LETTER

doi:10.1038/nature23309

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

Constanze Hoffmann[1]*, Fee Zimmermann[1,2]*, Roman Biek[3], Hjalmar Kuehl[4], Kathrin Nowak[1], Roger Mundry[4], Anthony Agbor[4], Samuel Angedakin[4], Mimi Arandjelovic[4], Anja Blankenburg[1], Gregory Brazolla[4], Katherine Corogenes[4], Emmanuel Couacy-Hymann[5], Tobias Deschner[4], Paula Dieguez[4], Karsten Dierks[4], Ariane Düx[1], Susann Dupke[2], Henk Eshuis[4], Pierre Formenty[6], Yisa Ginath Yuh[4], Annemarie Goedmakers[7], Jan F. Gogarten[1,4,8], Anne-Céline Granjon[4], Scott McGraw[9], Roland Grunow[2], John Hart[10], Sorrel Jones[4], Jessica Junker[4], John Kiang[11], Kevin Langergraber[12], Juan Lapuente[4], Kevin Lee[4], Siv Aina Leendertz[1], Floraine Léguillon[1], Vera Leinert[13], Therese Löhrich[1,4], Sergio Marrocoli[4], Kerstin Mätz-Rensing[14], Amelia Meier[4], Kevin Merkel[1], Sonja Metzger[1], Mizuki Murai[4], Svenja Niedorf[1], Hélène De Nys[1,4], Andreas Sachse[1], Joost van Schijndel[4], Ulla Thiesen[1], Els Ton[7], Doris Wu[1,4], Lothar H. Wieler[15], Christophe Boesch[4], Silke R. Klee[2], Roman M. Wittig[4], Sébastien Calvignac-Spencer[1] & Fabian H. Leendertz[1]

Anthrax is a globally important animal disease and zoonosis. Despite this, our current knowledge of anthrax ecology is largely limited to arid ecosystems, where outbreaks are most commonly reported[1–3]. Here we show that the dynamics of an anthrax-causing agent, *Bacillus cereus* biovar *anthracis*, in a tropical rainforest have severe consequences for local wildlife communities. Using data and samples collected over three decades, we show that rainforest anthrax is a persistent and widespread cause of death for a broad range of mammalian hosts. We predict that this pathogen will accelerate the decline and possibly result in the extirpation of local chimpanzee (*Pan troglodytes verus*) populations. We present the epidemiology of a cryptic pathogen and show that its presence has important implications for conservation.

Anthrax is a disease of wildlife, livestock and humans that predominantly affects low- and middle-income countries[2,4,5]. Although widely distributed, including some temperate regions, anthrax is most commonly associated with arid ecosystems, particularly African savannahs[1,3,6–11]. In these systems, major outbreaks typically cause high mortality in a few wild and domestic ungulate species at a time and usually exhibit strong seasonal and inter-annual variation[2,3,5,11,12]. For example, in Krüger National Park, South Africa, die-offs in kudus (*Tragelaphus strepsiceros*) and impalas (*Aepyceros melampus*) occur in the dry season with a ten-year periodicity coinciding with rainfall cycles[11]. In Etosha National Park, Namibia, mortality in elephants (*Loxodonta africana*) peaks at the start of the wet season, while plains ungulates (*Equus quagga*, *Conochaetes taurineus* and *Antidorcas marsupialis*) are most affected at the end of the wet season[3,13]. Such varying dynamics underline the importance of investigating the pathogen in close relation with its ecosystem, but so far anthrax research in Africa has been biased towards well-studied savannah regions.

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

We address this knowledge gap by testing a unique set of samples collected in TNP over 26 years. We started collecting bones in 1989 resulting in bones from 75 individual mammals (Supplementary Table 7 and Supplementary Information). From 1996 on, we investigated 204 fresh carcasses (Supplementary Table 2 and Supplementary Information). Because bone and carcass discovery was linked to the collection of chimpanzee behavioural data, we expected detection of Bcbva to be biased towards chimpanzees and other easily detectable medium-to-large-bodied mammals. We therefore tested whether carrion flies, which are relatively unbiased samplers of mammalian DNA[19], might also collect Bcbva or its genetic material while feeding and ovipositing on carcasses. Starting in 2008, we applied different horizontal and vertical sampling schemes to collect 1,634 flies (Supplementary Tables 1, 4 and Supplementary Information). We retrieved Bcbva isolates from all three sample types (bones, carcasses and flies). These allowed us to generate 178 whole-genome sequences spanning from 1996 to 2014 (Supplementary Table 8). To clarify the distribution of Bcbva on a larger scale, we sampled 1,089 flies and 136 bones from 16 other sites in 11 sub-Saharan countries from 2012 to 2014 (Fig. 1 and Supplementary Table 1).

