



HAL
open science

Resurgence of Ebola Virus in 2021 in Guinea Suggests a New Paradigm for Outbreaks

Alpha Kabinet Keita, Fara R. Koundouno, Martin Faye, Ariane Düx, Julia Hinzmann, Haby Diallo, Ahidjo Ayouba, Frédéric Le Marcis, Barré Soropogui, Kékoura Ifono, et al.

► **To cite this version:**

Alpha Kabinet Keita, Fara R. Koundouno, Martin Faye, Ariane Düx, Julia Hinzmann, et al.. Resurgence of Ebola Virus in 2021 in Guinea Suggests a New Paradigm for Outbreaks. *Nature*, 2021, 597, pp.539-543. 10.1038/s41586-021-03901-9 . halshs-03351513

HAL Id: halshs-03351513

<https://shs.hal.science/halshs-03351513v1>

Submitted on 21 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

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

Citation for published version:

Keita, AK, Koundouno, FR, Faye, M, Düx, A, Hinzmann, J, Diallo, H, Ayouba, A, Le Marcis, F, Soropogui, B, Ifono, K, Diagne, MM, Sow, MS, Bore, JA, Calvignac-Spencer, S, Vidal, N, Camara, J, Keita, MB, Renevey, A, Diallo, A, Soumah, AK, Millimono, SL, Mari-Saez, A, Diop, M, Doré, A, Soumah, FY, Kourouma, K, Vielle, NJ, Loucoubar, C, Camara, I, Kourouma, K, Annibaldi, G, Bah, A, Thielebein, A, Pahlmann, M, Pullan, ST, Carroll, MW, Quick, J, Formenty, P, Legand, A, Pietro, K, Wiley, MR, Tordo, N, Peyrefitte, C, McCrone, JT, Rambaut, A, Sidibé, Y, Barry, MD, Kourouma, M, Saouromou, CD, Condé, M, Baldé, M, Povogui, M, Keita, S, Diakite, M, Bah, MS, Sidibe, A, Diakite, D, Sako, FB, Traore, FA, Ki-Zerbo, GA, Lemey, P, Günther, S, Kafetzopoulou, LE, Sall, AA, Delaporte, E, Duraffour, S, Faye, O, Leendertz, FH, Peeters, M, Toure, A & Magassouba, NF 2021, 'Resurgence of Ebola virus in 2021 in Guinea suggests a new paradigm for outbreaks', *Nature*, vol. 597, no. 7877, pp. 539-543. <https://doi.org/10.1038/s41586-021-03901-9>

Digital Object Identifier (DOI):

[10.1038/s41586-021-03901-9](https://doi.org/10.1038/s41586-021-03901-9)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Peer reviewed version

Published In:

Nature

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



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

2

3 Alpha K. Keita^{1,2 a}, Fara R. Koundouno^{3,4 a}, Martin Faye^{5 a}, Ariane Düx^{6 a}, Julia
4 Hinzmann^{4,7,13 a}, Haby Diallo¹, Ahidjo Ayouba², Frederic Le Marcis^{1,2,8}, Barré Soropogui³,
5 Kékoura Ifono^{3,4}, Moussa M. Diagne⁵, Mamadou S. Sow^{1,9}, Joseph A Bore^{3,10}, Sebastien
6 Calvignac-Spencer⁶, Nicole Vidal², Jacob Camara³, Mamadou B. Keita¹¹, Annick Renevey^{4,7},
7 Amadou Diallo⁵, Abdoul K. Soumah¹, Saa L. Millimono^{3,4}, Almudena Mari-Saez⁶, Mamadou
8 Diop⁵, Ahmadou Doré³, Fodé Y. Soumah⁹, Kaka Kourouma¹¹, Nathalie J. Vielle^{4,12}, Cheikh
9 Loucoubar⁵, Ibrahima Camara¹, Karifa Kourouma^{3,4}, Giuditta Annibaldis^{4,12}, Assaïtou Bah³,
10 Anke Thielebein^{4,7}, Meike Pahlmann^{4,7}, Steven T. Pullan^{10,13}, Miles W. Carroll^{10,13}, Joshua
11 Quick¹⁴, Pierre Formenty¹⁵, Anais Legand¹⁵, Karla Pietro¹⁶, Michael R. Wiley^{16,17}, Noel
12 Tordo¹⁸, Christophe Peyrefitte⁵, John T. McCrone¹⁹, Andrew Rambaut¹⁹, Youssouf Sidibé²⁰,
13 Mamadou D. Barry²⁰, Madeleine Kourouma²⁰, Cé D. Saouromou²⁰, Mamadou Condé²⁰,
14 Moussa Baldé⁹, Moriba Povogui¹, Sakoba Keita²¹, Mandiou Diakite^{22,23}, Mamadou S. Bah²²,
15 Amadou Sidibe⁸, Dembo Diakite⁹, Fodé B. Sako⁹, Fodé A. Traore⁹, Georges A. Ki-Zerbo¹²,
16 Philippe Lemey²⁴, Stephan Günther^{4,7,12}, Liana E. Kafetzopoulou^{4,7,24}, Amadou A. Sall⁵, Eric
17 Delaporte^{2,25}, Sophie Duraffour^{4,7,12 b}, Ousmane Faye^{5 b}, Fabian Leendertz^{6 b}, Martine Peeters²
18 ^b, Abdoulaye Toure^{1,11 b}, N'Faly Magassouba^{3 b}.

