$\frac{1}{2}$	Genomic epidemiology of SARS-CoV-2 infections in The Gambia, March 2020 to Jan 2022
3	Abdoulie Kanteh ¹ , Haruna S. Jallow ² , Jarra Manneh ¹ , Bakary Sanyang ¹ , Mariama A. Kujabi ¹ , Sainab

- Abdoulie Kanteh¹, Haruna S. Jallow², Jarra Manneh¹, Bakary Sanyang¹, Mariama A. Kujabi¹, Sainabou Laye
 Ndure¹, Sheikh Jarju¹, Alhagie Papa Sey², Dabiri Damilare K¹, Yaya Bah¹, Sana Sambou³, Gibril Jarju³, Buba
 Manjang³, Abubacarr Jagne³, Sheikh Omar Bittaye³, Mustapha Bittaye³, Karen Forrest¹, Desta Alamerew
 Tiruneh⁷, Ahmadou Lamin Samateh³, Sheriffo Jange², Stéphane Hué^{4,5}, Nuredin Muhammed¹, Alfred
 Amambua-Ngwa¹, Beate Kampmann¹, Umberto D'Alessandro¹, Thushan I. de Silva^{1,6}, Anna Roca¹ and Abdul
 Karim Sesay¹
- 9

¹ Medical Research Council Unit The Gambia at London School of Hygiene and Tropical Medicine, Banjul, The
 Gambia (A Kanteh MSc, J Manneh MSc, B Sanyang BSc, M A Kujabi MSc, S L Ndure BSc, D Damilare K
 BSc, Y Bah BSc, S Jarju DVM, K Forrest MD, N Muhammed PhD, A Amambua-Ngwa PhD, Prof B
 Kampmann PhD, Prof U D'Alessandro PhD, T.I. de Silva PhD, Prof A Roca, PhD and AK Sesay PhD)

- ¹⁵ ² National Public Health Reference Laboratory, Ministry of Health, The Gambia (H Jallow MSc, A P Sey HND,
 ¹⁶ S Jagne MSc)
- 17
- ³ Ministry of Health, The Gambia (S Sambou MSc, G Jarju MSc, B Manjang PhD, A Jagne MD, S O Bittaye
 MD, M Bittaye MD, A L Samateh MD)
- 20

⁴ Centre for Mathematical Modelling of Infectious Diseases, London School of Hygiene and Tropical Medicine,
 London, UK (S Hue, PhD)

23 24

⁵Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, London,
 UK (S Hue, PhD)

26

⁶ The Florey Institute for Host-Pathogen Interactions & Department of Infection, Immunity and Cardiovascular
 Disease, Medical School, University of Sheffield, Sheffield, UK (T.I.de Silva PhD)

- 29
- 30 ⁷ World Health Organization, The Gambia (D A Tiruneh)
- 31

32 Corresponding author

- 33 Dr Abdul Karim Sesay
- 34 Medical Research Council Unit The Gambia at London School of Hygiene and Tropical Medicine
- 35 Banjul, The Gambia
- 36 Tel: +2202241521
- 37 38

39 Abstract

40

41 Background

42 COVID-19, caused by SARS-CoV-2, is one of the deadliest pandemics over the last 100 years. Sequencing is
 43 playing an important role in monitoring the evolution of the virus, including the detection of new viral variants.
 44 This study describes the genomic epidemiology of SARS-CoV-2 infections in The Gambia.

45

46 Methods

Nasopharyngeal and/or oropharyngeal swabs collected from suspected cases and travellers were tested for
 SARS-CoV-2 using standard RT-PCR methods. SARS-CoV-2 positive samples were sequenced following
 standard library preparation and sequencing protocols. Bioinformatic analysis was done using ARTIC pipelines
 and lineages assigned using Pangolin.

51

52 Findings

Between March 2020 to January 2022, there were almost 12,000 SARS-CoV-2 confirmed cases distributed into four waves, each of them lasting between 4 weeks and 4 months, with more cases during the rainy seasons (July- October). As shown by the 1643 sequenced samples, each wave occurred after new viral variants and/or lineages were introduced in The Gambia, generally those already established in Europe and/or in other African countries. Local transmission was higher during the first and third wave, with mostly B.1.416/Senegal/Gambian lineage and AY.34.1/Delta subtype, respectively. The second wave was driven by two variants, namely Alpha and Eta and B.1.1.420 lineage. The Omicron/fourth wave was the shortest.

60

61 Interpretation

62 Efficient surveillance, including strengthening entry points and screening asymptomatic individuals especially 63 during the rainy seasons would be important to promptly detect and control future waves in The Gambia and the 64 subregion.

65

66 Funding

67 Medical Research Unit The Gambia at LSHTM, UK Research and Innovation funding (grant reference 68 MC_PC_19084), MRC/UKRI MC_PC_19084 and World Health Organisation.

69 70

71 Introduction

The COVID-19 pandemic caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has claimed more than six million lives and infected over 526 million people globally [Accessed: 15-03-2022].¹ As of May 2022, Africa has contributed to 2.0% of cases and 2.7% deaths to the global COVID-19 burden, ² despite constituting 17% of the world's population.³ In contrast, Europe, with more than two million SARS-CoV-2 infections, contributed to 42% of cases and 32% of deaths. ² The pandemic has evolved differently around the globe starting later in Africa than in other regions ⁴ although sero-prevalence studies suggest that transmission has been much higher than reported. ⁵⁻⁷

79

80 Despite its low mutation rate, widespread transmission and high case numbers have increased the diversity of 81 SARS-CoV-2 during the pandemic.⁸ Prolonged infections in immunocompromised individuals ⁹ and high rate of 82 reinfections increase the likelihood of new mutations, leading to the emergence of new variants. More 83 importantly, introduction of new variants of concerns (VOCs) reported in different countries ^{10,11} is responsible for higher transmission, ¹² disease severity ^{13,14} and immune evasion ¹⁵. For example, the Alpha variant or 84 85 B·1·1·7 lineage, first identified in the UK in November 2020, was reported to be between 43% and 90% more transmissible than the wild-type SARS-CoV-2.¹⁶ The Delta variant/B·1·617·2, first reported in India in October 86 87 2020, is 4 times more likely to evade the immune system and 50 times more transmissible than the Alpha 88 variant.¹⁷ Indeed, the emergence and spread of the Delta variant changed the dynamic of the epidemic in many 89 African countries as it coincided with an increased numbers of reported cases and deaths, with a very long and 90 slow downward trend.¹⁸ The Omicron variant, first identified in South Africa in November 2021, was described 91 as the most transmissible variant, quickly spreading across the world and exhibiting a high degree of evading 92 antibody immunity.19

93

Effective strategies to quickly detect new variants, sources of infections, outbreaks and transmission patterns in different geographical settings are essential for controlling the spread of SARS-CoV-2. Whole genome sequencing enabled detailed understanding of pathogen's evolution and of its population structure. ²⁰ High resolution viral genome sequences linked with epidemiological data allowed identifying the origin of the virus ²¹, and monitoring its spread and evolution, both globally ²² and locally. ^{23,24} Moreover, the information obtained from sequencing is essential for guiding the design of the next generation vaccines. ²⁵ As of May 2022, more than 11 million genomes have been sequenced globally and submitted to GISAID. ²⁶ Unfortunately, only 1% of these sequences is from African countries ²⁶, a major obstacle to understanding the local evolution of the virus.

