



Hoffmann, C. et al. (2017) Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. *Nature*, 548(7665), pp. 82-86.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/145915/>

Deposited on: 22 August 2017

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest

Authors: Constanze Hoffmann^{1,†}, Fee Zimmermann^{1,2,†}, Roman Biek³, Hjalmar Kuehl⁴, Kathrin Nowak¹, Roger Mundry⁴, Anthony Agbor⁴, Samuel Angedakin⁴, Mimi Arandjelovic⁴, Anja Blankenburg¹, Gregory Brazolla⁴, Katherine Corogenes⁴, Emmanuel Couacy-Hymann⁵, Tobias Deschner⁴, Paula Dieguez⁴, Karsten Dierks⁴, Ariane Dux¹, Susann Dupke², Henk Eshuis⁴, Pierre Formenty⁶, Yisa Ginath Yuh⁴, Annemarie Goedmakers⁷, Jan F. Gogarten^{1,4,8}, Anne-Céline Granjon⁴, Scott McGraw⁹, Roland Grunow², John Hart¹⁰, Sorrel Jones⁴, Jessica Junker⁴, John Kiang¹¹, Kevin Langergraber¹², Juan Lapuente⁴, Kevin Lee⁴, Siv Aina Leendertz¹, Floraine Léguillon¹, Vera Leinert¹³, Therese Löhrich^{1,4}, Sergio Marrocoli⁴, Kerstin Mätz-Rensing¹⁴, Amelia Meier⁴, Kevin Merkel¹, Sonja Metzger¹, Mizuki Murai⁴, Svenja Niedorf¹, Hélène De Nys^{1,4}, Andreas Sachse¹, Joost van Schijndel⁴, Ulla Thiesen¹, Els Ton⁷, Doris Wu^{1,4}, Lothar H. Wieler¹⁵, Christophe Boesch⁴, Silke R. Klee², Roman M. Wittig⁴, Sébastien Calvignac-Spencer¹, Fabian H. Leendertz^{1,*}

[†]Equal contributions

*Corresponding author. E-mail: LeendertzF@rki.de

Affiliations:

¹Robert Koch Institute, P3: "Epidemiology of Highly Pathogenic Microorganisms", Seestraße 10-11, 13353 Berlin, Germany.

²Robert Koch Institute, ZBS 2: Centre for Biological Threats and Special Pathogens, Highly Pathogenic Microorganisms, Seestraße 10-11, 13353 Berlin, Germany.

³Institute of Biodiversity, Animal Health and Comparative Medicine, Boyd Orr Centre for Population and Ecosystem Health, College of Medical, Veterinary and Life Sciences, University of Glasgow, Glasgow, United Kingdom.

⁴Max Planck Institute for Evolutionary Anthropology (MPI EVAN), Deutscher Platz 6, 04103 Leipzig, Germany.

⁵LANADA/LCVB, Bingerville, 206, Côte d'Ivoire.

⁶World Health Organization, 1211 Geneva 27, Switzerland.

⁷Chimbo Foundation, Amstel 49, 1011 PW Amsterdam, Netherlands.

⁸McGill University, Department of Biology, 855 Sherbrooke Street West Montreal, QC, Canada, H3A 2T7.

40 ⁹The Ohio State University, Department of Anthropology, 4034 Smith Laboratory, 174
41 W. 18th Avenue, Columbus, OH, USA.

42
43 ¹⁰Lukuru Foundation, 1235 Avenue des Poids Lourds / Quartier de Kingabois, Kinshasa,
44 DRC.

45
46 ¹¹Limbe Wildlife Centre, Limbe, Cameroon.

47
48 ¹²Arizona State University, PO Box 872402, Tempe, AZ 85287-2402 USA.

49
50 ¹³Wild Chimpanzee Foundation (WCF), Deutscher Platz 6, 04103 Leipzig, Germany.

51
52 ¹⁴German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany.

53
54 ¹⁵Robert Koch Institute, Seestraße 10-11, 13353 Berlin, Germany.

55
56

57 **Anthrax is a globally significant animal disease and zoonosis. Despite this, current**
58 **knowledge of anthrax ecology is largely limited to arid ecosystems, where outbreaks**
59 **are most commonly reported¹⁻³. We reveal the dynamics of an anthrax causing**
60 **agent, *Bacillus cereus* biovar *anthracis*, in a tropical rainforest with severe**
61 **consequences for local wildlife communities. Using data and samples collected over**
62 **three decades we find that rainforest anthrax is a persistent and widespread cause**
63 **of death for a broad range of mammalian hosts. We predict that this pathogen will**
64 **accelerate the decline and possibly result in the extirpation of local chimpanzee (*Pan***
65 ***troglodytes verus*) populations. Our findings illuminate the epidemiology of a cryptic**
66 **pathogen and have important implications for conservation.**

67

68 Anthrax is a disease of wildlife, livestock and humans predominantly affecting low and
69 middle-income countries^{2,4,5}. Although widely distributed, including some temperate
70 regions, anthrax is most commonly associated with arid ecosystems, particularly African

71 savannas^{1,3,6-11}. In these systems, major outbreaks typically cause high mortality in a few
72 wild and domestic ungulate species at a time and usually exhibit strong seasonal and
73 inter-annual variation^{2,3,5,11,12}. For example, in Krüger National Park, South Africa, die-
74 offs in kudu (*Tragelaphus strepsiceros*) and impalas (*Aepyceros melampus*) occur in the
75 dry season with a ten year periodicity coinciding with rainfall cycles¹¹. In Etosha
76 National Park, Namibia, mortality in elephants (*Loxodonta Africana*) peaks at the start of
77 the wet season, while plains ungulates (*Equus quagga*, *Conochaetes taurineus*,
78 *Antidorcas marsupialis*) are most affected at the end of the wet season^{3,13}. Such varying
79 dynamics underline the importance of investigating the pathogen in close relation with its
80 ecosystem, but so far anthrax research in Africa has been biased towards well-studied
81 savanna regions.

82 In 2001, lethal anthrax-like cases in wild chimpanzees were reported in a rainforest
83 habitat: Taï National Park (TNP), Côte d'Ivoire (Fig. S1)¹⁴. The causative agent was a
84 bacterium combining the chromosomal background of *Bacillus cereus* with the virulence
85 plasmids of *B. anthracis* (*Bacillus cereus* biovar *anthracis*; *Bcbva*)¹⁵. Pathology and
86 histopathology of *Bcbva* cases were clearly suggestive of anthrax and in small animal
87 models *Bcbva* was as virulent as *B. anthracis*¹⁴⁻¹⁶. *Bcbva* cases have since been described
88 in animals in Cameroon (CM), Central African Republic (CAR) and the Democratic
89 Republic of Congo^{17,18}, suggesting a broad sub-Saharan distribution (Fig. 1). However,
90 the epidemiology of anthrax-like disease caused by *Bcbva* (hereafter anthrax), and to
91 what extent it matches that of classical anthrax, remain poorly understood.

92 We address this knowledge gap by testing a unique set of samples collected in TNP over
93 26 years. We started collecting bones in 1989 resulting in bones from 75 individual

94 mammals (Table S7, Supplementary information S4). From 1996 on, we investigated 204
95 fresh carcasses (Table S2, Supplementary information S2). Since bone and carcass
96 discovery was linked to the collection of chimpanzee behavioral data, we expected
97 detection of *Bcbva* to be biased towards chimpanzees and other easily detectable medium
98 to large-bodied mammals. We therefore tested whether carrion flies, which are relatively
99 unbiased samplers of mammalian DNA¹⁹, might also collect *Bcbva* or its genetic material
100 while feeding and ovipositing on carcasses. Starting in 2008, we applied different
101 horizontal and vertical sampling schemes to collect 1,634 flies (Table S1 and S4,
102 Supplementary information S3). We retrieved *Bcbva* isolates from all three sample types
103 (bones, carcasses, flies). These allowed us to generate 178 whole genome sequences
104 spanning from 1996 to 2014 (Table S8). To clarify the distribution of *Bcbva* on a larger
105 scale, we sampled 1089 flies and 136 bones from 16 other sites in 11 sub-Saharan
106 countries from 2012 to 2014 (Fig. 1, Table S1).