19

20 ^a These authors contributed equally

21 ^b These authors jointly supervised this work

22

23 Corresponding author: Alpha Kabinet Keita: 1. Centre de Recherche et de Formation en

24 Infectiologie de Guinée (CERFIG), Université de Conakry, Guinea.

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

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

- 50 19. Institute of Evolutionary Biology, University of Edinburgh, King's Buildings,
51 Edinburgh, EH9 3FL, UK
- 52 20. Hôpital Régional de Nzérékoré, Nzérékoré, Guinea
- 53 21. Agence Nationale de Sécurité Sanitaire, Conakry, Guinea
- 54 22. Direction Nationale des Laboratoires, Ministère de la Santé, Conakry, Guinée
- 55 23. Université Gamal Abdel Nasser de Conakry, Conakry, Guinea.
- 56 24. Department of Microbiology, Immunology and Transplantation, Rega Institute, KU
57 Leuven, Leuven, Belgium
- 58 25. Infectious Diseases Departement, University Hospital Montpellier, Montpellier,
59 France
- 60

61 **Abstract (150/150)**

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

74

75 **Main text**

76 At least 30 outbreaks of Ebola virus disease (EVD) have been identified since the late 1970s,
77 the deadliest hitting Guinea, Sierra Leone and Liberia from December 2013 to June 2016.^{1,2}
78 Guinea faced a new EVD outbreak in 2021, which started in Gouéké, a town about 200 km
79 away from the epicenter of the 2013-2016 outbreak. The probable index case was a 51-year-
80 old nurse, assistant of the hospital midwife in Gouéké. On the 21st of January 2021, she was
81 admitted to the hospital of Gouéké suffering from headache, asthenia, nausea, anorexia,
82 vertigo, and abdominal pain. She was diagnosed with malaria and salmonellosis and released
83 two days later. Feeling ill again once at home, she attended a private clinic in N'zérékoré, at
84 40 km, and visited a traditional healer, but died three days later. In the week following her
85 death, her husband and other family members, who attended her funeral, fell ill and four of

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

104

105 Genomic characterization of the virus causing the 2021 Guinean EVD epidemic was of
106 immediate public health importance. First, because diagnostic tools, therapeutics and
107 vaccines, with proven effectiveness in recent EVD outbreaks, i.e. in Guinea (2013-2016) and
108 in the Equateur and North-Kivu/Ituri provinces of the Democratic Republic of Congo (2018-
109 2020), have primarily been developed for EBOV.³⁻⁵ Secondly, to identify whether the
110 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 EBOV⁹, 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

