by the VOCs Alpha, Beta, Delta, and Omicron. These VOCs evolved to
379 evade overcome host immunity from infection with ancestral viruses and vaccination.
380 By doing so these newly emerging variants fueled renewed infectious waves and were
381 often associated with increased transmissibility compared to their genetic
382 predecessors [47-49]. The Omicron variant (specifically, lineage BA.1.1) had a
383 reproduction number four times higher than Delta. It is considered the fifth variant of
384 concern first detected in November/December 2021 [50-53].

385 One important lineage in the context of Ethiopia was the B.1.480, a distinct B lineage,
386 which emerged during the second wave. This variant was responsible for 2.77% of the
387 sampled cases. Notably, this variant was previously detected in passengers
388 originating from not only Ethiopia but also with those originating from European
389 countries such as the United Kingdom and France, as well as from travelers from
390 United Arab Emirates [50-53]). Unlike other VOCs, this lineage did not reach global
391 transmission and has only 473 submissions to the GISAID platform, however, it did
392 persist in Ethiopian populations for extended periods of time. Other worrisome
393 mutations carried by this lineage included the Spike mutation M1229I and D614G
394 [46,47,49,54].

395 We observed early introductions of the Delta variant as early as February 2021. To
396 rule out contamination as possibly a source of these sequences, we examine
397 mutational patterns within these sequences, nucleotide distributions in the sorted,

398 indexed and mapped bam files, and the timing of sequencing. Considering the timing
399 of sequencing and that these samples were sequenced along with several samples
400 that were also attributed to the Delta variant, contamination as a possible source could
401 not be entirely ruled out. However, we note that these sequences conform to
402 molecular clock estimations of the emergence of Delta globally estimated around 19
403 October 2020 with an earliest possible date estimated as 06 September 2020 [55],
404 Figure S5. A similar pattern of early emergence and slow rise to dominance is
405 observed in several countries across the continent [34] and, in a more detailed study
406 in Neighboring Kenya [56].

407 In conclusion, the changing nature of the pandemic over time in Ethiopia has been
408 driven by lineage turnovers that, while dominated by VOCs and B.1.480, include other
409 lineages (A, B.1, B.1.1). Results from this study, and that of others (ref listed above),
410 showed that despite pandemic mitigation strategies implemented by the country,
411 Ethiopia remained vulnerable to viral introduction events and served as a source of
412 viral introductions internationally. The COVID-19 pandemic is still ongoing and the
413 evolution of the virus is continuing to yield new variants some of which will likely reflect
414 altered viral behavior such as increased transmissibility, immune evasiveness, or
415 altered pathogenicity. It is therefore crucial that ongoing genomic surveillance be
416 strengthened and conducted to present real-time data generation to determine the
417 geographic distribution of these variants and enable the implementation of appropriate
418 disease mitigation strategies.

419

420 **Acronyms /abbreviations**

AAPHREML	Addis Ababa public health research and emergency management laboratory
AAU	Addis Ababa University
ACDC	Africa center for disease prevention and control
CERI	Centre for Epidemic Response and Innovation
CI	Confidence Interval
CT	Cycle Threshold
GISAID	Global Initiative on Sharing All Influenza Data
KRISP	KwaZulu-Natal Research Innovation and Sequencing Platform
NAAT	Nucleic Acid Amplification Tests
PCR	Polymerase Chain Reaction
RNA	Ribose Nucleic Acid
RT-PCR	Reverse Transcriptase Polymerase Chain Reaction
SARS-CoV-2	severe acute respiratory syndrome novel corona virus-2
VOCs	Variants of Concern
VTM	Virus Transport Medium
WHO	World Health Organization
SDG	sustainable development goal
COVID-19	Coronavirus disease 2019

421

422 **Funding**

423 This research was financed by Addis Ababa University through the COVID-19-AAU
424 research fund for staff with a reference #- PR/5.15/590/12/20 and adaptive research
425 and problem-solving projects with Ref # AR/043/2021 and RD/PY-148/2021. Also, this
426 work was realized by the generous financial and technical support of the AU through
427 Africa CDC, PGI. No more external funding was received. The APC was funded by
428 Africa Pathogen Genomics Initiative (Africa PGI).

429 The funders had no role in the study design, data collection, analysis, decision to
430 publish, or preparation of the manuscript. The contents are purely the responsibilities
431 of the authors and did not represent and reflect the view of the founder.

432 **Author Contributions**

433 AS and AFD conceived and designed the experiments; AS, DT and SEJ performed
434 the data analysis and experiments; AS, DT, AFD and SEJ analyzed the data; GM, SO,

435 SKT, CB, DM, RL, HT, MM, SvW, EW, TDO contributed materials/analysis tools; AT,
436 AFD and SEM supervise the work; AS, DT, SvW and SEJ wrote the paper. All
437 authors read the final version of the manuscript.

438 **Informed Consent Statement**

439 All participants gave written informed consent for participation and publication.

440 **Declaration of Competing Interest**

441 We all the authors declare that we have no known competing interests or personal
442 relationships that could have appeared to influence the work reported in this research
443 paper.

444 **Data availability**

445 All the generated sequences used for this manuscript have been deposited in the
446 GISAID sequence database and are purely publicly available. However, it's subject to
447 the terms and conditions of the GISAID database, <https://www.gisaid.org/>, GISAID
448 Identifier: EPI_SET_221214bg, DOI: [10.55876/gis8.221214bg](https://doi.org/10.55876/gis8.221214bg).

449 **Acknowledgment**

450 We would like to acknowledge Addis Ababa University for giving us the opportunity
451 and financing the research. This work is realized with the generous financial and
452 technical support of Africa CDC, PGI

453 We are very grateful and acknowledge the CERI and KRISP (South Africa) & ILRI
454 (Kenya) for supporting this work by performing the whole genomic sequencing.

455 We also wish to extend our gratitude to Arsho Medical Laboratories for supporting this
456 research by giving us a -80 °C refrigerator for centrally storing our specimens.

457 Last but not least we are very glad and thankful to Retina Pharmaceutical for
458 supporting this work by availing required pharmaceutical inquiries.

459

460

461

462

463 **Reference**

- 464 1. Ethiopia: the World Bank.
465 <https://www.worldbank.org/en/country/ethiopia/overview>. Accessed on 25 Jan
466 2023.
- 467 2. Ebssa Wakwaya. "The promise of regional projects for Africa's landlocked
468 countries: Focusing on Ethiopia." *African Journal of Political Science and*
469 *International Relations* 9.2 (2015): 67-75
- 470 3. Zhang W, Persoz L, Hakiza S, Biru L, Girmatsion L. Impact of COVID-19 on
471 Food Security in Ethiopia. *Epidemiologia*. 2022; 3(2):161-178.
472 <https://doi.org/10.3390/epidemiologia3020013>
- 473 4. [https://www.un.org/ohrlls/sites/www.un.org.ohrlls/files/impact_of_covid19_and](https://www.un.org/ohrlls/sites/www.un.org.ohrlls/files/impact_of_covid19_and_responses_in_llcs.pdf)
474 [responses_in_llcs.pdf](https://www.un.org/ohrlls/sites/www.un.org.ohrlls/files/impact_of_covid19_and_responses_in_llcs.pdf)
- 475 5. University JH. Johns Hopkins University JH. worldometers COVID-19 Daily
476 update 2022 [Available from: <https://www.worldometers.info/coronavirus/>].
- 477 6. World Health Organization (WHO Africa). First case of COVID-19 confirmed in
478 Ethiopia. [https://www.afro.who.int/news/first-case-covid-19-confirmed-](https://www.afro.who.int/news/first-case-covid-19-confirmed-ethiopia#:~:text=March%2013%2C%202020%2C%20The%20Federal,in%20C)
479 [ethiopia#:~:text=March%2013%2C%202020%2C%20The%20Federal,in%20C](https://www.afro.who.int/news/first-case-covid-19-confirmed-ethiopia#:~:text=March%2013%2C%202020%2C%20The%20Federal,in%20C)
480 [hina%20in%20December%202019](https://www.afro.who.int/news/first-case-covid-19-confirmed-ethiopia#:~:text=March%2013%2C%202020%2C%20The%20Federal,in%20C).
- 481 7. Lanyero B, Edea ZA, Musa EO, Watore SH, Mandalia ML, Livinus MC, et al.
482 Readiness and early response to COVID-19: achievements, challenges, and
483 lessons learnt in Ethiopia. *BMJ global health*. 2021; 6(6).
- 484 8. Ethiopian Public Health Institute, EPHI. COVID-19 PANDEMIC
485 PREPAREDNESS AND RESPONSE IN ETHIOPIA WEEKLY BULLETIN Nov
486 2021 [updated 2022. Available from: <https://ephi.gov.et/download/pheoc/>].
- 487 9. WHO. COVID-19 Preparedness Bulletin Ethiopia 2020 [Available from:
488 <https://extranet.who.int/sph/covid-19-preparedness-bulletin-ethiopia>].
- 489 10. Amhare AF, Tao Y, Li R, Zhang L. Early and Subsequent Epidemic
490 Characteristics of COVID-19 and Their Impact on the Epidemic Size in