107 In TNP we detected *Bcbva* DNA in 81 carcasses (40%; Fig. 2A, Extended Data Fig. S1,
108 Extended Data Fig. S2, Table S2), 26 bones (35%, Table S7) and 80 flies (5%; Fig. 2B,
109 Extended Data Fig. S3, Table S4). We could perform histopathological examinations on
110 15 positive carcasses and in all cases pathology was consistent with a lethal anthrax
111 infection (Table S2). Overall, 38% of observed local wildlife mortality was associated
112 with *Bcbva* (Tables S2 and S4), meeting the highest levels of mortality reported for
113 classical anthrax outbreaks in savanna ecosystems^{12,20}. We observed no obvious seasonal
114 variation in *Bcbva* carcass incidence, suggesting ongoing anthrax activity in the area
115 (Generalized Linear Mixed Model (GLMM), $\chi^2=6.3$, $df=10$, $P=0.789$, Supplementary
116 information S8a). However, *Bcbva* detection in flies peaked from December to March,

117 coinciding with the only distinct dry period in the park (GLMM, $\chi^2=6.9$, $df=2$, $P=0.032$,
118 Extended Data Fig. S4, Supplementary information S8b). This suggests climatic
119 conditions may influence *Bcbva* ecology in TNP, similar to observations from *B.*
120 *anthracis* in savannas¹, though seasonal mortality appears less pronounced.

121 *Bcbva* differed dramatically from *B. anthracis* in terms of host range. Ungulates
122 constitute the vast majority (> 99 %) of anthrax cases in savanna ecosystems^{11,12,20}. In
123 contrast, and in line with the more diverse fauna found in rainforests, we observed *Bcbva*
124 fatalities in a broader range of species in TNP, including chimpanzees (31/55), six
125 monkey species (21/81), duikers (26/40), mongooses (2/2) and porcupines (1/26 other
126 mammals) (Table S2). To further explore the host range of *Bcbva*, we analyzed the gut
127 content of all mammal and *Bcbva* positive flies (n=28, Table S1) using amplicon deep
128 sequencing. We detected sequences from most of the aforementioned species, and from
129 species belonging to 11 further mammalian genera, including carnivores, rodents and bats
130 (Table S5, Supplementary information S3e). This suggests that *Bcbva* may affect an even
131 broader range of mammals than inferred from carcass monitoring alone. Further, meal
132 compositions of mammal positive *Bcbva* positive flies (n=28) and mammal positive
133 *Bcbva* negative flies (n=29) did not differ significantly (GLMMs, Supplementary
134 information S8c), which may support the notion that there is no substantial difference in
135 *Bcbva* susceptibility among species.

136 To gain further insight into the ecology of *Bcbva*, we investigated 178 genomes derived
137 from isolates obtained from necropsy samples, bones and flies, collected between 1996
138 and 2014 (Table S8). Considering 126 chromosomal sequences originating from separate
139 hosts (mammals and flies) we detected 298 single nucleotide polymorphisms (SNP).

140 Plasmids contained negligible amounts of variation (Supplementary information S7a).
141 The maximum distance observed between isolates was 69 SNPs (median: 26 SNPs); the
142 most distant isolates originated in flies caught in two consecutive years only 6 km apart.
143 In comparison, a maximum distance of only 20 SNPs was observed in *B. anthracis*
144 isolates derived from cattle samples collected in the French Alps between 1997 and
145 2009²¹. The high genetic diversity observed in TNP is consistent with extensive *Bcbva*
146 activity in the area and suggests that this pathogen did not emerge recently (Fig.3,
147 Extended Data Fig. S5). In addition, considerably more divergence was seen compared to
148 isolates from other countries^{17,18}, supporting the notion that *Bcbva* has been circulating in
149 sub-Saharan Africa for an even much longer period than what we determined in TNP
150 (Extended Data Fig. S6, Supplementary information S7). To assess within-host diversity
151 we sequenced the genomes of two to six independent isolates for a subset of carcasses
152 and flies (Table S9). Two strains differing by 42 chromosomal SNPs were isolated from a
153 single fly, likely reflecting multiple carcass meals¹⁹, which further highlights the
154 commonness of *Bcbva* in TNP. Otherwise, the maximum distance observed within one
155 host was two chromosomal SNPs (mean: 0.35 SNPs). Within-host heterogeneity thus
156 seems negligible compared to the overall diversity observed for *Bcbva* suggesting strains
157 differing by more than two SNPs originate from separate carcasses.

158 *Bcbva* positive carcasses were broadly distributed throughout the TNP research area,
159 without the kind of geographic clustering described for anthrax in savanna
160 ecosystems^{12,22} (Fig. 2A). We determined *Bcbva* prevalence within and outside the
161 research area using a subset of 908 flies caught systematically according a grid system
162 within 19 days (Extended Data Fig. S7). We detected *Bcbva* positive flies in 16/83 traps

163 (Additional Data Table S1). Prevalence was higher in the research area (8/21 traps *Bcbva*
164 positive) than in the surrounding forest belt (8/62 traps *Bcbva* positive) (Fisher's Exact
165 Test, $P = 0.02$). Long-term research activity within the TNP research area has had a
166 protective effect on wildlife and led to an increased density of mammals²³, which might
167 explain higher *Bcbva* activity. Genome data revealed multiple contemporaneous
168 transmission chains caused by co-circulating strains (2 to 48 SNPs distance, median: 25
169 SNPs) in different areas of the park over the short time period of the fly snapshot
170 (Extended Data Fig. S8). For low genomic distances (≤ 35 SNPs), genomic and
171 geographic distances of all TNP isolates were positively correlated ($R^2 = 0.72$), providing
172 further indication of spatially restricted transmission (Extended Data Fig. S9), which
173 might reflect carcass-mediated spread of *Bcbva*. Since wildlife cases included exclusively
174 arboreal monkeys (Table S2), we explored the vertical distribution of *Bcbva* by catching
175 flies simultaneously on the ground and up to 30m into the canopy. We detected *Bcbva* in
176 12 of 103 canopy flies (11.7%) and retrieved isolates from five of these (Table S4,
177 Additional Data Table S1). While on the ground carcass deposition sites are likely to be
178 the source of *Bcbva* infections, flies may contribute to *Bcbva* transmission in the upper
179 strata of the rainforest²⁴.

180 Fly samples indicated a large proportion of undetected anthrax mortality. During 19 days
181 of focused fly sampling, we retrieved *Bcbva* isolates from 17 flies, with 13 strains being
182 more than two SNPs different from any other strain. Since two SNPs appear to be the
183 upper level of within-host diversity (Table S9), this implies the presence of at least 13
184 different *Bcbva* positive carcasses. Yet, during the same sampling period, only three
185 *Bcbva* positive carcasses were discovered and their isolates all corresponded to one of the

186 fly *Bcbva* lineages (≤ 2 SNPs difference). This suggests carcass monitoring alone
187 underestimates mortality by at least an order of magnitude.

188 We investigated the consequences of *Bcbva*-induced mortality on the species best studied
189 in this ecosystem, chimpanzees. Chimpanzees have a low reproduction rate²⁵ and are thus
190 particularly sensitive towards external changes to their environment. Based on
191 demographic data collected from habituated groups in TNP, we simulated population
192 viability at a 150 years horizon across a broad range of demographic models including
193 and excluding anthrax induced mortality (Fig. S7 and S8). Our simulations showed that,
194 with *Bcbva*, the TNP chimpanzee population would only have high chances to persist in
195 the case of an overall annual per capita mortality rate due to other causes of 1% (Fig. S7
196 and S8). Such a low mortality rate is, however, not even observed in captive
197 chimpanzees. In wild chimpanzees the lowest annual per capita mortality rate is 4% (in
198 early adults)²⁵. Under such a survival probability (0.96), the simulated presence of
199 anthrax invariably led to a clearly reduced survival probability of communities (Fig. 4).
200 For example, 76/84 models resulted in extirpation probability higher than 50%, while the
201 model which we consider the most realistic (community size 60, maximum age 46 years
202 and inter birth interval 6 years) resulted in an extirpation probability of 89% (Fig. 4). Our
203 simulations therefore suggest that anthrax induced mortality will result in deterministic
204 population declines and possible extirpation of TNP chimpanzees over the next 150
205 years. The risk of extirpation will increase if chimpanzee mortality due to hunting and
206 human-borne infectious diseases continues to rise^{23,26,27}.