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

139 Maximum likelihood phylogenetic reconstruction places the 12 genomes from the 2021
140 Guinea outbreak as a single cluster among the EBOV viruses responsible for the 2013-2016
141 EVD outbreak in West Africa (Figure 1 and Figure 2). The 2021 genomes share 10
142 substitutions accumulated during the 2013-2016 outbreak (compared to KJ660346), including
143 the A82V marker mutation for human adaptation in the glycoprotein that arose when the virus
144 spread to Sierra Leone.^{11,12} These patterns provide strong evidence for direct linkage to
145 human cases from the 2013-2016 outbreak rather than a new spillover from an animal
146 reservoir. The 2021 lineage is nested within a clade that predominantly consists of genomes
147 sampled from Guinea in 2014 (Figure 2). The branch by which the 2021 cluster diverges from
148 the previous outbreak exhibits only 12 substitutions, which is far fewer than expected from
149 EBOV evolution during 6 years of sustained human-to-human transmission (Figure 3). Using
150 a local molecular clock analysis, we estimate a 6.4-fold (95% Highest Posterior Density
151 interval (HPD): 3.3-fold,10.1-fold) lower rate along this branch. For comparison, we also
152 estimate a 5.5-fold (1.6-fold,10.8-fold) lower rate along the branch leading to the 2016 flare-
153 up that was linked to a survivor with virus persistence for more than 500 days.^{7,13} Rather than
154 a constant long-term low evolutionary rate, some degree of latency or dormancy during
155 persistent infection seems a more likely explanation for the low divergence of the 2021
156 genomes. We tested whether the 12 genomes from 2021, sampled over a time period of less
157 than one month, contained sufficient temporal signal to estimate the time to most recent
158 common ancestor (tMRCA) (Supplementary Figure 2), but did not identify statistical support
159 for sufficient divergence accumulation over this short time scale. We therefore calibrated our
160 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

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

197

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

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

224

225 Active genomic surveillance already showed resurgence of previous strains in other
226 outbreaks. For example, two EBOV variants circulated simultaneously within the same region
227 during the recent 2020 outbreak in the Equateur Province, DRC.³⁰ Moreover, strains from the
228 two consecutive outbreaks in Luebo, Democratic Republic of Congo (DRC), in 2007 and
229 2008, are also so closely related that it now appears difficult to exclude that the epidemic
230 observed in 2008 was possibly due to a resurgence event from a EVD survivor from the 2007
231 outbreak.^{31,32} However, the limited genomic sampling does not allow for a formal test of this
232 hypothesis.

233 While the majority of EVD outbreaks remained limited both in number of cases and
234 geographic spread, the two largest outbreaks in West Africa (December 2013 to June 2016)
235 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.

261

262

263

264

265 **References**

266 1. Baize, S. et al. Emergence of Zaire Ebola virus disease in Guinea. *N Engl J Med*

267 371:1418-25 (2014).

268

269 2. World Health Organization. 2016 Situation Report: Ebola Virus Disease. Geneva: World

270 Health Organization; 10 June 2016

271

272 3. Malvy, D., McElroy, A.K., de Clerck, H., Günther, S., van Griensven, J. Ebola virus

273 disease. *Lancet*. 393:936-948 (2019).

274

275 4. Mbala-Kingebeni, P. et al. Rapid Confirmation of the Zaire Ebola Virus in the Outbreak of

276 the Equateur Province in the Democratic Republic of Congo: Implications for Public Health

277 Interventions. *Clin Infect Dis*. 68:330-333 (2019)

278

279 5. Mbala-Kingebeni, P. et al. 2018 Ebola virus disease outbreak in Equateur Province,

280 Democratic Republic of the Congo: a retrospective genomic characterization. *Lancet Infect*

281 *Dis* 19: 641-647 (2019)

282

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

284 e04395 (2014)

285

- 286 7. Diallo, B. Resurgence of Ebola Virus Disease in Guinea Linked to a Survivor With Virus
287 Persistence in Seminal Fluid for More Than 500 Days. *Clin Infect Dis* 63: 1353-1356 (2016)
288
- 289 8. Sissoko, D. et al. Ebola Virus Persistence in Breast Milk After No Reported Illness: A
290 Likely Source of Virus Transmission From Mother to Child. *Clin Infect Dis* 64: 513-516
291 (2017)
292
- 293 9. Keita, A.K. et al. A 40-Month Follow-Up of Ebola Virus Disease Survivors in Guinea
294 (PostEbogui) Reveals Long-Term Detection of Ebola Viral Ribonucleic Acid in Semen and
295 Breast Milk. *Open Forum Infect Dis* 6: ofz482 (2019)
296
- 297 10. Quick, J. et al. Real-time, portable genome sequencing for Ebola surveillance. *Nature* 530,
298 228–232 (2016).
299
- 300 11. Diehl, W. E. et al. Ebola Virus Glycoprotein with Increased Infectivity Dominated the
301 2013-2016 Epidemic. *Cell* 167, 1088–1098.e6 (2016).
302
- 303 12. Urbanowicz, R. A. et al. Human Adaptation of Ebola Virus during the West African
304 Outbreak. *Cell* 167, 1079–1087.e5 (2016).
305
- 306 13. Blackley, D.J. et al. Reduced evolutionary rate in reemerged Ebola virus transmission
307 chains. *Sci Adv* 2:e1600378 (2016)
308

