

TITLE: Emergence and spread of two SARS-CoV-2 variants of interest in Nigeria.

Idowu B. Olawoye^{1,2*}, Paul E. Oluniyi^{1,2*}, Judith U. Oguzie^{1,2*}, Jessica N. Uwanibe^{1,2*}, Adeyemi T. Kayode^{1,2*}, Testimony J. Olumade^{1,2*}, Fehintola V. Ajogbasile^{1,2*}, Edyth Parker^{3*}, Philomena E. Eromon², Priscilla Abechi^{1,2}, Tope Sobajo^{1,2}, Chinedu Ugwu^{1,2}, George Uwem^{1,2}, Femi Ayoade^{1,2}, Kazeem Akano^{1,2}, Nicholas Oyejide², Iyanuoluwa Fred-Akintunwa², Kemi Adedotun-Sulaiman², Farida Brimmo², Benjamin Adegboyega², Courage Philip², Ayomide Adeleke², Grace C. Chukwu², Ahmed I. Muhammad², Oludayo O. Ope-ewe², Shobi Otitoola², Olusola A. Ogunsanya², Mudasiru F. Saibu², Ayotunde E. Sijuwola², Grace O. Ezekiel², Oluwagboadurami G. John^{1,2}, Julie O. Akin-John^{1,2}, Oluwasemilogo O. Akinlo², Olanrewaju O. Fayemi², Testimony O. Ipaye², Deborah C. Nwodo², Abolade E. Omoniyi², Iyobosa B. Omwanghe², Christabel A. Terkuma², Johnson Okolie², Olubukola Ayo-Ale², Odia Ikponmwosa⁴, Ebo Benevolence⁴, Okonofua Naregose⁴, Akhilomen Patience⁴, Osiemi Blessing⁴, Airende Micheal⁴, Agbukor Jacqueline⁴, Aiyepada John⁴, Paulson Ebhodaghe⁴, Omiunu Racheal⁴, Esumeh Rita⁴, Giwa Rosemary⁴, Ehikhametalor Solomon⁴, Ekanem Anieno⁴, Yerumoh Edna⁴, Aire Chris⁴, Adomeh Donatus⁴, Ephraim Ogbaini⁴, Mirabeau Y. Tattfeng⁵, Hannah E. Omunakwe⁶, Mienye Bob-Manuel⁶, Rahaman Ahmed⁷, Chika Onwuamah⁷, Joseph Shaibu⁷, Azuka Okwuraiwe⁷, Anthony E. Atage⁸, Andrew Bock-Oruma⁹, Funmi Daramola¹⁰, Akinwumi Fajola¹¹, Nsikak-Abasi Ntia¹², Julie J. Ekpo¹³, Anietie Moses¹³, Worbianueri B. Moore-Igwe¹⁴, Ibrahim F. Yusuf¹⁰, Enoch O. Fakayode¹⁵, Monilade Akinola¹⁶, Ibrahim Kida¹⁷, Bamidele S. Oderinde¹⁷, Zara Wudiri¹⁸, Olufemi O. Adeyemi¹⁹, Olusola Akanbi²⁰, Anthony Ahumibe²⁰, Afolabi Akinpelu²⁰, Oyeronke Ayansola²⁰, Olajumoke Babatunde²⁰, Adesuyi Omoare²⁰, Chimaobi Chukwu²⁰, Nwando Mba²⁰, Ewean C. Omoruyi²¹, Johnson A. Adeniji²², Moses O. Adewunmi²², Oluseyi Olayinka²², Olisa Olasunkanmi²³, Olatunji Akande²³, Ifeanyi Nwafor²⁴, Matthew Ekeh²⁴, Erim Ndoma²⁴, Richard Ewah²⁴, Rosemary Duruihuoma²⁴, Augustine Abu²⁴, Elizabeth Odeh²⁴, Venatious Onyia²⁴, Kingsley C. Ojide²⁴, Sylvanus Okoro²⁵, Daniel Igwe²⁵, Kamran Khan³⁴, Anthony N. Ajayi²⁶, Ebhodaghe Ngozi Ugwu²⁷, Collins N. Ugwu²⁶, Kingsley Ukwuaja²⁶, Emeka O. Ogah²⁵, Chukwuyem Abejegah²⁸, Nelson Adedosu²⁸, Olufemi Ayodeji²⁸, Rafiu O. Isamotu²⁹, Galadima Gadzama³⁰, Brittany Petros³¹, Katherine J. Siddle³¹, Stephen Schaffner³¹, George Akpede⁴, Cyril Oshomah Eramah⁴, Marycelin Baba^{16,17}, Femi Oladiji³², Rosemary Audu⁷, Nnaemeka Ndodo²⁰, Adeola Fowotade²³, Sylvanus Okogbenin⁴, Peter Okokhere⁴, Danny Park³¹, Bronwyn Mcannis³¹, Ifedayo Adetifa²⁰, Chikwe Ihekweazu²⁰, Babatunde L. Salako⁷, Oyewale Tomori², Anise N. Happi², Onikepe A. Folarin^{1,2}, Kristian G. Andersen³, Pardis C. Sabeti^{31,33}, Christian T. Happi^{1,2,33†}

1. Department of Biological Sciences, Faculty of Natural Sciences, Redeemer's University, Ede, Osun State, Nigeria.

2. African Centre of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University, Ede, Osun State, Nigeria.

3. Department of Immunology and Microbiology, The Scripps Research Institute, La Jolla, CA, USA.

4. Irrua Specialist Teaching Hospital, Edo State, Nigeria.

5. Department of Medical Laboratory Science, Niger Delta University, Bayelsa State, Nigeria.

6. Satellite Molecular Laboratory, Rivers State University Teaching Hospital, Port Harcourt, Rivers State, Nigeria.

7. The Nigerian Institute of Medical Research, Yaba, Lagos State, Nigeria.

- 50 8. Molecular Laboratory, Regional Centre for Biotechnology and Bioresources Research,
- 51 University of Port Harcourt, Port Harcourt, Rivers State, Nigeria.
- 52 9. Family Physician, SPDC, Port Harcourt, Rivers State, Nigeria.
- 53 10. Clinical Health, SPDC, Port Harcourt, Rivers State, Nigeria.
- 54 11. Regional Community Health, SPDC, Port Harcourt, Rivers State, Nigeria.
- 55 12. Occupational Health, SPDC, Port Harcourt, Rivers State, Nigeria.
- 56 13. Department of Medical Microbiology and Parasitology, University of Uyo, Akwa Ibom State,
- 57 Nigeria.
- 58 14. Rivers State University, Port Harcourt, Rivers State, Nigeria.
- 59 15. Department of Public Health, Ministry of Health, Ilorin, Kwara State, Nigeria.
- 60 16. WHO Polio Laboratory, University of Maiduguri Teaching Hospital, Maiduguri, Borno State,
- 61 Nigeria.
- 62 17. Department of Immunology, University of Maiduguri Teaching Hospital, Maiduguri, Nigeria.
- 63 18. Department of Community Medicine, University of Maiduguri Teaching Hospital, Maiduguri,
- 64 Borno State, Nigeria.
- 65 19. Department of Medical Microbiology and Parasitology. Faculty of Basic Clinical Sciences.
- 66 College of Health Sciences, University of Ilorin. Ilorin, Kwara State, Nigeria.
- 67 20. Nigeria Centre for Disease Control, Abuja, Nigeria.
- 68 21. Medical Microbiology and Parasitology Department, College of Medicine, University of
- 69 Ibadan, Ibadan, Nigeria.
- 70 22. Department of Virology, College of Medicine, University of Ibadan, Ibadan, Nigeria.
- 71 23. Biorepository Clinical Virology Laboratory, University of Ibadan, Ibadan, Nigeria.
- 72 24. Virology Laboratory, Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Ebonyi
- 73 State, Nigeria.
- 74 25. Alex Ekwueme Federal University Teaching Hospital, Abakaliki, Nigeria.
- 75 26. Internal Medicine Department, Alex Ekwueme Federal University Teaching Hospital,
- 76 Abakaliki, Nigeria.
- 77 27. Haematology Department, Alex Ekwueme Federal University Teaching Hospital, Abakaliki,
- 78 Nigeria.
- 79 28. Federal Medical Center, Owo, Ondo State, Nigeria
- 80 29. Ministry of Health, Osun State, Nigeria.
- 81 30. Department of Medical Microbiology, University of Maiduguri Teaching Hospital, Maiduguri,
- 82 Borno State, Nigeria.
- 83 31. Broad Institute of Harvard and MIT, Cambridge, MA, USA.
- 84 32. Department of Epidemiology and Community Health, Faculty of Clinical Sciences, College of
- 85 Health Sciences, University of Ilorin, Ilorin, Nigeria.
- 86 33. Department of Immunology and Infectious Diseases, Harvard TH Chan School of Public
- 87 Health, Boston, MA, USA.
- 88 34. Bluedot, Toronto, Canada

89

90 *These authors contributed equally.

91 ‡ Correspondence (happic@run.edu.ng)

92

93

94

95

96

97

98

99

100

101

102

103 **Abstract**

104

105 Identifying the dissemination patterns and impacts of a virus of economic or health importance
106 during a pandemic is crucial, as it informs the public on policies for containment in order to
107 reduce the spread of the virus. In this study, we integrated genomic and travel data to
108 investigate the emergence and spread of the B.1.1.318 and B.1.525 variants of interest in
109 Nigeria and the wider Africa region. By integrating travel data and phylogeographic
110 reconstructions, we find that these two variants that arose during the second wave emerged
111 from within Africa, with the B.1.525 from Nigeria, and then spread to other parts of the world.
112 Our results show how regional connectivity in downsampled regions like Africa can often
113 influence virus transmissions between neighbouring countries. Our findings demonstrate the
114 power of genomic analysis when combined with mobility and epidemiological data to identify the
115 drivers of transmission in the region, generating actionable information for public health decision
116 makers in the region.