207 To determine whether similar unrecognized effects on wildlife might be occurring
208 elsewhere, we tested 784 flies collected at eight different sites, as well as 136 bones from

209 twelve sites in five and nine sub-Saharan countries, respectively (Fig. S3, Table S1). All
210 sites had chimpanzee populations but none (nor the country) had previously reported
211 *Bcbva* cases. We only detected *Bcbva* genetic material in 2 of 105 flies and 1 of 8 bones
212 collected in the Grebo National Forest (GNF) in Liberia, about 40 km from TNP (Fig.
213 S1). The genome sequences of isolates from the two fly samples nested within the
214 diversity of *Bcbva* in TNP which may indicate an epidemiological link (Fig. 3). We did
215 not detect *Bcbva* in 305 flies from two sites where *Bcbva* cases have been previously
216 reported (Dja Reserve, CM, and Dzanga Sangha Protected Areas, CAR, Table S4). While
217 the lack of detection at other sites needs to be interpreted with caution due to variable fly
218 species composition (Extended Data Fig. S10, Supplementary information S3f), these
219 data suggest that *Bcbva* dynamics may also vary across rainforest ecosystems. It will be
220 important to further uncover the scale and environmental drivers behind *Bcbva*
221 prevalence. Such knowledge will be critical for mitigating against the detrimental effects
222 of *Bcbva* on wildlife and for better assessing human infection risk, which for anthrax in
223 rainforest ecosystems has, to date, been considered very low.

224

225 **References (1-27)**

- 226 1 Hampson, K. *et al.* Predictability of anthrax infection in the Serengeti, Tanzania. *J*
227 *Appl Ecol* **48**, 1333-1344, doi:10.1111/j.1365-2664.2011.02030.x (2011).
- 228 2 Hugh-Jones, M. & De Vos, V. Anthrax and wildlife. *Revue Scientifique et*
229 *Technique-Office International des Epizooties* **21**, 359-384 (2002).
- 230 3 Lindeque, P. M. & Turnbull, P. C. Ecology and epidemiology of anthrax in the
231 Etosha National Park, Namibia. *Onderstepoort J Vet Res* **61**, 71-83 (1994).
- 232 4 Beyer, W. & Turnbull, P. C. Anthrax in animals. *Mol Aspects Med* **30**, 481-489,
233 doi:10.1016/j.mam.2009.08.004 (2009).
- 234 5 Turnbull, P. C. B. *WHO Guidelines Approved by the Guidelines Review*
235 *Committee* (World Health Organization, Department of Communicable Diseases
236 Surveillance and Response, 2008).

- 237 6 Good, K. M., Houser, A., Arntzen, L. & Turnbull, P. C. Naturally acquired
238 anthrax antibodies in a cheetah (*Acinonyx jubatus*) in Botswana. *Journal of*
239 *wildlife diseases* **44**, 721-723 (2008).
- 240 7 Wafula, M. M., Patrick, A. & Charles, T. Managing the 2004/05 anthrax outbreak
241 in Queen Elizabeth and Lake Mburo National Parks, Uganda. *African Journal of*
242 *Ecology* **46**, 24-31 (2008).
- 243 8 Clegg, S., Turnbull, P., Foggin, C. & Lindeque, P. Massive outbreak of anthrax in
244 wildlife in the Malilangwe Wildlife Reserve, Zimbabwe. *The Veterinary Record*
245 **160**, 113-118 (2007).
- 246 9 Muoria, P. K. *et al.* Anthrax outbreak among Grevy's zebra (*Equus grevyi*) in
247 Samburu, Kenya. *African Journal of Ecology* **45**, 483-489 (2007).
- 248 10 Turnbull, P. *et al.* Anthrax in wildlife in the Luangwa Valley, Zambia. *The*
249 *Veterinary Record* **128**, 399-403 (1991).
- 250 11 de Vos, V. The ecology of anthrax in the Kruger National Park, South Africa.
251 *Salisbury Med Bull Suppl* **68**, 19-23 (1990).
- 252 12 Lembo, T. *et al.* Serologic surveillance of anthrax in the Serengeti ecosystem,
253 Tanzania, 1996-2009. *Emerg Infect Dis* **17**, 387-394, doi:10.3201/eid1703.101290
254 (2011).
- 255 13 Beyer, W. *et al.* Distribution and molecular evolution of *Bacillus anthracis*
256 genotypes in Namibia. *PLoS neglected tropical diseases* **6**, e1534 (2012).
- 257 14 Leendertz, F. H. *et al.* Anthrax kills wild chimpanzees in a tropical rainforest.
258 *Nature* **430**, 451-452, doi:10.1038/nature02722 (2004).
- 259 15 Klee, S. R. *et al.* The genome of a *Bacillus* isolate causing anthrax in chimpanzees
260 combines chromosomal properties of *B. cereus* with *B. anthracis* virulence
261 plasmids. *PLoS One* **5**, 0010986 (2010).
- 262 16 Brézillon, C. *et al.* Capsules, toxins and AtxA as virulence factors of emerging
263 *Bacillus cereus* biovar *anthracis*. *PLoS Negl Trop Dis* **9**, e0003455,
264 doi:10.1371/journal.pntd.0003455 (2015).
- 265 17 Antonation, K. S. *et al.* *Bacillus cereus* biovar *anthracis* causing anthrax in sub-
266 Saharan Africa - Chromosomal monophyly and broad geographic distribution.
267 *PLOS Neglected Tropical Diseases* (2016).
- 268 18 Leendertz, F. H. *et al.* Anthrax in Western and Central African great apes.
269 *American journal of primatology* **68**, 928-933, doi:10.1002/ajp.20298 (2006).
- 270 19 Calvignac-Spencer, S. *et al.* Carrion fly-derived DNA as a tool for comprehensive
271 and cost-effective assessment of mammalian biodiversity. *Molecular ecology* **22**,
272 915-924, doi:10.1111/mec.12183 (2013).
- 273 20 Berry, H. H. Surveillance and control of anthrax and rabies in wild herbivores and
274 carnivores in Namibia. *Rev Sci Tech* **12**, 137-146 (1993).
- 275 21 Vergnaud, G. *et al.* Comparison of French and worldwide *Bacillus anthracis*
276 strains favors a recent, Post-Columbian origin of the predominant North-
277 American Clade. *PLoS One* **11**, e0146216, doi:10.1371/journal.pone.0146216
278 (2016).
- 279 22 Smith, K. L. *et al.* *Bacillus anthracis* diversity in Kruger National Park. *Journal*
280 *of clinical microbiology* **38**, 3780-3784 (2000).

- 281 23 Campbell, G., Kuehl, H., Diarrassouba, A., N'Goran, P. K. & Boesch, C. Long-
282 term research sites as refugia for threatened and over-harvested species. *Biology*
283 *letters* **7**, 723-726, doi:10.1098/rsbl.2011.0155 (2011).
- 284 24 Blackburn, J. K., Van Ert, M., Mullins, J. C., Hadfield, T. L. & Hugh-Jones, M. E.
285 The necrophagous fly anthrax transmission pathway: empirical and genetic
286 evidence from wildlife epizootics. *Vector-Borne and Zoonotic Diseases* **14**, 576-
287 583 (2014).
- 288 25 Hill, K. *et al.* Mortality rates among wild chimpanzees. *Journal of human*
289 *evolution* **40**, 437-450, doi:10.1006/jhev.2001.0469 (2001).
- 290 26 Köndgen, S. *et al.* Pandemic human viruses cause decline of endangered great
291 apes. *Current biology : CB* **18**, 260-264, doi:10.1016/j.cub.2008.01.012 (2008).
- 292 27 Boesch, C. & Boesch-Achermann, H. *The chimpanzees of the Tai Forest:*
293 *Behavioural ecology and evolution.* (Oxford University Press, USA, 2000).

