Resurgence of Ebola virus in 2021 in Guinea suggests a new paradigm for outbreaks

Resurgence of Ebola virus in 2021, Guinea: A new paradigm about outbreaks

Alpha K. Keita [1,2 a], Fara R. Koundouno [3,4 a], Martin Faye [5 a], Ariane Düx[6 a], Julia Hinzmann[4,7,13 a], Haby Diallo[1], Ahidjo Ayouba[2], Frederic Le Marcis[1,2,8], Barré Soropogui[3], Kékoura Ifono[3,4], Moussa M. Diagne[5], Mamadou S. Sow[1,9], Joseph A Bore[3,10], Sebastien Calvignac-Spencer[6], Nicole Vidal[2], Jacob Camara[3], Mamadou B. Keita[11], Annick Renevey4[,7], Amadou Diallo[5], Abdoul K. Soumah[1], Saa L. Millimono[3,4], Almudena Mari-Saez[6], Mamadou Diop[5], Ahmadou Doré[3], Fodé Y. Soumah[9], Kaka Kourouma[11], Nathalie J. Vielle[**4,12**], Cheikh Loucoubar[5], Ibrahima Camara[1], Karifa Kourouma[3,4], Giuditta Annibaldis[4,12], Assaïtou Bah[3], Anke Thielebein[4,7], Meike Pahlmann[4,7], Steven T. Pullan[10,13], Miles W. Carroll[10.13], Joshua Quick[14], Pierre Formenty[15], Anais Legand[15], Karla Pietro[16], Michael R. Wiley[16,17], Noel Tordo[18], Christophe Peyrefitte[5], John T. McCrone[19], Andrew Rambaut[19], Youssouf Sidibé[20], Mamadou D. Barry[20], Madeleine Kourouma[20], Cé D. Saouromou[20], Mamadou Condé[20], Moussa Baldé[9], Moriba Povogui[1], Sakoba Keita[21], Mandiou Diakite[22,23], Mamadou S. Bah[22], Amadou Sidibe[8], Dembo Diakite[9], Fodé B. Sako[9], Fodé A. Traore[9], Georges A. Ki-Zerbo[12], Philippe Lemey[24], Stephan Günther[4,7,12], Liana E. Kafetzopoulou[4,7,24], Amadou A. Sall[5], Eric Delaporte[2,25], Sophie Duraffour[4,7,12 b], Ousmane Faye[5 b], Fabian Leendertz[6 b], Martine Peeters[2 b], Abdoulaye Toure[1,11 b], N'Faly Magassouba[3 b].

[a] These authors contributed equally

[b] These authors jointly supervised this work

Corresponding author: Alpha Kabinet Keita: 1. Centre de Recherche et de Formation en Infectiologie de Guinée (CERFIG), Université de Conakry, Guinea.

Email : alpha-kabinet.keita@cerfig.org, alpha-kabinet.keita@ird.fr

1. Centre de Recherche et de Formation en Infectiologie de Guinée (CERFIG), Université de Conakry, Conakry, Guinea
2. TransVIHMI, Montpellier University /IRD/INSERM, Montpellier, France
3. Laboratoire du Projet des Fièvres Hémorragiques de Guinée (PFHG), Conakry, Guinea
4. Bernhard Nocht Institute for Tropical Medicine (BNITM), Hamburg, Germany
5. Institut Pasteur de Dakar (IPD), Dakar, Senegal
6. Robert Koch Institute, (RKI) Berlin, Germany
7. German Center for Infection Research (DZIF), partner site Hamburg–Lübeck–Borstel–Riems, Hamburg, Germany
8. Ecole Nationale Supérieure de Lyon, France
9. Hôpital National Donka, Service des Maladies Infectieuses et Tropicales, Conakry, Guinea
10. Wellcome Trust Centre for Human Genetics, Nuffield Department of Medicine, Oxford University, Oxford, UK.
11. Institut National de Santé Publique de Guinée (INSP), Conakry, Guinea
12. World Health Organization (WHO), Conakry, Guinea
13. UK Health Security Agency, National infection Service, Porton Down, Salisbury, UK
14. Institute of Microbiology and Infection, University of Birmingham
15. Health Emergencies Program, World Health Organization (WHO), Geneva, Switzerland
16. PraesensBio, Nebraska, USA
17. University of Nebraska Medical Center, Nebraska, USA
18. Institut Pasteur de Guinée, Conakry, Guinea

19. Institute of Evolutionary Biology, University of Edinburgh, King's Buildings, Edinburgh, EH9 3FL, UK
20. Hôpital Régional de Nzérékoré, Nzérékoré, Guinea
21. Agence Nationale de Sécurité Sanitaire, Conakry, Guinea
22. Direction Nationale des Laboratoires, Ministère de la Santé, Conakry, Guinée
23. Université Gamal Abdel Nasser de Conakry, Conakry, Guinea.
24. Department of Microbiology, Immunology and Transplantation, Rega Institute, KU Leuven, Leuven, Belgium
25. Infectious Diseases Departement, University Hospital Montpellier, Montpellier, France

Abstract (150/150)

Seven years after the declaration of the first Ebola virus disease (EVD) epidemic in Guinea, the country faced a new outbreak in 2021 (February 14th to June 19th) near the epicenter of the previous epidemic [1,2]. Complete or near-complete Ebolavirus genomes were generated from samples from 12 different patients using next generation sequencing. The new Zaire Ebolavirus (EBOV) genomes formed a well-supported phylogenetic cluster with genomes from the previous outbreak, indicating that the new outbreak was not the result of a new spill-over event from an animal reservoir. The 2021 lineage shows considerably lower divergence than expected during sustained human-to-human transmission, suggesting a persistent infection with reduced replication or a period of latency. The resurgence of Ebola from humans five years after the end of the previous EVD outbreak reinforces the need for long-term medical and social care for survivors to reduce the risk of disease re-emergence and prevent further stigmatization.

Main text

At least 30 outbreaks of Ebola virus disease (EVD) have been identified since the late 1970s, the deadliest hitting Guinea, Sierra Leone and Liberia from December 2013 to June 2016.[1,2] Guinea faced a new EVD outbreak in 2021, which started in Gouéké, a town about 200 km away from the epicenter of the 2013-2016 outbreak. The probable index case was a 51-year-old nurse, assistant of the hospital midwife in Gouéké. On the 21st of January 2021, she was admitted to the hospital of Gouéké suffering from headache, asthenia, nausea, anorexia, vertigo, and abdominal pain. She was diagnosed with malaria and salmonellosis and released two days later. Feeling ill again once at home, she attended a private clinic in N'zérékoré, at 40 km, and visited a traditional healer, but died three days later. In the week following her death, her husband and other family members, who attended her funeral, fell ill and four of

them died. They were reported as the first suspect cases by the national epidemic alert system on 11th February. On February 12th, blood was taken from two suspect cases admitted at the hospital in N'zérékoré. On February 13th, both were confirmed with EVD by the laboratory in Guéckédou using a commercial real-time reverse transcription Polymerase Chain Reaction (RT-PCR) assay (RealStar Filovirus Screen Kit, Altona Diagnostics). On February 13th, the husband of the index case, who travelled for treatment from Gouéké to Conakry, the capital city of Guinea (>700 km distance), was admitted to the Centre de Traitement Epidémiologique (CTEpi) in Nongo, Ratoma Commune. He presented with fever, nausea, asthenia of abdominal and lumbar pain and was considered highly suspicious for EVD. A blood sample was analyzed on the same day and found positive for Ebola Zaire (*Zaire ebolavirus*; EBOV) on the GeneXpert molecular diagnostic platform (Xpert Ebola test, Cepheid) and by an in house qRT-PCR. The laboratory confirmation of EVD in the three suspect cases led to the official declaration of the epidemic on February 14th. At 5th March, 14 confirmed cases and 4 probable cases of EVD have been identified, leading to 9 deaths including five confirmed cases as reported by the Agence Nationale de la Sécurité Sanitaire (ANSS) of Guinea. After a period of 25 days without new cases, two new cases have been reported around N'zérékoré on April 1st and 3th and on 19th June 2021, the outbreak was declared over. Overall, 16 confirmed cases were reported, among them 12 people died.