In TNP we detected Bcbva DNA in 81 carcasses (40%; Fig. 2a, Extended Data Figs 1, 2 and Supplementary Table 2), 26 bones (35%, Supplementary Table 7) and 80 flies (5%; Fig. 2b, Extended Data Fig. 3 and Supplementary Table 4). We could perform histopathological examinations on 15 positive carcasses and in all cases pathology was consistent with a lethal anthrax infection (Supplementary Table 2). Overall, 38% of observed local wildlife mortality was associated with Bcbva (Supplementary Tables 2, 4), meeting the highest levels of mortality reported for classical anthrax outbreaks in savannah ecosystems[12,20]. We observed no obvious seasonal variation in Bcbva

[1]Robert Koch Institute, P3: "Epidemiology of Highly Pathogenic Microorganisms", Seestraße 10–11, 13353 Berlin, Germany. [2]Robert Koch Institute, ZBS 2: Centre for Biological Threats and Special Pathogens, Highly Pathogenic Microorganisms, Seestraße 10–11, 13353 Berlin, Germany. [3]Institute 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 G12 8QQ, UK. [4]Max Planck Institute for Evolutionary Anthropology (MPI EVAN), Deutscher Platz 6, 04103 Leipzig, Germany. [5]LANADA/LCVB, Bingerville, 206, Côte d'Ivoire. [6]World Health Organization, 1211 Geneva 27, Switzerland. [7]Chimbo Foundation, Amstel 49, 1011 PW Amsterdam, The Netherlands. [8]McGill University, Department of Biology, 855 Sherbrooke Street, West Montreal, Quebec H3A 2T7, Canada. [9]The Ohio State University, Department of Anthropology, 4034 Smith Laboratory, 174 West 18th Avenue, Columbus, Ohio 43210, USA. [10]Lukuru Foundation, 1235 Avenue des Poids Lourds/Quartier de Kingabois, Kinshasa, Democratic Republic of the Congo. [11]Limbe Wildlife Centre, Limbe, Cameroon. [12]Arizona State University, PO Box 872402, Tempe, Arizona 85287-2402, USA. [13]Wild Chimpanzee Foundation (WCF), Deutscher Platz 6, 04103 Leipzig, Germany. [14]German Primate Center, Kellnerweg 4, 37077 Göttingen, Germany. [15]Robert Koch Institute, Seestraße 10–11, 13353 Berlin, Germany.
*These authors contributed equally to this work.

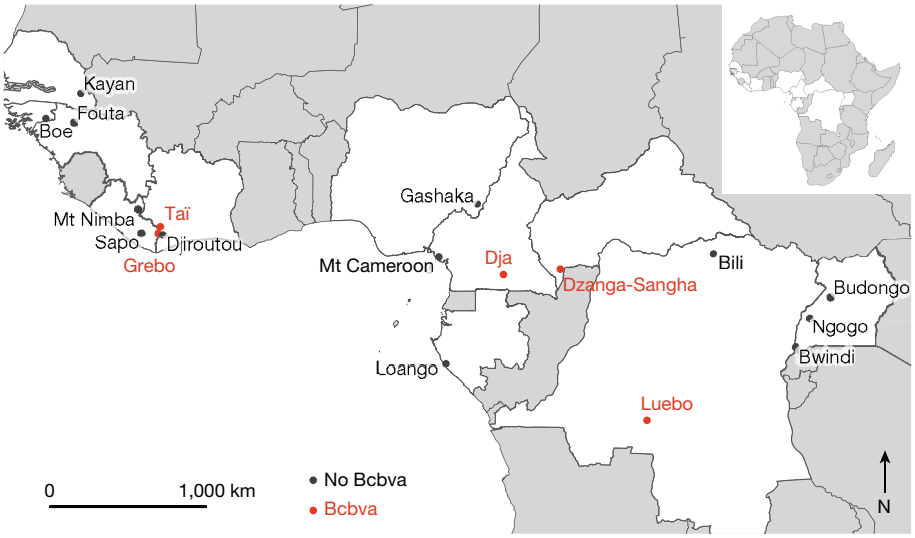


Figure 1 | Bcbva occurrence and study sampling sites in sub-Saharan Africa. Sites with known Bcbva occurrence are indicated in red. Detection of Bcbva in TNP (Taï), Dja Reserve, Dzanga-Sangha Protected Areas and Luebo has been described in previous studies. For all Bcbva sites, except Luebo, samples were available. Within this study we could identify Grebo as a new site of Bcbva occurrence. Bcbva was not detected at the other tested sub-Saharan sites (indicated in black).

carcass incidence, suggesting ongoing anthrax activity in the area (generalized linear mixed model (GLMM), $\chi^2 = 6.3$, degrees of freedom (d.f.) = 10, $P = 0.789$; Supplementary Information). However, Bcbva detection in flies peaked from December to March, coinciding with the only distinct dry period in the park (GLMM, $\chi^2 = 6.9$, d.f. = 2, $P = 0.03$; Extended Data Fig. 4 and Supplementary Information). This suggests climatic conditions may influence Bcbva ecology in TNP, similar to observations from *B. anthracis* in savannahs[1], although seasonal mortality appears less pronounced.