294

295 **Supplementary information**

296 Supplementary information: this file contains a more detailed method section as well as
297 additional tables (Tables S1-10) and figures (Fig. S1-8).

298 Additional Data Table S1: Results derived from the analyses of flies caught in TNP
299 analyzed in this study. This file includes results from PCR and culture as well as flymeal
300 analysis results for a selection of flies.

301 Additional Data Table S2: Results of fly meal analysis with taxonomic assignment at
302 genus level. This file provides the number of sequences per amplicon assigned at genus
303 level.

304 Additional Data Table S3: Results of fly meal analysis with taxonomic assignment at
305 order level. This file provides the number of sequences per amplicon assigned at order
306 level.

307

308 **Acknowledgements**

309 We thank the authorities in Côte d'Ivoire for long-term support, especially the Ministry of
310 the Environment and Forests, the Ministry of Research, the directorship of the Tai

311 National Park, and the CSRS in Abidjan. We also thank the national authorities from all
312 other countries for providing permissions for our research (MINFoF, MINRESI, the
313 Service de la Conservation de la Réserve du Dja, Cameroon, in CAR the Ministère d'Eaux
314 et Fôret, Chasse et Peche and the Ministère de l'Education Nationale, de
315 l'Alphabetisation, de l'Enseignement Superieur, et de la Recherche, the Agence
316 Nationale des Parcs Nationaux, Gabon, Centre National de la Recherche Scientifique et
317 Technologique, Gabon, Direction des Eaux, Forêts et Chasses, Senegal, Forestry
318 Development Authority, Liberia, Institut Congolais pour la Conservation de la Nature,
319 DR-Congo, Ministère de l'Agriculture de l'Elevage et des Eaux et Forêts, Guinea,
320 Instituto da Biodiversidade e das Áreas Protegidas (IBAP), Guinea-Bissau, Ministère de
321 la Recherche Scientifique, DR-Congo, Ministère de le Recherche Scientifique et
322 Technologique, R-Congo, Nigeria National Park Service, Nigeria, Uganda National
323 Council for Science and Technology, Ugandan Wildlife Authority, Uganda).

324 We thank the WWF CAR for their logistical support. For the collection of samples in
325 TNP we thank the field assistants, Arthur Henlin, Katerina Albrechtova and Alexander
326 Lang. We are also grateful to the field assistants from all other sites for their support. For
327 laboratory we are grateful to Silke Becker, Tatjana Franz, Sabine Howaldt, Angelika
328 Lander, Petra Lochau, Herbert Nattermann and Andy Schneider, for sequencing we thank
329 the RKI Central Sequencing Laboratory, particularly Julia Hinzmann, Andreas Nitsche
330 and Julia Tesch and for bioinformatic support to Piotr Wojciech Dabrowski and Torsten
331 Semmler from RKI, as well as Graham Hamilton at Glasgow Polyomics. We would
332 further like to thank Teresa Börding, Thurston Hicks, Yasmin Moebius, Volker Sommer,
333 Klaus Zuberbühler and Martine Peeters, and for administrative support Maja Kovacev-

334 Wegener. For funding we thank the German Research Council DFG KL 2521/1-1 and the
335 Sonnenfeld-Stiftung. We would like to thank the Max-Planck-Society and Krekeler
336 Foundation for funding of the *Pan African Programme*.

337

338 **Author contributions**

339 CH, FZ, AA, SA, MA, GB, KC, PD, KD, HE, PF, YG, AG, AG, SMG, JH, SJ, JJ, JK,
340 KL, JL, KL, VL, TL, SM, AM, SM, MM, JvS and ET collected flies, bones and
341 according field data. Necropsies on wildlife found dead were performed by FZ, KN, AB,
342 ECH, AD, PF, SAL, TL, SM, SN, HDN and FHL and laboratory analyses performed by
343 CH, FZ, KN, SD, KMR, KM, SM, HDN, AS, UT, SK, SC and FHL. The data were
344 analyzed by CH, FZ, RB, HK, RM and SC and the manuscript prepared by CH, FZ, RB,
345 HK, RM, JG, SC and FHL. The manuscript was revised and approved by all authors. The
346 study was supervised by CB, RW, SC and FHL.

347

348 **Author information**

349 Reprints and permissions information is available at www.nature.com/reprints. The
350 authors declare no competing financial interests. Correspondence and requests for
351 materials should be addressed to leendertzf@rki.de.

352

353 **Figure legends**

354 **Fig. 1. *Bcbva* occurrence and study sampling sites in sub-Saharan Africa.** Sites with
355 known *Bcbva* occurrence are indicated in red. Detection of *Bcbva* in Tai National Park,
356 Dja Reserve, Dzanga-Sangha Protected Areas and Luebo has been described in previous

357 studies. For all *Bcbva* sites, except Luebo, samples were available. Within this study we
358 could identify Grebo as a new site of *Bcbva* occurrence. *Bcbva* was not detected at the
359 other tested sub-Saharan sites (indicated in black).

360

361 **Fig. 2. *Bcbva* cases in Taï National Park.** (A) *Bcbva* positive and negative carcasses.

362 38% of the observed wildlife mortality in Taï National Park is due to *Bcbva*. *Bcbva*

363 positive carcasses were broadly distributed throughout the research area with no obvious

364 pattern identifiable. GPS data was available for 113 of 204 detected carcasses and not for

365 those detected before 2001. (B) *Bcbva* positive and negative fly traps. Five percent of all

366 analyzed flies contained *Bcbva* genetic material. Flies were also caught outside the

367 research area. A systematic snapshot sampling revealed higher prevalence of *Bcbva*

368 positive fly traps within the research area.

369

370 **Fig. 3. Phylogenomic tree of *Bcbva* isolates.** Maximum likelihood tree based on

371 chromosomal sequences of *Bcbva* isolates from TNP (Côte d'Ivoire, n=124) and Grebo

372 (Liberia, n=2). One sequence per host (mammals/fly, two divergent isolates for fly 600)

373 was included and the final alignment of variant sites measured 298bp. Internal branches

374 with bootstrap values lower than 90 are colored in grey. The colored strip represents

375 different host species. The tree was rooted using the heuristic residual mean squared

376 function in *TempEst* v 1.5. The scale bar is in substitution per chromosomal site.

377

378 **Fig. 4. Proportions of simulated chimpanzee communities surviving 150 years with**

379 **and without presence of anthrax.** Shown are results for different community sizes and

380 anthrax being absent (a, blue boxes) or present (p, red boxes). Bars represent median

381 estimates and boxes quartiles across a range of simulation models assuming different
382 inter birth intervals and maximum ages. All models summarized here assumed an annual
383 per capita survival rate of 0.96.

384

385 **Methods**

386 *Study sites*

387 TNP covers an area of 3,300 km² and an additional 200 km² buffer zone. Since 2001 a
388 veterinary program conducts outbreak investigations in wildlife. We defined the research
389 area as the habitat ranges of the three habituated chimpanzee groups plus a 500 m buffer
390 zone (103 km²; Fig. S2).

391 Samples belonging to the large-scale data set were collected at 16 sites in 11 sub-Saharan
392 countries stretching from Senegal to Uganda (Fig. S3, Table S1). Most sites (14 out of
393 16) were temporary research sites of the *Pan African Programme*
394 (www.panafrican.eva.mpg.de) where *Bcbva* has not been described. Additional samples
395 were obtained from Dja Faunal Reserve (DJR), Cameroon¹⁸ and Dzanga-Sangha
396 Protected Areas (DSPA), Central African Republic¹⁷, where *Bcbva* cases have been
397 previously described. Study sites are described in detail in Supplementary information
398 S1.