Genomic characterization of the virus causing the 2021 Guinean EVD epidemic was of immediate public health importance. First, because diagnostic tools, therapeutics and vaccines, with proven effectiveness in recent EVD outbreaks, i.e. in Guinea (2013-2016) and in the Equateur and North-Kivu/Ituri provinces of the Democratic Republic of Congo (2018-2020), have primarily been developed for EBOV.[3-5] Secondly, to identify whether the outbreak resulted from a new zoonotic transmission event or from the resurgence of a viral

111 strain that had circulated in a previous EBOV outbreak - EBOV can persist in body fluids of
112 EVD survivors and be at the origin of new transmission chains.[6-8] Although the Xpert Ebola
113 test has been developed to detect only EBOV strains and the in-house qRT-PCR assay uses a
114 probe specifically designed to detect EBOV [9], additional confirmation by sequence analysis
115 was sought by targeting a short fragment in the viral protein (VP) 35 region on the sample
116 from the patient hospitalized in Conakry. The phylogenetic tree (Supplementary Figure 1)
117 underscores that this highly conserved region can discriminate between Ebola virus species
118 and confirmed that the new strain belongs to the species *Zaire ebolavirus* (EBOV). This
119 confirmed that available vaccines and the vast majority of molecular diagnostics tools and
120 therapeutics could be immediately applied.
121
122 To gain further insight into the genomic make-up of the viruses causing this outbreak, eleven
123 complete or near-complete (>95% recovery), eight partial (>65% recovery) genomic
124 sequences from 12 of all 14 confirmed cases were obtained by three different laboratories
125 using different next generation sequencing technologies (Table 1). To facilitate the public
126 health response and evaluation of existing medical countermeasures, sequencing results were
127 made publicly available on March 12th through joint posting
128 (https://virological.org/c/ebolavirus/guinea-2021/44). Blood and swab samples from 14
129 confirmed EVD patients, sampled from February 12th to March 4, were processed by the
130 following methods; i.e. hybridization capture technology and sequencing on Illumina
131 iSeq100, amplicon-based protocol with EBOV-specific primer pools and sequencing on
132 MinIon (Oxford Nanopore Technologies, Oxford, UK) and a hybrid-capture based approach
133 using a probe panel that included Ebola virus (EBOV) specific targets followed by TruSeq
134 Exome Enrichment, as previously described.[5] Data generated between the three groups were
135 pooled and the sequence with the highest quality was chosen for each patient. This allowed us

to reconstruct twelve high quality EBOV genomes which cover 82.9%-99.9% of the reference genome (KR534588) (Table 1). The consensus EBOV sequences with the highest genome recovery (>82.9%) from 12 different patients were used in further analyses.

Maximum likelihood phylogenetic reconstruction places the 12 genomes from the 2021 Guinea outbreak as a single cluster among the EBOV viruses responsible for the 2013-2016 EVD outbreak in West Africa (Figure 1 and Figure 2). The 2021 genomes share 10 substitutions accumulated during the 2013-2016 outbreak (compared to KJ660346), including the A82V marker mutation for human adaptation in the glycoprotein that arose when the virus spread to Sierra Leone.[11,12] These patterns provide strong evidence for direct linkage to human cases from the 2013-2016 outbreak rather than a new spillover from an animal reservoir. The 2021 lineage is nested within a clade that predominantly consists of genomes sampled from Guinea in 2014 (Figure 2). The branch by which the 2021 cluster diverges from the previous outbreak exhibits only 12 substitutions, which is far fewer than expected from EBOV evolution during 6 years of sustained human-to-human transmission (Figure 3). Using a local molecular clock analysis, we estimate a 6.4-fold (95% Highest Posterior Density interval (HPD): 3.3-fold,10.1-fold) lower rate along this branch. For comparison, we also estimate a 5.5-fold (1.6-fold,10.8-fold) lower rate along the branch leading to the 2016 flare-up that was linked to a survivor with virus persistence for more than 500 days.[7,13] Rather than a constant long-term low evolutionary rate, some degree of latency or dormancy during persistent infection seems a more likely explanation for the low divergence of the 2021 genomes. We tested whether the 12 genomes from 2021, sampled over a time period of less than one month, contained sufficient temporal signal to estimate the time to most recent common ancestor (tMRCA) (Supplementary Figure 2), but did not identify statistical support for sufficient divergence accumulation over this short time scale. We therefore calibrated our analysis using an evolutionary rate that reflects EBOV evolution under sustained human-to-

161 human transmission (as estimated by the local molecular clock analysis). This resulted in a
162 tMRCA estimate of January 22nd 2021 [95% Highest Posterior Density interval: December
163 29th, 2020; February 10th, 2021].
164
165 These results open a new perspective on the relatively rare observation of EBOV re-
166 emergence. It is assumed that all known filovirus outbreaks in humans are the result of
167 independent zoonotic transmission events from bat reservoir species or from intermediate or
168 amplifying hosts like apes and duikers.[6] Here we clearly show that, even almost five years
169 after the declaration of the end of an epidemic, new outbreaks could also be the result of
170 transmission from humans infected during a previous epidemic. The viruses from the 2021
171 outbreak fall within the lineage of EBOV viruses obtained from humans during the 2014-
172 2016 outbreak, therefore it is thus very unlikely that this new outbreak has an animal origin or
173 is the result from a new cross-species transmission with the same lineage that remained latent
174 in this natural host which in that scenario would be at the basis of the west african cluster. The
175 limited genomic divergence between 2014-2015 and 2021 is compatible with a slow long-
176 term evolutionary rate. However, a relatively long phase of latency may be more likely than
177 continuous slow replication. Independent of the mechanistic explanation, the virus most likely
178 persisted at low level in a human survivor. Plausible scenarios of EBOV transmission to the
179 index case include (i) sexual transmission by exposure to EBOV in semen from a male
180 survivor, (ii) contact to body fluids from a survivor with relapse to symptomatic EVD, for
181 example during health care – the index case was a healthcare worker, or (iii) relapse of EVD
182 disease in the index case, although she was not a known survivor, she may have had an
183 asymptomatic or pauci-symptomatic EBOV infection during the previous outbreak. Detailed
184 investigation by anthropologists on the family of the index case revealed that she was not
185 known as an EVD survivor, nor her husband or close relatives. However, among more

distantly related family, 25 individuals had EVD during the previous outbreak. Only five survived, but the index case apparently had no recent contacts with this part of the family. Consultation of the hospital registers in Gouécké, showed that all patients seen by the index case in January 2021 were in good health and were still in good health in March 2021. However, the index case also performed informal consultations outside the hospital environment which could not be verified. Alternatively, the nurse was not the actual index case but part of a small unrecognized chain of human-to-human transmission in this area of Guinea. However, the diversity of the currently available genomes is limited and molecular clock analysis suggests a recent time to the most recent common ancestor, with a mean estimate close to the time the nurse was first hospitalized and 95% HPD boundary around the turn of the year. This provides some reassurance that the outbreak was detected early.