103

104 The Gambia, a small country in West Africa, identified its first imported COVID-19 case from the UK 27 in 105 March 2020. By February, 2022, The Gambia had recorded almost 12,000 cases and 365 deaths.²⁸. More than 106 1,600 SARS-CoV-2 genomes have already been sequenced and analysed with the aim of providing real time 107 genomic data to monitor circulating variants within The Gambia. Moreover, this information is essential for 108 monitoring local transmission and will significantly contribute to the global genomic dataset. We have already 109 described the origin of the first five cases diagnosed in The Gambia during the early phase of the pandemic 110 and used phylogenetic analysis to confirm SARS-CoV-2 reinfection in two healthy individuals.²⁹ The study 111 presented here aims to extensively describe the genomic epidemiology of the first four SARS-CoV-2 waves 112 observed in The Gambia over a 23-month study period. 113

114 Methods 115

116 The Gambia, population demographics and climate

The Gambia has a population of about 2.5 million people, predominantly Muslims, with a median age of 17.8 years. ³⁰ More than half of the population live in urban areas, mainly at the coast. The climate is subtropical with two seasons, i.e., a rainy season between June and October and a dry season for the rest of the year, although November and December are "wet months" during which the malaria transmission peak occurs. ³¹ Average temperatures range between 23°C and 33°C during the rainy season with high humidity (>80%) and between 18°C and 30°C during the dry season. ³¹

123

124 The Gambia Healthcare system

The Gambian Government is the main healthcare provider, operating seven referral hospitals, eight major health centres and 16 minor ones across the country. Healthcare delivery is mostly through the primary or local health posts. In addition, there are about 30 private and nongovernmental organisation clinics across the country. ³²

129 COVID-19 response in The Gambia

Measures implemented to control the spread of the COVID-19 epidemic have already been described elsewhere.
 ³³ Briefly, international borders were closed in March 2020 and a state of emergency declared (Figure 1). Initial
 SARS-CoV-2 testing by PCR was focused on identifying imported cases and tracing, plus isolating case
 contacts, especially among travellers from Senegal and their contacts.

134 135

Figure 1. Timeline of events and measures put in place during the pandemic.

136137 Sample collection

Nasopharyngeal, oropharyngeal or both swabs were collected for SARS-CoV-2 detection using FLOQSwabs. ³⁴ Although there were different reasons for testing, most samples (>85%) were collected as a pre-travel testing requirement (FigureS1).³⁵ Swabs were placed in single tubes containing a universal transport medium. ³⁴ and delivered to the laboratory within 24 hours. Demographic data including gender, age, address, travel history and any other relevant information were collected whenever possible.

143

144 Figure S1. Distribution of sample collection (reason for testing) in The Gambia over the 23 months of study 145 period.

146

147 **RNA extraction and PCR**

Samples were transported for analysis to the National Public Health Laboratory (NPHL) or the MRC Unit The Gambia at LSHTM molecular diagnostic laboratory. Sample processing, RNA extraction and real-time reverse transcription PCR (RT-PCR) were done following standard WHO guidelines.^{33,36} From August 2020 to date, PCR positive SARS-CoV-2 required PCR amplification of both the screening (E) and confirmatory (N) genes. Results and associated metadata collected for each sample were recorded in RedCap ³⁷ and DHIS2 ³⁸, respectively.

155 Sample selection profile

Samples without basic metadata such as date of collection, age and sex were not processed (Figure S2). Criteria for sequencing changed over time. From March to June 2020, all samples were processed and sequenced, regardless of cycle threshold (Ct) values. However, due to the high proportion of failed sequencing among samples with Ct values above 30, in June 2020 it was decided to sequence only samples with a Ct value of \leq 30. In addition, samples with low DNA concentration ((\leq 4ng/µ1) after library preparation were also excluded.

161

162 Figure S2. Total count of samples that were dropped from the study and reason for dropping.163

164 Library preparation and sequencing

cDNA conversion and multiplex PCR were prepared following ARTIC nCoV-2019 protocol. ³⁹ Updated primer
 schemes were always used to ensure optimum sequencing of emerging VOCs. ³⁹⁻⁴² Due to variable Ct values, all
 RNA samples were run for 35 cycles. Pooled PCR products were either cleaned up using 1x SPRI/AMPure XP
 beads or library prepared directly without bead clean-up following manufacturer's instructions. Libraries were
 prepared for sequencing on the Oxford Nanopore platform as described previously. ⁴⁴

171 Quality Control

172 Samples were prepared and sequenced in 96-well plates with one cDNA and RNA extraction negative control 173 each, per plate. Consensus sequence was defined as passing quality control if greater than 50% of the genome 174 was covered by confident calls or there was at least one contiguous sequence of more than 10,000 bases and no 175 evidence of contamination in the negative control. A confident call was defined as having 10x depth of 176 coverage. If the coverage fell below these thresholds, the bases were masked with the character N, indicating 177 that the base at that position is unknown or not available. Low quality SNP were also masked with Ns. For 178 phylogenetic analysis, only genomes with less than 10% missing bases were used to construct robust 179 phylogenetic trees.

180

181 Clustering and phylogenetic analyses

Raw reads were analysed using the ARTIC bioinformatics pipeline.⁴⁵ FASTQ files were trimmed to remove 182 183 adapters and mapped to a reference genome (GenBank accession number MN908947-3). Consensus genomes were further analysed to assigned lineages and clades using the Pangolin online tool ⁴⁶ and Nextclade, ⁴⁷ 184 185 respectively. To find the closest global sequences to that of Gambian samples from GISAID database, we first 186 stratified our sequences into different waves (wave 1 to 4). We then used Uvaia (GitHub, last accessed: 14-05-187 2022) to find the closest sequences to our samples against from global sequences deposited in GISAID 188 [accessed 14-05-2022]. Only sequences (genomes in GISAID) with high nucleotide similarity (raked as 1 in 189 Uvaia) to Gambian samples were retained for phylogenetic analysis. Ambiguous bases and low coverage

regions were masked prior to multiple sequence alignment using MAFFT (v7·453).⁴⁸ Maximum likelihood phylogenetic trees were constructed using IQTREE (v1·3·11·1) ⁴⁹ under the HKY model of nucleotide substitution, and branch supports were estimated using ultra-fast bootstrapping, as implemented in IQTREE, with 1,000 replicates. The phylogenies were visualised and edited using microreact. ⁵⁰

- 195 Results
- 196

197 Samples, cases, and sequences

Between March 2020 and January 2022, among the 75,554 samples tested for SARS-CoV-2, 15-8% (11,911) were positive [Accessed:13-02-2022].² Among the 3,254 samples received for sequencing, 1643 (50.5%) passed all QCs and were included in this study (Figure 2). Reasons for missing sequenced data are detailed in figure S2

Figure 2. Number of SARS-CoV-2 positive samples sequenced by month. The bars represent the total number of cases (blue), total number of samples received for sequencing (orange), total samples sequenced (grey) and total passed genotyping (yellow). The blue line represents the proportion of samples sequenced that passed genotyping.