399 *Necropsies*

400 Carcass monitoring was performed in TNP by a veterinarian, performing necropsies on
401 every carcass reported by researchers working in the forest (n=173). Samples of all inner
402 organs were collected, as far as carcass decomposition allowed. Necropsies followed a

403 standardized protocol, including use of full personal protective equipment. Carcass sites
404 were decontaminated according to World Health Organization (WHO) guidelines^{5,28}. For
405 each sample aliquots were stored in liquid nitrogen and formalin in the field. Frozen
406 samples were transported on dry ice and subsequently stored at -80°C. We received
407 additional tissue samples from carcasses sampled by the WHO in TNP between 1996 and
408 2000 (n=31) (Table S2).

409 Rather than using serology, which would also detect animals that survived non-lethal
410 infections, we used PCRs to detect the presence of anthrax in internal organs to confirm
411 that anthrax was the likely cause of death. DNA was extracted from various tissues per
412 animal (liver, spleen and lung when available) using the DNeasy Blood and Tissue Kit
413 (Qiagen, Hilden, Germany); extracts were quantified using a Nanodrop (Thermo Fisher
414 Scientific, Waltham, MA, USA) and stored at -20°C. Two hundred ng DNA or 5 µl of
415 DNA extract (if DNA concentration was below 40 ng/µl) were tested for anthrax in
416 duplicate real-time PCR reactions (details in Supplementary Methods S2c). The full
417 anthrax assay used includes three real-time PCRs, each targeting one of the following
418 gene markers: *pag* (gene for protective antigen) located on the pXO1 plasmid²⁹, *capB*
419 (gene for capsule synthesis) located on pXO2 and *Island IV*, a chromosomal marker
420 specific for *Bcbva*^{15,17} (Table S3). Samples were first tested for *pag* and samples positive
421 in duplicate for *pag* were tested for *capB* and *Island IV* (Extended Data Fig. S1 and
422 Extended Data Fig. S2).

423 Culture under BSL3 conditions was attempted for all PCR positive necropsy samples
424 collected until the end of 2013 (June 2014 for duikers) (Table S2). A native and heat-
425 treated (65°C for 30 min, to assess presence of spores) aliquot were plated onto the

426 following agar plates: Columbia blood agar (Oxoid, Wesel, Germany), blood-
427 trimethoprim agar (1.6 mg trimethoprim, 6.4 mg sulfamethoxazole, 20 mg polymyxin B
428 per liter agar medium) and Cereus Ident agar (Heipha Diagnostica, Eppelheim, Germany)
429 with the chromogenic substrate 5-bromo-4-chloro-3-indoxyl-myoinositol-1-phosphate³⁰.
430 Cultures were incubated at 37°C and monitored daily. Morphologically suspicious
431 colonies were sub-cultured and tested in real-time PCR. *Bcbva* was cultured from native
432 and heat-treated samples indicating the presence of heat-resistant spores. Isolates were
433 frozen in Microbank tubes (Mast Diagnostica, Reinfeld, Germany) at -80°C.

434 Histopathology was performed on a subset of necropsy samples, including 15 *Bcbva* PCR
435 positive necropsy samples (Table S2). No signs of anthrax infection were detected in
436 carcasses that were PCR negative for anthrax, while for PCR positive carcasses the most
437 consistent histopathologic finding was per-acute to acute anthrax related pneumonia
438 characterized by mild lymphohistiocytic infiltrates and intraalveolar eosinophilic and
439 proteinaceous or fibrinous material. Numerous bacilli were found intravascular and
440 intraalveolar. Multifocal alveolar and peribronchiolar hemorrhages were present in all
441 animals. Lymph node changes consisted of sinus histiocytosis, cortical hemorrhages and
442 edema especially in the mediastinal, tracheobronchiolar and mesenteric lymph nodes.
443 Huge amounts of bacilli were demonstrable in the sinusoids. Within the abdominal cavity
444 the spleen was the organ most affected, with myriads of bacilli visible in the splenic
445 sinusoids, partly embedded in fibrin deposition. There was moderate lymphoid depletion,
446 lymphocytolysis and histiocytosis. The liver parenchyma was severely congested with
447 masses of bacilli within the hepatic sinusoids. All anthrax PCR positive carcasses were

448 also tested for filoviruses³¹ and respiratory diseases²⁶ to rule out co-infection with other
449 common causes of death in this ecosystem.

450 *Blow flies*

451 Flies were caught on the ground and in the canopy using custom-made traps (Fig. S4 and
452 S5, Supplementary information S3a). Trapping was done for 60 min or until 20 flies were
453 collected. Flies were euthanized with ether and stored at -20°C in 2 ml Cryotubes (Carl
454 Roth) containing up to 10 flies or at ambient temperature on silica in 50 ml Falcon tubes
455 (Thermo Fisher Scientific) containing up to 20 flies. In TNP, 726 flies were randomly
456 collected within the research area in 2008, 2009, 2012 and 2013 (Table S4). Another 908
457 flies were collected over 19 days in May and June 2014 according to a 2x2 km grid
458 system covering the research area and 225 km² surrounding the research area (referred to
459 as “snapshot flies”; Extended Data Fig. S7, Table S4). At a larger scale, 784 flies were
460 collected at 8 sites within 5 sub-Saharan countries (*Pan African Programme*) from 2012
461 to 2014 (Table S4) and 305 flies were analyzed from two sub-Saharan sites, DJR (n=105)
462 and DSPA^{17,18} (n=200) (Table S4). In total, 2,723 flies were analyzed (Table S4).

463 DNA extraction of individual flies was performed using the GeneMATRIX Stool DNA
464 Purification Kit (Roboklon, Berlin, Germany). We followed manufacturer’s instructions
465 except that each fly was first cut into small pieces using sterilized scissors before being
466 homogenized using a Fast Prep® (MP Biomedicals, Santa Ana, CA, USA). DNA
467 concentration measurements and anthrax testing by real-time PCR were performed as
468 described for necropsy samples (Table S4).

469 A subset of 50 flies containing high *pag* copy numbers underwent bacterial culture (Table
470 S4, Additional Data Table S1). Half of the fly mush remaining after DNA extraction was

471 plated directly onto the same culture media described for necropsy samples. Additionally,
472 a 10 µl aliquot of the mush was diluted 1:10 in sterile NaCl, heat treated for 30 min at
473 65°C and plated. *Bcbva* was retrieved from native and heat-treated samples, indicating
474 the presence of heat-resistant spores in flies. An on-site study in TNP also used direct
475 culture of 204 flies without preceding PCR testing. Flies were homogenized and plated
476 directly onto Cereus Ident agar. Suspicious colonies were sub-cultured on blood-
477 trimethoprim agar and tested in real-time PCR. This approach yielded another 21 *Bcbva*
478 isolates.

479 To examine whether certain mammals were preferentially affected by *Bcbva*, we tested
480 for differences in fly meal composition of anthrax positive and negative flies. We
481 screened a subset of 750 TNP flies for mammalian DNA using a real time PCR targeting
482 a 130 bp fragment of mammalian 16S mitochondrial DNA (described in Calvignac-
483 Spencer et al.¹⁹). We chose a subset of mammal and anthrax positive (n=28) and the
484 according number of mammal positive but anthrax negative flies (n=29) from the same
485 traps (Additional Data Table S1). To dissect fly meal composition, we used a
486 metabarcoding approach, whereby 16S amplicons were deep-sequenced, adapting the
487 amplicon preparation protocol provided by Illumina (San Diego, CA, US)
488 (Supplementary information S3e). We used a custom pipeline to determine taxonomic
489 assessment of each read to the genus and order level described in the Supplementary
490 information S3e (Table S5, Additional Data Tables S1 and S2). Sequences assigned to
491 domestic animals were regarded as contamination as it was shown that even stringent
492 anti-contamination procedures do not prevent the amplification of human and domestic
493 animal sequences present in the environment and reagents³².