- 309 14. Diallo, M.S.K. et al. Understanding the long-term evolution and predictors of sequelae of
310 Ebola virus disease survivors in Guinea: A 48-month prospective, longitudinal cohort study
311 (PostEboGui). *Clin Infect Dis*:ciab168 (202). doi: 10.1093/cid/ciab168. Online ahead of print.
312
- 313 15. Etard, J.F. et al. Multidisciplinary assessment of post-Ebola sequelae in Guinea
314 (Postebogui): an observational cohort study. *Lancet Infect Dis* 17:545-552 (2017)
315
- 316 16. PREVAIL III Study Group. A Longitudinal Study of Ebola Sequelae in Liberia. *N Engl J*
317 *Med* 380:924-934 (2019)
318
- 319 17. Diallo, M.S.K. et al. Prevalence of infection among asymptomatic and paucisymptomatic
320 contact persons exposed to Ebola virus in Guinea: a retrospective, cross-sectional
321 observational study. *Lancet Infect Dis* 19:308-316 (2019)
322
- 323 18. Glynn, J.R. et al. Asymptomatic infection and unrecognised Ebola virus disease in Ebola-
324 affected households in Sierra Leone: a cross-sectional study using a new non-invasive assay
325 for antibodies to Ebola virus. *Lancet Infect Dis* 17:645-653 (2017)
326
- 327 19. Camara, I. et al. Unrecognized Ebola virus infection in Guinea: complexity of surveillance
328 in a health crisis situation: case report. *Pan Afr Med J* 36:201 (2020)
329
- 330 20. Thorson, A.E. et al. Persistence of Ebola virus in semen among Ebola virus disease
331 survivors in Sierra Leone: A cohort study of frequency, duration, and risk factors. *PLoS Med*
332 18:e1003273 (2021)
333

- 334 21. Fischer, W.A. et al. Ebola Virus Ribonucleic Acid Detection in Semen More Than Two
335 Years After Resolution of Acute Ebola Virus Infection. *Open Forum Infect Dis* 4:ofx155
336 (2017)
337
- 338 22. Sissoko, D. et al. Persistence and clearance of Ebola virus RNA from seminal fluid of
339 Ebola virus disease survivors: a longitudinal analysis and modelling study. *Lancet Glob*
340 *Health* 5:e80-e88 (2017).
341
- 342 23. Mate, S.E. et al. Molecular Evidence of Sexual Transmission of Ebola Virus. *N Engl J*
343 *Med* 373:2448-54 (2015)
344
- 345 24. Dokubo, E.K. et al. Persistence of Ebola virus after the end of widespread transmission in
346 Liberia: an outbreak report. *Lancet Infect Dis* 18:1015-1024 (2018)
347
- 348 25. Liu, W.J. et al. Comprehensive Clinical and Laboratory Follow-up of a Female Patient
349 With Ebola Virus Disease: Sierra Leone Ebola Virus Persistence Study. *Open Forum Infect*
350 *Dis* 6:ofz068 (2019)
351
- 352 26. Wiedemann, A. et al. Long-lasting severe immune dysfunction in Ebola virus disease
353 survivors. *Nat Commun* 11:3730 (2020)
- 354 27. Adaken, C. et al. Ebola virus antibody decay-stimulation in a high proportion of survivors.
355 *Nature* 590:468-472 (2021)
356
- 357 28. Thom, R. et al. Longitudinal antibody and T cell responses in Ebola virus disease
358 survivors and contacts: an observational cohort study. *Lancet Infect Dis* 21:507-516 (2021)

359

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

362

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

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

368

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

371

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

375

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

378

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

381

382 36. Farmer, P. Aids and Accusation. Haiti and the Geography of Blame. *University of
383 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).