117 Introduction

118 With a population of over 200 million, Nigeria is the most populous country in Africa¹. Since
119 the first report of SARS-CoV-2 in Nigeria on the 27th of February 2020, the cumulative
120 number of confirmed COVID-19 cases in Nigeria has risen to more than 265 000 as of mid-
121 September 2022². However, this burden is extremely low relative to SARS-Cov-2 infections
122 in the rest of the world. Adjusting for population size, there have been 129 cases per 100
123 000 people in Nigeria as of mid-September, compared to more than 34 700 and 28 700 in
124 the UK and USA, and 2340 and 711 in Indonesia and Pakistan, respectively. Concurrently,
125 Nigeria has also only had a little over 3000 reported deaths (2 per 100 000), compared to the
126 UK and USA, which had 303 and 315 cumulative deaths per 100 000 people². This relatively
127 low but heterogeneous incidence in Nigeria and across the wider African region has been
128 the cause of persistent speculation, including on the putative central role of case
129 underascertainment^{3,4,5}. The underlying cause remains understudied and is likely
130 multifactorial, with many other factors speculated to contribute, including skewed age
131 structures of populations, more restricted human mobility in certain regions, host genetics,
132 environmental factors, and potential pre-existing population immunity to related viruses^{6,7}.
133 Before the emergence of globally sweeping variants such as Delta and Omicron, it was also
134 unclear if the specific genetic diversity circulating in the African region may have contributed
135 to the heterogeneous incidence and mortality observed^{8,9,10}. It is therefore important to
136 characterise the genomic epidemiology of SARS-CoV-2 in Nigeria over the course of the
137 pandemic, to improve our understanding of what variants emerged to dominate the different
138 epidemic waves. It is also important to improve our understanding of the drivers of
139 transmission in the region, as this remains understudied in Africa and findings from other
140 regions may not be generalizable^{9,11}. Nigeria is highly connected to its neighbouring
141 countries and the wider region, potentially acting as a dominant source of transmission via
142 the high volume of movement across land as well as air borders¹².

143 Several SARS-CoV-2 variants, which have a constellation of mutations that are biologically
144 significant to the virus, emerged during the pandemic^{9,13,14}. The B.1.525 (Eta) and B.1.1.318
145 variants of interest (VOIs) dominantly co-circulated with the Alpha variant during the second
146 wave in Nigeria from December 2020 to March 2021². B.1.525 and B.1.1.318 were
147 suggested to have emerged in Nigeria based on epidemiological reports from travellers in
148 countries such as the UK, India, Mauritius, Canada, and Brazil^{15,16,17,18,19}. Notably, B.1.525
149 and B.1.1.318 both showed a significant increase in the infectivity of human and monkey cell
150 lines experimentally, raising concerns of intrinsic increased transmissibility²⁰. Both of these
151 lineages also share the well-characterised E484K substitution in the Spike protein receptor-
152 binding domain (RBD), which effectively reduces antibody neutralisation²¹.

153 In this study, we examine the genomic epidemiology of SARS-CoV-2 in Nigeria from March
154 2020 to September 2021, across the first three epidemic waves. In particular, we investigate
155 the timing and origin of the emergence of the B.1.525 and B.1.1.318 VOIs. In
156 phylogeographic reconstructions, we characterized the source-sink profiles of these best-
157 sampled VOIs to better understand the bidirectional transmission dynamics of SARS-CoV-2
158 in Nigeria and the wider region. We compare our findings from our genomic data to
159 integrated travel and epidemiological data to explore the role of Nigeria's regional and
160 intercontinental connectivity in the estimated import and export dynamics.

161

162 Results

163

164 Lineage dynamics of SARS-CoV-2 in Nigeria over the first three epidemic waves

165

166 We generated a total of 1577 genomes from samples obtained between March 2020 and
167 September 2021, across the first three waves of the pandemic (Fig. 1A and B). We collected
168 samples from 25 of the 36 states and the Federal Capital Territory of Nigeria (Fig. 1C). The
169 northern states were highly undersampled relative to the southern states, with sampling highly
170 uneven across time (Fig. 1A and B).

171

172 To investigate the lineage dynamics across the different outbreak waves, we used the pangolin
173 nomenclature tool to assign all sequences to corresponding lineages. We detected more than
174 35 lineages across the first three waves (Fig. 1D). We found that a number of ancestral lineages
175 circulated during the first wave, including A, B, B.1, and B.1.1 (Fig. 1D). The onset of community
176 transmission in Nigeria was initially delayed, with the first wave initiated two months after the
177 first case detection in March 2020. A nationwide lockdown and ban on local and international air
178 travel were enforced, with the first wave ending in August 2020 after four months (Fig. 1B).
179 During the lockdown, we found that there was up to a 50% decline in activity at workplaces,
180 retail, supermarkets, and public transport stations, and up to a 25% increase in residential
181 activity (Fig. 1E).

182

183 After this strictly adhered-to lockdown, Nigeria opened its airspace to international traffic in
184 October 2020. This was shortly followed by an increase in new cases in November 2020,
185 marking the start of the second wave (Fig. 1B), which was characterised by B.1.1.7, B.1.525,
186 B.1.1.318 (VOC and VOI), and other variants such as L.3 and B.1.1.487 (Fig. 1D).

187

188 Our secondary analysis focused on the lineages (B.1.525 and B.1.1.318) that contributed to the
189 surge during the second wave and have also been shown to possess genotypic traits for
190 increased transmissibility and also increased virulence in human and animal cell lines, such as
191 the E484K spike protein RBD mutation. We excluded B.1.1.7 from our analysis as it has already
192 been well studied^{22,23}, and it was not the dominant variant in Nigeria during the second wave,
193 unlike in other places where it was reported, such as the UK and the USA^{24,25,26}. Moreso, the
194 second wave began after travel restrictions were lifted and human mobility was back to normal,
195 thus the need to investigate the emergence of B.1.1.318 and B.1.525 in Nigeria. The third wave
196 was initiated in June 2021, with the Delta variant (B.1.617.2) and its sub-lineages sweeping to
197 dominance in sampling (Fig. 1B and D). In the secondary analysis described in this work, we
198 also sought to identify the timings of emergence and number of introductions of these VOI.

It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/).

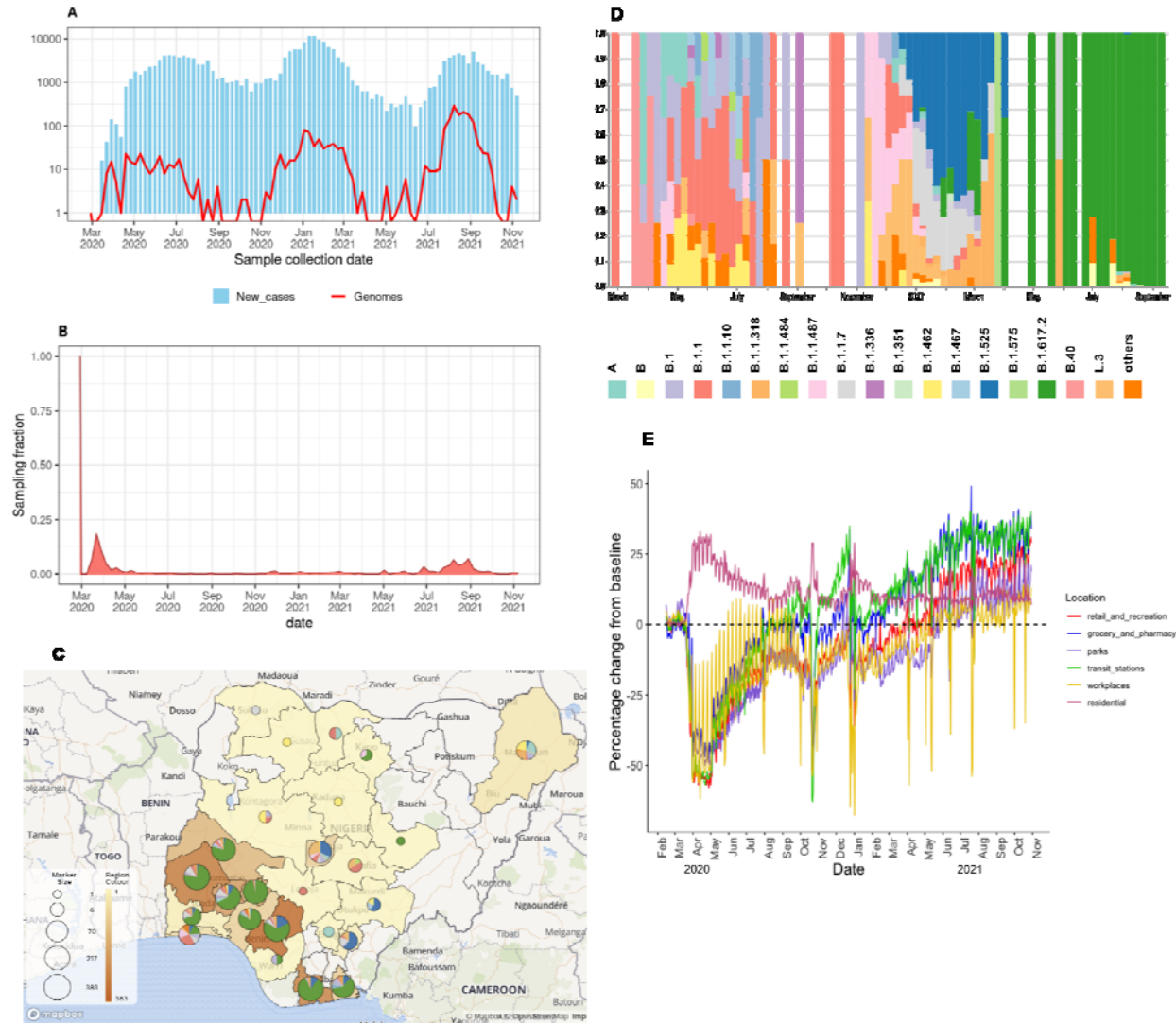


Fig. 1: Genomic epidemiology of SARS-CoV-2 in Nigeria. **A)** Reported daily news Epicurve of SARS-CoV-2 cases in Nigeria (blue bars) in Nigeria and genomes produced assembled during the first three waves (red line) with the y-axis transformed to a log-scale. **B)** Time-varying sampling fraction of genomes produced in this study per new cases **C)** Geographic distribution of sequences generated in the current study. The map of Nigeria shows the number of genomes from states across the country as region colour and marker size display lineages per state. **D)** Lineage frequency profile for the study period. The frequency of PANGO lineages across the country over a period of 80 weeks. Lineages that are not VOCs or VOIs, or that do not appear more than three times over the course of the pandemic, are grouped as "others". **E)** Human mobility data from Google for retail shops, pharmacies, parks, public transport, workplaces, and residential activity places from January 2020 to November 2021, as compared to the pre-pandemic baseline (January to February, 2021).