- 491 Ethiopia. *Front Public Health*. 2022;10: 834592. Published 2022 May 11.
492 doi:10.3389/fpubh.2022.834592
- 493 11. Leulseged TW, Abebe KG, Hassen IS, et al. COVID-19 disease severity and
494 associated factors among Ethiopian patients: A study of the millennium COVID-
495 19 care center. *PLoS One*. 2022;17(1):e0262896. Published 2022 Jan 27.
496 doi:10.1371/journal.pone.0262896
- 497 12. Abdella S, Riou S, Tessema M, et al. Prevalence of SARS-CoV-2 in urban and
498 rural Ethiopia: Randomized household serosurveys reveal level of spread
499 during the first wave of the pandemic. *EClinicalMedicine*. 2021; 35:100880.
500 Published 2021 May 7. doi:10.1016/j.eclinm.2021.100880
- 501 13. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high
502 throughput. *Nucleic acids research*. 2004;32(5):1792-7.
- 503 14. Sisay A, Abera A, Dufera B, Endrias T, Tasew G, Tesfaye A, et al. Diagnostic
504 accuracy of three commercially available one step RT-PCR assays for the
505 detection of SARS-CoV-2 in resource limited settings. *PloS one*. 2022;
506 17(1):e0262178.
- 507 15. MOLBIOL T. RT-PCR test kits and VirSNiP Mutation Assays for strain
508 surveillance [Available from: <https://www.tib-molbiol.de/covid-19>.
- 509 16. World Health Organization W. Emergency Use Listing Procedure for in vitro
510 diagnostics 2020 [Available from: [https://www.who.int/teams/regulation-](https://www.who.int/teams/regulation-prequalification/eul/in-vitro-emergency-use-listing-procedure)
511 [prequalification/eul/in-vitro-emergency-use-listing-procedure](https://www.who.int/teams/regulation-prequalification/eul/in-vitro-emergency-use-listing-procedure).
- 512 17. BGI. Real-Time Fluorescent RT-PCR Kit for Detecting SARS-CoV-2 2020
513 [Available from: [https://www.bgi.com/us/sars-cov-2-real-time-fluorescent-rt-pcr-](https://www.bgi.com/us/sars-cov-2-real-time-fluorescent-rt-pcr-kit-ivd/)
514 [kit-ivd/](https://www.bgi.com/us/sars-cov-2-real-time-fluorescent-rt-pcr-kit-ivd/)
- 515 18. Tshiabuila D, Giandhari J, Pillay S, et al. Comparison of SARS-CoV-2
516 sequencing using the ONT GridION and the Illumina MiSeq. *BMC Genomics*.
517 2022; 23(1):319. Published 2022 Apr 22. doi:10.1186/s12864-022-08541-5
- 518 19. Pembaur A, Sallard E, Weil PP, Ortelt J, Ahmad-Nejad P, Postberg J.
519 Simplified Point-of-Care Full SARS-CoV-2 Genome Sequencing Using
520 Nanopore Technology. *Microorganisms*. 2021; 9(12).

- 521 20. Vilsker M, Moosa Y, Nooij S, et al. Genome Detective: an automated system for
522 virus identification from high-throughput sequencing data. *Bioinformatics*.
523 2019;35(5):871-873. doi:10.1093/bioinformatics/bty695
- 524 21. Cleemput S, Dumon W, Fonseca V, et al. Genome Detective Coronavirus
525 Typing Tool for rapid identification and characterization of novel coronavirus
526 genomes. *Bioinformatics*. 2020;36(11):3552-3555.
527 doi:10.1093/bioinformatics/btaa145
- 528 22. Larsson A. AliView: a fast and lightweight alignment viewer and editor for large
529 datasets. *Bioinformatics (Oxford, England)*. 2014; 30(22):3276-8.
- 530 23. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high
531 throughput. *Nucleic acids research*. 2004; 32(5):1792-7.
- 532 24. Shu Y, McCauley J. GISAID: Global initiative on sharing all influenza data -
533 from vision to reality. *Euro surveillance: bulletin Europeen sur les maladies*
534 *transmissibles = European communicable disease bulletin*. 2017;22(13).
- 535 25. Nextclade. Next clade. <https://clades.nextstrain.org/> 2021 [Available from:
536 <https://clades.nextstrain.org/>.
- 537 26. Aksamentov I., Neher R. Nextalign: Viral Genome Reference Alignment.
538 [(accessed on 16 Jan 2023)]. Available
539 online: <https://github.com/neherlab/nextalign>
- 540 27. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, von
541 Haeseler A, Lanfear R. IQ-TREE 2: New Models and Efficient Methods for
542 Phylogenetic Inference in the Genomic Era. *Mol Biol Evol*. 2020 May
543 1;37(5):1530-1534. doi: 10.1093/molbev/msaa015.
- 544 28. Gatto L, Catanzaro D, Milinkovitch MC. Assessing the applicability of the GTR
545 nucleotide substitution model through simulations. *Evol Bioinform Online*.
546 2007;2:145-155. Published 2007 Feb 4.
- 547 29. Pavel Sagulenko, Vadim Puller, Richard A Neher, TreeTime: Maximum-
548 likelihood phylodynamic analysis, *Virus Evolution*, Volume 4, Issue 1, January
549 2018, vex042, <https://doi.org/10.1093/ve/vex042>