Bcbva differed markedly from *B. anthracis* in terms of host range. Ungulates constitute the vast majority (>99%) of anthrax cases in savannah ecosystems[11,12,20]. By contrast, and in line with the more diverse fauna found in rainforests, we observed Bcbva fatalities in a broader range of species in TNP, including chimpanzees (31 out of 55), six monkey species (21 out of 81), duikers (26 out of 40), mongooses (2 out of 2) and porcupines (1 out of 26 other mammals) (Supplementary Table 2). To further explore the host range of Bcbva, we analysed the gut content of all mammalian DNA- and Bcbva-positive flies ($n = 28$, Supplementary Table 1) using amplicon deep sequencing. We detected sequences from most of the aforementioned species, and from species belonging to 11 further mammalian genera, including carnivores, rodents and bats (Supplementary Table 5 and Supplementary Information). This suggests that Bcbva may affect an even broader range of mammals than inferred from carcass

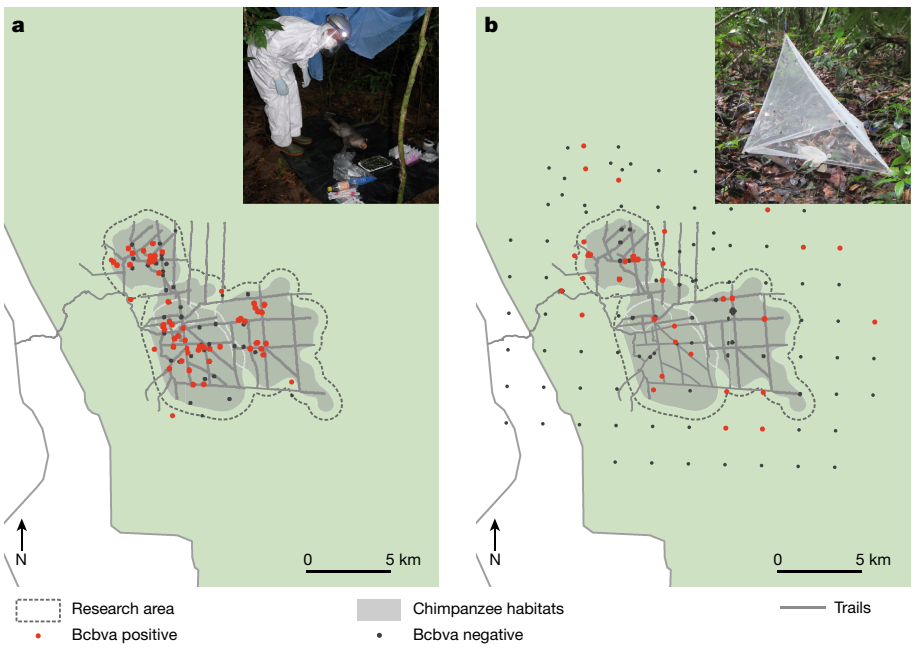


Figure 2 | Bcbva cases in TNP. a, Bcbva-positive and -negative carcasses. In TNP, 38% of the observed wildlife mortality is due to Bcbva. Bcbva-positive carcasses were widely distributed throughout the research area with no obvious pattern identifiable. GPS data was available for 113 out of 204 detected carcasses and was not available for those detected before 2001. **b**, Bcbva-positive and -negative fly traps. Five per cent of all analysed flies contained genetic material of Bcbva. Flies were also caught outside the research area. A systematic snapshot sampling revealed higher prevalence of Bcbva-positive fly traps within the research area.