Regional connectivity drove the introduction of B.1.1.318 into Nigeria.

Using Bayesian phylogeographic reconstructions, we investigated the timing and origin of the B.1.1.318 emergence to better understand Nigeria's connectivity in the global SARS-CoV-2 transmission networks (see Methods). B.1.1.318 was first detected in Nigeria in Lagos State in

199
200

201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218

219 December 2020. The lineage was detected in multiple countries within and outside Africa,
220 notably resulting in a large outbreak in Mauritius¹⁹.

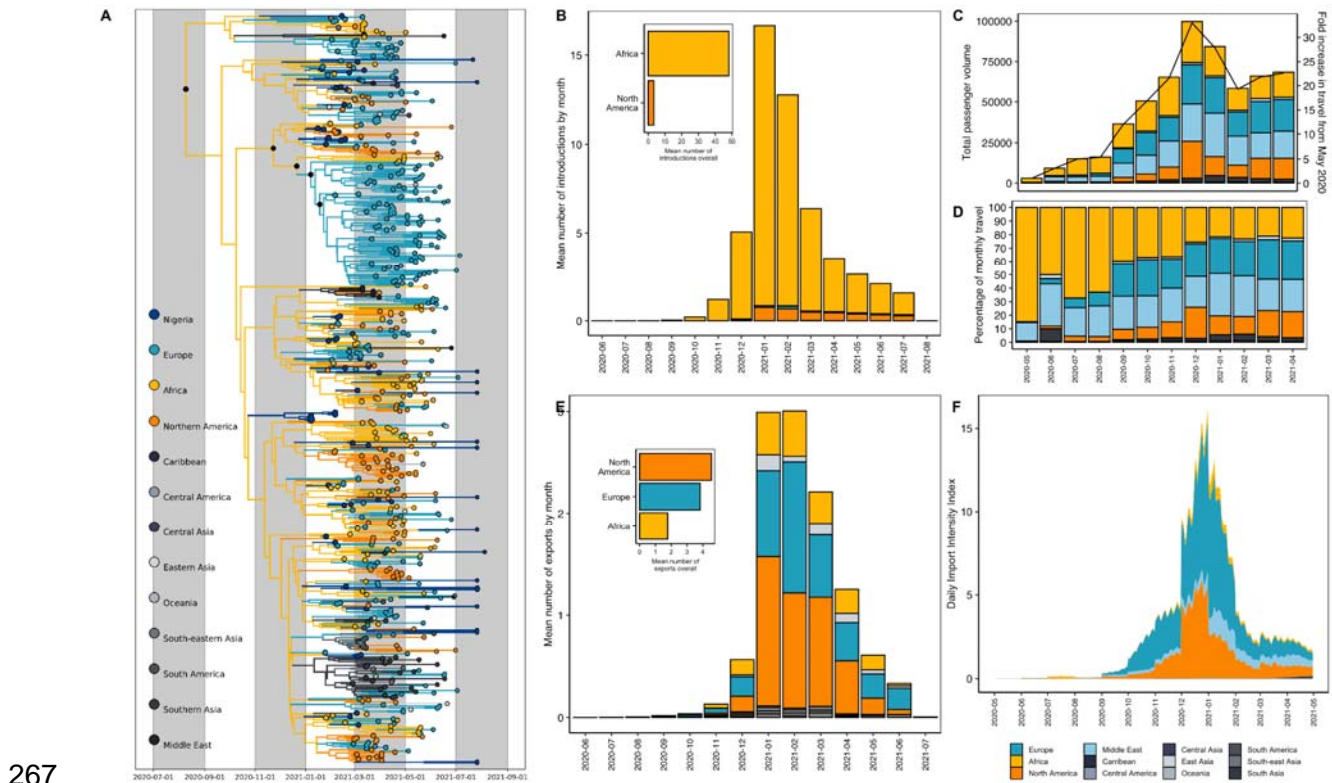
221
222 In our phylogeographic reconstruction, we found that B.1.1.318 emerged in Africa (root state
223 posterior support = 0.95-6) in early August [mean tMRCA = 5 August, 95% HPD 25 June to 20
224 September] across two replicates (Fig. 2A). We estimated that B.1.1.318 was introduced to
225 Nigeria on at least 53 independent occasions [mean introductions, 95% HPD 50-59] beginning
226 in November 2020, after travel restrictions were lifted in October (Fig. 2B). We found that the
227 majority of introductions originated from other African nations (Fig. 2B), indicating Nigeria's
228 strong connectivity to the region. However, the number and origin of introductions and exports
229 estimated from genomic data is sample dependent, with bidirectional transmission to and from
230 undersampled regions obscured by uneven sampling globally. To overcome this, we quantified
231 air travel patterns to and from Nigeria over time to better understand which countries were most
232 likely to act as transmission sources or sinks based on their connectivity to Nigeria.

233
234 Before travel restrictions were lifted in October 2020, we found that the low levels of incoming
235 air travel volume predominantly originated from other African nations, followed by the Middle
236 East (driven by the UAE) and Europe (driven by the UK) (Fig. 2C-D, SFigure 2). This suggests
237 that the introduction risk was predominantly driven by regional connectivity during the period
238 when travel restrictions were in place. We sought to investigate this further by quantifying the
239 introduction risk for all countries connected to Nigeria by air travel. We estimated the
240 introduction intensity index (III), which accounts for the number of cases in the source countries
241 as well as the travel volume to Nigeria (see Methods). Overall, we found that the introduction
242 risk from air travel was very low during the period of travel restrictions (May 2020 to October
243 2020), as incoming travellers originated largely from other African nations with low incidence
244 (Fig. 2F).

245
246 We found that the number of incoming air passengers peaked in December 2020 - January
247 2021 (Fig. 2C-D). We also observed that the number of inbound passengers from Europe, the
248 Middle East, and Northern America increased relative to other African nations after restrictions
249 were lifted, resulting in comparable volumes from Africa, Europe, and the Middle East (Fig. 2C-
250 D). The 30-fold increase in incoming passengers by December 2020 was reflected in the
251 corresponding peak in the III, attributable to increased introduction risk from Europe and North
252 America from November 2020 into early 2021 (Fig. 2F). The surge was predominantly driven by
253 the high-incidence second waves of the UK and USA, respectively (SFigure 1).

254
255 We found that the III peak but not the estimated sources of risk were consistent with the number
256 and origin of introductions of B.1.1.318 in our phylogeographic reconstructions (Fig. 2B). We
257 estimated the highest introduction risk for Europe after travel restrictions were lifted, whereas
258 the majority of introductions estimated from genomic data originated in other African countries.
259 The III from other African nations was relatively low, despite comparable incoming air travel
260 volume. This is likely both a factor of the relatively low incidence in most African countries
261 during this period (Fig. 2F, SFigure 3) as well as the fact that our travel data does not account
262 for land-based travel, which will expectedly severely underestimate the introduction risk from

263 surrounding countries. The ILL from the African region was largely driven by South Africa, which
264 was experiencing a peaked incidence from its second wave (SFigure 3). In our ILL analyses, we
265 found negligible introduction risk from outside of Europe and North America.
266



267

268 **Fig. 2: A)** Time-resolved B.1.1.318 phylogeny. Branches are coloured by region-level geographic state
269 reconstruction. Internal nodes annotated with black points represent posterior support > 0.75. **B)** Mean
270 number of introductions (Markov jumps) of B.1.1.318 into Nigeria from all source regions, binned by
271 month. The mean number of introductions overall (inset). **C)** Volume of monthly air passengers inbound
272 to Nigeria across 2020-2021 by source region (bar, left axis) and fold increase in travel volume compared
273 to May 2020 baseline. **D)** Estimated Introduction Intensity Index for Nigeria from 2020-2021. **E)**
274 Percentage of inbound air travel to Nigeria by source region across 2020-2021. See Supplementary
275 Figure 2 for country-level data.