The 2013-2016 outbreak in West Africa was the largest and most complex outbreak of EBOV with more than 28,000 cases, 11,000 deaths and an estimated 17,000 survivors, notably in Guinea, Liberia and Sierra Leone.[2] The large outbreak provided new information about the disease itself as well as about the medical, social and psychological implications for EVD survivors.[14-16] It was also possible to estimate to some extent the proportions of asymptomatic or pauci-symptomatic infections and to identify their role in specific unusual transmission chains.[17-19] While the main route of human-to-human EBOV transmission is direct contact with infected body fluids from symptomatic or deceased patients, some transmission chains were associated with viral persistence in semen.[3] Several studies demonstrated viral persistence in more than 50% of male survivors at 6 months after discharge from Ebola Treatment Units (ETU) and the maximal duration of persistence in semen has been reported to last up to 500 - 700 days post ETU discharge in a handful of male EVD survivors.[9,20-22] Transmission through other body fluids (breast milk, cervicovaginal fluids) is also

suspected.[8,23-25] Furthermore, some immunological studies among survivors suggest a continuous or intermittent EBOV antigenic stimulation due to persistence of an EBOV reservoir in some survivors [26,27], although this was not confirmed in another study. [28] Cases of relapse of EVD have also been sporadically reported and can be at the origin of large transmission chains as recently reported in the North-Kivu outbreak in DRC.[29] For example, we recently reported presence of EBOV RNA in breast milk 500 days after ETU discharge in a woman who was not pregnant when she developed EVD. She attended the hospital due to complications at 8 months of pregnancy and a breast milk sample taken 1 month after delivery tested positive for EBOV RNA.[9] These examples illustrate that health care workers can be exposed to EBOV when taking care of patients who survived EVD, but have an unrecognized relapse of their infection. The 2021 outbreak now highlights that viral persistence and reactivation is not limited to a two-year period but can also occur much longer with late reactivation.

Active genomic surveillance already showed resurgence of previous strains in other outbreaks. For example, two EBOV variants circulated simultaneously within the same region during the recent 2020 outbreak in the Equateur Province, DRC.[30] Moreover, strains from the two consecutive outbreaks in Luebo, Democratic Republic of Congo (DRC), in 2007 and 2008, are also so closely related that it now appears difficult to exclude that the epidemic observed in 2008 was possibly due to a resurgence event from a EVD survivor from the 2007 outbreak.[31,32] However, the limited genomic sampling does not allow for a formal test of this hypothesis.

While the majority of EVD outbreaks remained limited both in number of cases and geographic spread, the two largest outbreaks in West Africa (December 2013 to June 2016) and in Eastern DRC (August 2018 to June 2020) infected thousands of individuals over wide

236 geographic areas leading to large numbers of EVD survivors. This means that the risk of
237 resurgence is higher than ever before. Continued surveillance of EVD survivors is therefore
238 warranted to monitor reactivation and relapse of EVD infection and potential presence of
239 virus in body fluids. This work and associated communications have to be conducted with the
240 utmost care towards the well-being of EVD survivors. During the 2013-2016 Ebola outbreak
241 in Guinea, Ebola survivors had a mixed experience after discharge from ETUs. On the one
242 hand, they were considered as heroes by NGOs and became living testimonies of a possible
243 recovery.[33,34] On the other hand, they experienced different forms of stigmatization such as
244 rejection by family and friends, refusal to be involved in collective work, loss of jobs and
245 housing and sometimes self-isolation from social life and workplaces.[35] The human origin of
246 the current EVD outbreak and the associated shift in our perception of EBOV emergence, call
247 for careful attention to survivors. The concrete danger that survivors will be stigmatized as a
248 source of danger should be a matter of scrupulous attention.[36] This is especially true for the
249 area of Gouécké which is only 9 km away from Womey, a village emblematic of the violent
250 reaction of the population toward the EVD response team during the 2013-2016 epidemic.[37]
251
252 Since the 2013–2016 EVD outbreak in Western Africa, genome sequencing became a major
253 component of the outbreak response.[10,38-41] Establishment of in-country sequencing and
254 capacity building allowed for a timely characterization of EBOV strains in this new outbreak
255 in Guinea. In addition to the importance of appropriate health care measures focused on
256 survivors, late resurgence also highlights the urgent need for further research on potent
257 antiviral agents that can eradicate the latent virus reservoir in EVD patients and on efficient
258 vaccines with long-term protection. In parallel, vaccination could also be considered to boost
259 protective antibody responses in survivors.[27] Vaccination of populations in areas with
260 previous EBOV outbreaks could also be promoted to prevent secondary cases.

References

1. Baize, S. et al. Emergence of Zaire Ebola virus disease in Guinea**.** *N Engl J Med* 371:1418-25 (2014).

2. World Health Organization. 2016 Situation Report: Ebola Virus Disease. Geneva: World Health Organization; 10 june 2016

3. Malvy,D. , McElroy, A.K., de Clerck, H., Günther, S., van Griensven, J. Ebola virus disease. *Lancet*. 393:936-948 (2019).

4. Mbala-Kingebeni, P. et al. Rapid Confirmation of the Zaire Ebola Virus in the Outbreak of the Equateur Province in the Democratic Republic of Congo: Implications for Public Health Interventions. *Clin Infect Dis*. 68:330-333 (2019)

5.Mbala-Kingebeni, P. et al. 2018 Ebola virus disease outbreak in Equateur Province, Democratic Republic of the Congo: a retrospective genomic characterization**.** *Lancet Infect Dis* 19: 641-647 (2019)

6. Pigott, D.M. et al. Mapping the zoonotic niche of Ebola virus disease in Africa. *Elife*. 3: e04395 (2014)

7. Diallo, B. Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus Persistence in Seminal Fluid for More Than 500 Days. *Clin Infect Dis* 63: 1353-1356 (2016)

8. Sissoko, D. et al. Ebola Virus Persistence in Breast Milk After No Reported Illness: A Likely Source of Virus Transmission From Mother to Child. *Clin Infect Dis* 64: 513-516 (2017)

9. Keita, A.K. et al. A 40-Month Follow-Up of Ebola Virus Disease Survivors in Guinea (PostEbogui) Reveals Long-Term Detection of Ebola Viral Ribonucleic Acid in Semen and Breast Milk. *Open Forum Infect Dis* 6: ofz482 (2019)

10. Quick, J. et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530, 228–232 (2016).

11. Diehl, W. E. et al. Ebola Virus Glycoprotein with Increased Infectivity Dominated the 2013-2016 Epidemic. *Cell* 167, 1088–1098.e6 (2016).

12. Urbanowicz, R. A. et al. Human Adaptation of Ebola Virus during the West African Outbreak. *Cell* 167, 1079–1087.e5 (2016).

13. Blackley, D.J. et al. Reduced evolutionary rate in reemerged Ebola virus transmission chains. *Sci Adv* 2:e1600378 (2016)

14. Diallo, M.S.K. et al. Understanding the long-term evolution and predictors of sequelae of Ebola virus disease survivors in Guinea: A 48-month prospective, longitudinal cohort study (PostEboGui). *Clin Infect Dis*:ciab168 (202). doi: 10.1093/cid/ciab168. Online ahead of print.

15. Etard, J.F. et al. Multidisciplinary assessment of post-Ebola sequelae in Guinea (Postebogui): an observational cohort study. *Lancet Infect Dis* 17:545-552 (2017)

16. PREVAIL III Study Group. A Longitudinal Study of Ebola Sequelae in Liberia. *N Engl J Med* 380:924-934 (2019)

17. Diallo, M.S.K. et al. Prevalence of infection among asymptomatic and paucisymptomatic contact persons exposed to Ebola virus in Guinea: a retrospective, cross-sectional observational study. *Lancet Infect Dis* 19:308-316 (2019)

18. Glynn, J.R. et al. Asymptomatic infection and unrecognised Ebola virus disease in Ebola-affected households in Sierra Leone: a cross-sectional study using a new non-invasive assay for antibodies to Ebola virus. *Lancet Infect Dis* 17:645-653 (2017)

19. Camara, I. et al. Unrecognized Ebola virus infection in Guinea: complexity of surveillance in a health crisis situation: case report. *Pan Afr Med J* 36:201 (2020)

20. Thorson, A.E. et al. Persistence of Ebola virus in semen among Ebola virus disease survivors in Sierra Leone: A cohort study of frequency, duration, and risk factors. *PLoS Med* 18:e1003273 (2021)

21. Fischer, W.A. et al. Ebola Virus Ribonucleic Acid Detection in Semen More Than Two Years After Resolution of Acute Ebola Virus Infection. *Open Forum Infect Dis* 4:ofx155 (2017)

22. Sissoko, D. et al. Persistence and clearance of Ebola virus RNA from seminal fluid of Ebola virus disease survivors: a longitudinal analysis and modelling study. *Lancet Glob Health* 5:e80-e88 (2017).