- 550 30. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y. ggtree: an r package for
551 visualization and annotation of phylogenetic trees with their covariates and
552 other associated data. *Methods in Ecology and Evolution*. 2017;8(1):28-36.
- 553 31. Ellingjord-Dale, M., Kalleberg, K.T., Istre, M.S. *et al.* The use of public transport
554 and contraction of SARS-CoV-2 in a large prospective cohort in Norway. *BMC*
555 *Infect Dis* 22, 252 (2022). <https://doi.org/10.1186/s12879-022-07233-5>
- 556 32. Mekuriaw W, Kinde S, Kindu B, Mulualem Y, Hailu G, Gebresilassie A, et.,al.
557 Epidemiological, Entomological, and Climatological Investigation of the 2019
558 Dengue Fever Outbreak in Gewane District, Afar Region, North-East
559 Ethiopia. *Insects*. 2022; 13(11):1066. <https://doi.org/10.3390/insects13111066>
- 560 33. Mengesha Tsegaye, M., Beyene, B., Ayele, W. *et al.* Sero-prevalence of yellow
561 fever and related Flavi viruses in Ethiopia: a public health perspective. *BMC*
562 *Public Health* **18**, 1011 (2018). <https://doi.org/10.1186/s12889-018-5726-9>
- 563 34. Tegally H, San JE, Cotten M, et al. The evolving SARS-CoV-2 epidemic in
564 Africa: Insights from rapidly expanding genomic surveillance. *Science*.
565 2022;378(6615):eabq5358. doi:10.1126/science.abq5358
- 566 35. Abagero A, Ragazzoni L, Hubloue I, Barone-Adesi F, Lamine H, Addissie A,
567 Della Corte F, Valente M. A Review of COVID-19 Response Challenges in
568 Ethiopia. *Int. J. Environ. Res. Public Health*. 2022; 19(17):11070.
569 <https://doi.org/10.3390/ijerph191711070>
- 570 36. Tessema GA, Kinfu Y, Dachew BA, Tesema AG, Assefa Y, Alene KA, et al.
571 The COVID-19 pandemic and healthcare systems in Africa: a scoping review of
572 preparedness, impact and response. *BMJ global health*. 2021;6(12):e007179.
- 573 37. Nigussie H. The Coronavirus Intervention in Ethiopia and the Challenges for
574 Implementation. *Front. Commun.*, 21 May 2021 Sec. Health Communication
575 Volume 6 – 2021. <https://doi.org/10.3389/fcomm.2021.562512>
- 576 38. Ngeh S, Vogt F, Sikazwe CT, et al. Travel-associated SARS-CoV-2
577 transmission documented with whole genome sequencing following a long-haul
578 international flight. *J Travel Med*. 2022;29(6):taac057. doi:10.1093/jtm/taac057

- 579 39. Gelanew T, Seyoum B, Mulu A, et al. High seroprevalence of anti-SARS-CoV-2
580 antibodies among Ethiopian healthcare workers. *BMC Infect Dis.*
581 2022;22(1):261. Published 2022 Mar 16. doi:10.1186/s12879-022-07247-z
- 582 40. Gudina EK, Ali S, Girma E, et al. Seroepidemiology and model-based
583 prediction of SARS-CoV-2 in Ethiopia: longitudinal cohort study among front-
584 line hospital workers and communities. *Lancet Glob Health.* 2021;9(11):e1517-
585 e1527. doi:10.1016/S2214-109X(21)00386-7
- 586 41. COVID-19 Vaccine tracker. [https://www.reuters.com/graphics/world-](https://www.reuters.com/graphics/world-coronavirus-tracker-and-maps/vaccination-rollout-and-access/)
587 [coronavirus-tracker-and-maps/vaccination-rollout-and-access/](https://www.reuters.com/graphics/world-coronavirus-tracker-and-maps/vaccination-rollout-and-access/)
- 588 42. Kemp, S.A., Collier, D.A., Datir, R.P. *et al.* SARS-CoV-2 evolution during
589 treatment of chronic infection. *Nature* 592, 277–282 (2021).
590 <https://doi.org/10.1038/s41586-021-03291-y>
- 591 43. Wilkinson SAJ, Richter A, Casey A, Osman H, Mirza JD, Stockton J, et al.
592 Recurrent SARS-CoV-2 mutations in immunodeficient patients. *Virus evolution.*
593 2022;8(2):veac050.
- 594 44. Weigang, S., Fuchs, J., Zimmer, G. *et al.* Within-host evolution of SARS-CoV-2
595 in an immunosuppressed COVID-19 patient as a source of immune escape
596 variants. *Nat Commun* 12, 6405 (2021). [https://doi.org/10.1038/s41467-021-](https://doi.org/10.1038/s41467-021-26602-3)
597 [26602-3](https://doi.org/10.1038/s41467-021-26602-3)
- 598 45. Moghadas SM, Vilches TN, Zhang K, et al. The impact of vaccination on
599 COVID-19 outbreaks in the United States. Preprint. *medRxiv.*
600 2021;2020.11.27.20240051. Published 2021 Jan 2.
601 doi:10.1101/2020.11.27.20240051
- 602 46. Korber B, Fischer WM, Gnanakaran S, et al. Tracking Changes in SARS-CoV-2
603 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. *Cell.*
604 2020;182(4):812-827.e19. doi:10.1016/j.cell.2020.06.043
- 605 47. Harvey, W.T., Carabelli, A.M., Jackson, B. *et al.* SARS-CoV-2 variants, spike
606 mutations and immune escape. *Nat Rev Microbiol* 19, 409–424 (2021).
607 <https://doi.org/10.1038/s41579-021-00573-0>
- 608 48. Liu Y, Rocklöv J. The reproductive number of the Delta variant of SARS-CoV-2
609 is far higher compared to the ancestral SARS-CoV-2 virus. *J Travel Med.* 2021

610 Oct 11;28(7):taab124. doi: 10.1093/jtm/taab124. PMID: 34369565; PMCID:
611 PMC8436367.

612 49. Thakur S, Sasi S, Pillai SG, et al. SARS-CoV-2 Mutations and Their Impact on
613 Diagnostics, Therapeutics and Vaccines. *Front Med (Lausanne)*. 2022;9:
614 815389. Published 2022 Feb 22. doi:10.3389/fmed.2022.815389

615 50. Shrivastava S, Mhaske ST, Modak MS, Virkar RG, Pisal SS, Mishra AC, et al.
616 Emergence of two distinct variants of SARS-CoV-2 and an explosive second
617 wave of COVID-19: the experience of a tertiary care hospital in Pune, India.
618 *Archives of virology*. 2022; 167(2):393-403.

619 51. Dhama K, Nainu F, Frediansyah A, Yatoo MI, Mohapatra RK, Chakraborty S, et
620 al. Global emerging Omicron variant of SARS-CoV-2: Impacts, challenges and
621 strategies. *Journal of Infection and Public Health*. 2023; 16(1):4-14.

622 52. Zhou Y, Zhi H, Teng Y. The outbreak of SARS-CoV-2 Omicron lineages,
623 immune escape, and vaccine effectivity. *J Med Virol*. 2022;10.1002/jmv.28138.
624 doi:10.1002/jmv.28138

625 53. Corriero A, Ribezzi M, Mele F, Angrisani C, Romaniello F, Daleno A, et., al.
626 COVID-19 Variants in Critically Ill Patients: A Comparison of the Delta and
627 Omicron Variant Profiles. *Infectious Disease Reports*. 2022; 14(3):492-500.
628 <https://doi.org/10.3390/idr14030052>

629 54. Dubey, A., Choudhary, S., Kumar, P. et al. Emerging SARS-CoV-2 Variants:
630 Genetic Variability and Clinical Implications. *Curr Microbiol* **79**, 20 (2022).
631 <https://doi.org/10.1007/s00284-021-02724-1>

632 55. McCrone, J.T., Hill, V., Bajaj, S. et al. Context-specific emergence and growth
633 of the SARS-CoV-2 Delta variant. *Nature* **610**, 154–160 (2022).
634 <https://doi.org/10.1038/s41586-022-05200-3>

635 56. Githinji G, Lambisia AW, Omah I, O'Toole A, Mohamed KS, de Laurent ZR, et
636 al. The genomic epidemiology of SARS-CoV-2 variants of concern in Kenya.
637 medRxiv. 2022:2022.10.26.22281446.

638

639

640

641

642 **Tables, Figures and Supplementary Files**

643 **Table 1.** Socio-demographic characteristics of study participants along with VOC of SARS-CoV-2 in
644 Ethiopia, 2020-2022.