206

207 Description of epidemic waves

The Gambia has experienced four different waves since the first identified case of SARS-CoV-2 infection (Figure 3). The first wave, predominantly associated with the B·1·416 lineage started in July 2020 and lasted for 3 months. The second wave started in January 2021 and ended in May 2021 and was milder (less cases and deaths) and lasted longer than the first wave. Variants Eta/B·1·525 and Alpha/B·1·1·7, and a lineage (B·1·1·420), dominated during the second wave. The third wave, with a higher number of cases and deaths, occurred between July and September 2021, and was driven by the Delta/B·1·617·2 variant and its subtypes. The fourth wave, between December 2021 and January 2022, was dominated by the Omicron variant/subtypes.

215

Figure 3. Number of reported SARS-CoV-2 overtime (a) reported cases and deaths (b) trend of lineages and major variants in The Gambia from March 2020 – January 2022.

219

220 Changes in variant prevalence

Between March 2020 and January 2022, 55 different lineages were identified (Figure 4). During the first wave,
lineage B·1·416 represented 49% (173/354) of the total samples sequenced followed by B·1·1 16% (57/354)
(Figure S3a). Six lineages were observed in only one sample (B·1·525, B, B·1·177, A·27, B·1·1·7 and B·1·582).

Figure 4. The number of different lineages identified each month during each wave from March 2020 – January
 2022.

Figure S3. The proportion of lineages by wave during the epidemic.

Some lineages identified during the first wave disappeared during the second wave, while new lineages emerged. For example, B·1·416 represented only 1% (8/555) of lineage sequenced from the second wave, with the last identified in February 2021 (Figure S3b). In March 2021, almost one year after the identification of the first case in The Gambia, B·1·1·420 became the dominant lineage constituing 23% (126/555) of isolates, followed by Eta/B·1·525 (22%, 124/555) and Alpha/B·1·1·7 (21%, 119/555) respectively (S3b). Fifteen lineages were represented by only one sample.

236

The Delta/B·1·617·2 variant was first confirmed in The Gambia in early April 2021 and became the dominant variant within a few weeks (Figure S3c). Between July and August 2021 (peak of the third wave), the Delta variant subtype AY·34·1 represented more than half (256/470) of all the samples sequenced, while the rest consisted of other lineages such as B·1·617·2 (31%, 146/470) and AY·34 (9%, 41/470), among others. Six lineages were observed in only one sample (AY·22, AY·4, AY·42, B·1·466, B·1·17 and Q·4).

242

The omicron variant and its subtypes dominated the fourth wave (Figure S3d). Subtype BA.1.1 accounted for 79% (124/156) of all the genomes sequenced. The remaining samples were BA·1 representing approximately 7% (6/156) and the Delta variant subtypes, representing 8% (12/156) of all samples, with subtype AY.34 being the most common. Five lineages were observed in only one sample (AY·122, BA·1·15, BA·1·17, BA·1·17·2, and BA·1·18).

248

249 Clustering and Phylogenetic analysis

Phylogenetic analysis of Gambian samples revealed eight distinct, strongly supported clades, which clustered based on lineage types. (Figure 5). Phylogenetic clustering using closest global sequences showed high local transmission events during the first and third wave. (Figure 6a and 6c). Lineage B·1·416 clustered with genomes from Senegal, suggesting spread of infections between these two countries (Figure 6a). Similarly, high clustering among Gambian Delta samples during the third wave showed possible importation and subsequent spread via local transmission. Overall, cluster analysis showed possible importation of the virus mainly from Europe and Africa.

257

267

Figure 5. Maximum likelihood tree of 1,313 SARS-CoV-2 genomes sampled in The Gambia between March
 2020 – January 2022. Clades are labelled A through H with lineages or variants associated with each clade
 shown.

Figure 6. Maximum likelihood tree of the closest SARS-CoV-2 genomes to those sampled from The Gambia between March 2020 – January 2022. Each tree represents a specific wave, waves 1 - 4 (labelled a - d) with branch tips representing the country where closest global sequences were sampled from. All trees were rooted on the SARS-CoV-2 reference genome (GenBank accession number MN908947·3).

Discussion

268 269 We have presented a comprehensive description of the genomic epidemiology of the SARS-CoV-2 epidemic in 270 The Gambia. Following the first detected case diagnosed in March 2020, there were four different waves, two 271 waves in the rainy season and two in the dry season, each lasting between 4 weeks and 4 months. All waves 272 ended despite the lack of major restrictions to stop the epidemic. New waves occurred when new variants and/or 273 lineages entered the country, generally those already prevalent in Europe, and other countries at the time. 274 Phylogenetic clustering revealed high local transmission during the summer (first and third wave) with possibly 275 more imported cases during the second wave. 276

277 278 Both the first and the third waves were short, lasting less than three months but intense. They occurred during the rainy seasons and started almost at the same time in 2020 (B·1 416 lineage) and 2021 (Delta variant), when 279 humidity is more than 80% and maximum temperatures above 33°C. ³¹ The effect of temperature and humidity 280 on SARS-CoV-2 tranmission is unclear, ^{51,52} but this observation is consistent with transmission dynamics 281 described for other respiratory viral pathogens for which the rainy season coincides with the period of highest 282 incidence.⁵³ Even if the virus infectivity is not modified by climate parameters, people are more likely to spend 283 more time indoors during heavy rains and this may increase the likelihood of household transmission. ³³ The 284 initial hypothesis that lower temperatures could be associated with higher SARS-CoV-2 transmission ⁵⁴ with 285 increased susceptibility to infection due to irritation of nasal mucosa especially in temperate regions ⁵⁵ does not 286 fit with the occurrence of the second wave in The Gambia. Despite the introduction of very transmissible 287 variants, Alpha and Eta. 56, the epidemic peak was lower but lasted longer in this second wave when compared 288 to the first and third waves.

289

290 In contrast with Senegal and Burkina Faso, where a sharp increase in cases was seen just after the first few 291 confirmed cases ³³, the first epidemic wave in The Gambia did not immediately followed the diagnosis of the 292 first case. This is probably due to the declaration of the state of emergency closing The Gambia's air and land 293 entry points. Although genomic analysis from the first five cases in The Gambia confirmed importation of the 294 virus from Europe and Asia³³, the close genomic relatedness of Gambian sequences during the first wave with 295 viruses sampled in Senegal suggests that importation of the new strains occurred mainly by land as air enforced 296 travel restrictions lasted until the end of the first wave. Enforcing land border closures is also materially more 297 difficult than air travel restrictions. These genomic observations are epidemiologically probable as the sudden 298 increase in the number of cases in July 2020 coincided with the preparations for the major Muslim feast of Eid-299 al Adha, where an inflow of undetected cases through the porous Senegal-Gambia border may have occurred. 300 This concurs with the high prevalence of the so-called Senegal/Gambia lineage or B·1·416¹¹, which also 301 dominated the first wave in Senegal (March – August 2020).⁵⁷ This lineage is characterised by few mutations 302 across the genome, the most notable being the D614G on the spike protein gene linked to increased viral load 303 and high transmission, but not to increased disease severity.⁵⁸ So far, The Gambia and Senegal have reported the 304 highest number of B 1 416 cases.²⁶ Other African countries such as Burkina Faso and Morocco have also reported this lineage, but only two genomes each.²⁶ The low number of isolates of this lineage in other West 305 306 African or neighbouring countries could be due to lack of sequencing facilities in the region, emphasising the 307 importance of genomic sequencing for monitoring the evolution of the virus during an epidemic. Nevertheless, 308 South Africa sequenced more genomes than any other country in Africa²⁶ and did not identify this lineage. 309 Importantly, high clustering among Gambian sequences suggests high local transmission, mostly associated

310 with B·1·416 lineage. Consistent with our findings, importations of SARS-CoV-2 within Africa increased as 311 more cases were reported in Africa due to high local transmission. ⁵⁹ On the other hand, phylogenetic analysis of 312 Gambian genomes along with global sequences also showed relatedness with samples from Europe.