23. Mate, S.E. et al. Molecular Evidence of Sexual Transmission of Ebola Virus**.** *N Engl J Med* 373:2448-54 (2015)

24. Dokubo, E.K. et al. Persistence of Ebola virus after the end of widespread transmission in Liberia: an outbreak report. *Lancet Infect Dis* 18:1015-1024 (2018)

25. Liu, W.J. et al. Comprehensive Clinical and Laboratory Follow-up of a Female Patient With Ebola Virus Disease: Sierra Leone Ebola Virus Persistence Study. *Open Forum Infect Dis* 6:ofz068 (2019)

26. Wiedemann, A. et al. Long-lasting severe immune dysfunction in Ebola virus disease survivors. *Nat Commun* 11:3730 (2020)

27. Adaken, C. et al. Ebola virus antibody decay-stimulation in a high proportion of survivors. *Nature* 590:468-472 (2021)

28. Thom, R. et al. Longitudinal antibody and T cell responses in Ebola virus disease survivors and contacts: an observational cohort study. *Lancet Infect Dis* 21:507-516 (2021)

29. Mbala Kingebeni, P. et al. Ebola Virus Transmission Initiated by Relapse of Systemic Ebola Virus Disease. *N Engl J Med* 384:1240-1247 (2021)

30. Muyembe-Tamfum, J.J. et al. Two Ebola virus variants circulating during the 2020 Equateur Province outbreak. https://virological.org/t/two-ebola-virus-variants-circulating-during-the-2020-equateur-province-outbreak/538.

31. Leroy, E.M. et al. Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne Zoonotic Dis* 9:723-8 (2009)

32. Grard, G. et al. Emergence of divergent Zaire ebola virus strains in Democratic Republic of the Congo in 2007 and 2008. *J Infect Dis* 204 (Suppl 3):S776-84 (2011)

33. Enria, L. What Crisis Produces: Dangerous Bodies, Ebola Heroes and Resistance in Sierra Leone' *'BathPapers in International Development and Wellbeing*, Centre for Development Studies, University of Bath no. 53 (2017)

34. Ronse, M. et al. What motivates Ebola survivors to donate plasma during an emergency clinical trial? The case of Ebola-Tx in Guinea. *PLoS Negl Trop Dis* 12: e0006885 (2018)

35. Sow, S., Desclaux, A., Traverne, B. Ebola en Guinée : formes de la stigmatisation des acteurs de Sante survivants. *Bulletin de la Société de pathologie exotique* 109: 309-313 (2016)

36. Farmer, P. Aids and Accusation. Haiti and the Geography of Blame. *University of California Press* (2006.)

384

385 37. Wilkinson, A., Fairhead, J. Comparison of social resistance to Ebola response in Sierra
386 Leone and Guinea suggests explanations lie in political configurations not culture. *Crit Public*
387 *Health*. 27:14-27 (2017)

388

389 38. Gire, S. K. et al. Genomic surveillance elucidates Ebola virus origin and transmission
390 during the 2014 outbreak. Science 345, 1369–1372 (2014)

391

392 39. Arias, A. et al. Rapid outbreak sequencing of Ebola virus in Sierra Leone identifies
393 transmission chains linked to sporadic cases. Virus Evol 2:vew016 (2016)

394

395 40. Pini, A. et al. Field investigation with real-time virus genetic characterisation support of a
396 cluster of Ebola virus disease cases in Dubréka, Guinea, April to June 2015. *Euro Surveill*
397 23:17-00140 (2018)

398

399 41. Mbala-Kingebeni, P. et al. Medical countermeasures during the 2018 Ebola virus disease
400 outbreak in the North Kivu and Ituri Provinces of the Democratic Republic of the Congo: a
401 rapid genomic assessment. Lancet Infect Dis 19: 648-657 (2019

402

403 42. Guindon, S. et al. New algorithms and methods to estimate maximum-likelihood
404 phylogenies: assessing the performance of PhyML 3.0. Syst Biol 59: 307-321 (2010)..

405

406 43. Nguyen, L.T., Schmidt, H.A., von Haeseler, A., Minh, B.Q. IQ-TREE: a fast and
407 effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol
408 32: 268-274 (2015).

44. Li, H. et al. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079 (2009)

45. Weidmann M, Mühlberger E, Hufert FT. Rapid detection protocol for filoviruses. J Clin Virol 30:94-9 (2004)

46. Katoh, K., Asimenos, G.,Toh, H. Multiple alignment of DNA sequences with MAFFT. Methods Mol Bio. 537: 39–64 (2009).

47. Dudas, G. et al. Virus genomes reveal factors that spread and sustained the Ebola epidemic. Nature 544: 309–315 (2017).

48. Minh, B. Q. et al. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol Evol 37:1530–1534 (2020).

49. Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol 2, vew007 (2016).

50. Suchard, M.A. et al. Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol 4, vey016 (2018).

51. Worobey, M., Han, G.Z., Rambaut, A.A. Synchronized global sweep of the internal genes of modern avian influenza virus. Nature 508: 254–257 (2014).

52. Gill, M. S. et al. Improving Bayesian population dynamics inference: a coalescent-based model for multiple loci. Mol Biol Evol 30, 713–724 (2013).

53. Duchene, S., et al. Temporal signal and the phylodynamic threshold of SARS-CoV-2. Virus Evol 6:veaa061 (2020)

Acknowledgements

The authors thank the Agence Nationale de Sécurité Sanitaire (ANSS) and the Ministry of Health of the Republic of Guinea, the healthcare workers (medical doctors, nurses and laboratory technicians) from the treatment centers in Nzérékoré and Nongo, Conakry; the laboratory personnel in Guékédou, Institut National de Santé Publique (INSP, Conakry) and Centre de Recherche et Formation en Infectiologie de Guinéé (CERFIG, Conakry). CERFIG also thanks Jan Gogarten and the bioinformatics support of RKI. UK Health Security Agency would like to thank Oxford Nanopore for the donation of reagents and equipment to support the setting up of the sequencing capacity at PFHG. BNITM thanks the ARTIC network (https://artic.network/).

Funding

The work of CERFIG and TransVIHMI was supported in part by grants EBO-SURSY project funded by European Union, International Mixt Laboratory "RESPIRE" of IRD (Institut de Recherche pour le Developpement), Montpellier Université d'Excellence (EBOHEALTH; I-Site MUSE, ANR-16-IDEX-0006) and Institut National de la Santé et de la Recherche Médicale (INSERM)/REACTing (REsearch and ACTion targeting emerging infectious diseases). The work of RKI was partly funded by the German Ministry of Health "Global Protection Program" project TRICE. Frédéric Le Marcis received funding through the program « EBOVAC3 Bringing a prophylactic Ebola vaccine to licensure » funded by IMI (grant agreement number 800176) and run by LSTHM and INSERM.

The work of PFHG and BNITM was supported by the German Federal Ministry of Health through support of the WHO Collaborating Centre for Arbovirus and Hemorrhagic Fever Viruses at the Bernhard-Nocht-Institute for Tropical Medicine (agreement ZMV I1-2517WHO005), and through the Global Health Protection Programme (GHPP, agreements

ZMV I1-2517GHP-704 and ZMVI1-2519GHP704), and by the Coalition for Epidemic Preparedness Innovations (CEPI). Work of BNITM and the UK Health Security Agency has been further supported by the Research and Innovation Programme of the European Union under Horizon 2020 grant agreement no. 871029-EVA-GLOBAL.

The European Mobile Laboratory (EMLab) coordinated by BNITM is a technical partner of the WHO Global Outbreak Alert and Response Network (GOARN) and the deployment of EMLab experts and sequencing capacities to Guinea has been coordinated and supported by the GOARN Operational Support Team at WHO/HQ and the WHO country office in Guinea. The research leading to these results has received funding from the European Research Council under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 725422-ReservoirDOCS). The Artic Network receives funding from the Wellcome Trust through project 206298/Z/17/Z. PL acknowledges support by the Research Foundation -- Flanders (`Fonds voor Wetenschappelijk Onderzoek -- Vlaanderen', G066215N, G0D5117N and G0B9317N).