276 **The Eta variant of interest (B.1.525) likely emerged in Nigeria.**

277

278 The B.1.525 (Eta) lineage drove the second wave from January to March 2021. B.1.525 was
279 initially reported to have originated in Nigeria or the UK^{16,17,18}. We investigate the timing, origin,
280 and transmission dynamics of B.1.525 with Bayesian phylogeographic reconstructions (see
281 Methods). We estimated that the B.1.525 lineage emerged in Nigeria (root state posterior
282 support = 0.998) in late July [mean tMRCA 23 July, 95% HPD 7 June to 4 September], during
283 the final weeks of Nigeria's first wave (Fig. 3A). This suggests that this lineage circulated
284 cryptically for several months before its first detection in Nigeria on 12 December 2020, likely
285 owing to low-incidence-associated sparse sampling (Fig. 1A).
286

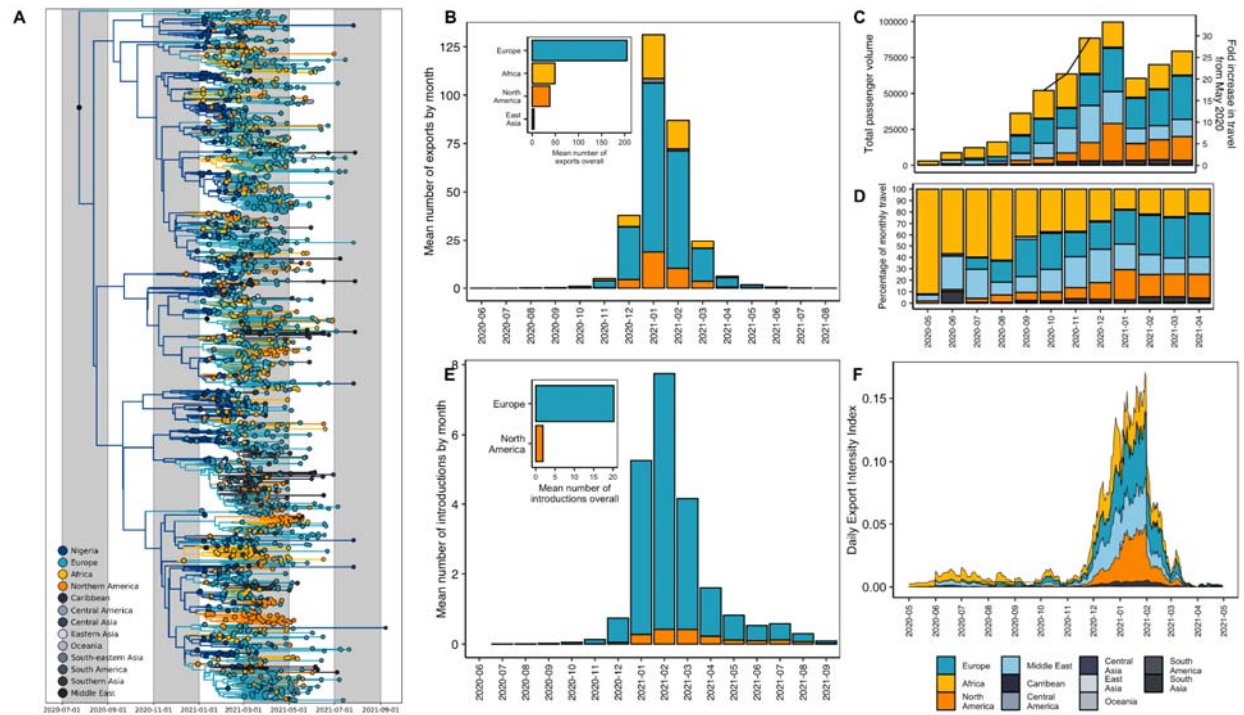
287 We sought to better understand Nigeria's connectivity in global SARS-CoV-2 transmission
288 networks by identifying the most likely destinations of B.1.525's exportation, as the lineage 1)
289 emerged in Nigeria and 2) is Nigeria's most sampled lineage, excluding the Delta variant

290 sublineages. To investigate the spread of B.1.525 from Nigeria, we reconstructed the timing and
291 pattern of geographic transitions out of and into Nigeria across the full posterior of our Bayesian
292 phylogeographic reconstructions. We estimated that B.1.525 was exported from Nigeria a lower
293 bound of 295 times [mean Markov jumps, 95% HPD 259-335] (Fig. 3B). The majority of
294 sampled exports were destined for Europe, followed by the rest of Africa and Northern America
295 from December 2020 to March 2021 (encompassing Nigeria's second wave) (Fig. 3B). During
296 this period, we also found that B.1.525 was re-introduced from Europe a lower bound of 20
297 [95% HPD 6-37] times (Fig. 3E). These estimated source-sink dynamics support Nigeria's
298 strong connectivity, particularly to Europe, but will underestimate exports to undersampled
299 regions such as neighbouring African countries. To mitigate the effect of global sampling biases
300 on genomic estimates of transmission dynamics, we again analysed changes in air travel to and
301 from Nigeria over time to understand the sources and seeds of Nigeria's bidirectional
302 transmission dynamics.

303
304 African nations dominated the reduced level of outbound travel during the period of travel
305 restrictions (May to October 2020) (Fig. 3C-D), suggesting that regional connectivity drove
306 Nigeria's export risk during this period. We integrated the air travel data with Nigeria's epidemic
307 incidence data to quantify an exportation intensity index (EII)²⁷. The EII quantifies the temporal
308 trend in the daily estimated number of viral exports from Nigeria to sink countries (see
309 Methods). Given the data available, Nigeria's estimated export intensity was low overall
310 compared to countries with high air traffic data to Nigeria, as the country's SARS-CoV-2
311 incidence remained comparatively low on a global-scale for the entire period under investigation
312 (Fig. 3F). Nigeria's first wave had relatively low incidence, peaking at about 4000 cases in a
313 week in late June (week 18) (SFigure 4B). Combined with reduced travel volumes, the period
314 from May 2020 to September 2020 was characterised by negligible export intensity (Fig. 3F).

315
316 Outbound air travel recovered gradually over 2020, peaking in December 2020 - January 2021
317 (Fig. 3C-D). We found that this peak corresponded with the peak of the larger second wave in
318 Nigeria and therefore also the EII, as well as the distribution of exports of B.1.1.525 from Nigeria
319 (Fig. 3B, F). The proportion of outbound air travel destined for countries in Europe, the Middle
320 East, and North America increased relative to travel destined for African countries after travel
321 restrictions were lifted (Fig. 3C-D). We estimated the highest EII for Europe (driven
322 predominantly by travel to the UK) and Northern America (driven by travel to the USA) from
323 December 2020 to February 2021 (SFigure 4). The export intensity to Europe is consistent with
324 the high number of estimated exports in the genomic data, though North America was
325 disproportionately underestimated in the genomic data (Fig. 3B). Overall, air travel volume
326 destined for Europe was only moderately higher (10%) than Africa, North America, and the
327 Middle East from December (after first detection and at the start of the second wave) to April
328 2021 (end of the second wave) (Fig. 3C-D). However, Europe had 4-fold more sampled
329 introductions of B.1.525 than African nations or the USA. Notably, the comparatively high EII for
330 the Middle East region (representing 20% of outbound volume overall) was not represented in
331 our phylogeographic reconstructions (Fig. 3B). This highlights how unevenly distributed
332 surveillance capacity can lead to an underestimation of transmission events (SFigure 3A).

333
334



335
336 **Fig. 3: A)** Time-resolved B.1.525 phylogeny. Branches are coloured by region-level geographic state
337 reconstruction. Internal nodes annotated with black points represent posterior support > 0.75. **B)** Mean
338 number of exports (Markov jumps) of B.1.525 from Nigeria to destination regions, binned by month. Mean
339 number of exports overall (inset). **C)** Volume of monthly air passengers outbound from Nigeria across
340 2020 - 2021 by destination region (bar, left axis) and fold increase in travel volume compared to baseline
341 (May 2020). **D)** Percentage of outbound air travel from Nigeria by destination region across 2020-2021.
342 See Supplementary Figure 1 for country-level data. **E)** Mean number of introductions (Markov jumps) of
343 B.1.525 into Nigeria from all origin regions, binned by month. The mean number of introductions overall
344 (inset). **F)** Estimated Exportation Intensity Index (EII) for Nigeria from 2020-2021.

345

346

347 Discussion

348

349 In this study, we combined genomic, travel and epidemiological data to characterise the
350 emergence and bidirectional transmission of two focal variants of interest in Nigeria. The focal
351 VOIs, B.1.525 and B.1.1.318, were suggested to have emerged in Nigeria before spreading
352 globally, resulting in large-scale outbreaks in Brazil and Mauritius^{15,19}. In our phylogeographic
353 reconstructions, we found that the B.1.1.318 lineage most likely emerged in the African region in
354 early August 2020, with its dominance in Nigeria driven by multiple introductions during the
355 second wave from November 2020 onwards. Our import risk index underestimates the
356 importance of regional connectivity in driving these introduction estimates, as they likely result
357 from shorter distance connectivity to surrounding countries, which we could not collect data on.
358 We also found that the Eta variant (B.1.525) emerged in Nigeria in late July, with high levels of
359 export to Europe reflected in the genomic and estimated export index.

360

361 The true extent of bidirectional transmission with other African nations will be severely
362 underestimated in phylogeographic reconstructions, as Africa remains disproportionately
363 undersampled despite heroic efforts²⁸. Our findings should be interpreted in the light of our own
364 sampling fraction, which was 0.026%. We attempted to mitigate these global surveillance biases
365 by supplementing genomic estimates with sampling-independent metrics like the introduction

366 and export intensity indices. However, these metrics are dependent on epidemiological
367 incidence and will be underestimated if there is large-scale underascertainment in source
368 countries, as previously shown for African (and other) nations based on excess mortality
369 data^{29,30,31}.

370

371 At an unprecedented speed (72 hours from receiving the sample), ACEGID generated the first
372 whole-genome sequence of the virus in Africa in March 2020. This catalysed and built
373 confidence in SARS-CoV-2 genomic sequencing on the African continent. This immense
374 collaborative effort, involving seventeen partner laboratories in Nigeria, helped to characterise
375 the pandemic in near real-time during its first three waves in the country. As we have learnt the
376 hard way in recent years, novel variants of interest or concern can emerge from anywhere
377 around the world, notably including undersampled regions such as Africa^{13,14}. In a highly
378 connected world, there is a proactive need to adopt early warning systems for pandemic
379 preemption and response in Africa, such as SENTINEL³². This would enable equality in global
380 genomic surveillance, especially in countries with fragile public health systems, in order to
381 effectively detect and curb emerging outbreaks before they spread.