313

314 The second wave in The Gambia was probably favoured by the reopening of borders and relaxation of 315 lockdown measures in September 2020, and the influx of tourists from Europe and Africa during and 316 immediately after Christmas. The introduction of more lineages/variants in the country and indeed the clustering 317 of Gambian sequences with those from different parts of the world further confirms the effect of 318 travelling/tourism in transmission.⁶⁰ This wave was predominantly caused by two variants, Alpha, and Eta and 319 B·1·1·420 lineage. In contrast to reports from the UK and elsewhere¹⁶, the Alpha variant, although more 320 transmissible than pre-existing lineages, was associated with slightly lower number of sequenced cases during 321 the wave compared to Eta and B·1·1·420 lineage. A similar pattern was also reported in Nigeria where Alpha 322 and Eta co-circulated.⁶¹ Using mathematical modelling, Zhao and his colleagues suggested that Eta seemed 323 slightly more infectious than Alpha in Nigeria. Phylogenetic analysis showed high clustering of most of the 324 Alpha Gambia samples with samples from Germany, Spain, England as well as Ghana while Eta samples 325 clustered more with UK, Germany and Belgium. The B·1·1·420 lineage, reported to originate from Senegal ⁵⁷, 326 dominated the second wave in both countries. Overall, sequences from The Gambia clustered closely mainly 327 with sequences from Senegal and from Europe.

328

329 The third wave was caused by the more contagious Delta variant and its subtypes, resulting in higher number of 330 daily cases.⁶² In The Gambia, the peak was in July-August 2021 and was associated with higher mortality, 331 similar to other African countries, with deaths peaking during the week of 19 July 2021.⁶² According to our 332 phylogenetic analysis, most of the Gambian Delta sequences were highly related, suggesting high local 333 transmission. Few genomes clustered with sequences from Europe, Africa (Senegal only), Asia as well as North 334 America suggesting possible importation from these countries. In contrast to other circulating variants, Delta 335 evolved into two distinct clades, suggesting that the variant may have gone through positive selection due to the 336 acquisition of many mutations. This divergence was driven by AY-34-1 sub-lineage forming a single cluster 337 associated with more than half (54%) of the samples sequenced in The Gambia during this wave. High 338 prevalence of this lineage was reported in Senegal during the same period, with only few seen in other parts of 339 Africa ^{26,63}, highlighting again the proximity of these two countries. 340

The fourth wave caused by the Omicron variant was characterised by the sharpest increase in cases to date followed by a drastic decrease in less than a month. ⁶⁴ In South Africa, where Omicron was first detected in November 2021, the daily cases peaked in December and quickly after, declined.⁶⁵ A similar trend was seen in many other countries, including Eastern and Central Africa, and Europe. ⁶⁵ The rapid decline of cases may be due to the high infectivity of the virus (Omicron) leading to high percentage of infection in the population, subsequently decreasing the number of susceptible individuals. ⁶⁶

348 The fourth wave coincided with the Gambian tourist season (October - February) when fully vaccinated 349 individuals were allowed into the country, without needing a negative PCR result. This probably resulted in new 350 introductions followed by local transmission. As expected, phylogenetic analysis showed more Gambian 351 omicron sequences clustering with those from Europe, the origin of more than half of all tourist. ⁶⁷ In addition, 352 BA:1.1 subtypes, associated with high infection rates in Europe and North America was the most dominant 353 subtype in the Gambia. The omicron variant encodes more than 30 amino acid substitutions in the spike protein, 354 enhancing its ability to evade the immune system 52 and thus spreading faster and more likely to re-infect 355 previously infected and/or vaccinated individuals. A significant number of cases may have been missed as the 356 requirement for a negative PCR for outgoing travellers, who represent almost 90% of the overall SARS-CoV-2 357 cases, was waived. Interestingly, Delta variant subtypes were still circulating during the fourth wave, most 358 notably the AY 43 sub-lineage reported in more than 130 countries and 54 US states¹¹.

359

360 Our study has some limitations to be considered when interpreting the results. Systematic testing of individuals 361 with flu-like symptoms in The Gambia has been limited throughout the pandemic and therefore most sequences 362 were obtained from samples collected from asymptomatic outgoing travellers. There is probably a substantial 363 under-estimation of the overall case numbers and some lineages circulating in the country may have been 364 missed. Most samples were collected at the coast and lineages circulating exclusively upcountry may have been 365 missed. The inference of higher transmission and severity of the third wave, mainly Delta variant, although 366 expected as per international data may have been biased by a potential change in behaviour towards testing sick 367 individuals (severity) or the increased number of outgoing travellers coinciding with the summer months of 368 2021 when international flights were open. Most of the detected cases in The Gambia were outgoing travellers.

369 In addition, most positive samples failed our quality checks and hence were either not sequenced or failed 370 analysis.

371

372 Our analysis of SARS-CoV-2 epidemic in The Gambia shows an annual trend, with higher transmission during 373 the rainy season, in line with transmission patterns of other respiratory viruses. Phylogenetic analysis confirmed 374 multiple introductions of the virus from different parts of the world, mostly from Senegal during the first wave, 375 and subsequently from Europe and Asia. As expected, most of the lineages described were imported, later 376 spreading locally. New waves coincided with the introduction of new lineages, highlighting the importance of 377 implementing well-structured and active genomic surveillance at national level to detect and monitor emerging 378 and circulating variants. In-depth analysis using phylodynamic inference could elucidate the transmission 379 dynamics and importation patterns of the virus over the course of this and future pandemics.

380

381 Declaration of interest

382 The authors declare no conflict of interest

383384 Acknowledgement

385 The genomic COVID-19 surveillance in The Gambia is conducted by Ministry of Health (National public health 386 lab) in collaboration with Medical Research Council Unit The Gambia at the London School of Hygiene & 387 Tropical Medicine (MRCG at LSHTM). The authors wisht to acknowledge the support of all staff at National 388 Public Health laboratory, Directorate of Planning and Information and Epidemiology and Disease Control. We 389 would also thank the lab team at MRCG at LSHTM, field workers, nurses, and data management. AK wish to 390 thank Dr Andrew Page and team at the Quadram Institute for their support. We are grateful to all participants 391 whose samples were used in this study. Finally, we thank all laboratories that deposited their sequences in 392 GISAID for public access.