Variable	Class	Frequency (%)	Variant (VOC)			
			Alpha	Beta	Delta	Omicron
Sex	Female	181 (51.3)	10	2	99	57
	Male	172 (48.7)	7	1	102	46
Age Range (years)	1-20	40 (11.3)	7	1	19	8
	21-30	121 (34.3)	1	1	80	25
	31-40	84 (23.8)	5	1	43	30
	41-50	41 (11.6)	2	0	25	12
	51-60	21 (10.2)	2	0	18	15
	>60	31 (8.8)	0	0	16	13
Reason for testing (during sampling)	Suspect	107 (30.3)	6	1	133	2
	Contacts of confirmed cases	79 (22.4)	8	2	45	1
	Community Surveillance	167 (47.3)	3	0	23	100
Clinical Status of the clients while sampling	Asymptomatic*	53 (15)	3	0	37	5
	Mild	273 (77.3)	14	3	142	94
	Sever	27 (7.7)	0	0	22	4

645 *Asymptomatic – a person has tested positive for COVID-19 but never exhibits any signs or symptoms of the
646 disease
647

648 *Figure 1: SARS-CoV-2 epidemiology and variant turnover in Ethiopia: (a) 7-day rolling*
649 *average for the number of new SARS-CoV-2 cases. Others, represented in blue, comprise*
650 *mostly B lineages including B, B.1, and B.1.1 lineages. The size of the dot is indicative of the*
651 *variant proportion in the population. (b) The proportion of individuals that have received at*
652 *least one vaccine dose in Ethiopia compared to the number of new deaths that occurred. (c)*
653 *Summary of the mean number of detected introductions into Ethiopia binned by month for*
654 *different regions across the dataset. (d) Summary of the mean number of detected*
655 *exportations out of Ethiopia to different regions binned by month.*

656 *Figure 2: SARS-CoV-2 lineage diversity in Ethiopia. (A) Proportions of SARS-CoV-2 lineages,*
657 *classified by Pangolin nomenclature, circulating between June 2020 and February 2022 in*
658 *Ethiopia. (B) Absolute counts of SARS-CoV-2 genomes throughout the study period. Only*
659 *VOCs and major lineages are listed in the legend. (C) Maximum-Likelihood phylogeny of 353*
660 *SARS-CoV-2 genomes. Branches are colored by lineage.*

661

662

663

664 **List of Supplementary Files**

665 Table S1- Timeline of COVID-19 and major events following the first confirmed cases
666 in Ethiopia

667 Figure S2- Consortium Diagram of Sampling for Molecular Epidemiology and Diversity
668 of SARS-CoV-2 in Ethiopia

669 Figure S3- The sampling area covers in Ethiopia

670 Figure S4-Major SARS-CoV-2 lineages of detected imported and exported Ethiopia
671 between June 2020 and February 2022

672 Figure S5- SARS-CoV-2 Delta VOC sub tree molecular clock, Ethiopia.

673 Figure S6-Major SARS-CoV-2 lineages identified Ethiopia between June 2020 and
674 February 2022

675

676

677

678

679

680

681

682

683

684

685

686 **Table S1 Summary of the COVID-19 timeline and major events following the first**
 687 **confirmed cases in Ethiopia**

Timeline	Major events	Remark
December 2019/January 2020	Wuhan City, Hubei Province, China, which was reported to the WHO on 31 December 2019. Wuhan, China, was reported in January 2020	
27 January 2020	Prior to the WHO's declaration of a global pandemic, the federal democratic republic of Ethiopia Council of Ministers 'activated' a National Public Health Emergency Preparedness Centre, EOC and began preparations to deal with a potential outbreak of Covid-19	
30 January 2020	the Director-General of the WHO declared COVID-19 a Public Health Emergency of International Concern	
7 February 2020	Ethiopia set up the national Influenza and Arboviruses reference laboratory as the first COVID-19 testing laboratory in the country with the support from WHO and ACDC.	
March 11, 2020-	WHO declared COVID-19 as a Pandemic	
March 13, 2020, ***	On 12 March 2020, a week after entering the country from Burkina Faso, a 48-year-old Japanese national presented himself at a public health center in the capital city, Addis Ababa, and was diagnosed as having COVID-19. Since Ethiopia did not have the capacity to test for COVID-19 at the onset. Thus, samples were initially transported for testing to the National Institute for Communicable Diseases (NICD) Laboratory in South Africa	
15 March 2020	The second confirmed case from Ethiopian nationalities (three additional)cases of the COVID-19 were reported	
20 March 2020	Anyone entering the country/international travelers/ should stay in quarantine for up to 14 days at their costs	
23 March 2020	Ethiopia closed all land borders and deployed security forces to halt the movement of people along the borders (National lockdown....)	
24 March 2020	<ul style="list-style-type: none"> ● banned all public gatherings and sports events ● The closure of bars and clubs ● all federal employees were to work from home, except those designated by each ministry and federal agency as essential workers 	
05 April 2020	The First COVID-19-related confirmed death	
8 April 2020	A state of emergency was declared giving authorities across-the-board powers to battle the disease	
06 March 2021 <i>*after a year of pathogen occurrence</i>	The first genome sequence of SARS-CoV-2 sampled from Ethiopia was submitted to the GISAID database	
13March, 2021	First vaccination started	

688 *** Form 13 March-Mid April 2020-Sample were tested by referral to NICD, the country had no RT-PCR testing
 689 center during this time
 690

691 **Figure S2--**Consortium Diagram of Sampling for Molecular Epidemiology and Diversity of SARS-CoV-2
692 in Ethiopia