382

383 **Methods**

384

385 **Sample preparation**

386

387 RNA was extracted from nasopharyngeal swabs in viral transport media and saliva/sputum in
388 PBS using the QiAmp viral RNA mini kit (Qiagen) according to the manufacturer's instructions
389 and MagMax pathogen RNA/DNA kit (Applied Biosystems) using a Kingfisher Flex purification
390 system (ThermoFisher Scientific) according to the manufacturer's instructions.

391

392 **Probe-based SARS-CoV-2 real-time RT-PCR**

393

394 RT-qPCR screening of suspected samples was carried out targeting N, ORF1ab and RdRP
395 genes of the virus using commercially available kits: genesig® (Primerdesign Ltd, UK),
396 DaAnGene (Daan Gene Co., Ltd, China), Liferiver (Shanghai ZJ Bio-Tech Co., Ltd, China),
397 Genefinder (OSANG Healthcare Co., Ltd, Korea), and Sansure (Sansure Biotech Inc., China).

398

399 **SARS-CoV-2 whole-genome sequencing**

400

401 Samples with moderate to high viral load detected with RT-qPCR (Ct value < 30) were selected
402 for sequencing, which were samples collected from genomic surveillance and hospitalised
403 individuals spatially and temporally. The sampling gap was the inability to sequence samples
404 from every state during all different waves of the pandemic in Nigeria. The Illumina COVIDseq
405 protocol was used for sequencing preparation (<https://www.illumina.com/products/by-type/clinical-research-products/covidseq-assay.html>). RNA was converted to cDNA, followed by
406 tiling amplification (400 bp) with ARTIC V3 primer pools, which covers the whole genome of the
407 virus. Nextera DNA flex libraries were made from the amplicons. Libraries were pooled and
408 sequenced on the Illumina MiSeq, NextSeq 2000, and NovaSeq 6000 platforms at the African
409 Center of Excellence for Genomics of Infectious Diseases (ACEGID), Redeemer's University,
410 Ede, Nigeria.

411

413 **Genome Assembly**

414

415 We used the viral-ngs pipeline v2.1.19 (<https://github.com/broadinstitute/viral-ngs>) for
416 demultiplexing, quality control, and genome assembly, consistent with prior work^{33,34,35}. The
417 demux_plus applet of viral-ngs was used for demultiplexing basecalls to BAM files. Genome
418 assembly was carried out using the assemble_refbased workflow on the viral-ngs pipeline which
419 maps each unmapped BAM file to the SARS-CoV-2 reference genome (NC_045512.2) to
420 generate coverage plots and FASTA files of successfully assembled genomes.

421

422

423 **Lineage/clade assignment**

424

425 We used Phylogenetic Assignment of Named Global Outbreak LINEages (PANGOLIN)³⁶ to
426 identify the lineages of SARS-CoV-2 circulating within Nigeria and to identify the lineages
427 circulating in specific states of Nigeria. Nextclade (<https://clades.nextstrain.org/>) was used to
428 assign the sequences to globally circulating viral clades and to investigate mutations in the
429 genome that could affect primer amplification during PCR.

430

431 **Maximum likelihood phylogenetic analysis**

432

433 1,577 genomes with sequence length $\geq 20,930$ (covering 70% or more of the reference -
434 NC_045512.2) were assembled and aligned using MAFFT³⁷. We used IQTREE³⁸ and the GTR
435 (generalised time-reversible) model with a bootstrap value of 1000 to construct the phylogenetic
436 tree. Treetime³⁹ was used for quantifying the molecular clock and identifying ancestral
437 phylogeny in the augur pipeline of Nextstrain⁴⁰.

438

439 **Bayesian phylogenetic analysis**

440

441 All B.1.1.318 (n = 3,858 with major locations: Europe, USA, and Africa accounting for 94%) and
442 B.1.525 (n = 8,278 with major locations: Europe, USA, and Africa accounting for 86%)
443 sequences were downloaded from GISAID⁴¹ on 18 August 2021. Sequences with > 5%
444 ambiguous nucleotides, a length < 95%, or incomplete dates were discarded. Sequences were
445 aligned to the reference (NC_045512.2) using minimap2, with the 5' and 3' UTRs and known
446 problematic sites (GitHub - W-L/ProblematicSites_SARS-CoV2 3) masked⁴². Both lineage-
447 specific datasets were downsampled for representativeness and computational tractability with
448 a phylogenetic-informed downsampling scheme. A maximum-likelihood phylogeny was
449 reconstructed for each lineage with automatic model selection in IQTREE⁴³. The phylogeny was
450 downsampled by root-to-tip tree-traversal, with all internal nodes subject to two rules: 1) if 95%
451 of the leaves subtended by the internal node represented a single country, the earliest
452 representative was retained alongside a random sequence; 2) if leaves from the same location
453 were separated by a zero-branch length, the earliest representative was retained alongside a
454 random sequence. All countries with more than the median number of sequences across
455 countries were then downsampled randomly to the median, this yielded a total number of 1,118
456 and 1,746 genomes for B.1.1.318 and B.1.525 respectively. All Nigerian sequences were
457 retained (n = 73 for B.1.1.318; n = 256 for B.1.525).

458

459 We reconstructed Bayesian time-scaled phylogenies for each lineage using BEAST 1.10.5⁴⁴.
460 For both lineages, the time-scaled phylogenies were reconstructed under an HKY substitution
461 model with a gamma-distributed rate for variation among sites⁴⁵, a relaxed molecular clock with
462 a log-normal prior⁴⁶, and an exponential growth coalescent tree prior⁴⁷. For each dataset, we
463 combined two independent Markov Chain Monte Carlo (MCMC) chains of 200 million states, run
464 with the BEAGLE package⁴⁸ to improve run time. Parameters and trees were sampled every
465 20,000 steps, with the first 20% of steps discarded as burn-in. Convergence and mixing of the

466 MCMC chains were assessed in Tracer v1.7, to ensure the effective sample size (ESS) of all
467 estimated parameters was $> 200^{49}$.

468

469 We performed an asymmetric discrete trait analysis using BEAST version 1.10.5 to reconstruct
470 the location-transition history across an empirical distribution of 4000 time-calibrated trees
471 (sampled from each of the posterior tree distributions estimated above). We aggregated the
472 country of sampling on a regional level for computational tractability. We used Bayesian
473 stochastic search variable selection (BSSVS) to infer non-zero migration rates and identify the
474 statistically supported transition routes into and out of Nigeria by a Bayes factor test⁵⁰. In
475 addition to the discrete trait analysis, we used a Markov jump counting approach to estimate the
476 timing and origin of geographic transitions into Nigeria to account for uncertainty in
477 phylogeographic reconstruction associated with sparse sampling and low sequence variability⁵¹.
478 We used the TreeMarkovJumpHistoryAnalyzer from the pre-release version of BEAST 1.10.5 to
479 extract the Markov jumps from posterior tree distributions⁵². We used TreeAnnotator 1.10 to
480 construct Maximum clade credibility (MCC) trees for all datasets. Trees were visualised using
481 baltic (<https://github.com/evogytis/baltic>).

482

483 **Air traffic and human movement data**

484

485 In order to associate international air travel and local human movement with variant movement
486 across borders, we used air traffic data from the International Air Transportation Association
487 (IATA) obtained from BlueDot (<https://bluedot.global/>) to quantify the volume of international
488 travellers to and from Nigeria. The dataset included the number of travellers by origin and
489 destination country, aggregated by month, across all international airports in Nigeria. Data was
490 obtained for the period May 2020 to April 2021, encompassing the emergence and spread of
491 B.1.525 and B.1.1.318. We also curated human movement data within Nigeria from Google
492 Mobility (<https://www.google.com/covid19/mobility/>) from January 2020 to November 2021,
493 which reflects a baseline before lockdown (January - March 2020) and the study's period of
494 interest. We used R to display movement patterns after grouping human movement into six
495 divisions: retail & recreation, grocery & pharmacy, parks, transport stations, workplaces, and
496 residential areas.