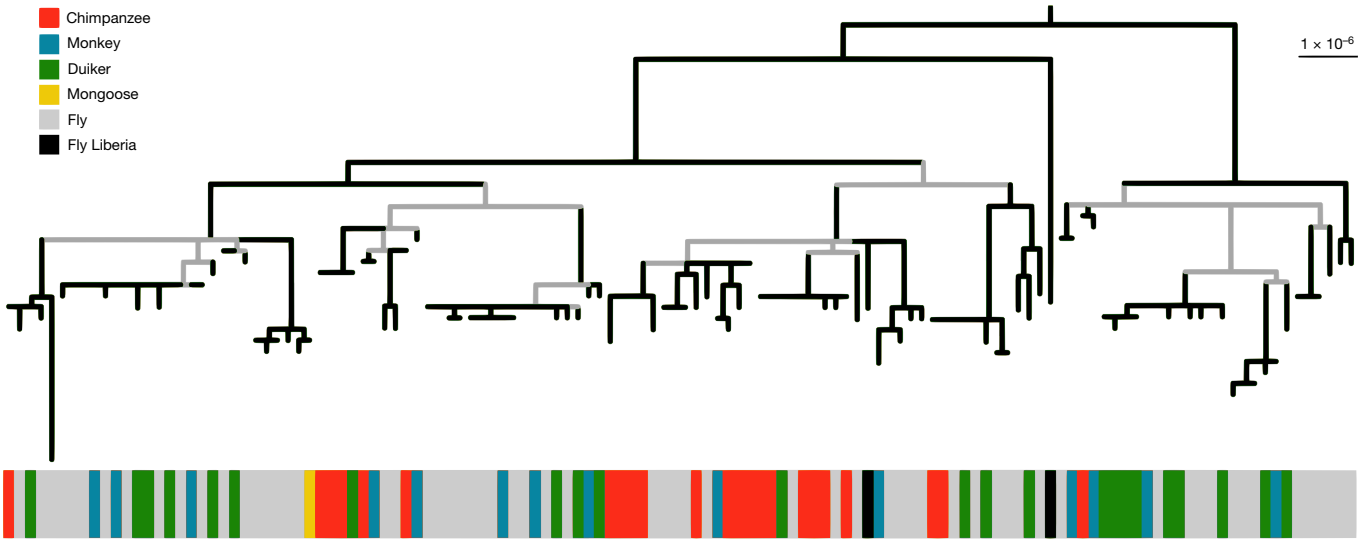


Figure 3 | Phylogenomic tree of Bcbva isolates. Maximum likelihood tree based on chromosomal sequences of Bcbva isolates from TNP (Côte d'Ivoire, $n = 124$) and Grebo (Liberia, $n = 2$). One sequence per host (mammals or flies, two divergent isolates for fly 600) was included and the final alignment of variant sites measured 298 bp. Internal branches with bootstrap values lower than 90 are coloured in grey. The coloured strip represents different host species. The tree was rooted using the heuristic residual mean squared function in TempEst version 1.5. The scale bar indicates substitution per chromosomal site.

monitoring alone. Further, meal compositions of mammalian DNA-positive, Bcbva-positive flies ($n = 28$) and mammalian DNA-positive, Bcbva-negative flies ($n = 29$) did not differ significantly (GLMMs, Supplementary Information), which may support the notion that there is no substantial difference in Bcbva susceptibility among species.

To gain further insight into the ecology of Bcbva, we investigated 178 genomes derived from isolates obtained from necropsy samples, bones and flies, collected between 1996 and 2014 (Supplementary Table 8). Considering 126 chromosomal sequences originating from separate hosts (mammals and flies), we detected 298 single-nucleotide polymorphisms (SNPs). Plasmids contained negligible amounts of variation (Supplementary Information). The maximum distance observed between isolates was 69 SNPs (median: 26 SNPs); the most distant isolates originated in flies caught in two consecutive years only 6 km apart. In comparison, a maximum distance of only 20 SNPs was observed in *B. anthracis* isolates derived from cattle samples collected in the French Alps between 1997 and 2009 (ref. 21). The high genetic diversity observed in TNP is consistent with extensive Bcbva activity in the area and suggests that this pathogen did not emerge recently (Fig. 3 and Extended Data Fig. 5). In addition, considerably more divergence was seen compared to isolates from other countries[17,18], supporting the notion that Bcbva has been circulating in sub-Saharan Africa for an even longer period than what we determined in TNP (Extended Data Fig. 6 and Supplementary Information). To assess within-host diversity, we sequenced the genomes of two to six independent isolates for a subset of carcasses and flies (Supplementary Table 9). Two strains differing by 42 chromosomal SNPs were isolated from a single fly, probably reflecting multiple carcass meals[19], which further highlights the commonness of Bcbva in TNP. Otherwise, the maximum distance observed within one host was two chromosomal SNPs (mean: 0.35 SNPs). Within-host heterogeneity therefore seems negligible compared to the overall diversity observed for Bcbva, suggesting strains differing by more than two SNPs originate from separate carcasses.

Bcbva-positive carcasses were broadly distributed throughout the TNP research area, without the kind of geographic clustering described for anthrax in savannah ecosystems[12,22] (Fig. 2a). We determined Bcbva prevalence within and outside the research area using a subset of 908 flies caught systematically according to a grid system within 19 days (Extended Data Fig. 7). We detected Bcbva-positive flies in 16 out of 83 traps (Supplementary Table 11). Prevalence was higher in the research area (8 out of 21 traps were Bcbva-positive) than in the surrounding forest belt (8 out of 62 traps were Bcbva-positive) (Fisher's exact test, $P = 0.02$). Long-term research activity within the TNP research area has had a protective effect on wildlife and led to an increased density of mammals[23], which might explain higher Bcbva activity. Genome data revealed multiple contemporaneous transmission chains caused by co-circulating strains (2–48 SNPs distance, median: 25 SNPs) in different areas of the park over the short time period of the fly snapshot (Extended Data Fig. 8). For low genomic distances ($\leq$35 SNPs), genomic and geographic distances of all TNP isolates were positively correlated ($R^2 = 0.72$), providing further indication of spatially restricted transmission (Extended Data Fig. 9), which might reflect carcass-mediated spread of Bcbva. Because wildlife cases included exclusively arboreal monkeys (Supplementary Table 2), we explored the vertical distribution of Bcbva by catching flies simultaneously on the ground and up to 30 m in the canopy. We detected Bcbva in 12 out of 103 canopy flies (11.7%) and retrieved isolates from five of these (Supplementary Tables 4, 11). While on the ground carcass-deposition sites are likely to be the source of Bcbva infections, flies may contribute to Bcbva transmission in the upper strata of the rainforest[24].

Fly samples indicated a large proportion of undetected anthrax mortality. During 19 days of focused fly sampling, we retrieved Bcbva isolates from 17 flies, with 13 strains being more than two SNPs different from any other strain. Because two SNPs appear to be the upper level of within-host diversity (Supplementary Table 9), this implies the presence of at least 13 different Bcbva-positive carcasses. Yet, during the same sampling period, only three Bcbva-positive carcasses were discovered and their isolates all corresponded to one of the fly Bcbva lineages ($\leq$2 SNPs difference). This suggests carcass monitoring alone underestimates mortality by at least an order of magnitude.