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

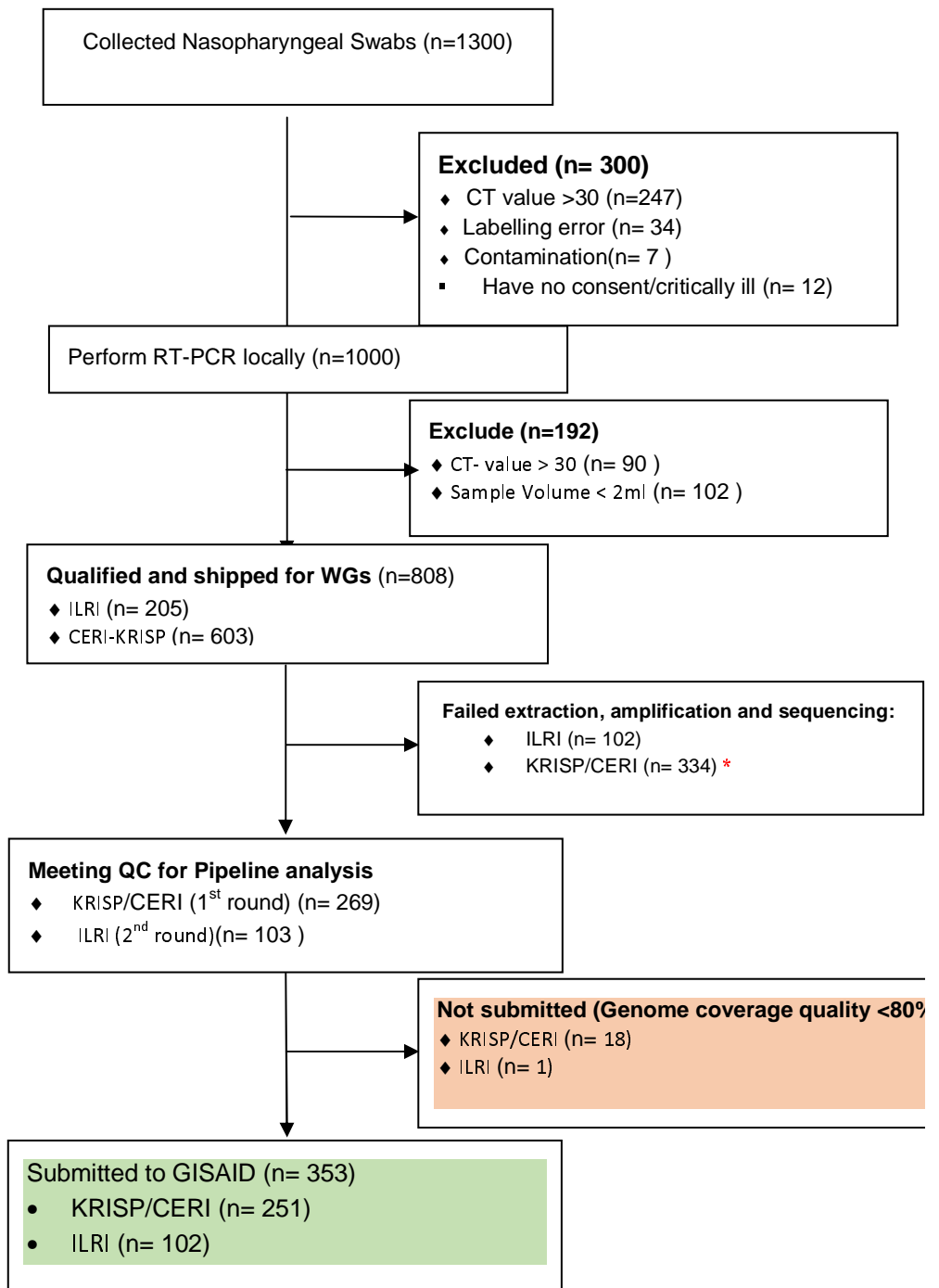
711

712

713

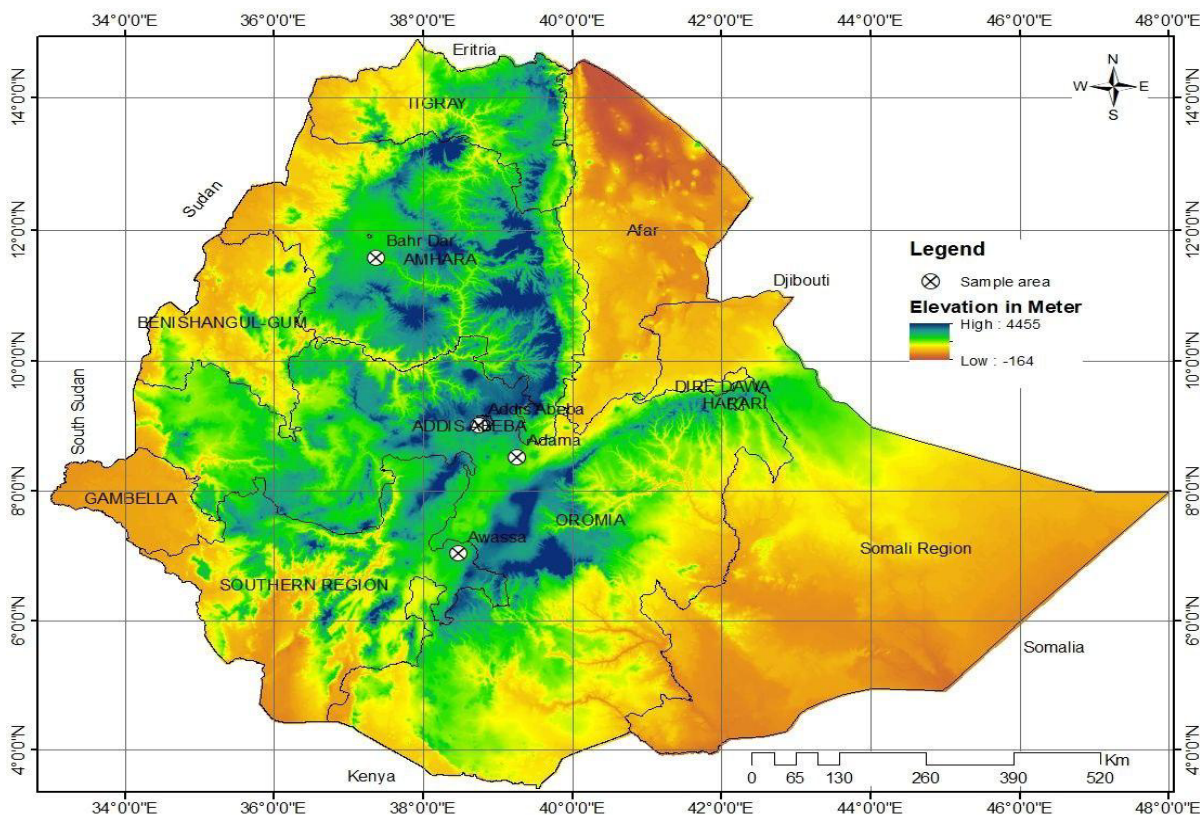
714

715



*because of the old specimen, long shipping time, backlog seq., temp not maintained ...

716 Figure - S3a- Sampling area for Molecular Epidemiology and Diversity of SARS-CoV-2 in Ethiopia



717

718 Table S3b: Sequence proportion of the regional government

Name of Regional	Proportion of sequencing (%)
Addis Ababa	80%
Oromia	8%
Amhara	5%
SNNPR and Sidama	5%
Others	2%