- 393
- 394 395

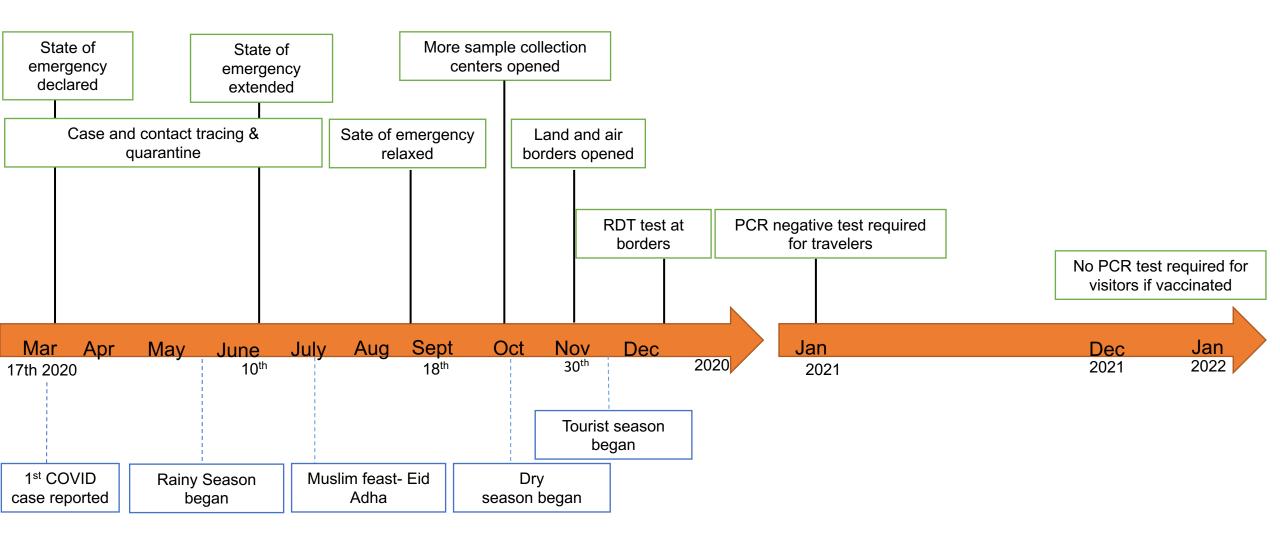
396 <u>Reference:</u>

397 Worldometer. COVID Live - Coronavirus Statistics - Worldometer [Internet]. Worldometer. 2022 [cited 1. 398 2022 Mar 15]. Available from: https://www.worldometers.info/coronavirus/ 399 2. World Health Organization. WHO Coronavirus Disease (COVID-19) Dashboard With Vaccination Data 400 WHO Coronavirus (COVID-19) Dashboard With Vaccination Data [Internet]. World Health 401 Organization. 2021 [cited 2022 Feb 15]. p. 1–5. Available from: https://covid19.who.int/ 402 3. Population of Africa [Internet]. 2020 [cited 2022 Feb 28]. Available from: 403 https://www.worldometers.info/world-population/africa-population/ 404 4. Martinez-Alvarez M, Jarde A, Usuf E, Brotherton H, Bittaye M, Samateh AL, et al. COVID-19 405 pandemic in west Africa. Lancet Glob Heal [Internet]. 2020;8(5):e631-2. Available from: 406 http://dx.doi.org/10.1016/S2214-109X(20)30123-6 407 5. Uyoga S, Adetifa IMO, Karanja HK, Nyagwange J, Tuju J, Wanjiku P, et al. Seroprevalence of anti-408 SARS-CoV-2 IgG antibodies in Kenyan blood donors. Science (80-) [Internet]. 2021;371(6524):79-82. 409 Available from: https://www.science.org 410 6. Chibwana MG, Jere KC, Kamng'ona R, Mandolo J, Katunga-Phiri V, Tembo D, et al. High SARS-411 CoV-2 seroprevalence in health care workers but relatively low numbers of deaths in urban Malawi. 412 Wellcome Open Res [Internet]. 2021 [cited 2022 May 26];5:1-15. Available from: 413 https://pubmed.ncbi.nlm.nih.gov/32766597/ 414 7. Shaw JA, Meiring M, Cummins T, Chegou NN, Claassen C, Du Plessis N, et al. Higher SARS-CoV-2 415 seroprevalence in workers with lower socioeconomic status in Cape Town, South Africa. PLoS One 416 [Internet]. 2021 Feb 1 [cited 2022 Mar 9];16(2 February):e0247852. Available from: 417 https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0247852 418 8. Chen J, Lu H. New Challenges to Fighting COVID-19: Virus Variants, Potential Vaccines, and 419 Development of Antivirals. Biosci Trends [Internet]. 2021;15(2):126-8. Available from: 420 www.biosciencetrends.com 421 9. Corey L, Beyrer C, Cohen MS, Michael NL, Bedford T, Rolland M. SARS-CoV-2 Variants in Patients 422 with Immunosuppression. N Engl J Med. 2021;385(6):562-6. 423 10. World Health Organization. Tracking SARS-CoV-2 variants [Internet]. Who. 2021 [cited 2022 Feb 15]. 424 p. https://www.who.int/en/activities/tracking-SARS-Co. Available from: 425 https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/ 426 11. O'Toole Á, Hill V, Pybus OG, Watts A, Bogoch II, Khan K, et al. Tracking the international spread of 427 SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2 with grinch. Wellcome Open Res [Internet]. 2021 428 Sep 17 [cited 2022 Feb 15];6:121. Available from: /pmc/articles/PMC8176267/