We investigated the consequences of Bcbva-induced mortality on the species best studied in this ecosystem, chimpanzees. Chimpanzees have a low reproduction rate[25] and are thus particularly sensitive towards external changes to their environment. On the basis of demographic data collected from habituated groups in TNP, we simulated population viability at a 150 years horizon across a broad range of demographic models, which included or excluded anthrax-induced mortality (Supplementary Figs 7, 8). Our simulations showed that, with Bcbva, the TNP chimpanzee population would only have high chances to persist in the case of an overall annual per capita mortality rate owing to

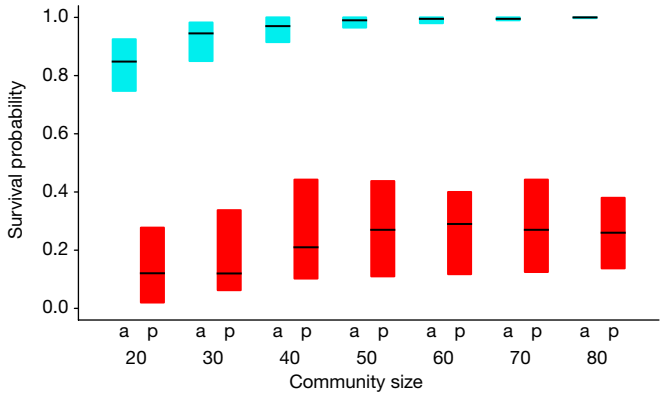


Figure 4 | Proportions of simulated chimpanzee communities surviving 150 years with and without presence of anthrax. Shown are results for different community sizes and anthrax being absent (a, blue boxes) or present (p, red boxes). Bars represent median estimates and boxes represent quartiles across a range of simulation models assuming different inter-birth intervals and maximum ages. All models summarized here assumed an annual per capita survival rate of 0.96.

other causes of 1% (Supplementary Figs 7, 8). Such a low mortality rate is, however, not even observed in captive chimpanzees. In wild chimpanzees the lowest annual per capita mortality rate is 4% (in early adults)[25]. Under such a survival probability (0.96), the simulated presence of anthrax invariably led to a clearly reduced survival probability of communities (Fig. 4). For example, 76 out of 84 models resulted in an extirpation probability that was higher than 50%, whereas the model, which we consider the most realistic (community size 60, maximum age 46 years and inter-birth interval 6 years), resulted in an extirpation probability of 89% (Fig. 4). Our simulations therefore suggest that anthrax-induced mortality will result in deterministic population declines and possible extirpation of TNP chimpanzees over the next 150 years. The risk of extirpation will increase if chimpanzee mortality due to hunting and human-borne infectious diseases continues to rise[23,26,27].

To determine whether similar unrecognized effects on wildlife might be occurring elsewhere, we tested 784 flies collected at eight different sites, as well as 136 bones from twelve sites in five and nine sub-Saharan countries, respectively (Supplementary Fig. 3 and Supplementary Table 1). All sites had chimpanzee populations but none (nor the country) had previously reported Bcbva cases. We only detected Bcbva genetic material in 2 out of 105 flies and 1 out of 8 bones collected in the Grebo National Forest in Liberia, about 40 km from TNP (Supplementary Fig. 1). The genome sequences of isolates from the two fly samples nested within the diversity of Bcbva in TNP, which may indicate an epidemiological link (Fig. 3). We did not detect Bcbva in 305 flies from two sites where Bcbva cases have previously been reported (Dja Reserve, Cameroon, and Dzanga Sangha Protected Areas, Central African Republic, Supplementary Table 4). Although the lack of detection at other sites needs to be interpreted with caution owing to variable fly species composition (Extended Data Fig. 10 and Supplementary Information), these data suggest that Bcbva dynamics may also vary across rainforest ecosystems. It will be important to further uncover the scale and environmental drivers behind Bcbva prevalence. Such knowledge will be critical for mitigating against the detrimental effects of Bcbva on wildlife and for better assessing human infection risk, which for anthrax in rainforest ecosystems has, to date, been considered very low.

Online Content Methods, along with any additional Extended Data display items and Source Data, are available in the online version of the paper; references unique to these sections appear only in the online paper.

Received 25 October 2016; accepted 15 June 2017.