Author contributions

Conceived and designed the study: A.A.S, A.K.K, A.To., E.D., F.H.L, F.R.K., L.E.K., M.Pe, N.M., O.F, S.D., S.G.

collected data and/or performed and/or performed medical examinations and/or laboratory diagnostics: A.B., A.Do., A.K.K, A.K.S., A.M.S., A.S., B.S., C.D.S, D.D., F.L, F.R.K., F.Y.S., F.A.T, F.B.S, G.A., H.D., I.C., J.A.B., J.C. ,K.I. ,Kb.K., Kr.K., L.M., M.B., M.B.K., M.C., M.D.B., M.K., M.Po., M.S.S., N.V., N.J.V., S.D., Y.S.

Performed sequencing and/or sequence validation: A.A., A.Du., A.Re., B.S., H.D., J.A.B., J.H., K.I., K.P, M.F, M.M.D, S.C-S

Formal phylogenetic analysis: A.A., A.Ra., J.H., J.T.M, L.E.K., M.R.W, P.L., S.C-S, S.D.

Funding acquisition: E.D., F.H.L, M.Pe., M.W.C, S.D., S.G.

Provided reagents: F.H.L., K.P., M.R.W, S.C-S.

Project implementation: A.K.K., A.L., A.Th., A.To., C.P, E.D., F.H.L, G.A.K, J.Q., M.Dia., M.Pa., M.Pe, M.S.B, M.W.C, N.M., N.T, P.F., S.D., S.G., S.K., S.T.P.

Wrote the first draft of the manuscript: A.Du, A.K.K, M.Pe, S.C-S.

wrote and edited the manuscript: A.A.S, A.Du., A.K.K, A.M.S., A.Re., A.Ra, A.To., E.D., F.H.L, F.L, L.E.K., M.F, M.M.D, M.Pe, N.M., O.F, P.L., S.C-S, S.D., S.G.

All authors read and approved the contents of the manuscript.

Competing interests: MWC received materials for this study from Oxford Nanopore. All other authors declare no competing interests.

Figure legends.

Figure 1. **Maximum likelihood phylogenetic reconstruction for 55 representative genomes from previous outbreaks of Zaire ebolavirus and 12 genomes from the 2021 outbreak in Guinea.** Most clades for single or multiple closely related outbreaks are collapsed and internal node support is proportional to the size of the internal node circles. The clades or tip circles are labelled with the locations and years of the outbreaks, and colored according to the (first) year of detection.

Figure 2. **Maximum likelihood phylogenetic reconstruction for 1065 genomes sampled during the 2013-2016 West African outbreak and 12 genomes from the 2021 outbreak in Guinea.** A color gradient is used to color the tip circles. The 2021 genomes are shown with a larger circle in yellow.

Figure 3. **Temporal divergence plot of genetic divergence from the root against time of sampling for the tree shown in Figure 2.** The regression is exclusively fitted to genomes sampled between 2014 and 2015. The same colors are used for the data points as in Figure 2. The dashed yellow lines highlight how the 2021 data points deviate from the relationship between sampling time and sequence divergence. According to this relationship, about 95 substitutions (95% prediction interval: 88-101) are expected on the branch ancestral to the 2021 cluster, whereas only 12 are inferred on this branch.

Table 1. Patient and sample characteristics and sequencing results obtained by the laboratories involved in the study.

				CERFIG		PFHG		IPD	
Patient	Sex	Age (years)	Date of sampling	Ct value[1]	% genome recovery	Ct value[1]	% genome recovery	Ct value[1]	% genome recovery
1	F	54	12-Feb-2021[a]	–[2]	–	22.4	87.8	29.3	99.6
			19-Feb-2021[b]	33.1	0.3	-	–	–	–
2	F	70	12-Feb-2021[a]	–	–	25.9	67.8	37.1	98.7
3	M	61	13-Feb-2021[a]	29.4	5.0	neg.	7.5	neg.	2.5
4	M	46	20-Feb-2021[b]	24.3	12.6	–	–	–	–
5	M	22	22-Feb-2021[a]	32.5	99.4	23.2	93.3	–	–
6	M	65	23-Feb-2021[b]	–	–	20.5	97.3	–	–
7	F	75	26-Feb-2021[a]	–	–	19.5	95.5	–	–
8	M	29	26-Feb-2021[a]	–	–	18.8	98.1	–	–
9	**M**	**32**	**26-Feb-2021[a]**	**24.6**	**99.9**	**27.8**	**77.6**	**–**	**–**
10	F	30	26-Feb-2021[a]	26.0	99.7	23.0	82.2	–	–
11	F	55	26-Feb-2021[a]	36.4	75.4	28.8	82.9	–	–
12	M	45	26-Feb-2021[a]	–	–	20.5	96.1	–	–
13	M	42	26-Feb-2021[a]	25.0	99.9	29.3	70.4	–	–
14	F	40	4-Mar-2021[a]	–	–	22.0	97.5	–	–

[1] Ct value was measured in the sequencing laboratory before starting the sequencing process.

[2] Sample was not tested in this laboratory

The values differ between the laboratories due to possible degradation of the sample or the RNA during transport and storage.

[a] patient samples at diagnosis

[b] follow-up samples from patients

All patient samples were whole EDTA blood except for patient 12 for whom a swab was used

Materials and Methods

Ethics Statement

Diagnostic specimens were collected as part of the emergency response from the Ministry of public health from Guinea, and therefore consent for sample collection was waived. All preparation of samples for sequencing, genomic analysis and data analysis was performed on anonymized samples identifiable only by their laboratory or epidemiological identifier.

Confirmation of Ebola virus species by sequence analysis of VP 35 fragment at CERFIG

Viral RNA was extracted from 140 ul of whole blood collected from the samples from the patient hospitalized in Conakry, with the Nuclisens kit (Biomerieux, France) and following manufacturer's instructions. Amplification of a small fragment of VP35 region was attempted in a semi-nested PCR with a modified protocol as previously described.[4] First-round VP35 PCR-products from positive samples were barcoded and pooled using the Native Barcoding Kit EXP-NBD104 (Oxford Nanopore Technologies, Oxford, UK). Sequencing libraries were generated from the barcoded products using the Genomic DNA Sequencing Kit SQK-LSK109 (Oxford Nanopore Technologies) and were loaded onto a R9 flow cell on a MinIon (Oxford Nanopore Technologies). Genetic data were collected for one hour. Basecalling, adapter removal and demultiplexing of fastq files were performed with MinKNOW, version 4.1.22. Fastq reads >Q11 were used for mapping a virus database with the Genome Detective tool (https://www.genomedetective.com/app/typingtool/virus/). The generated consensus sequence was used for further analysis. For phylogenetic inference, we retrieved one sequence per outbreak from the Haemorhagic Fever Virus (HFV) database to which we added the newly generated VP35 sequence of the novel outbreak. Phylogenetic analyses were done using Maximum Likelihood methods using IQ-Tree with 1,000 bootstraps for branch

support.[42,43] The GTR model plus a discrete gamma distribution were used as nucleotide substitution models.