429	12.	Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, et al. Emergence and
430		rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage
431		with multiple spike mutations in South Africa. medRxiv. 2020;
432	13.	Lin L, Liu Y, Tang X, He D. The Disease Severity and Clinical Outcomes of the SARS-CoV-2 Variants
433		of Concern. Front Public Heal. 2021 Nov 30;9:1929.
434	14.	Becerra-Flores M, Cardozo T. SARS-CoV-2 viral spike G614 mutation exhibits higher case fatality rate.
435		Int J Clin Pract. 2020 Aug 1;74(8).
436	15.	Thomson EC, Rosen LE, Shepherd JG, Spreafico R, da Silva Filipe A, Wojcechowskyj JA, et al.
437		Circulating SARS-CoV-2 spike N439K variants maintain fitness while evading antibody-mediated
438		immunity. Cell [Internet]. 2021 Mar 4 [cited 2022 Feb 15];184(5):1171-1187.e20. Available from:
439		http://www.cell.com/article/S0092867421000805/fulltext
440	16.	Angius F, Pala G, Manzin A. SARS-CoV-2 and Its Variants: The Pandemic of Unvaccinated. Front
441		Microbiol. 2021 Sep 24;12:2695.
442	17.	Davies NG, Abbott S, Barnard RC, Jarvis CI, Kucharski AJ, Munday JD, et al. Estimated
443		transmissibility and impact of SARS-CoV-2 lineage B.1.1.7 in England. Science (80-) [Internet]. 2021
444		Apr 9 [cited 2022 Feb 15];372(6538). Available from:
445		https://www.science.org/doi/abs/10.1126/science.abg3055
446	18.	World Healh Organization. WHO Regional Office for Africa - WHO Regional Office for Africa
447		[Internet]. World Health Organization. 2021 [cited 2022 Feb 15]. Available from:
448		https://www.afro.who.int/news/covid-19-variants-prolong-africas-pandemic-wave
449	19.	Karim SSA, Karim QA. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic
450		[Internet]. Vol. 398, The Lancet. Elsevier B.V.; 2021 [cited 2022 May 24]. p. 2126-8. Available from:
451		http://www.thelancet.com/article/S0140673621027586/fulltext
452	20.	Bryant J, Chewapreecha C, Bentley SD. Developing insights into the mechanisms of evolution of
453		bacterial pathogens from whole-genome sequences [Internet]. Vol. 7, Future Microbiology. 2012 [cited
454		2022 Feb 15]. p. 1283–96. Available from:
455		https://www.futuremedicine.com/doi/epdf/10.2217/fmb.12.108?src=getftr
456	21.	Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019
457		novel coronavirus: implications for virus origins and receptor binding. Lancet. 2020 Feb
458		22;395(10224):565–74.
459	22.	Alm E, Broberg EK, Connor T, Hodcroft EB, Komissarov AB, Maurer-Stroh S, et al. Geographical and
460		temporal distribution of SARS-CoV-2 clades in the WHO European Region, January to June 2020.
461		Eurosurveillance [Internet]. 2020 Aug 13 [cited 2022 Feb 15];25(32):2001410. Available from:
462		https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.32.2001410
463	23.	Page AJ, Mather AE, Le-Viet T, Meader EJ, Alikhan NF, Kay GL, et al. Large-scale sequencing of sars-
464		cov-2 genomes from one region allows detailed epidemiology and enables local outbreak management.
465		Microb Genomics [Internet]. 2021 [cited 2022 Mar 10];7(6). Available from: https://www.
466	24.	Gutierrez B, Márquez S, Prado-Vivar B, Becerra-Wong M, Guadalupe JJ, da Silva Candido D, et al.
467		Genomic epidemiology of SARS-CoV-2 transmission lineages in Ecuador. Virus Evol. 2021;7(2).
468	25.	Seib KL, Dougan G, Rappuoli R. The key role of genomics in modern vaccine and drug design for
469		emerging infectious diseases [Internet]. Vol. 5, PLoS Genetics. 2009. Available from:
470		http://www.1000genomes.org/
471	26.	GISAID. GISAID Initiative [Internet]. Vol. 2008, Advances in Virus Research. 2020 [cited 2022 Feb
472		15]. p. 1–7. Available from: https://www.gisaid.org/
473	27.	Kanteh A, Manneh J, Jabang S, Kujabi MA, Sanyang B, Oboh MA, et al. Origin of imported SARS-
474		CoV-2 strains in the Gambia identified from whole genome sequences. PLoS One [Internet]. 2021;16(8
475		August). Available from: https://doi.org/10.1371/journal.pone.0241942
476	28.	COVID-19 - MINISTRY OF HEALTH, THE GAMBIA [Internet]. 2022 [cited 2022 Feb 28]. Available
477		from: https://www.moh.gov.gm/
478	29.	Sanyang B, Kanteh A, Usuf E, Nadjm B, Jarju S, Bah A, et al. COVID-19 reinfections in The Gambia
479		by phylogenetically distinct SARS-CoV-2 variants-first two confirmed events in west Africa
480		[Internet]. Vol. 9, The Lancet Global Health. Elsevier Ltd; 2021 [cited 2022 Feb 15]. p. e905–7.
481		Available from: https://www.
482	30.	Demographics G. Gambia Demographics 2020 (Population, Age, Sex, Trends) - Worldometer
483		[Internet]. 2020 [cited 2022 Mar 10]. Available from:
484		https://www.worldometers.info/demographics/gambia-demographics/
485	31.	Access Gambia. Gambia Climate [Internet]. 2021 [cited 2022 Mar 8]. Available from:
486		https://www.accessgambia.com/information/climate-weather.html
487	32.	Health Care System & Centres in Gambia. Health Care System & Centres in Gambia [Internet]. 2020
488		[cited 2022 Mar 10]. Available from: https://www.accessgambia.com/information/health-care.html