494 Details on blow fly analyses and results are in Supplementary information S3.

495 ***Bones***

496 Bones were collected in TNP and 12 *Pan African Programme* sites in 9 countries (Fig.
497 S3, Table S1 and S7). Bones were transported and stored at ambient temperature. DNA
498 was extracted using a silica-based method^{33,34} (Supplementary information S4b). Bone
499 extracts were tested by real-time PCR as described for necropsy samples (Table S1 and
500 S7). Powder from PCR positive bones was also used for bacterial culture attempts after
501 homogenization in sterile NaCl (Table S7). We processed the homogenates as described
502 above for necropsy samples with one native aliquot and one heat- treated aliquot. Details
503 on bone analyses and additional results are in Supplementary information S4.

504 ***Whole-genome sequencing of Bcbva isolates and SNP calling***

505 Table S8 contains a complete list of all *Bcbva* isolates sequenced (Fig. S6). Isolate
506 preparation and extraction is described in the Supplementary Methods S6a. Libraries for
507 whole-genome sequencing were prepared with the Nextera XT DNA Library Preparation
508 Kit (Illumina). Libraries were pooled and sequenced on the HiSeq 1500 platform
509 (Illumina) in rapid run mode using either v1 (2x150 bp) or v2 (2x250 bp) chemistry.
510 Illumina adapters were removed using *scythe* v0.993³⁵ and trimmed with *sickle* v1.33³⁶
511 applying a quality threshold of 25. Quality trimmed reads were aligned to the reference
512 genome (*Bcbva* strain CI, Accession numbers CP001746-749) with the BWA-MEM
513 algorithm implemented in *bwa* v0.7.12-r1039³⁷. For conversion to bam format, sorting,
514 deduplication and indexing of aligned reads, we used the *picard tools* 1.136³¹ software
515 package applying the commands *SortSam*, *MarkDuplicates* and *BuildBamIndex*.
516 Subsequent variant calling was performed using the *Genome Analysis Toolkit* (*GATK*)

517 v3.4³⁸⁻⁴⁰. We realigned bam files with the tools *RealignerTargetCreator* and
518 *IndelRealigner*. Variants were called with *UnifiedGenotyper* with a minimum phred
519 scaled confidence threshold of 30 for SNPs to be called. Hard filtering of SNP sites was
520 done with the *VariantFiltration* command using recommended filter settings. With the
521 *SelectVariants* command, only SNP sites that passed the filter were selected for further
522 processing. *SelectVariants* was also used to exclude all SNPs with a coverage < 5x, a
523 minor allele frequency of > 0.1 and a *GATK* Genotype Quality value < 99. Final
524 consensus sequences were composed with *FastaAlternateReferenceMaker*. We assessed
525 coverages of all samples with the *GATK* tools *DepthOfCoverage* and
526 *CoveredByNSamplesSites*. Details and further analysis of whole-genome sequencing of
527 *Bcbva* isolates and SNP calling is Supplementary information S6.

528 ***Phylogenetic analyses***

529 126 genome sequences (one isolate per mammal/fly) from TNP and GNF (Table S8)
530 were aligned and stripped of non-variant sites with *Geneious Pro* v8.1.3 (Biomatters
531 ltd.)⁴¹. Resulting alignments of variable sites were 298, 18 and 11 bp long for the
532 chromosome, pXO1 and pXO2 respectively. Given the low number of variable sites in
533 pXO1 and pXO2, we only performed phylogenetic analyses on the chromosome
534 alignment. *jModelTest* v2.1.4⁴² was used for determination of the best nucleotide
535 substitution model in a maximum likelihood (ML) framework, resulting in the choice of
536 TVMef⁴³.
537 ML analysis was performed with *PhyML* v20131022⁴⁴ using a combination of subtree-
538 pruning-grafting (SPR) and nearest-neighbor-interchange (NNI) tree search algorithms.
539 Branch support was estimated using non-parametric bootstrapping with 100 pseudo-

540 replicates. The tree was rooted using the heuristic residual mean squared function in
541 *TempEst v 1.5*⁴⁵, placing the root at the position resulting in the most clock-like structure
542 of the data(Fig. 3) .

543 We also performed phylogenetic analyses using the Bayesian Markov Chain Monte Carlo
544 (BMCMC) sampling approach implemented in *BEAST v1.8.2*⁴⁶ specifying a constant
545 population coalescent tree prior and assuming an uncorrelated lognormal relaxed
546 molecular clock⁴⁷ (Supplementary information S7c). The maximum clade credibility tree
547 derived from this analysis was very similar to the ML tree (Fig. 3).

548 Another data set was assembled to compare *Bcbva* from TNP to other strains from sub-
549 Saharan Africa. It included the chromosomal sequences from a representative TNP
550 genome, GNF ones, as well as previously published genomes determined from isolates
551 derived from *Bcbva* cases in CAR and CM^{15,17} (Extended Data Fig. S6). The alignment
552 was compiled as described above and contained 1,016 variable positions. Model selection
553 with *jModelTest v2.1.4*⁴² selected a TPM1⁴⁸ nucleotide substitution model. We performed
554 ML analyses as described above.

555 ***Statistical analyses***

556 To test the effect of season on the probability of a carcass or fly, respectively, being
557 anthrax positive, we used a Generalized Linear Mixed Model (GLMM)⁴⁹ with binomial
558 error structure and logit link function⁵⁰. As predictors we included the species (monkeys,
559 chimps, duikers, others, blow flies), season and their interaction. 'Season' was modelled
560 by first turning the sampling date into a circular variable and including its sine and cosine
561 into the model. As random intercept effects we included trap id (i.e., GPS location) and
562 the combination of sampling date and GPS location, the latter accounting for potential

563 non-independence of flies sampled on the same day from the same trap. We further
564 included random slopes of season within trap id^{51,52}. To test the effect of season we
565 compared the full model with a null model lacking the fixed effects of season and its
566 interaction with species⁵³, using a likelihood ratio test⁵⁴. Sample size for this model was
567 1803 samples (carcasses and flies), collected at 352 locations and 328 combinations of
568 sampling date and location including necropsy samples and flies.

569 In a second model we tested whether the probability of a fly to be tested anthrax positive
570 was influenced by season and the amount of mammalian DNA within in the fly. We used
571 a Generalized Linear Mixed Model (GLMM)⁴⁹ with binomial error structure and logit
572 link function⁵⁰. Into this we included the amount of mammalian DNA found within the
573 fly (determined with real time PCR described above) and season as fixed effects. 'Season'
574 we modeled by first turning the sampling date into a circular variable and then including
575 it sine and cosine into the model. Since the amount of mammalian DNA within the fly
576 was highly skewed, we log transformed it before fitting the model. As random effects
577 (random intercepts) we included the ID of the trap and the date of sampling. To avoid
578 overconfident estimates we included random slopes of the amount of mammal DNA
579 within trap ID and trapping date^{51,52}. As an overall test of the effects of the amount of
580 mammal DNA and season we compared the full model with a null model lacking these
581 effects⁵³ using a likelihood ratio test⁵⁴. We also used likelihood ratio tests to test for the
582 individual predictors (comparing the full model with a respective reduced model lacking
583 the predictor to be tested⁵¹). We fitted the model in R⁵⁵ using the function *glmer* of the R
584 package *lme4* (version 1.1-10⁵⁶). To estimate model stability we excluded levels of the
585 random effects one at a time which did not indicate influential levels to exist. The total

586 sample size for this model was a total of 474 flies caught on 43 days in 33 traps.
587 (Extended Data Fig. S4, Table S10, Supplementary information S8b).