1. Hampson, K. *et al.* Predictability of anthrax infection in the Serengeti, Tanzania. *J. Appl. Ecol.* **48,** 1333–1344 (2011).
2. Hugh-Jones, M. E. & de Vos, V. Anthrax and wildlife. *Rev. Sci. Tech.* **21,** 359–383 (2002).
3. Lindeque, P. M. & Turnbull, P. C. Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. *Onderstepoort J. Vet. Res.* **61,** 71–83 (1994).
4. Beyer, W. & Turnbull, P. C. Anthrax in animals. *Mol. Aspects Med.* **30,** 481–489 (2009).
5. Turnbull, P. C. B. *WHO Guidelines Approved by the Guidelines Review Committee* (World Health Organization, Department of Communicable Diseases Surveillance and Response, 2008).
6. Good, K. M., Houser, A., Arntzen, L. & Turnbull, P. C. Naturally acquired anthrax antibodies in a cheetah (*Acinonyx jubatus*) in Botswana. *J. Wildl. Dis.* **44,** 721–723 (2008).
7. Wafula, M. M., Patrick, A. & Charles, T. Managing the 2004/05 anthrax outbreak in Queen Elizabeth and Lake Mburo National Parks, Uganda. *Afr. J. Ecol.* **46,** 24–31 (2008).
8. Clegg, S. B., Turnbull, P. C., Foggin, C. M. & Lindeque, P. M. Massive outbreak of anthrax in wildlife in the Malilangwe Wildlife Reserve, Zimbabwe. *Vet. Rec.* **160,** 113–118 (2007).
9. Muoria, P. K. *et al.* Anthrax outbreak among Grevy's zebra (*Equus grevyi*) in Samburu, Kenya. *Afr. J. Ecol.* **45,** 483–489 (2007).
10. Turnbull, P. C. *et al.* Anthrax in wildlife in the Luangwa Valley, Zambia. *Vet. Rec.* **128,** 399–403 (1991).
11. de Vos, V. The ecology of anthrax in the Kruger National Park, South Africa. *Salisbury Med. Bull.* **68** (Suppl), 19–23 (1990).
12. Lembo, T. *et al.* Serologic surveillance of anthrax in the Serengeti ecosystem, Tanzania, 1996–2009. *Emerg. Infect. Dis.* **17,** 387–394 (2011).
13. Beyer, W. *et al.* Distribution and molecular evolution of *Bacillus anthracis* genotypes in Namibia. *PLoS Negl. Trop. Dis.* **6,** e1534 (2012).
14. Leendertz, F. H. *et al.* Anthrax kills wild chimpanzees in a tropical rainforest. *Nature* **430,** 451–452 (2004).
15. Klee, S. R. *et al.* The genome of a *Bacillus* isolate causing anthrax in chimpanzees combines chromosomal properties of *B. cereus* with *B. anthracis* virulence plasmids. *PLoS ONE* **5,** e10986 (2010).
16. Brézillon, C. *et al.* Capsules, toxins and AtxA as virulence factors of emerging *Bacillus cereus* biovar *anthracis*. *PLoS Negl. Trop. Dis.* **9,** e0003455 (2015).
17. Antonation, K. S. *et al. Bacillus cereus* biovar *anthracis* causing anthrax in sub-Saharan Africa—chromosomal monophyly and broad geographic distribution. *PLoS Negl. Trop. Dis.* **10,** e0004923 (2016).
18. Leendertz, F. H. *et al.* Anthrax in western and central African great apes. *Am. J. Primatol.* **68,** 928–933 (2006).
19. Calvignac-Spencer, S. *et al.* Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. *Mol. Ecol.* **22,** 915–924 (2013).
20. Berry, H. H. Surveillance and control of anthrax and rabies in wild herbivores and carnivores in Namibia. *Rev. Sci. Tech.* **12,** 137–146 (1993).
21. Vergnaud, G. *et al.* Comparison of French and worldwide *Bacillus anthracis* strains favors a recent, post-Columbian origin of the predominant North-American clade. *PLoS ONE* **11,** e0146216 (2016).
22. Smith, K. L. *et al. Bacillus anthracis* diversity in Kruger National Park. *J. Clin. Microbiol.* **38,** 3780–3784 (2000).
23. Campbell, G., Kuehl, H., Diarrassouba, A., N'Goran, P. K. & Boesch, C. Long-term research sites as refugia for threatened and over-harvested species. *Biol. Lett.* **7,** 723–726 (2011).
24. Blackburn, J. K., Van Ert, M., Mullins, J. C., Hadfield, T. L. & Hugh-Jones, M. E. The necrophagous fly anthrax transmission pathway: empirical and genetic evidence from wildlife epizootics. *Vector Borne Zoonotic Dis.* **14,** 576–583 (2014).
25. Hill, K. *et al.* Mortality rates among wild chimpanzees. *J. Hum. Evol.* **40,** 437–450 (2001).
26. Köndgen, S. *et al.* Pandemic human viruses cause decline of endangered great apes. *Curr. Biol.* **18,** 260–264 (2008).
27. Boesch, C & Boesch-Achermann, H. *The Chimpanzees of the Taï Forest: Behavioural Ecology and Evolution* (Oxford Univ. Press, 2000).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank the authorities in Côte d'Ivoire for long-term support, especially the Ministry of the Environment and Forests, the Ministry of Research, the directorship of the TNP and the CSRS in Abidjan; and national authorities from all other countries for providing permissions for our research (MINFoF, MINRESI, the Service de la Conservation de la Réserve du Dja, Cameroon, in Central African Republic; the Ministère des Eaux et Fôrets, Chasse et Peche and the Ministère de l'Education Nationale, de l'Alphabetisation, de l'Enseignement Superieur, et de la Recherche, the Agence Nationale des Parcs Nationaux, Gabon; Centre National de la Recherche Scientifique et Technologique, Gabon; Direction des Eaux, Forêts et Chasses, Senegal; Forestry Development Authority, Liberia; Institut Congolais pour la Conservation de la Nature, Democratic Republic of the Congo; Ministère de l'Agriculture de l'Elevage et des Eaux et Forêts, Guinea; Instituto da Biodiversidade e das Áreas Protegidas (IBAP), Guinea-Bissau; Ministère de la Recherche Scientifique, Democratic Republic of the Congo; Ministère de le Recherche Scientifique et Technologique, Democratic Republic of the Congo; Nigeria National Park Service, Nigeria, Uganda National Council for Science and Technology, Ugandan