497

498 **Estimated introduction and exportation intensity index**

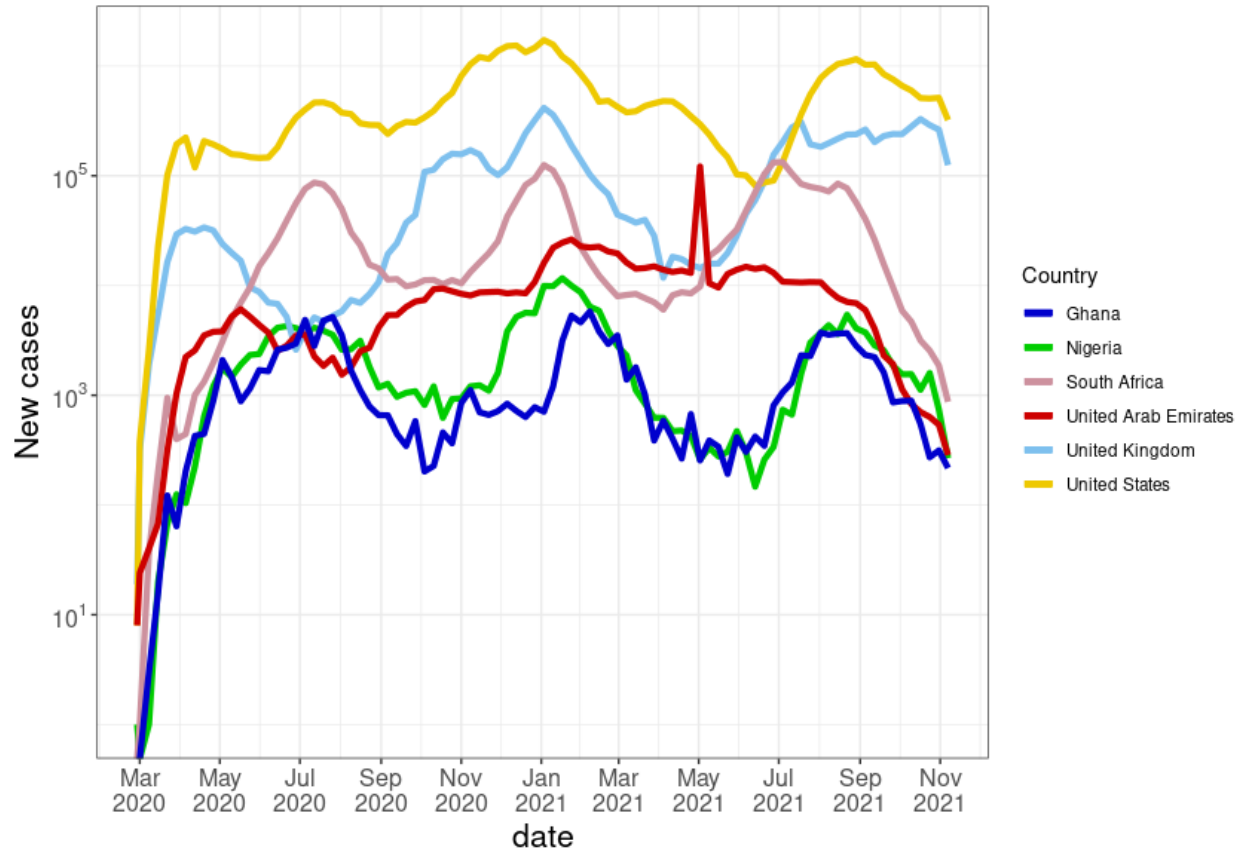
499

500 From genomic data, estimates of the number and origin/destination of bidirectional
501 transmissions with Nigeria are dependent on the sample's representativeness of the population
502 and are therefore limited by global and local sampling biases. To limit these sampling biases,
503 we supplemented our phylogeographic analyses by estimating the introduction (III) and export
504 intensity index (EII) according to Du Plessis et al²⁷. The III estimates the daily risk of introductions
505 into Nigeria from each country as a product of the number of asymptotically infected
506 individuals in each source country on that day (estimated from the time series of deaths) and
507 their likelihood of traveling to Nigeria (based on the volume of inbound air travel from the source
508 country). The EII is calculated similarly, based on the number of asymptotically infected
509 individuals in Nigeria and their likelihood of travel to each destination by air. Notably, we could
510 not obtain data for land-based travel. Connectivity to regional and neighbouring countries and
511 the associated introduction and export intensity are therefore likely severely underestimated. A
512 time series of reported deaths for each country were collected from the outbreak.info R

513 package⁵³. As air travel data was aggregated by month, we conservatively assumed that travel
514 was uniform across all days of the month. We quantified the IIR and EIR for the period May 2020
515 to April 2021, encompassing the emergence and spread of B.1.525 and B.1.1.318.

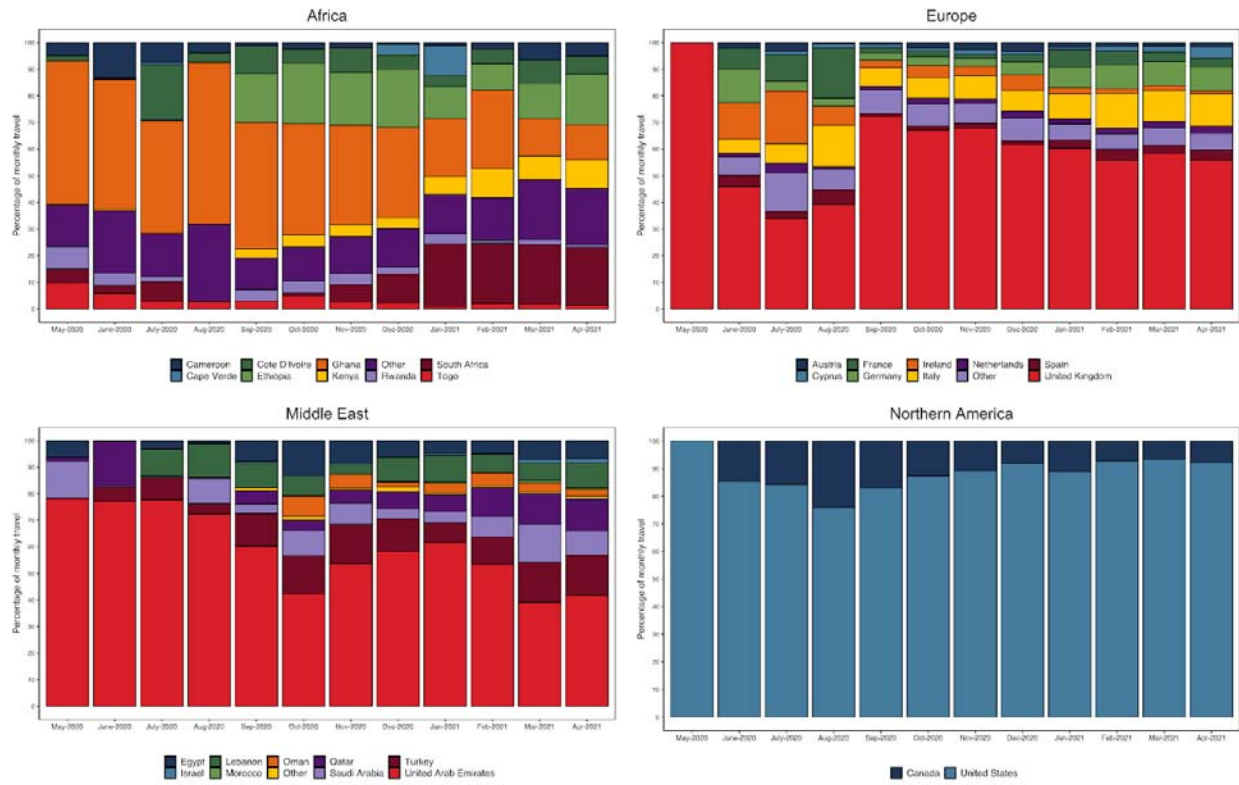
516
517
518
519
520
521
522
523
524
525

526 Supplementary figures



527
528 **SFigure 1:** COVID-19 new cases per week as seen in Nigeria and countries with highest flight traffic in
529 and out of Nigeria from the four major connected regions with the y-axis transformed to a log scale.
530
531

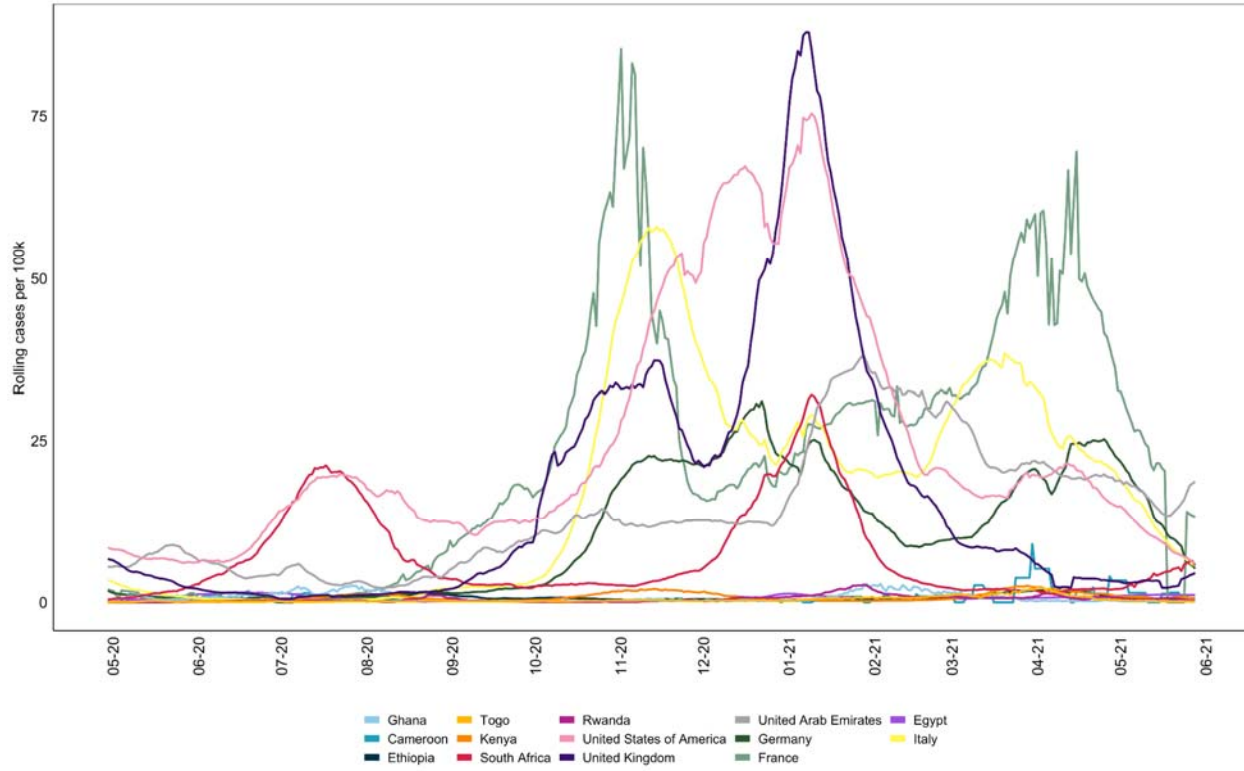
532
533



534
535
536
537
538
539
540

Figure 2: Percentage of incoming travel from the four major connected regions by country as per Figure 3