Full-length genome sequencing of the new Ebola viruses

Genome sequencing at CERFIG

Whole genome sequencing was attempted on viral extracts for samples that were positive for EBOV NP and GP on the GeneXpert molecular diagnostic platform (Xpert Ebola Assay) with the glycoprotein (GP) and the Nucleoprotein (NP) of the Zaire Ebola virus. We extracted full nucleic acid using the QIAamp® Viral RNA Mini Kit (Qiagen). After DNase treatment with TURBO DNA-free™ Kit (Ambion) and clean-up with RNA Clean & Concentrator Kit (Zymo Research), RNA was converted to double-stranded cDNA (ds-cDNA) using the SuperScript™ IV First-Strand Synthesis System (Invitrogen) and NEBNEXT® mRNA Second Strand Synthesis Module (New England Biolabs). Resulting ds-cDNA was enzymatically fragmented with NEBNext® dsDNA Fragmentase® (New England Biolabs) and converted to dual indexed libraries with the NEBNext® Ultra™ II DNA Library Prep Kit for Illumina® (New England Biolabs) and NEBNext® Multiplex Oligos for Illumina® (New England Biolabs). To enrich EBOV in the libraries, we performed two rounds of hybridization capture (16 hours at 65°C) with custom made biotinylated RNA baits (120 nucleoties, 2-fold tiling; Arbor Biosciences) covering representative genomes for *Zaire ebolavirus* (KC242801), *Sudan ebolavirus* (KC242783), *Reston ebolavirus* (NC_004161), *Taï Forest ebolavirus* (NC_014372), *Bundibugyo ebolavirus* (KC545395) and *Marburg marburgvirus* (FJ750956), following the myBaits Hybridization Capture for Targeted NGS protocol (Version 4.01). After the second round, capture products were quantified using the Qubit 3.0 Fluorometer with Qubit™ dsDNA HS Assay Kit (Invitrogen), and pooled equimolarly for sequencing on an Illumina iSeq using iSeq 100 i1 Reagents (2 x 150-cycle).

Sequencing reads were filtered (adapter removal and quality filtering) with Trimmomatic (Bolger, 2014) (settings: LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30 MINLEN:40), merged with ClipAndMerge (https://github.com/apeltzer/ClipAndMerge), and mapped to the *Zaire ebolavirus* RefSeq genome (NC_002549) using BWA-MEM.[44] Mapped reads were sorted and deduplicated with SortSam and MarkDuplicates from the Picard suite (Broad Institute, Picard; http://broadinstitute.github.io/picard). We generated consensus sequences using Geneious Prime 2020.2.3 (https://www.geneious.com) where unambiguous bases were called when at least 90% of at least 20 unique reads were in agreement (20x, 90%). For samples with few mapped reads (0001, 0002, 0010, 0030), we also called a consensus at 2X, 90% and 5X, 90%.

Genome sequencing at PFHG

Sequencing at PFHG was performed using a mobile MinION facility deployed by BNITM to Guinea beginning of March 2021. A total of 13 EBOV positive initial diagnostic samples processed at the "Laboratoire des Fièvres Hémorragiques Virales de Gueckédou", the "Laboratoire Régional de l'Hôpital de N'Zérékoré" were used for sequencing. If RNAs from diagnostic procedures performed by the peripheral laboratories was not sent to PFHG, samples were inactivated and RNA was extracted from 50 µl for whole blood EDTA, 70 µl of plasma from EDTA blood or from 140 µl of wet swabs using the QIAamp Viral RNA Mini Kit (Qiagen) following the manufacturer's instructions. Tiled primers generating overlapping products combined with a highly multiplexed PCR protocol were used for the amplicon generation.[10] At start of deployment, three different primers pools (V3 or pan_10_EBOV, V4 or pan_EBOV and Zaire-PHE or EBOV-Zaire-PHE) were tested and results were combined for optimal recovery of consensus. A new primer pool V5 (EBOV-Makona-V5) was further designed and implemented to increase consensus recovery. Primer pools V3, V4 and V5 were

designed by the ARTIC network and Zaire-PHE primer pools by Public Health England (PHE). For V3, 62 primers were used, while for V4 and V5, 61 primers pairs were used, to amplify products of ~400 nt length. For Zaire-PHE, 71 primer pairs were used to amplify products of ~350 nt length for the ~20 kb viral genome. All primer pools used can be found in Supplementary Table S1. The multiplex PCR was performed as described by the most up-to-date ARTIC protocol for nCoV-2019 amplicon sequencing (nCoV-2019 sequencing protocol V3 (LoCost) V.3 (Artic Network. https://artic.network/ncov-2019), adapted to include the EBOV specific primer sets. Briefly, RNA was directly used for cDNA synthesis using the LunaScript RT SuperMix (New England Biolabs) and the cDNA generated was used as template in the multiplex PCR, which was performed in two reaction pools using Q5 Hot Start DNA Polymerase (New England Biolabs). The resulting amplicons from the two PCR pools were pooled in equal volumes and the pooled amplicons were diluted 1:10 with nuclease-free water.

Sequencing libraries were prepared, barcoded and multiplexed using the Oxford Nanopore Technologies (ONT) Ligation Sequencing Kit (SQK-LSK109) combined with the Native Expansion pack (EXP-NDB104, EXP-NBD114, EXP-NBD196) following the ARTIC Network's library preparation protocol (nCoV-2019 sequencing protocol v3 (LoCost) V.3 (Artic Network. https://artic.network/ncov-2019). For the preparation of less than 11 samples, each sample was prepared in multiples to achieve the library concentration required for sequencing. Briefly, the diluted pooled amplicons were end-repaired using the Ultra II End Prep Module (New England Biolabs) followed by barcode ligation using the Blunt/TA Ligase Master Mix and one unique barcode per sample. Equal volumes from each native barcoding reaction were pooled and subsequently bead cleaned-up using 0.4x AMPure beads. The pooled barcoded amplicons were quantified using the Qubit Fluorometer (Thermo Fisher Scientific) and AMII adapter ligation was performed using the Quick T4 DNA Ligase (New

England Biolabs) followed by an additional bead clean-up. The adaptor ligated barcoded amplicon pool was quantified using the Qubit Fluorometer (Thermo Fisher Scientific) aiming for a minimum recovery of 15 ng sequencing library to load onto the flow cell.

Sequencing libraries were sequenced using R9.4.1 Flow Cells (FLO-MIN106D, ONT) on the Mk1C device (ONT) using MinKNOW version 21.02.2 with real-time high accuracy basecalling and stringent demultiplexing (minimum barcoding score = 60). Within the barcoding options, barcoding on both ends and mid-read barcodes were both switched on. Reads were demultiplexed and binned in a barcode specific folder only if a barcode above the minimum barcoding score was identified on both read ends and if mid-read barcodes were not identified. Sequencing runs were stopped after ~24hr and basecalling was allowed to finish prior to data handling.

Bioinformatics data analysis was done as per ARTIC protocol using a combination of the ARTIC EBOV (Artic Network. https://artic.network/ebov/ebov-bioinformatics-sop.html) and ARTIC SARS-CoV-2 (Artic Network. https://artic.network/ncov-2019/ncov2019-bioinformatics-sop.html) pipelines. A few minor modifications to the ARTIC bioinformatics protocol were incorporated. The two initial steps described, basecalling with guppy and demultiplexing, were skipped as these were both done on the Mk1C device in real-time during the sequencing run, subsequently, the bioinformatics analysis was initiated from the read filtering step (artic guppyplex). Briefly, the artic guppyplex program was used to collect reads for each barcode into a single fastq file, in the presence of a length filter to remove chimeric reads. Reads were filtered based on length with a minimum (option: --min-length) and maximum (option: --max-length) length cut-off based on the amplicon size used (For V3, V4 and V5 primer pools: --min-length 400 and --max-length 700, for Zaire-PHE primer pool: --min-length 350 and --max-length 650). The quality check was skipped since only reads above a quality score of 7 were processed. Following merging and filtering, the arctic minion

pipeline was used to obtain the consensus sequences. The data was normalized to 200 and using the --scheme-directory option the pipeline was directed to the respective primer scheme used for each barcode. Reads were aligned to the NCBI reference KJ660347 (Zaire ebolavirus isolate H.sapiens-wt/GIN/2014/Makona-Gueckedou-C07) for data generated using V3, V4, and V5 primer pools and to NC_002549.1 (Zaire ebolavirus isolate Ebola virus/H.sapiens-tc/COD/1976/Yambuku-Mayinga) for data generated using Zaire-PHE primer pools.

Sequencing at IPD

Viral RNA was extracted from 140 µl of whole blood samples using the QIAamp Viral RNA Mini Kit (Qiagen, Heiden, Germany) according to manufacturer's instructions and eluted in nuclease-free water for a final volume of 60 µl. Extracted RNA was tested by real-time reverse transcription-polymerase chain reaction (RT-PCR) as previously described.[45] Briefly, the DNA library were prepared and enriched using the Illumina RNA Prep with enrichment, (L) Tagmentation kit (Illumina, San Diego, CA, USA) according to the manufacturer's recommendations with a pan viral probe panel that included EBOV specific targets.[5] The purified libraries were pooled and sequenced on the Illumina MiSeq platform using the Miseq reagents kit v3 (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Illumina sequence reads were quality trimmed by Prinseq-lite and consensus EBOV genome sequences were generated using an in-house de novo genome assembly pipeline.