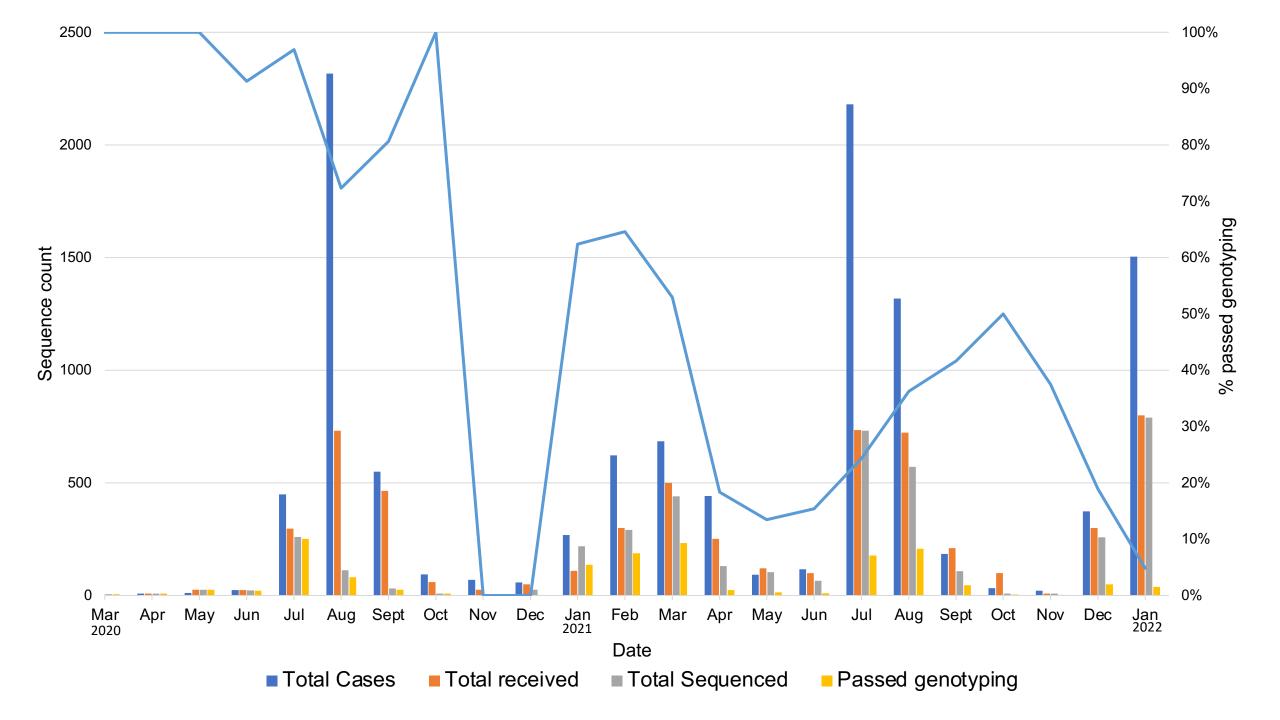
400		
489	33.	Abatan B, Agboghoroma O, Akemoke F, Antonio M, Awokola B, Bittaye M, et al. Intense and mild
490	24	first epidemic wave of Coronavirus disease, the Gambia. Emerg Infect Dis. 2021;27(8):2064–72.
491	34.	COPAN Diagnostics Inc. COPAN Diagnostics Inc. Collection and Transport Systems Leader
492 493	25	[Internet]. 2022 [cited 2022 Mar 10]. Available from: https://www.copanusa.com/
495 494	35.	The Gambia Health Information System. The Gambia Health Information System. 2022 [cited 2022]
494	36.	May 24]; Available from: https://covid19.moh.gm/tracker/dhis-web-commons/security/login.action#/ World Health Organization. Laboratory testing for coronavirus disease 2019 (COVID-19) in suspected
496	50.	human cases: interim guidance, 2 March 2020 [Internet]. World Health Organization. 2020 [cited 2022
497		Feb 16]. Available from: https://apps.who.int/iris/handle/10665/331329
498	37.	REDCap. REDCap [Internet]. 2021 [cited 2022 Feb 16]. Available from: https://www.project-
499	57.	redcap.org/
500	38.	DHIS2. Home - DHIS2 [Internet]. 2021 [cited 2022 Feb 16]. Available from: https://dhis2.org/
501	39.	Quick J, Barnes K. nCoV-2019 sequencing protocol v3 (LoCost). [Internet]. Vol. 3, Protocols.Io. 2020
502		[cited 2022 Feb 16]. Available from: https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-
503		locost-bh42j8ye?version_warning=no
504	40.	Freed N, Silander O, Freed N. SARS-CoV2 genome sequencing protocol (1200bp amplicon "midnight"
505		primer set, using Nanopore Rapid kit). 2021 [cited 2022 Feb 16];2:1–14. Available from:
506		https://www.protocols.io/view/sars-cov2-genome-sequencing-protocol-1200bp-amplic-bwyppfvn
507	41.	Davis JJ, Long SW, Christensen PA, Olsen RJ, Olson R, Shukla M, et al. Analysis of the ARTIC
508		Version 3 and Version 4 SARS-CoV-2 Primers and Their Impact on the Detection of the G142D Amino
509		Acid Substitution in the Spike Protein. [cited 2022 Feb 16]; Available from:
510		https://community.artic.network/t/sars
511	42.	Gautreau I. NEBNext® Varskip Short ARTIC SARS-CoV-2 RT-PCR Module E7626 [Internet]. 2021
512		[cited 2022 Feb 16]. Available from: https://www.protocols.io/view/nebnext-varskip-short-artic-sars-
513	10	cov-2-rt-pcr-modu-byfdpti6
514	43.	Xp AA. Instructions for Use Instructions for Use - EN. 第13回衛生工学シンポジウム [Internet].
515		2018 [cited 2022 Feb 16];3.6(August):4–4. Available from:
516 517	4.4	www.beckmancoulter.com/customersupport/support.
517	44.	Protocol N. Introduction to the protocol Overview of the protocol Rapid Barcoding Sequencing (SQK- PRK004) 2021/05286272011 10
519	45.	RBK004). 2021;(05386273):1–19. Josh Quick NL and AC. ARTIC Network [Internet]. 2019 [cited 2022 May 23]. p.
520	45.	Https://artic.network/ncov-2019. Available from: https://artic.network/ncov-2019/ncov2019-
521		bioinformatics-sop.html
522	46.	O'Toole Á, Scher E, Underwood A, Jackson B, Hill V, McCrone JT, et al. Assignment of
523		epidemiological lineages in an emerging pandemic using the pangolin tool. Virus Evol. 2021;7(2).
524	47.	Aksamentov I, Roemer C, Hodcroft E, Neher R. Nextclade: clade assignment, mutation calling and
525		quality control for viral genomes. J Open Source Softw [Internet]. 2021 Nov 30 [cited 2022 Feb
526		17];6(67):3773. Available from: https://joss.theoj.org/papers/10.21105/joss.03773
527	48.	Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: Improvements in
528		performance and usability. Mol Biol Evol [Internet]. 2013;30(4):772-80. Available from:
529		https://academic.oup.com/mbe/article/30/4/772/1073398
530	49.	Letunic I, Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of
531	~0	phylogenetic and other trees. Nucleic Acids Res [Internet]. 2016;44. Available from: http://itol.embl.de
532	50.	Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing
533		and sharing data for genomic epidemiology and phylogeography. Microb genomics [Internet]. 2016 Nov
534 535	51	1 [cited 2022 Apr 12];2(11):e000093. Available from: /pmc/articles/PMC5320705/
536	51.	Wu Y, Jing W, Liu J, Ma Q, Yuan J, Wang Y, et al. Effects of temperature and humidity on the daily new cases and new deaths of COVID-19 in 166 countries. Sci Total Environ [Internet]. 2020 Aug 10
537		[cited 2022 Mar 10];729. Available from: https://doi.org/10.1016/j.scitotenv.2020.139051
538	52.	Mecenas P, da Rosa Moreira Bastos RT, Rosário Vallinoto AC, Normando D. Effects of temperature
539	52.	and humidity on the spread of COVID-19: A systematic review. PLoS One [Internet]. 2020;15(9
540		September). Available from: https://doi.org/10.1371/journal.pone.0238339
541	53.	Jarju S, Greenhalgh K, Wathuo M, Banda M, Camara B, Mendy S, et al. Viral Etiology, Clinical
542		Features and Antibiotic Use in Children <5 Years of Age in the Gambia Presenting with Influenza-like
543		Illness. Pediatr Infect Dis J [Internet]. 2020 [cited 2022 Mar 8];925-30. Available from:
544		https://journals.lww.com/pidj/Fulltext/2020/10000/Viral_Etiology,_Clinical_Features_and_Antibiotic.1
545		3.aspx
546	54.	Moriyama M, Hugentobler WJ, Iwasaki A. Annual review of virology seasonality of respiratory viral
547		infections. Annu Rev Virol [Internet]. 2020;7:83–101. Available from: https://doi.org/10.1146/annurev-
548		virology-012420-