409

410 44. Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25,
411 2078–2079 (2009)

412

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

415

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

418

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

421

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

424

425 49. Rambaut, A., Lam, T.T., Max Carvalho, L., Pybus, O.G. Exploring the temporal structure
426 of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol* 2, vew007
427 (2016).

428

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

431

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

434

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

437

438 53. Duchene, S., et al. Temporal signal and the phylodynamic threshold of SARS-CoV-2.

439 *Virus Evol* 6:veaa061 (2020)

440 **Acknowledgements**

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

450

451 **Funding**

452 The work of CERFIG and TransVIHMI was supported in part by grants EBO-SURSY project
453 funded by European Union, International Mixt Laboratory “RESPIRE” of IRD (Institut de
454 Recherche pour le Developpement), Montpellier Université d'Excellence (EBOHEALTH; I-
455 Site MUSE, ANR-16-IDEX-0006) and Institut National de la Santé et de la Recherche Médicale
456 (INSERM)/REACTing (REsearch and ACTion targeting emerging infectious diseases). The work
457 of RKI was partly funded by the German Ministry of Health “Global Protection Program”
458 project TRICE. Frédéric Le Marcis received funding through the program « EBOVAC3
459 Bringing a prophylactic Ebola vaccine to licensure » funded by IMI (grant agreement number
460 800176) and run by LSTHM and INSERM.

461 The work of PFHG and BNITM was supported by the German Federal Ministry of Health
462 through support of the WHO Collaborating Centre for Arbovirus and Hemorrhagic Fever
463 Viruses at the Bernhard-Nocht-Institute for Tropical Medicine (agreement ZMV I1-
464 2517WHO005), and through the Global Health Protection Programme (GHPP, agreements

465 ZMV 11-2517GHP-704 and ZMVI1-2519GHP704), and by the Coalition for Epidemic
466 Preparedness Innovations (CEPI). Work of BNITM and the UK Health Security Agency has
467 been further supported by the Research and Innovation Programme of the European Union
468 under Horizon 2020 grant agreement no. 871029-EVA-GLOBAL.

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

479

480 **Author contributions**

481 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,
482 N.M., O.F, S.D., S.G.

483

484 collected data and/or performed and/or performed medical examinations and/or laboratory
485 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.,
486 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.,
487 M.C., M.D.B., M.K., M.Po., M.S.S., N.V., N.J.V., S.D., Y.S.

488

489 Performed sequencing and/or sequence validation: A.A., A.Du., A.Re., B.S., H.D., J.A.B.,
490 J.H., K.I., K.P, M.F, M.M.D, S.C-S
491
492 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.
493
494 Funding acquisition: E.D., F.H.L, M.Pe., M.W.C, S.D., S.G.
495
496 Provided reagents: F.H.L., K.P., M.R.W, S.C-S.
497
498 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.,
499 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.
500
501 Wrote the first draft of the manuscript: A.Du, A.K.K, M.Pe, S.C-S.
502
503 wrote and edited the manuscript: A.A.S, A.Du., A.K.K, A.M.S., A.Re., A.Ra, A.To., E.D.,
504 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.
505
506
507 All authors read and approved the contents of the manuscript.
508
509
510 **Competing interests:** MWC received materials for this study from Oxford Nanopore. All
511 other authors declare no competing interests.

512 **Figure legends.**

513

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

520

521

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

526

527 **Figure 3. Temporal divergence plot of genetic divergence from the root against time of**
528 **sampling for the tree shown in Figure 2.** The regression is exclusively fitted to genomes
529 sampled between 2014 and 2015. The same colors are used for the data points as in Figure 2.
530 The dashed yellow lines highlight how the 2021 data points deviate from the relationship
531 between sampling time and sequence divergence. According to this relationship, about 95
532 substitutions (95% prediction interval: 88-101) are expected on the branch ancestral to the
533 2021 cluster, whereas only 12 are inferred on this branch.