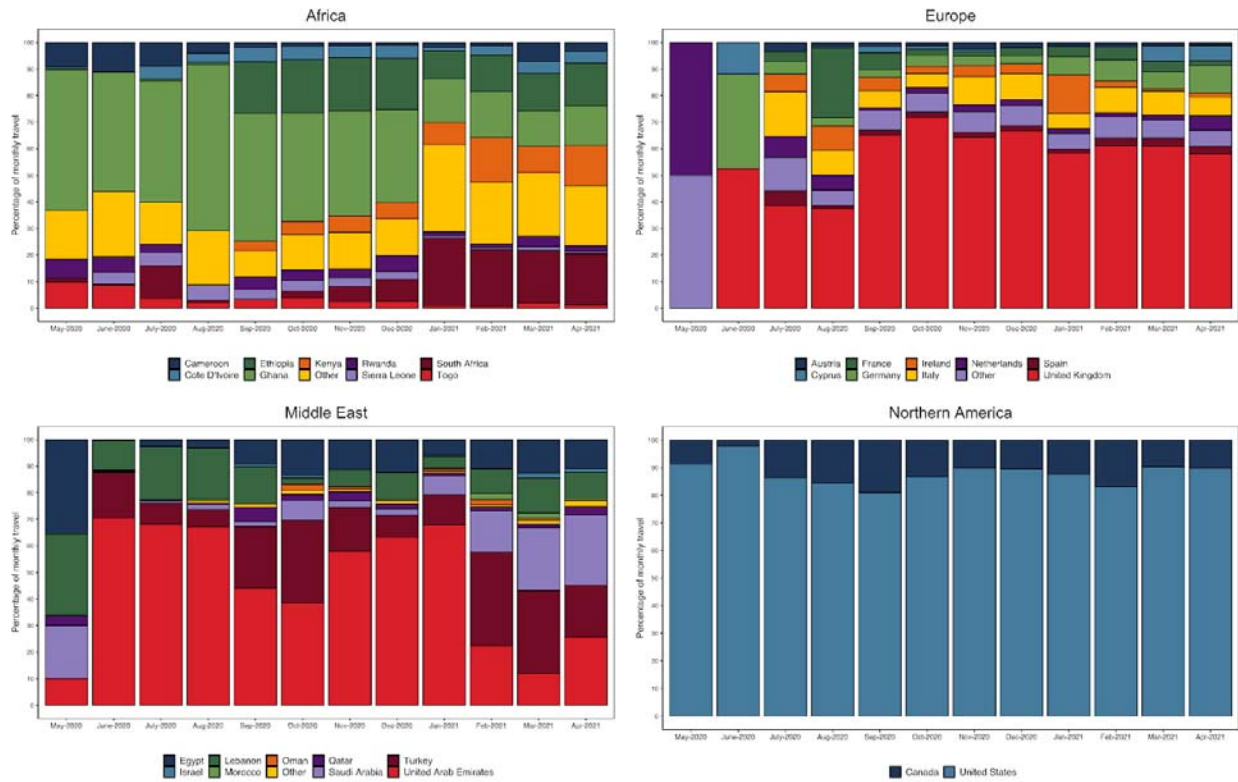
It is made available under a [CC-BY-NC-ND 4.0 International license](https://creativecommons.org/licenses/by-nc-nd/4.0/) .



541
542
543
544
545

Figure 3: Population adjusted rolling average of new daily cases for countries with the highest incoming air travel as per Figure S2.

546



547
548
549
550
551
552
553
554
555
556
557
558
559

Figure 4: Percentage of outgoing travel to the four major connected regions by country as per Figure 4

560 Acknowledgements

561

562 This work is made possible by the support provided to ACEGID by a cohort of generous donors
563 through TED's Audacious Project, including the ELMA Foundation, MacKenzie Scott, the Skoll
564 Foundation, and Open Philanthropy. This work was also partly supported by grants from the
565 National Institute of Allergy and Infectious Diseases (<https://www.niaid.nih.gov>), NIH-H3Africa
566 (<https://h3africa.org>) (U01HG007480 and U54HG007480), the World Bank (projects ACE-019
567 and ACE-IMPACT), the Rockefeller Foundation (Grant #2021 HTH), the Africa CDC through the
568 African Society of Laboratory Medicine [ASLM] (Grant #INV018978), the Wellcome Trust
569 (Project 216619/Z/19/Z), the WARN-ID/CREID grant # U01AI151812, and the Science for Africa
570 Foundation.

571 We want to especially thank the Redeemer's University Management, the State Government of
572 Osun and NCDC for their support during the course of the COVID-19 pandemic.

573

574

575 References

576

577 1. World Bank, World Development Indicators (2022). Available at: <http://wdi.worldbank.org/tables>
578 (Accessed: 29th April 2022).

579

580 2. Tsueng G, Mullen J, Alkuzweny M, Cano M, Benjamin HR, Emily O, Curators L, Alaa A, Zhou X, Qian
581 Z, Andersen KG. Outbreak. info Research Library: A standardized, searchable platform to discover and
582 explore COVID-19 resources and data. *BioRxiv*. 2022 Feb 1.

583

584 3. Nordling L. The pandemic appears to have spared Africa so far. Scientists are struggling to explain
585 why. *Science* (80-). 2020 Aug 11;

586

587 4. Kreier F. Morgue data hint at COVID's true toll in Africa. *Nature*. 2022 Mar 23.

588

589 5. Tessema SK, Nkengasong JN. Understanding COVID-19 in Africa. *Nature Reviews Immunology*. 2021
590 Aug;21(8):469-70.

591

592 6. Gilbert, M., Pullano, G., Pinotti, F., Valdano, E., Poletto, C., Boëlle, P.Y., d'Ortenzio, E., Yazdanpanah,
593 Y., Eholie, S.P., Altmann, M. and Gutierrez, B., 2020. Preparedness and vulnerability of African countries
594 against importations of COVID-19: a modelling study. *The Lancet*, 395(10227), pp.871-877.

595

596 7. Borrega, R., Nelson, D.K., Koval, A.P., Bond, N.G., Heinrich, M.L., Rowland, M.M., Lathigra, R., Bush,
597 D.J., Aimukanova, I., Phinney, W.N. and Koval, S.A., 2021. Cross-reactive antibodies to SARS-CoV-2
598 and MERS-CoV in pre-COVID-19 blood samples from Sierra Leoneans. *Viruses*, 13(11), p.2325.

599

600 8. Bugembe DL, Phan MV, Ssewanyana I, Semanda P, Nansumba H, Dhaala B, Nabadda S, O'Toole
601 ÁN, Rambaut A, Kaleebu P, Cotten M. Emergence and spread of a SARS-CoV-2 lineage A variant (A.
602 23.1) with altered spike protein in Uganda. *Nature microbiology*. 2021 Aug;6(8):1094-101.

603

604 9. Butera Y, Mukantwari E, Artesi M, O'Toole ÁN, Hill V, Rooke S, Hong SL, Dellicour S, Majyambere O,
605 Bontems S, Boujemla B. Genomic sequencing of SARS-CoV-2 in Rwanda reveals the importance of
606 incoming travelers on lineage diversity. *Nature communications*. 2021 Sep 29;12(1):1-1.

607

- 608 10. Wilkinson E, Giovanetti M, Tegally H, San JE, Lessells R, Cuadros D, Martin DP, Rasmussen DA,
609 Zekri AR, Sangare AK, Ouedraogo AS. A year of genomic surveillance reveals how the SARS-CoV-2
610 pandemic unfolded in Africa. *Science*. 2021 Oct 22;374(6566):423-31.
611
- 612 11. Parker E, Anderson C, Zeller M, Tibi A, Havens JL, Laroche G, Benlarbi M, Ariana A, Robles-Sikisaka
613 R, Latif AA, Watts A. Regional connectivity drove bidirectional transmission of SARS-CoV-2 in the Middle
614 East during travel restrictions. *Nature communications*. 2022 Aug 15;13(1):1-4.
615
- 616 12. Bogoch II, Creatore MI, Cetron MS, Brownstein JS, Pesik N, Miniota J, Tam T, Hu W, Nicolucci A,
617 Ahmed S, Yoon JW. Assessment of the potential for international dissemination of Ebola virus via
618 commercial air travel during the 2014 west African outbreak. *The Lancet*. 2015 Jan 3;385(9962):29-35.
619
- 620 13. Tegally H, Wilkinson E, Giovanetti M, Iranzadeh A, Fonseca V, Giandhari J, Doolabh D, Pillay S, San
621 EJ, Msomi N, Mlisana K. Detection of a SARS-CoV-2 variant of concern in South Africa. *Nature*. 2021
622 Apr;592(7854):438-43.
623
- 624 14. Viana, R., Moyo, S., Amoako, D.G., Tegally, H., Scheepers, C., Althaus, C.L., Anyaneji, U.J., Bester,
625 P.A., Boni, M.F., Chand, M. and Choga, W.T., 2022. Rapid epidemic expansion of the SARS-CoV-2
626 Omicron variant in southern Africa. *Nature*, 603(7902), pp.679-686.
627
- 628 15. Pereira, F. et al. Genomic surveillance activities unveil the introduction of the SARS-CoV-2 B.1.525
629 variant of interest in Brazil: Case report. *Journal of Medical Virology*, 93, 9, 5523–5526 (2021).
630
- 631 16. Dresch, M. Nigerian Covid-19 variant under investigation with 38 new cases found.
632 <https://www.mirror.co.uk/news/uk-news/breaking-nigerian-covid-19-variant-23511168> (2021) (Accessed
633 24th March, 2022).
- 634 17. Vadlapatla, S. B.1.525 Variant Linked to Nigeria now Found in Telangana -
635 [https://timesofindia.indiatimes.com/city/hyderabad/b-1-525-variant-linked-to-nigeria-now-found-in-](https://timesofindia.indiatimes.com/city/hyderabad/b-1-525-variant-linked-to-nigeria-now-found-in-telangana/articleshow/83133727.cms)
636 [telangana/articleshow/83133727.cms](https://timesofindia.indiatimes.com/city/hyderabad/b-1-525-variant-linked-to-nigeria-now-found-in-telangana/articleshow/83133727.cms). (2021) (Accessed 24th March, 2022).
- 637 18. Takeuchi C. COVID-19 in B.C.: New variant from Nigeria detected, ski awareness campaign
638 launched, and more. [https://www.straight.com/covid-19-pandemic/february-12-coronavirus-update-bc-](https://www.straight.com/covid-19-pandemic/february-12-coronavirus-update-bc-vancouver-new-variant-from-nigeria-detected-ski-awareness-campaign-and-more)
639 [vancouver-new-variant-from-nigeria-detected-ski-awareness-campaign-and-more](https://www.straight.com/covid-19-pandemic/february-12-coronavirus-update-bc-vancouver-new-variant-from-nigeria-detected-ski-awareness-campaign-and-more)Takeuchi. (2021)
640 (Accessed 24th March, 2022).
- 641
- 642 19. Tegally, H. et al. A novel and expanding SARS-CoV-2 Variant, B. 1.1. 318, dominates infections in
643 Mauritius. *medRxiv* <https://doi.org/10.1101/2021.06.16.21259017> (2021).
644
- 645 20. Zhang L, Cui Z, Li Q, Wang B, Yu Y, Wu J, Nie J, Ding R, Wang H, Zhang Y, Liu S. Ten emerging
646 SARS-CoV-2 spike variants exhibit variable infectivity, animal tropism, and antibody neutralization.
647 *Communications biology*. 2021 Oct 13;4(1):1-0.
648
- 649 21. Weisblum Y, Schmidt F, Zhang F, DaSilva J, Poston D, Lorenzi JC, Muecksch F, Rutkowska M,
650 Hoffmann HH, Michailidis E, Gaebler C. Escape from neutralizing antibodies by SARS-CoV-2 spike
651 protein variants. *Elife*. 2020 Oct 28;9:e61312.
652
- 653 22. Galloway SE, Paul P, MacCannell DR, Johansson MA, Brooks JT, MacNeil A, Slayton RB, Tong S,
654 Silk BJ, Armstrong GL, Biggerstaff M. Emergence of SARS-CoV-2 b. 1.1. 7 lineage—united states,
655 december 29, 2020–january 12, 2021. *Morbidity and Mortality Weekly Report*. 2021 Jan 1;70(3):95.
656
- 657 23. Emary KR, Golubchik T, Aley PK, Ariani CV, Angus B, Bibi S, Blane B, Bonsall D, Cicconi P, Charlton
658 S, Clutterbuck EA. Efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine against SARS-CoV-2 variant of