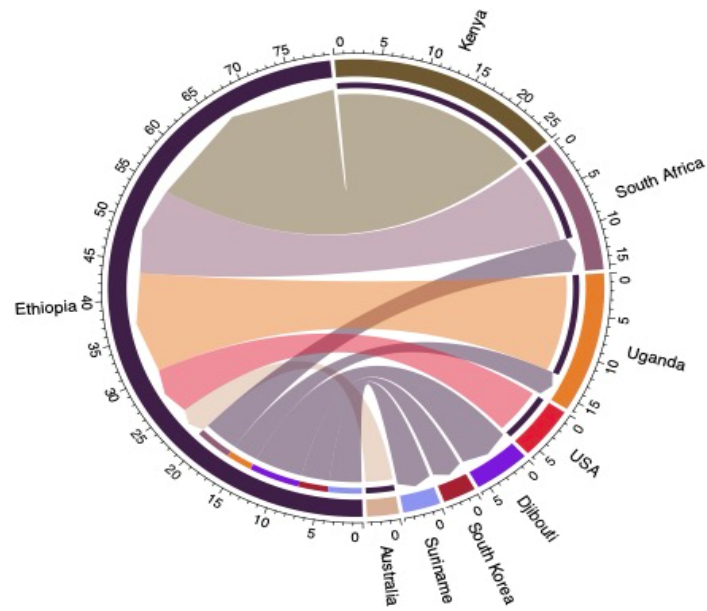
719

720

721

722

723

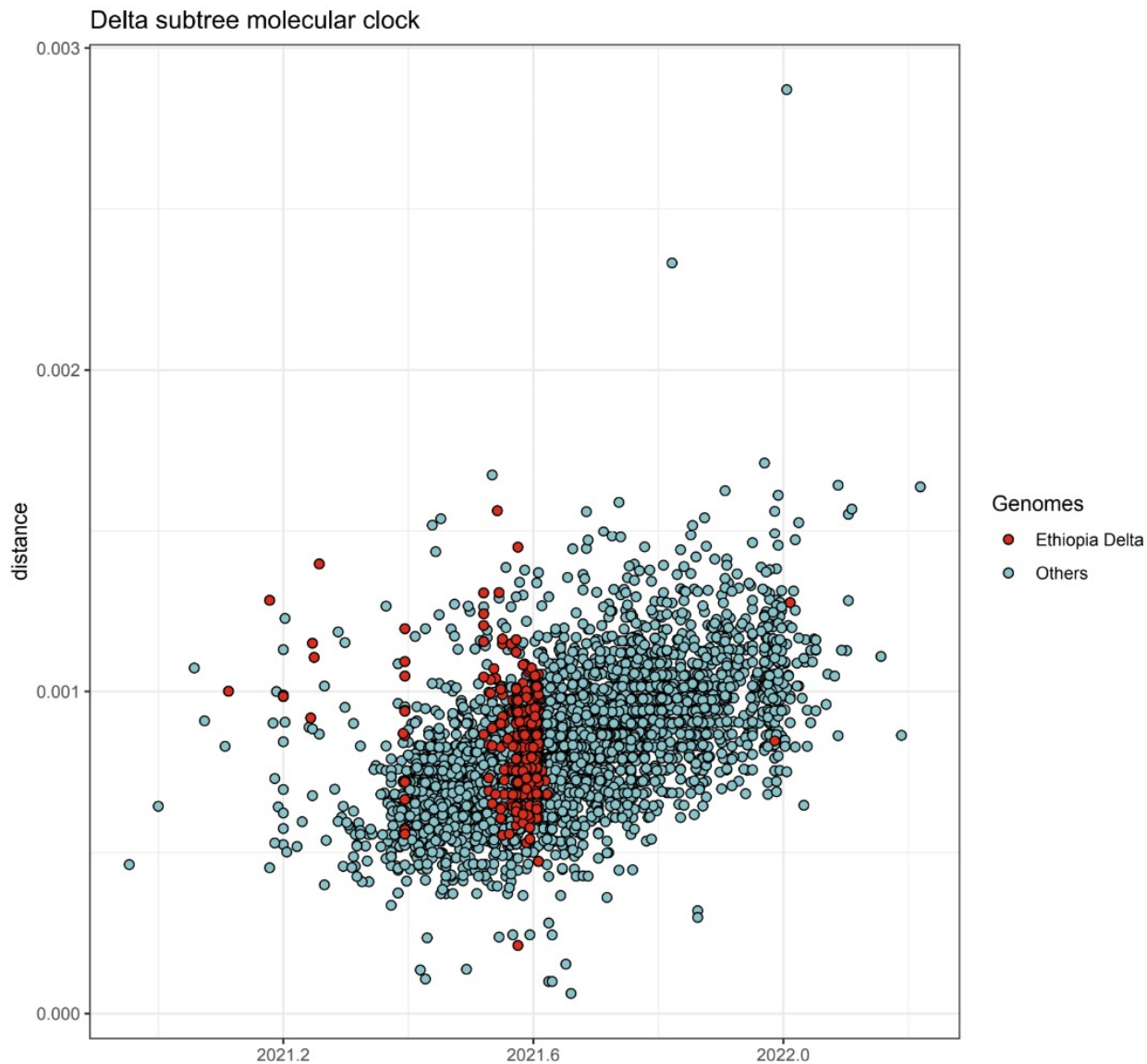


724

725 **Figure S 4**-Major SARS-CoV-2 lineages of detected imported and exported Ethiopia between June
726 2020 and February 2022

727

728



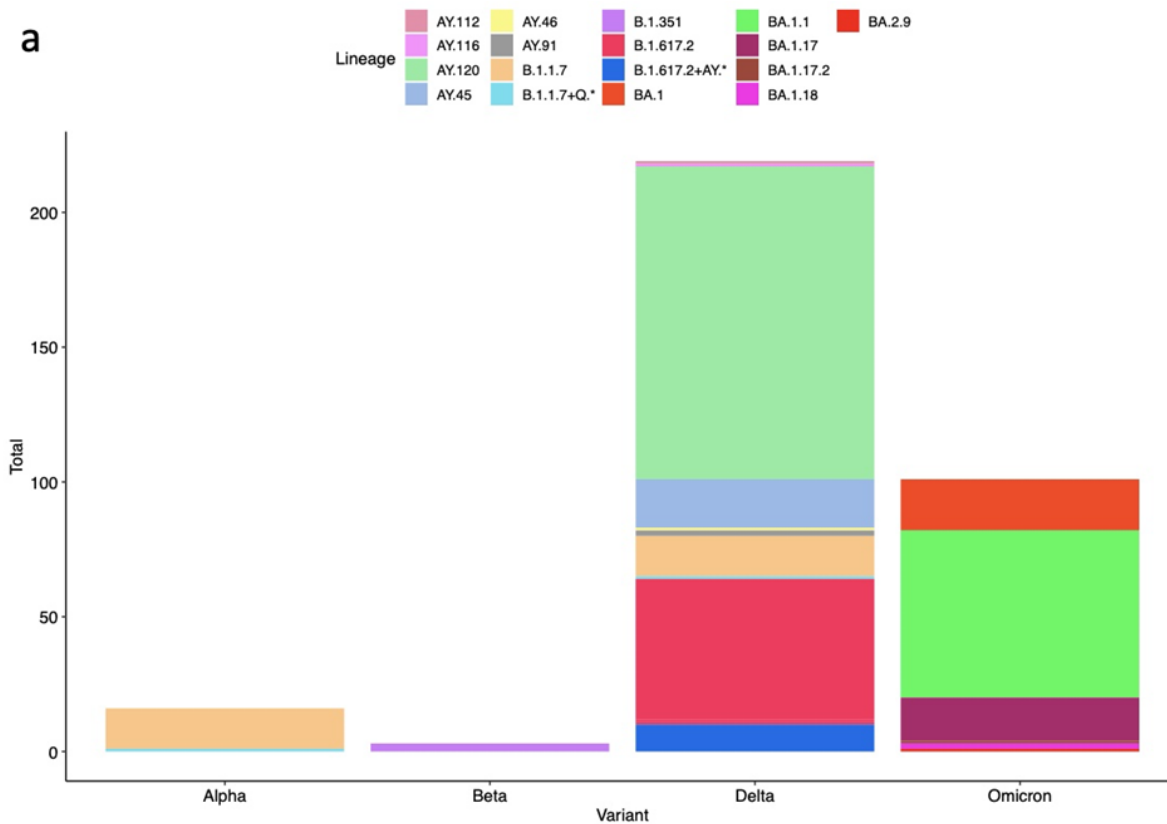
729

730 **Figure S 5** - SARS-CoV-2 Delta VOC sub tree molecular clock, Ethiopia.

731

732

733



734

735

Figure S 5-Major SARS-CoV-2 lineages circulating in Ethiopia between June 2020 and February 2022