Wildlife Authority, Uganda). We thank the WWF Central African Republic, T. Börding, T. Hicks, Y. Moebius, V. Sommer, K. Zuberbühler and M. Peeters for their logistical support; the field assistants A. Henlin, K. Albrechtova and A. Lang for the collection of samples in TNP; and the field assistants from all other sites for their support; S. Becker, T. Franz, S. Howaldt, A. Lander, P. Lochau, H. Nattermann and A. Schneider for the laboratory work; J. Hinzmann, A. Nitsche and J. Tesch for sequencing; P. Wojciech Dabrowski and T. Semmler from RKI, as well as G. Hamilton at Glasgow Polyomics, for bioinformatic support; and M. Kovacev-Wegener for administrative support. We thank the German Research Council DFG KL 2521/1-1 and the Sonnenfeld-Stiftung for funding; and the Max-Planck-Society and Krekeler Foundation for funding of the Pan African Programme.

Author Contributions C.H., F.Z., A.A., S.A., M.A., G.B., K.C., T.D., P.D., K.D., H.E., P.F., Y.G.Y., A.G., A.-C.G., S.McG., J.H., S.J., J.J., J.K., K.La., J.L., K.Le., F.L., V.L., T.L., S.Ma., A.M., S.Me., M.M., J.v.S., E.T. and D.W. collected flies, bones and associated field data. Necropsies on wildlife that was found dead were performed by F.Z., K.N., A.B., E.C.-H., A.D., P.F., S.A.L., T.L., S.Me., S.N., H.D.N. and F.H.L. and laboratory analyses were performed by C.H., F.Z., K.N., S.D., R.G., K.M.-R., K.M., S.Me., H.D.N., A.S., U.T., S.R.K., L.H.W., S.C.-S. and F.H.L. The data were analysed by C.H., F.Z., R.B., H.K., R.M. and S.C.-S. and the manuscript was prepared by C.H., F.Z., R.B., H.K., R.M., J.F.G., S.C.-S. and F.H.L. The manuscript was revised and approved by all authors. The study was supervised by C.B., R.M.W., S.C.-S. and F.H.L.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations. Correspondence and requests for materials should be addressed to F.H.L. (leendertzf@rki.de).

METHODS

Study sites. TNP covers an area of 3,300 km^2 and an additional 200 km^2 buffer zone. Since 2001 a veterinary program conducts outbreak investigations in wildlife. We defined the research area as the habitat ranges of the three habituated chimpanzee groups plus a 500 m buffer zone (103 km^2; Supplementary Fig. 2).

Samples belonging to the large-scale dataset were collected at 16 sites in 11 sub-Saharan countries stretching from Senegal to Uganda (Supplementary Fig. 3 and Supplementary Table 1). Most sites (14 out of 16) were temporary research sites of the Pan African Programme (http://panafrican.eva.mpg.de/) where Bcbva has not been described. Additional samples were obtained from the Dja Faunal Reserve, Cameroon[18] and Dzanga-Sangha Protected Areas, Central African Republic[17], where Bcbva cases have been previously described. Study sites are described in detail in the Supplementary Information.

Necropsies. Carcass monitoring was performed in TNP by a veterinarian, performing necropsies on every carcass reported by researchers working in the forest ($n = 173$). Samples of all inner organs were collected, as far as carcass decomposition allowed. Necropsies followed a standardized protocol, including use of full personal protective equipment. Carcass sites were decontaminated according to World Health Organization (WHO) guidelines[5,28]. For each sample, aliquots were stored in liquid nitrogen and formalin in the field. Frozen samples were transported on dry ice and subsequently stored at −80 °C. We received additional tissue samples from carcasses sampled by the WHO in TNP between 1996 and 2000 ($n = 31$) (Supplementary Table 2).

Rather than using serology, which would also detect animals that survived non-lethal infections, we used PCR to detect the presence of anthrax in internal organs to confirm that anthrax was the probable cause of death. DNA was extracted from various tissues of each animal (liver, spleen and lung when available) using the DNeasy Blood and Tissue Kit (Qiagen); extracts were quantified using a Nanodrop (Thermo Fisher Scientific) and stored at −20 °C. Subsequently, 200 ng DNA or 5 µl of DNA extract (if the DNA concentration was below 40 ng µl^{-1}) was tested for anthrax in duplicate real-time PCR reactions (see Supplementary Methods). The full anthrax assay used includes three real-time PCR assays, each targeting one of the following gene markers: *pag* (gene for protective antigen) located on the pXO1 plasmid[29], *capB* (gene for capsule synthesis) located on pXO2 and *Island IV*, a chromosomal marker specific for Bcbva[15,17] (Supplementary Table 3). Samples were first tested for *pag* and samples positive in duplicate for *pag* were tested for *capB* and *Island IV* (Extended Data Figs 1, 2).