534

535

536

537

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

540

Patient	Sex	Age (years)	Date of sampling	CERFIG		PFHG		IPD	
				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	–	–

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

542 ² Sample was not tested in this laboratory

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

545 ^a patient samples at diagnosis

546 ^b follow-up samples from patients

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

548 **Materials and Methods**

549 **Ethics Statement**

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

554

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

556 Viral RNA was extracted from 140 ul of whole blood collected from the samples from the
557 patient hospitalized in Conakry, with the Nuclisens kit (Biomerieux, France) and following
558 manufacturer's instructions. Amplification of a small fragment of VP35 region was attempted
559 in a semi-nested PCR with a modified protocol as previously described.⁴ First-round VP35
560 PCR-products from positive samples were barcoded and pooled using the Native Barcoding
561 Kit EXP-NBD104 (Oxford Nanopore Technologies, Oxford, UK). Sequencing libraries were
562 generated from the barcoded products using the Genomic DNA Sequencing Kit SQK-
563 LSK109 (Oxford Nanopore Technologies) and were loaded onto a R9 flow cell on a MinIon
564 (Oxford Nanopore Technologies). Genetic data were collected for one hour. Basecalling,
565 adapter removal and demultiplexing of fastq files were performed with MinKNOW, version
566 4.1.22. Fastq reads >Q11 were used for mapping a virus database with the Genome Detective
567 tool (<https://www.genomedetective.com/app/typingtool/virus/>). The generated consensus
568 sequence was used for further analysis. For phylogenetic inference, we retrieved one sequence
569 per outbreak from the Haemorrhagic Fever Virus (HFV) database to which we added the
570 newly generated VP35 sequence of the novel outbreak. Phylogenetic analyses were done
571 using Maximum Likelihood methods using IQ-Tree with 1,000 bootstraps for branch

572 support.^{42,43} The GTR model plus a discrete gamma distribution were used as nucleotide
573 substitution models.

574

575 **Full-length genome sequencing of the new Ebola viruses**

576 *Genome sequencing at CERFIG*

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

597 Sequencing reads were filtered (adapter removal and quality filtering) with Trimmomatic
598 (Bolger, 2014) (settings: LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30
599 MINLEN:40), merged with ClipAndMerge (<https://github.com/apeltzer/ClipAndMerge>), and
600 mapped to the *Zaire ebolavirus* RefSeq genome (NC_002549) using BWA-MEM.⁴⁴ Mapped
601 reads were sorted and deduplicated with SortSam and MarkDuplicates from the Picard suite
602 (Broad Institute, Picard; <http://broadinstitute.github.io/picard>). We generated consensus
603 sequences using Geneious Prime 2020.2.3 (<https://www.geneious.com>) where unambiguous
604 bases were called when at least 90% of at least 20 unique reads were in agreement (20x,
605 90%). For samples with few mapped reads (0001, 0002, 0010, 0030), we also called a
606 consensus at 2X, 90% and 5X, 90%.

607

608 ***Genome sequencing at PFHG***

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

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

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

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

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

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

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

678

679 ***Sequencing at IPD***

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

691

692 **Phylogenetic analysis of full-length genome sequences**

693 ***Phylogenetic inference***

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

697 1065 genomes sampled from Guinea, Sierra Leone and Liberia between 2014 and 2015.
698 Multiple sequence alignment was performed using mafft.⁴⁶ We identified 6 T-to-C mutations
699 in the genome from patient 11 that were indicative of mutations induced by adenosine
700 deaminases acting on RNA (ADARs). According to the recommendations by Dudas et al.⁴⁷,
701 we masked these positions in this genome in all further analyses. Maximum likelihood trees
702 were reconstructed using IQ-tree under the general time-reversible (GTR) model with gamma
703 (G) distributed rate variation among sites.⁴⁸ Temporal divergence plots of genetic divergence
704 from the root of phylogenies against sampling time were constructed using TempEst.⁴⁹ To
705 construct the temporal divergence plot for the Guinean 2021 genome data, we used a tree
706 reconstructed under an HKY+G model.

707

708 ***Local molecular clock model analysis***

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

714

715 ***Guinea 2021 tMRCA estimation***

716 Temporal signal was evaluated using the BETS procedure.⁵³ We estimated a slightly lower
717 log marginal likelihood for a model that uses tip dates (-26063.6) compared to a model that
718 assumes sequences are sampled at the same time (-26062.1). These BEAST analyses were
719 performed using an exponential growth model, a strict molecular clock model and an HKY+G
720 substitution model. We specified a lognormal prior with mean of 1 and standard deviation of
721 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](https://github.com/kabinet1980/Ebov_Guinea2021/blob/main/EBOV_Guinea_2021_genomes_CERFIG.fasta)
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/14dfGdNjWw17TkjrEQKLCrwIJ4WBBHI6K>.

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