736

737

738

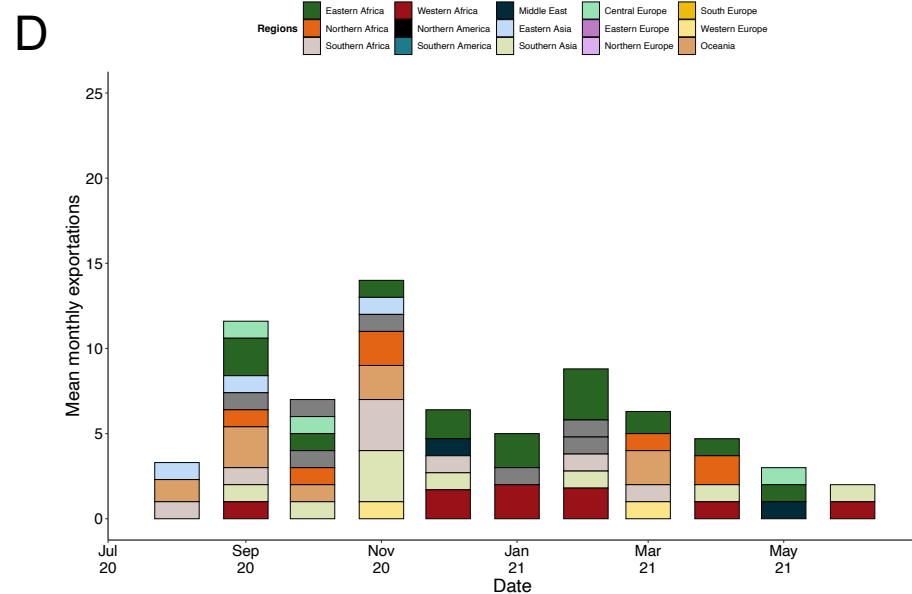
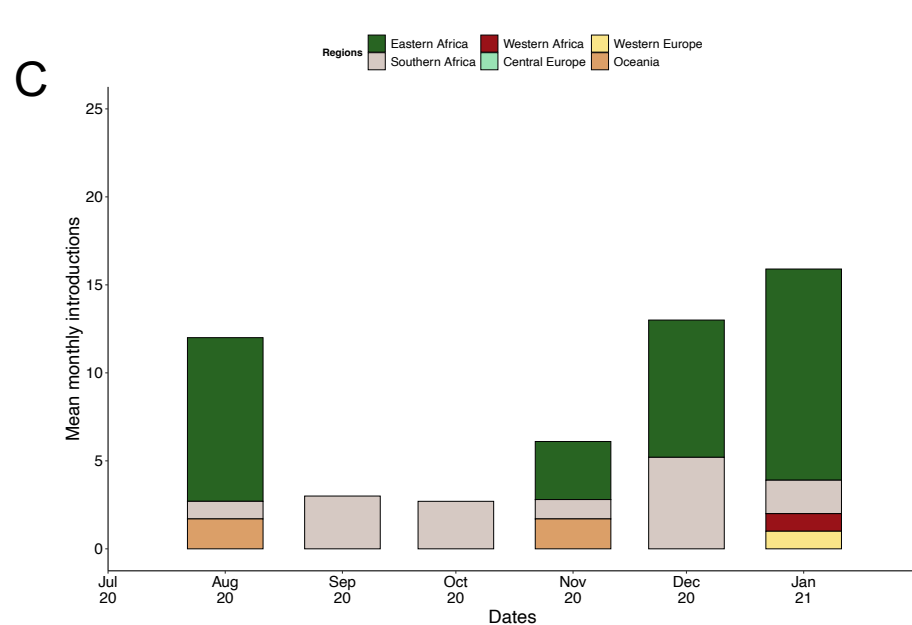
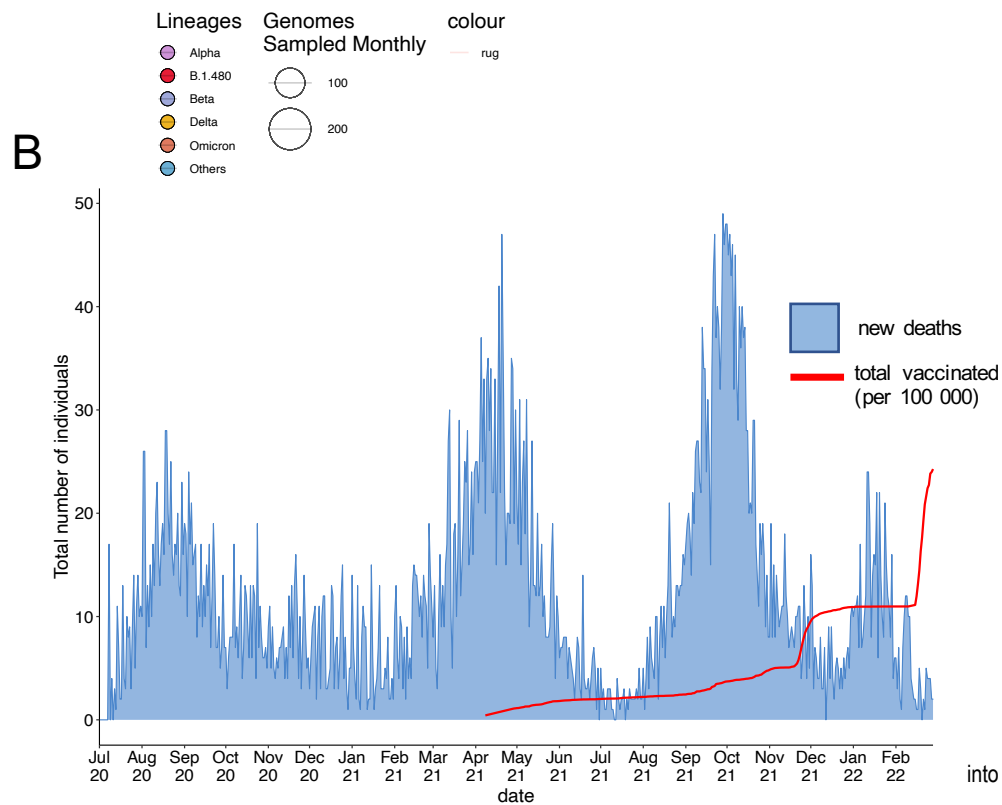
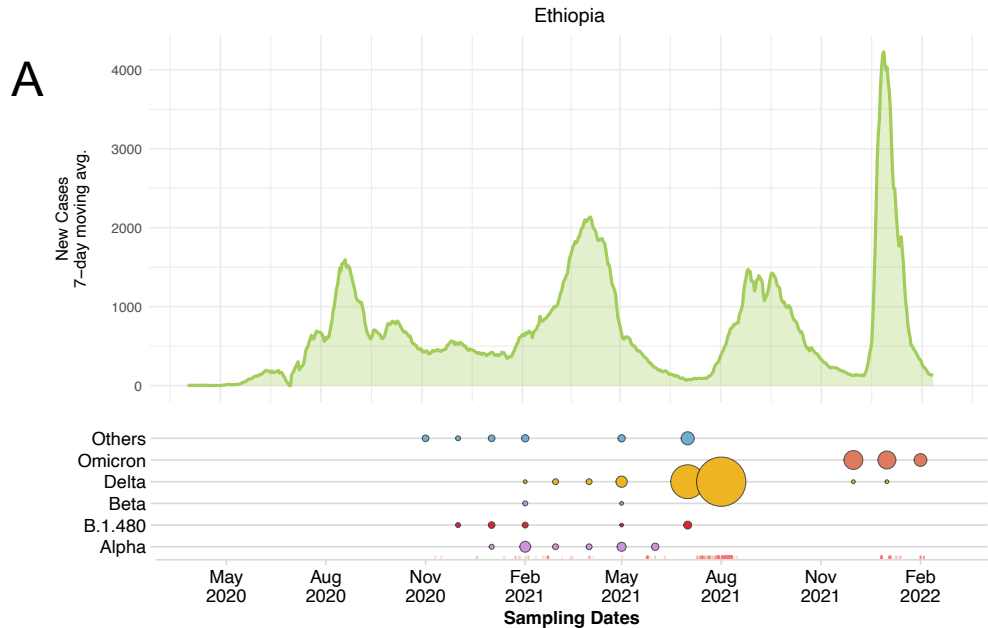


Figure 1: SARS-CoV-2 epidemiology and variant turnover in Ethiopia: (a) 7-day rolling average for the number of new SARS-CoV-2 cases. Size of the dot is indicative of the variant proportion in the population. (b) The proportion of individuals that have received at least one vaccine dose in Ethiopia compared to the number of new deaths that occurred. (c) Summary of the mean number of detected introductions into Ethiopia binned by month (d) Summary of the mean number of detected exportations out of Ethiopia to different regions binned by month.