588 To evaluate the reproducibility of fly meal identification for each fly we correlated the
589 proportion of sequence counts per amplicon (two amplicons per fly) that was assigned to
590 different mammalian genera using Spearman correlation. To test whether there were
591 differences in fly meal composition of anthrax positive and anthrax negative flies, we
592 tested whether detection of a given mammal taxon in a fly sample was associated with
593 anthrax positivity. We used GLMMs⁴⁹ applied separately for each mammal genus
594 identified in the flies. The response was whether the fly was anthrax positive and the key
595 predictor with fixed effect was mammal presence. We considered a mammal to be
596 present when it was detected in at least one of the two amplicons per fly. We included
597 only those mammal genera in the model that were detected in at least five of all generated
598 amplicons (two per fly). In addition to mammal presence, we included tid and the factor
599 sampling date as random effects (random intercepts)^{51,52}. Models were fitted with
600 binomial error structure and logit link function⁵⁰. Sample size for all models was 57 flies,
601 caught in 22 different traps on 13 days. To test whether mammal presence had an impact
602 on anthrax positivity, we dropped mammal presence from the model⁵³ and compared the
603 models using a likelihood ratio test⁵⁴. Model stability was assessed as above. We fitted
604 models at two different taxonomic resolutions: one with taxonomic assignment at genus
605 level and the other at order level. GLMMs were fit in R ⁵⁵ using the function *glmer* of the
606 R package *lme4* v1.1-10⁵⁶.

607 To evaluate geographic distribution of *Bcbva* in TNP we checked whether, due to higher
608 mammal density²³, *Bcbva* was more likely to occur inside the research area. To test this

609 hypothesis we analyzed 908 flies from 83 different traps (Extended Data Fig. S7, Table
610 S4, Additional Data Table S1). 21 traps were located within the research area and 62
611 traps in the adjoining forest belt. 8/21 traps within the research area were anthrax positive
612 and 8/62 outside the research area. We compared the two groups using Fisher Exact's
613 Test (Supplementary information S8d).

614 To learn more about the spatial dynamics of *Bcbva* in TNP, we investigated the
615 correlation between genetic and geographic distances. To correct for genetic and spatial
616 autocorrelation, we excluded strains from the data set that originated from the same fly
617 catching point (in a 1 km² radius) on the same day or from the same followed-up outbreak
618 in mammals. Only one strain was kept per outbreak or fly catching point, the selection
619 criterion being high average coverage of the genome (Table S8, Supplementary
620 information S8e). Geographic distances (in km) were derived from GPS data using
621 *GeographicDistanceMatrixGenerator* v1.2.3⁵⁷. Genetic distances were approximated
622 using the relative distances drawn from a Maximum Likelihood Tree built in *PhyML*
623 v20131022⁴⁴ with the R package *ape*⁵⁸ using the *cophenetic* function. Multiple regression
624 on distance matrices (MRM) as implemented in the R *ecodist* package⁵⁹ using 1000
625 permutations and Spearman correlations was performed on genetic and geographic
626 distance matrices. To examine variation within genetic lineages, we binned our data by
627 genetic distance (bin size=relative genetic distance of 0.03, approx. 2.5 SNPs) and
628 focused on groups with low genetic distance (max relative genetic distance <0.5) and
629 their mean geographic distance (Extended Data Fig. S9). Homogeneity of variance
630 between groups was assured with the Fligner Killeen test (p=0.07; >0.05 as requested).

631 To evaluate the impact *Bcbva* could have on the TNP chimpanzee population, we
632 conducted a simulation (Supplementary information S8f). We first defined a series of
633 population parameters for the simulation²⁷. We simulated the survival prospects of
634 chimpanzee communities of a given size, with individuals reproducing at certain regular
635 intervals after maturation, having a maximum age, and an annual survival probability.
636 Since most of these parameters are associated with considerable uncertainty and since we
637 wanted to assess to what extent the simulation results depend on the particular parameters
638 chosen we parameterized the simulations as follows: Initial community size: 20 to 80
639 individuals (increment: 10); inter-birth interval: 4 to 7 years (increment: 1); interval after
640 death of infant: 1 year; maximum age: 40 to 50 years (increment: 2); age of first
641 reproduction of males and females: 10 and 14 years, respectively. Since per capita annual
642 survival probability without the influence of anthrax is unknown (mortality cases due to
643 anthrax may not be detected in all cases, in particular before necropsies were made
644 systematically), we simulated per capita annual survival probabilities from 0.93 to 0.99
645 (increment: 0.03). In addition, we made survival probability density dependent, as this is
646 a common characteristic observed in many species including chimpanzees⁶⁰. For this we
647 introduced a logistic function ($1/(1+\exp(-(20-0.08*\text{community size})))$) that increased or
648 reduced mortality rate as a function of chimpanzee community size. At the beginning of
649 each simulation run we generated a community of the simulated size by randomly
650 allocating sexes (proportion of females: 0.7) and ages (uniformly distributed between 10
651 and the simulated maximum age) to individuals. To avoid stochastic effects of the
652 initially generated community, we let the simulation run for 50 time steps (i.e., 'years')
653 without anthrax presence before the evaluated time period began.

654 We estimated the risk of annual anthrax outbreak probability, its dependence on
655 community size and the number of individuals affected as $\exp(-1.83+0.039*\text{community}$
656 $\text{size})$ from a Poisson regression (null, full model comparison, $\chi^2= 7.89$, $df=1$, $p < 0.01$).
657 We simulated both an anthrax and a non-anthrax scenarios for 150 time steps (i.e., 'years')
658 with 100 replications each and for each possible combination of the simulated
659 parameters. Communities were considered to be extinct, when no reproducing females
660 were present.

661 All R scripts are available upon request. Details on statistical analyses and additional
662 results are in Supplementary information S8.

663 ***Data availability***

664 Raw reads of 16S amplicons are available in the European Nucleotide Archive (ENA)
665 under project accession number PRJEB14554, sample accession numbers ERS1217219-
666 336. Raw reads for all 178 *Bcbva* isolates from TNP and GNF are available in the ENA
667 under project accession number PRJEB14616, sample accession numbers ERS1222903-
668 3080. Variable position alignments are available from the Dryad Digital Repository:
669 <http://dx.doi.org/10.5061/dryad.v8bn7>.

670

671 **Extended data figure legends**

672 **Extended Data Fig. S1. Necropsies performed since 1996.** Shown is the total amount
673 of necropsies performed per year in TNP from 1996 to 2015. Grey bars indicate the
674 number and according proportion of *Bcbva* positive necropsies. In the years 2003 and
675 2010 only limited veterinary service was available at TNP due to political insecurity in
676 the region.

677

678 **Extended Data Fig. S2. Geographic location of *Bcbva* positive carcasses in Tai**
679 **National Park.** Shown are *Bcbva* positive tested necropsies in TNP since 2001. GPS data
680 was available for 70 of all positive tested (n=81) necropsies.

681

682 **Extended Data Fig. S3. Effect of mammalian DNA content on anthrax positivity in**
683 **flies.** . Shown is the probability of *Bcbva* positivity (PA, pag respectively) as a function
684 of the amount of mammalian DNA (copies) found in a fly. The amount of mammal DNA
685 was binned (bin width 0.25) and the area of the points depicts the number of flies (range:
686 1 to 206) in the respective bins. The dashed line indicates the fitted model and the dotted
687 lines the 95% confidence interval.

688

689 **Extended Data Fig. S4. Effect of season on anthrax positivity in flies.** Shown is the
690 probability of *Bcbva* (PA) positivity over the course of a year (binned in 10 day periods).
691 The area of the points depicts the number of flies in the respective ten days period. The
692 dashed line indicates the fitted model and the dotted lines the 95% confidence interval.

693

694 **Extended Data Fig. S 5. Maximum clade credibility tree based on chromosomal**
695 **sequences of *Bcbva* isolates from TNP (Côte d’Ivoire, n=124) and Grebo (Liberia,**
696 **n=2).** One sequence per host hosts (mammals/flies, two divergent isolates for fly 600)
697 was included and the final alignment of variant sites measured 298bp. Size of nodes
698 represents posterior probability values. The location of the root received a posterior
699 probability of 1.

700

701 **Extended Data Fig. S 6. Maximum likelihood tree for sub-Saharan *Bcbva* strains.**
702 ML tree based on chromosomal sequences of *Bcbva* strains from Côte d’Ivoire,
703 Cameroon, Central African Republic and Liberia. The alignment of variant sites
704 measured 1016bp. Bootstrap values are shown above branches and the scale bar reflects
705 the genome-wide substitution rate. The tree was rooted using TempEst v 1.5.