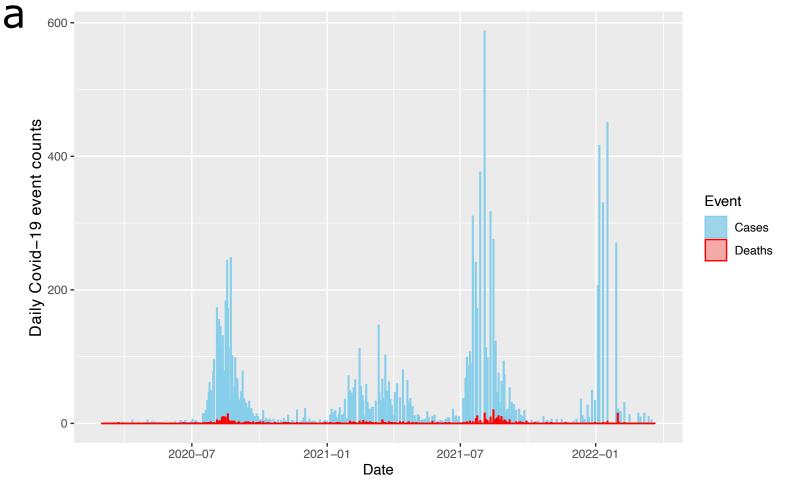
549 550 551	55.	Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell [Internet]. 2020 Aug 20 [cited 2022 Mar 9];182(4):812-827.e19. Available from:
552		https://doi.org/10.1016/j.cell.2020.06.043
553	56.	Zhao Y, Huang J, Zhang L, Chen S, Gao J, Jiao H. The global transmission of new coronavirus variants.
554 555	57	Environ Res. 2022 Apr 15;206.
555 556 557	57.	Perez LJ, Orf GS, Berg MG, Rodgers MA, Meyer T V, Mohaimani A, et al. The early SARS-CoV-2 epidemic in Senegal was driven by the local emergence of B.1.416 and the introduction of B.1.1.420
557 558	50	from Europe. Virus Evol. 2022;8(1):1–12.
558 559	58.	Korber B, Fischer WM, Gnanakaran S, Yoon H, Theiler J, Abfalterer W, et al. Tracking Changes in
560		SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID-19 Virus. Cell [Internet]. 2020 Aug 20 [cited 2022 Feb 15];182(4):812-827.e19. Available from:
561		http://www.cell.com/article/S0092867420308205/fulltext
562	59.	Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, et al. A year of genomic
563	57.	surveillance reveals how the SARS-CoV-2 pandemic unfolded in Africa. Science (80-) [Internet]. 2021
564		Oct 22 [cited 2022 Mar 8];374(6566):423–31. Available from: https://www.science.org
565	60.	Russell TW, Wu JT, Clifford S, Edmunds WJ, Kucharski AJ, Jit M. Effect of internationally imported
566		cases on internal spread of COVID-19: a mathematical modelling study. Lancet Public Heal. 2021 Jan
567		1;6(1):e12–20.
568	61.	Zhao S, Musa SS, Chong MK, Ran J, Javanbakht M, Han L, et al. Correspondence to: The co-
569		circulating transmission dynamics of SARS-CoV-2 Alpha and Eta variants in Nigeria: A retrospective
570		modeling study of COVID-19. 2021 [cited 2022 Mar 29];11:5028. Available from: www.jogh.org
571	62.	WHO. COVID-19 variants prolong Africa's pandemic wave WHO Regional Office for Africa
572		[Internet]. 2021 [cited 2022 Mar 10]. Available from: https://www.afro.who.int/news/covid-19-variants-
573		prolong-africas-pandemic-wave
574	63.	O'Toole Á, Kraemer MUG, Hill V, Pybus OG, Watts A, Bogoch II, et al. Tracking the international
575		spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2. Wellcome Open Res [Internet]. 2021
576		May 19 [cited 2022 Mar 12];6:121. Available from: https://wellcomeopenresearch.org/articles/6-121
577	64.	World Health Organization. Enhancing response to Omicron SARS-CoV-2 variant : Technical brief
578		and priority actions for Member States (https://www.who.int/publications/m/item/enhancing-readiness-
579		for-omicron-(b.1.1.529)-technical-brief-and-priority-actions-for-member-states). 2022;(January):1–28.
580	65.	WHO. Omicron-fuelled COVID-19 surge in Africa plateaus [Internet]. 2022 [cited 2022 Mar 10].
581		Available from: https://www.afro.who.int/news/omicron-fuelled-covid-19-surge-africa-plateaus
582	66.	Vox. Omicron surge: Why Covid-19 cases often spike sharply and fall rapidly - Vox [Internet]. 2022
583 584		[cited 2022 Mar 31]. Available from: https://www.vox.com/22905020/omicron-wave-surge-covid-19-
584 585	67	cases-vaccines
585 586	67.	Gambia Tourism Statistics. Gambia Tourism Statistics [Internet]. 2022 [cited 2022 May 24]. Available
580 587		from: https://www.accessgambia.com/information/tourism-statistics.html
507		

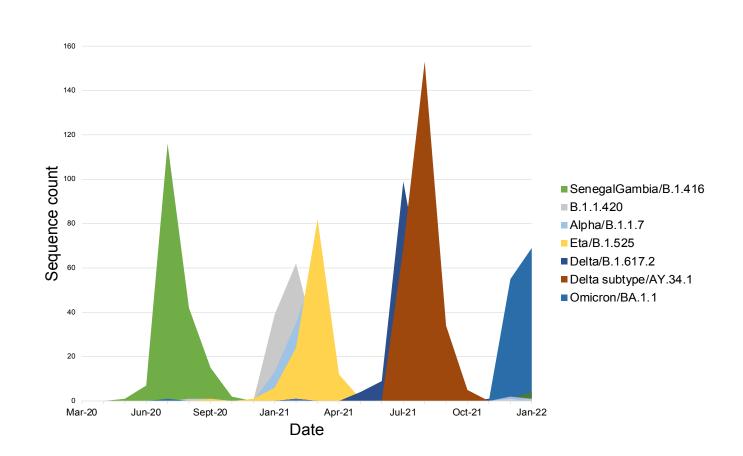


Measures put in place during the pandemic

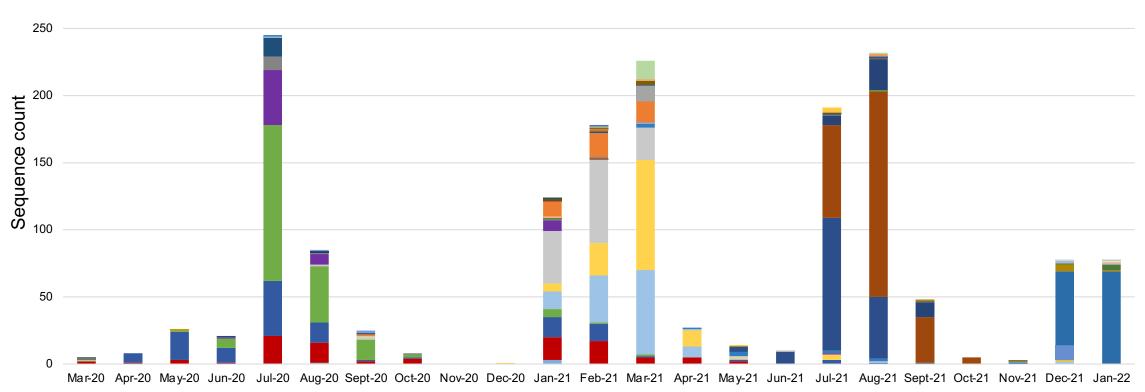
Events







b



Date

B	■ B.1	■ B.1.1	B.1.416/Senegambia linea	ge 🔳 Alpha/ B. 1. 1.7	Eta/B.1.525
B.1.1.420	B.1.1.466	■ B.1.1.242	B.1.1.10	B.1.1.359	B.1.1.271
B.1.1.318	B.1.1.8	B.1.1.362	B.1.1.368	■ B.1.177	A.27
A.29	A	Delta/B.1.617.2	Delta/AY.34.1	Delta/AY. 122	Delta/AY. 133
Delta/AY.22	Delta/AY. 34	Delta/AY.34.1.1	Delta/AY.36	Delta/AY.4	Delta/AY.42
Delta/AY.43	Delta/AY.46	Delta/AY.5	Delta/AY.55	B.1.177.44	■ B.1.201
B.1.214.3	B.1.292	■ B.1.36.1	B.1.378	B.1.390	B.1.582
B.1.609	B.1.619	■ B.6	Omicron/BA.1.1	Omicron/BA.1	■ Omicron/BA.1.1.14
Omicron/BA.1.1.18	Omicron/BA.1.15	Omicron/BA.1.17	Omicron/BA.1.17.2	Omicron/BA.1.18	Q.4

300