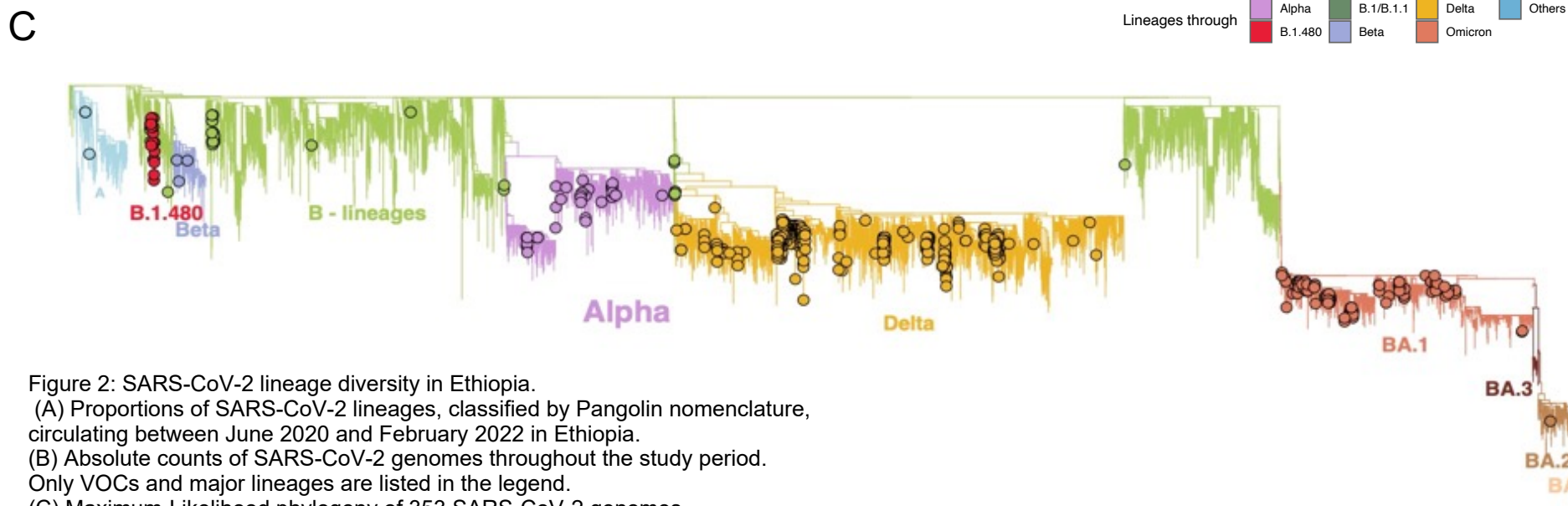
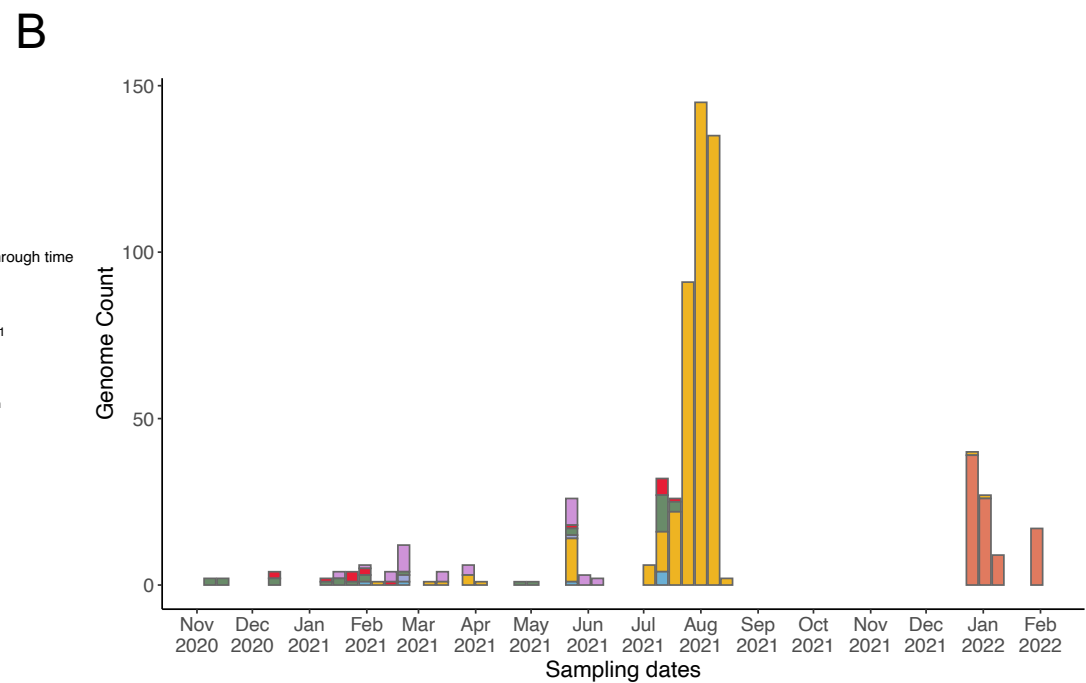
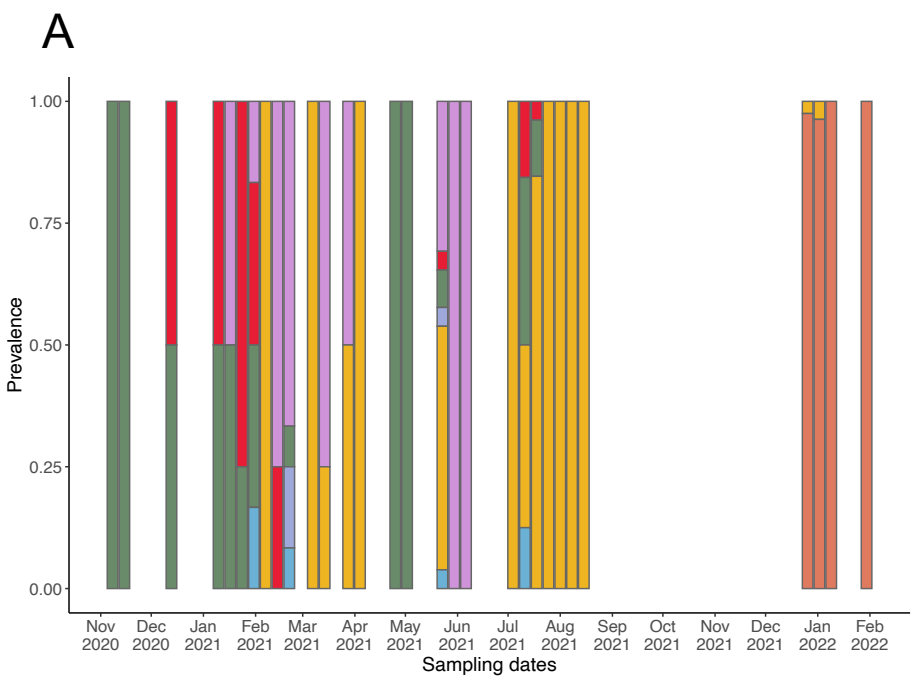


Figure 2: SARS-CoV-2 lineage diversity in Ethiopia.

(A) Proportions of SARS-CoV-2 lineages, classified by Pangolin nomenclature, circulating between June 2020 and February 2022 in Ethiopia.

(B) Absolute counts of SARS-CoV-2 genomes throughout the study period.

Only VOCs and major lineages are listed in the legend.

(C) Maximum-Likelihood phylogeny of 353 SARS-CoV-2 genomes.

Branches are colored by lineage.