Phylogenetic analysis of full-length genome sequences

Phylogenetic inference

The new EBOV genome sequences were embedded in different data sets for subsequent analyses. For phylogenetic reconstruction, we use a Zaire Ebola virus data set consisting of 55 representative genomes from previous outbreaks and a Makona virus data set consisting of

1065 genomes sampled from Guinea, Sierra Leone and Liberia between 2014 and 2015. Multiple sequence alignment was performed using mafft.[46] We identified 6 T-to-C mutations in the genome from patient 11 that were indicative of mutations induced by adenosine deaminases acting on RNA (ADARs). According to the recommendations by Dudas et al.[47], we masked these positions in this genome in all further analyses. Maximum likelihood trees were reconstructed using IQ-tree under the general time-reversible (GTR) model with gamma (G) distributed rate variation among sites.[48] Temporal divergence plots of genetic divergence from the root of phylogenies against sampling time were constructed using TempEst.[49] To construct the temporal divergence plot for the Guinean 2021 genome data, we used a tree reconstructed under an HKY+G model.

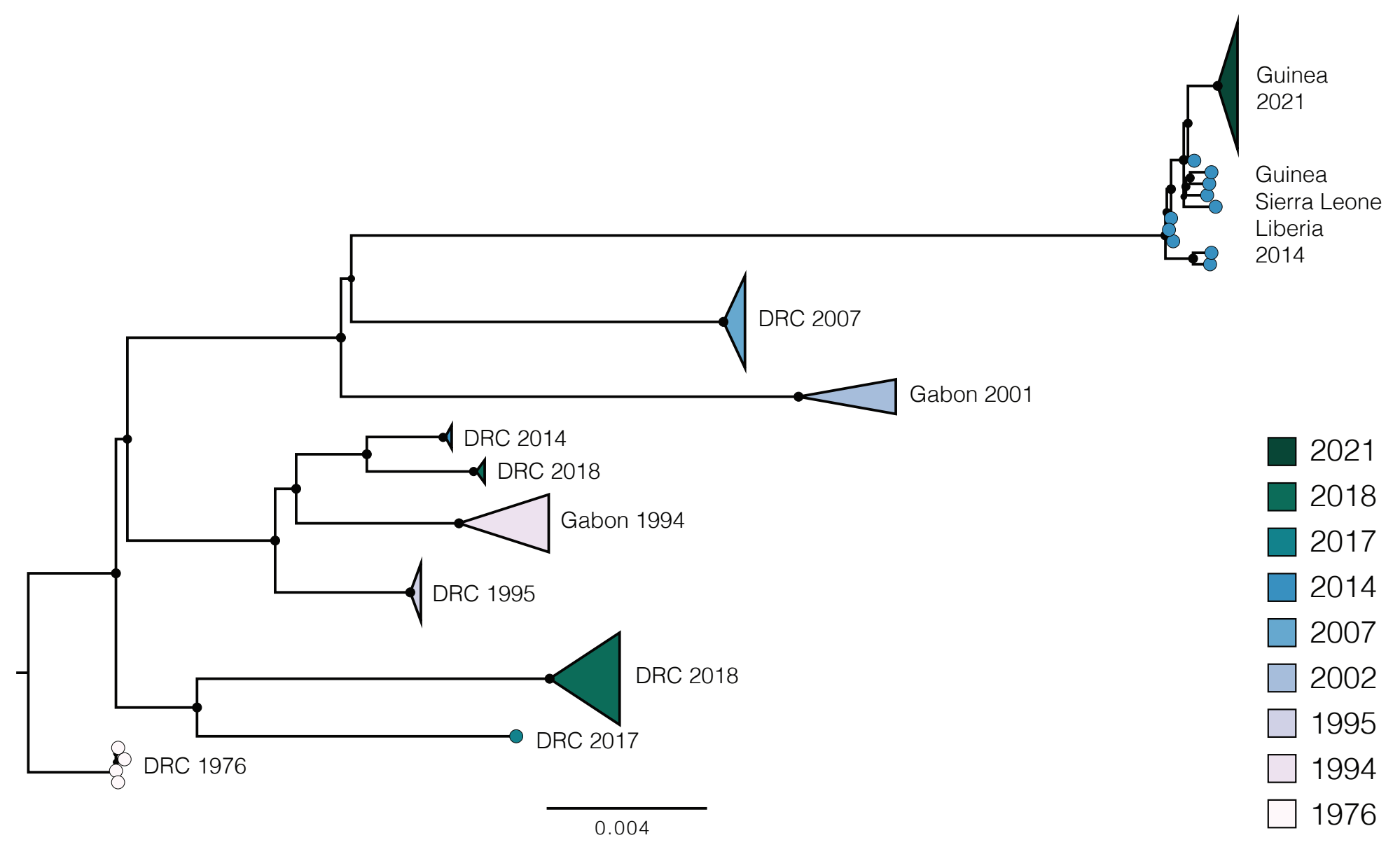
Local molecular clock model analysis

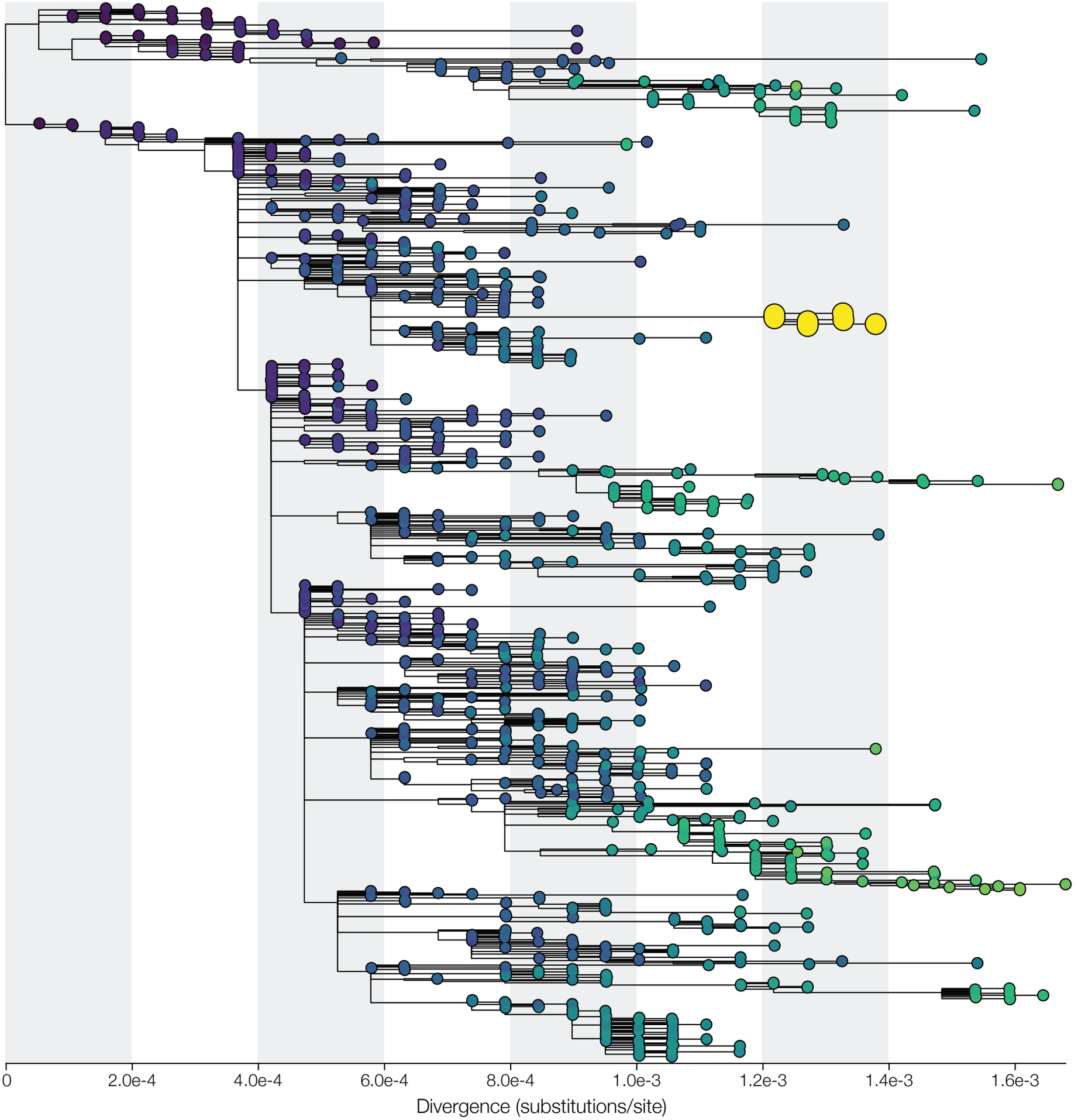
We used BEAST to fit a local molecular clock model to a data set consisting of 1020 dated Makona virus genomes and one of the 2021 genomes (patient 1).[50,51] We specified a separate rate on the tip branch for this genome as well as on the tip branch for a genome in a 2016 flare-up. We used the skygrid coalescent model as a flexible nonparametric tree prior and an HKY+G substitution model.[52]