- 659 concern 202012/01 (B. 1.1. 7): an exploratory analysis of a randomised controlled trial. *The Lancet*. 2021
660 Apr 10;397(10282):1351-62.
661
- 662 24. Chaillon, A. & Smith, D. M. Phylogenetic analyses of SARS-CoV-2 B.1.1.7 lineage suggest a single
663 origin followed by multiple exportation events versus convergent evolution. *Clinical Infectious Diseases*
664 **73**, 12, 2314–2317 (2021).
665
- 666 25. Washington, N. L. et al. (2021). Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the
667 United States. *Cell* **184**, 10, 2587-2594.
668
- 669 26. Davies, N. G. et al. Estimated transmissibility and impact of SARS-CoV-2 lineage B. 1.1. 7 in England.
670 *Science* **372**, 6538, eabg3055 (2021).
671
- 672 27. du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK.
673 *Science (New York, N.Y.)*, **371**, 6530, 708–712 (2021).
674
- 675 28. Tegally H, San JE, Cotten M, Tegomoh B, Mboowa G, Martin DP, Baxter C, Moir M, Lambisia A,
676 Diallo A, Amoako DG. The evolving SARS-CoV-2 epidemic in Africa: Insights from rapidly expanding
677 genomic surveillance. medRxiv. 2022 Jan 1.
678
- 679 29. Kleynhans J, Tempia S, Wolter N, von Gottberg A, Bhiman JN, Buys A, Moyes J, McMorro ML,
680 Kahn K, Gómez-Olivé FX, Tollman S. SARS-CoV-2 Seroprevalence in a rural and urban household
681 cohort during first and second waves of infections, South Africa, July 2020–March 2021. *Emerging*
682 *infectious diseases*. 2021 Dec;27(12):3020.
683
- 684 30. Wang H, Paulson KR, Pease SA, Watson S, Comfort H, Zheng P, Aravkin AY, Bisignano C, Barber
685 RM, Alam T, Fuller JE. Estimating excess mortality due to the COVID-19 pandemic: a systematic analysis
686 of COVID-19-related mortality, 2020–21. *The Lancet*. 2022 Apr 16;399(10334):1513-36.
687
- 688 31. Cohen C, Kleynhans J, von Gottberg A, McMorro ML, Wolter N, Bhiman JN, Moyes J, du Plessis M,
689 Carrim M, Buys A, Martinson NA. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an
690 urban setting: results of the PHIRST-C cohort study, South Africa, 2020–21. *The Lancet Infectious*
691 *Diseases*. 2022 Mar 14.
692
- 693 32. Botti-Lodovico, Y. et al. The Origins and Future of Sentinel: An Early-Warning System for Pandemic
694 Preemption and Response. *Viruses* **13**, 8, 1605 (2021).
- 695 33. Siddle, K. J. et al. Genomic Analysis of Lassa Virus during an Increase in Cases in Nigeria in
696 2018. *The New England Journal of Medicine* **379**, 18, 1745–1753 (2018).
- 697 34. Ajogbasile, F. V. et al. Real-time metagenomic analysis of undiagnosed fever cases unveils a
698 yellow fever outbreak in Edo State, Nigeria. *Scientific Reports* **10**, 1, 1-6 (2020).
- 699 35. Oluniyi, P. First African SARS-CoV-2 genome sequence from Nigerian COVID-19 case.
700 *Virological* [https://virological.org/t/first-african-sars-cov-2-genome-sequence-from-nigerian-covid-19-](https://virological.org/t/first-african-sars-cov-2-genome-sequence-from-nigerian-covid-19-case/421)
701 [case/421](https://virological.org/t/first-african-sars-cov-2-genome-sequence-from-nigerian-covid-19-case/421) (2020).
702
- 703 36. Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic
704 epidemiology. *Nat Microbiol* <https://doi.org/10.1038/s41564-020-0770-5> (2020).
705
- 706 37. Katoh, K., Misawa, K., Kuma, K., & Miyata, T. MAFFT: a novel method for rapid multiple sequence
707 alignment based on fast Fourier transform. *Nucleic Acids Research*, **30**, 14, 3059–3066 (2002).
708
- 709 38. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A. &
710 Lanfear, R. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era.
711 *Mol. Biol. Evol.* **37**, 1530-1534 (2020).

- 712
713 39. Sagulenko, P., Puller, V. & Neher, R.A. TreeTime: Maximum-likelihood phylodynamic analysis. *Virus*
714 *Evol.* **4**, 1, vex042 (2018).
715
716 40. Hadfield, J., Megill, C., Bell, S. M., Huddleston, J., Potter, B., Callender, C., Sagulenko, P., Bedford,
717 T. & Neher, R. A. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics (Oxford, England)*,
718 **34**, 23, 4121–4123 (2018).
719
720 41. Elbe, S. & Buckland-Merrett, G. Data, disease and diplomacy: GISAID’s innovative contribution to
721 global health. *Global Challenges*, **1**, 33-46 (2017).
722
723 42. Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics*, **34**, 18, 3094-3100
724 (2018).
725
726 43. Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A. & Jermini, L. S. ModelFinder: fast
727 model selection for accurate phylogenetic estimates. *Nature methods*, **14**, 587-589 (2017).
728
729 44. Suchard, M. A., Lemey, P., Baele, G., Ayres, D. L., Drummond, A. J. & Rambaut, A. Bayesian
730 phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus evolution*, **4**, 1, vey016 (2018).
731
732 45. Hasegawa, M., Kishino, H. & Yano, T. A. . Dating of the human-ape splitting by a molecular clock of
733 mitochondrial DNA. *Journal of Molecular Evolution*, **22**, 2, 160-174 (1985).
734
735 46. Drummond, A. J., Ho, S. Y. W., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with
736 confidence. *PLoS Biology*, **4**, 5, e88 (2006).
737
738 47. Drummond, A. J., Rambaut, A., Shapiro, B. E. T. H. & Pybus, O. G. Bayesian coalescent inference of
739 past population dynamics from molecular sequences. *Molecular Biology and Evolution*, **22**, 5, 1185-1192
740 (2005).
741
742 48. Ayres, D. L. et al. BEAGLE: an application programming interface and high-performance computing
743 library for statistical phylogenetics. *Systematic Biology*, **61**, 1, 170-173 (2012).
744
745 49. Rambaut, A., Drummond, A. J., Xie, D., Baele, G. & Suchard, M. A. Posterior summarization in
746 Bayesian phylogenetics using Tracer 1.7. *Systematic Biology*, **67**(5), 901 (2018).
747
748 50. Lemey, P., Rambaut, A., Drummond, A. J. & Suchard, M. A. Bayesian phylogeography finds its roots.
749 *PLoS Computational Biology*, **5**, 9, e1000520 (2009).
750
751 51. Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science*, **370**,
752 6516, 564-570 (2020).
753
754 52. Lemey, P. et al. Untangling introductions and persistence in COVID-19 resurgence in Europe. *Nature*,
755 **595**, 7869, 713-717 (2021).
756
757 53. Gangavarapu K, Latif AA, Mullen JL, Alkuzweny M, Hufbauer E, Tsueng G, Haag E, Zeller M, Aceves
758 CM, Zaiets K, Cano M. Outbreak. info genomic reports: scalable and dynamic surveillance of SARS-CoV-
759 2 variants and mutations. Research square. 2022 Jan 1.
760
761