706

707 **Extended Data Fig. S 7. Fly snapshot sampling scheme.** For the fly snapshot flies were
708 caught following a 2x2 km grid system within and outside the research area within 19
709 days. In total 908 snapshot flies were analyzed.

710

711 **Extended Data Fig. S 8. Genetic and geographic distances of *Bcbva* isolates from the**
712 **fly snapshot.** (A) Maximum Likelihood Tree based on chromosomal sequences of *Bcbva*
713 isolates from the 19 day fly snapshot. Each dot represents one fly isolate. Colors were
714 chosen to illustrate the distribution of genetically clustering isolates on the map presented
715 in panel B. The final alignment of variant sites measured 123bp. Bootstrap values are
716 shown above all internal branches. The tree was rooted using the “best-fit” option in

717 Path-O-Gen v1.2. The scale bar is in substitution per site. (B) Geographic origin of *Bcbva*
718 isolates collected during the fly snapshot. Colors correspond to ML tree (A). Big circles
719 represent two isolates.

720

721 **Extended Data Fig. S 9.** Boxplot of genetic and mean geographic distances. *Bcbva*
722 isolates from TNP were binned by relative genetic distance (bin size = 0.03, approx. 2.5
723 SNPs). The two most genetically distant isolates received a value of 1 and all other
724 distances were scaled accordingly. Diamonds indicate the geographic distance means of
725 the groups. To examine variation within genetic lineages, we analyzed isolates with low
726 genetic distance (max relative genetic distance < 0.5, marked in blue) and their mean
727 geographic distance. For low genomic distances, genetic and mean geographic distances
728 are correlated ($p= 4 \times 10^{-5}$, $R^2= 0.72$).

729

730 **Extended Data Fig. S10. Fly species composition based on GMYC analysis.** Fly
731 species composition for three sites with known *Bcbva* occurrence: TNP (Côte d'Ivoire)
732 (A), DJR (Cameroon) (B), DSPA (Central African Republic) (C). Shown are the
733 proportions of flies per site in % belonging to one single fly species identified with
734 GMYC models. Different colors indicate different taxonomic fly families.

735

736 **References (28-60)**

- 737 28 Turnbull, P. C. B. (World Health Organization, Department of Communicable
738 Diseases Surveillance and Response, 1998).
- 739 29 Ellerbrok, H. *et al.* Rapid and sensitive identification of pathogenic and
740 apathogenic *Bacillus anthracis* by real-time PCR. *FEMS Microbiol Lett* **214**, 51-
741 59 (2002).
- 742 30 Klee, S. R. *et al.* Characterization of *Bacillus anthracis*-like bacteria isolated from
743 wild great apes from Cote d'Ivoire and Cameroon. *J Bacteriol* **188**, 5333-5344
744 (2006).
- 745 31 Panning, M. *et al.* Diagnostic reverse-transcription polymerase chain reaction kit
746 for filoviruses based on the strain collections of all European biosafety level 4
747 laboratories. *Journal of Infectious Diseases* **196**, S199-S204 (2007).
- 748 32 Leonard, J. A. *et al.* Animal DNA in PCR reagents plagues ancient DNA
749 research. *Journal of Archaeological Science* **34**, 1361-1366 (2007).
- 750 33 Gamba, C. *et al.* Comparing the performance of three ancient DNA extraction
751 methods for high-throughput sequencing. *Molecular ecology resources* **16**, 459-
752 469, doi:10.1111/1755-0998.12470 (2016).
- 753 34 Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth.
754 *Nature protocols* **2**, 1756-1762, doi:10.1038/nprot.2007.247 (2007).
- 755 35 Buffalo, V. Scythe: A 3'-end adapter contaminant trimmer [Software]. Available
756 at <https://github.com/vsbuffalo/scythe>.
- 757 36 Joshi, N. A. & Fass, J. N. Sickle: A sliding-window, adaptive, quality-based
758 trimming tool for FastQ files [Software]. Available at
759 <https://github.com/najoshi/sickle>. (2011).
- 760 37 Li, H. Aligning sequence reads, clone sequences and assembly contigs with
761 BWA-MEM. *arXiv preprint arXiv:1303.3997* (2013).
- 762 38 Auwera, G. A. *et al.* From FastQ data to high-confidence variant calls: the
763 genome analysis toolkit best practices pipeline. *Current Protocols in*
764 *Bioinformatics*, 11.10. 11-11.10. 33 (2013).
- 765 39 DePristo, M. A. *et al.* A framework for variation discovery and genotyping using
766 next-generation DNA sequencing data. *Nature genetics* **43**, 491-498 (2011).
- 767 40 McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for
768 analyzing next-generation DNA sequencing data. *Genome Res* **20**, 1297-1303,
769 doi:10.1101/gr.107524.110 (2010).
- 770 41 Kears, M. *et al.* Geneious Basic: an integrated and extendable desktop software
771 platform for the organization and analysis of sequence data. *Bioinformatics* **28**,
772 1647-1649 (2012).
- 773 42 Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models,
774 new heuristics and parallel computing. *Nat Methods* **9**, 772,
775 doi:10.1038/nmeth.2109 (2012).
- 776 43 Posada, D. Using MODELTEST and PAUP* to select a model of nucleotide
777 substitution. *Current protocols in bioinformatics*, 6.5. 1-6.5. 14 (2003).
- 778 44 Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood
779 phylogenies: assessing the performance of PhyML 3.0. *Syst Biol* **59**, 307-321,
780 doi:10.1093/sysbio/syq010 (2010).

781 45 Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the
782 temporal structure of heterochronous sequences using TempEst (formerly Path-O-
783 Gen). *Virus Evolution* **2** (2016).

784 46 Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian
785 phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* **29**, 1969-1973,
786 doi:10.1093/molbev/mss075 (2012).

787 47 Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics
788 and dating with confidence. *PLoS biology* **4**, e88,
789 doi:10.1371/journal.pbio.0040088 (2006).

790 48 Kimura, M. Estimation of evolutionary distances between homologous nucleotide
791 sequences. *Proceedings of the National Academy of Sciences* **78**, 454-458 (1981).

792 49 Baayen, R. H. *Analyzing linguistic data: A practical introduction to statistics
793 using R*. (Cambridge University Press, 2008).

794 50 McCullagh, P. & Nelder, J. A. *Generalized linear models*. Vol. 37 (CRC press,
795 1989).

796 51 Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for
797 confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* **68**,
798 doi:10.1016/j.jml.2012.11.001 (2013).

799 52 Schielzeth, H. & Forstmeier, W. Conclusions beyond support: overconfident
800 estimates in mixed models. *Behavioral Ecology* **20**, 416-420 (2009).

801 53 Forstmeier, W. & Schielzeth, H. Cryptic multiple hypotheses testing in linear
802 models: overestimated effect sizes and the winner's curse. *Behavioral Ecology
803 and Sociobiology* **65**, 47-55 (2011).

804 54 Dobson, A. J. & Barnett, A. *An introduction to generalized linear models*. (CRC
805 press, 2008).

806 55 Team, R. C. R: *A language and environment for statistical computing*., (R
807 Foundation for Statistical Computing, Vienna, Austria, 2015).

808 56 Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects
809 models using lme4. *Journal of Statistical Software* **67**, 1-48 (2014).

810 57 Ersts, P. Geographic distance matrix generator (version 1.2. 3). American
811 Museum of Natural History. Available at
812 http://biodiversityinformatics.amnh.org/open_source/gdmg. *Center for
813 Biodiversity and Conservation* (2011).

814 58 Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and
815 Evolution in R language. *Bioinformatics* **20**, 289-290 (2004).

816 59 Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis
817 of ecological data. *Journal of Statistical Software* **22**, 1-19 (2007).

818 60 Kuehl, H., Elzner, C., Möbius, Y., Boesch, C. & Walsh, P. D. The price of play:
819 Self-organized infant mortality cycles in chimpanzees. *PLoS One* **3**,
820 doi:10.1371/journal.pone.0002440 (2008).

821