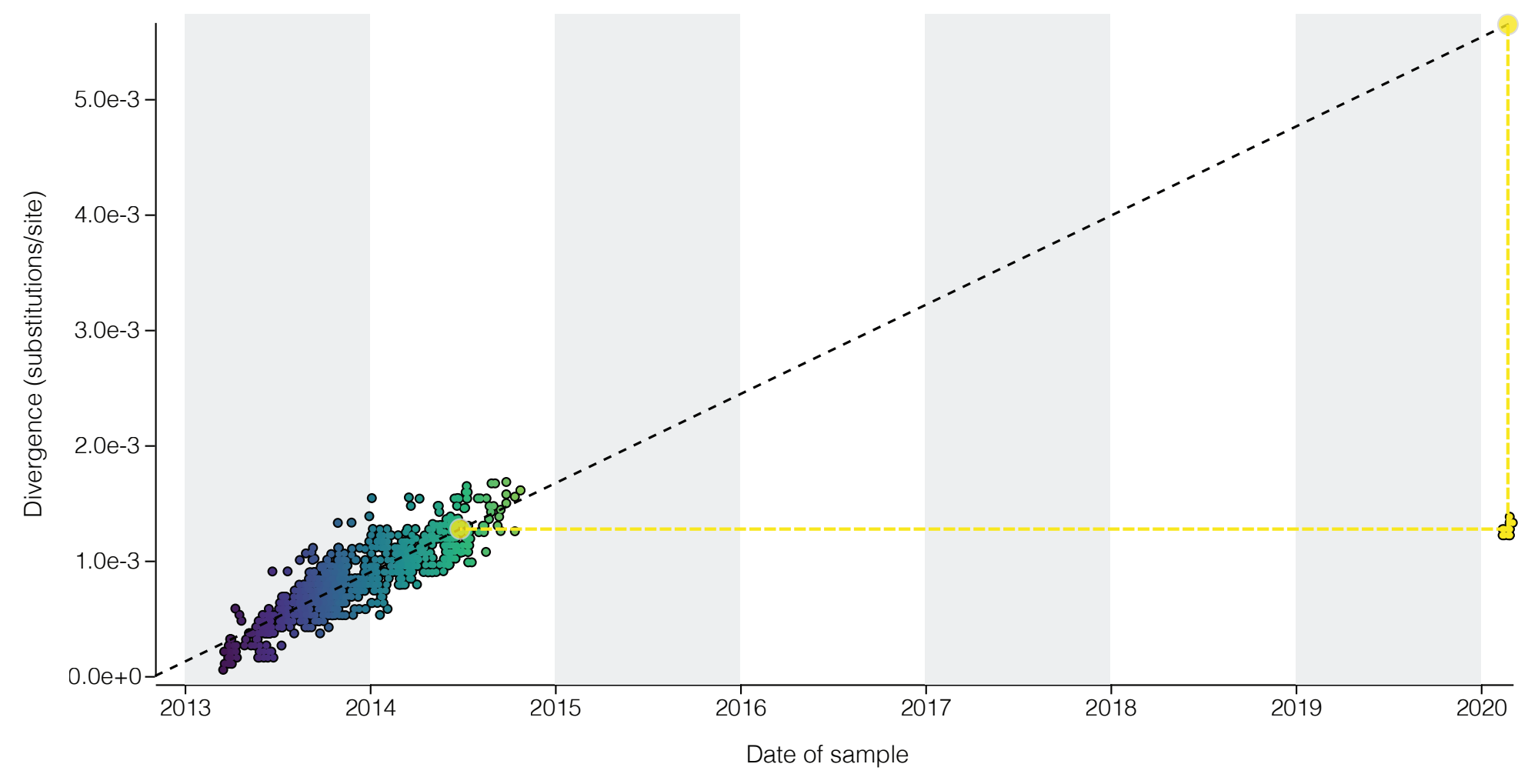
Guinea 2021 tMRCA estimation

Temporal signal was evaluated using the BETS procedure.[53] We estimated a slightly lower log marginal likelihood for a model that uses tip dates (-26063.6) compared to a model that assumes sequences are sampled at the same time (-26062.1). These BEAST analyses were performed using an exponential growth model, a strict molecular clock model and an HKY+G substitution model. We specified a lognormal prior with mean of 1 and standard deviation of 5 on the population size and a Laplace prior with a scale of 100 on the growth rate. Default

722 priors were used for all other parameters. For the divergence time estimation, we used a
723 normal prior on the substitution rate with a mean of 0.001 and a standard deviation of 0.00004
724 based on the background EBOV rate estimated by the local molecular clock analysis.
725
726 **Data availability**
727 Sequencing results were made publicly available on March 12th through joint posting on
728 https://virological.org/c/ebolavirus/guinea-2021/44. The sequences generated at CERFIG
729 have been deposited under GitHub project link:
730 https://github.com/kabinet1980/Ebov_Guinea2021/blob/main/EBOV_Guinea_2021_genomes
731 _CERFIG.fasta and The European Nucleotide Archive (ENA) project number : PRJEB43650
732 (https://www.ebi.ac.uk/ena/browser/view/PRJEB43650); The sequences generated at PFHVG
733 have been deposited under GitHub project link: https://github.com/PFHVG/EBOVsequencing
734 and the genome sequences for the two samples at IPD are available under link:
735 https://drive.google.com/drive/folders/14dfGdNjWw17TkjrEQKLCrwlJ4WBBHI6K .
736 Genome sequences are also available on NCBI GenBank with the following accession
737 numbers: ERX5245591 to ERX5245598; MZ424849 to MZ424862; MZ605320 and
738 MZ605321
739
740 **Code availability:**
741 All the codes for the analyses presented in this paper, including the analysis_pipeline is
742 described in detail in methods and is available in published papers, public websites or for in-
743 house pipelines available upon request.
744
745


Guinea
2021
Guinea
Sierra Leone
Liberia
2014
DRC 2007
Gabon 2001
DRC 2014
DRC 2018
Gabon 1994
DRC 1995
DRC 2018
DRC 2017
DRC 1976
0.004
2021
2018
2017
2014
2007
2002
1995
1994
1976


0
2.0e-4
4.0e-4
6.0e-4
8.0e-4
1.0e-3
1.2e-3
1.4e-3
1.6e-3
Divergence (substitutions/site)


Divergence (substitutions/site)
5.0e-3
4.0e-3
3.0e-3
2.0e-3
1.0e-3
0.0e+0
2013
2014
2015
2016
2017
2018
2019
2020
Date of sample