Culture under BSL3 conditions was attempted for all PCR-positive necropsy samples collected until the end of 2013 (June 2014 for duikers) (Supplementary Table 2). A native and heat-treated (65 °C for 30 min, to assess presence of spores) aliquot were plated onto the following agar plates: Columbia blood agar (Oxoid), blood-trimethoprim agar (1.6 mg trimethoprim, 6.4 mg sulfamethoxazole, 20 mg polymyxin B per litre agar medium) and Cereus Ident agar (Heipha Diagnostica) with the chromogenic substrate 5-bromo-4-chloro-3-indoxyl-myo-inositol-1-phosphate[30]. Cultures were incubated at 37 °C and monitored daily. Morphologically suspicious colonies were sub-cultured and tested by real-time PCR. Bcbva was cultured from native and heat-treated samples indicating the presence of heat-resistant spores. Isolates were frozen in Microbank tubes (Mast Diagnostica) at −80 °C.

Histopathology was performed on a subset of necropsy samples, including 15 Bcbva PCR-positive necropsy samples (Supplementary Table 2). No signs of anthrax infection were detected in carcasses that were PCR-negative for anthrax, while for PCR-positive carcasses the most consistent histopathologic finding was per-acute to acute anthrax-related pneumonia characterized by mild lymphohistiocytic infiltrates and intra-alveolar eosinophilic and proteinaceous or fibrinous material. Numerous intravascular and intra-alveolar bacilli were found. Multifocal alveolar and peribronchiolar haemorrhages were present in all animals. Lymph node changes consisted of sinus histiocytosis, cortical haemorrhages and oedema especially in the mediastinal, tracheobronchiolar and mesenteric lymph nodes. Large amounts of bacilli were found in the sinusoids. Within the abdominal cavity the spleen was the most affected organ, with myriads of bacilli visible in the splenic sinusoids, partly embedded in fibrin deposition. There was moderate lymphoid depletion, lymphocytolysis and histiocytosis. The liver parenchyma was severely congested with masses of bacilli within the hepatic sinusoids. All anthrax PCR-positive carcasses were also tested for filoviruses[31] and respiratory diseases[26] to rule out co-infection with other common causes of death in this ecosystem.

Blow flies. Flies were caught on the ground and in the canopy using custom-made traps (Supplementary Figs 4, 5 and Supplementary Information). Trapping was done for 60 min or until 20 flies were collected. Flies were euthanized with ether and stored at −20 °C in 2-ml Cryotubes (Carl Roth) containing up to 10 flies or at ambient temperature on silica in 50-ml Falcon tubes (Thermo Fisher Scientific) containing up to 20 flies. In TNP, 726 flies were randomly collected within the research area in 2008, 2009, 2012 and 2013 (Supplementary Table 4). Another 908 flies were collected over 19 days in May and June 2014 according to a 2 × 2-km grid system covering the research area and 225 km^2 surrounding the research area (referred to as 'snapshot flies'; Extended Data Fig. 7 and Supplementary Table 4). At a larger scale, 784 flies were collected at eight sites within five sub-Saharan countries (Pan African Programme) from 2012 to 2014 (Supplementary Table 4) and 305 flies were analysed from two sub-Saharan sites, Dja Faunal Reserve, Cameroon[18] ($n = 105$) and Dzanga-Sangha Protected Areas, Central African Republic[17] ($n = 200$) (Supplementary Table 4). In total, 2,723 flies were analysed (Supplementary Table 4).

DNA extraction of individual flies was performed using the GeneMATRIX Stool DNA Purification kit (Roboklon). We followed manufacturer's instructions except that each fly was first cut into small pieces using sterilized scissors before being homogenized using a Fast Prep (MP Biomedicals). DNA concentration measurements and anthrax testing by real-time PCR were performed as described for necropsy samples (Supplementary Table 4).

A subset of 50 flies containing high *pag* copy numbers underwent bacterial culture (Supplementary Tables 4, 11). Half of the fly mush remaining after DNA extraction was plated directly onto the same culture medium as described for the necropsy samples. Additionally, a 10-µl aliquot of the mush was diluted 1:10 in sterile NaCl, heat-treated for 30 min at 65 °C and plated. Bcbva was retrieved from native and heat-treated samples, indicating the presence of heat-resistant spores in the flies. An on-site study in TNP also used direct culture of 204 flies without preceding PCR testing. Flies were homogenized and plated directly onto Cereus Ident agar. Suspicious colonies were sub-cultured on blood-trimethoprim agar and tested by real-time PCR. This approach yielded another 21 Bcbva isolates.

To examine whether certain mammals were preferentially affected by Bcbva, we tested for differences in fly meal composition of anthrax-positive and -negative flies. We screened a subset of 750 TNP flies for mammalian DNA using a real-time PCR targeting a 130-bp fragment of the mammalian 16S mitochondrial DNA (described in ref. 19). We chose a subset of mammal and anthrax-positive flies ($n = 28$) and a comparable number of mammal-positive but anthrax-negative flies ($n = 29$) from the same traps (Supplementary Table 11). To dissect fly meal composition, we used a metabarcoding approach, for which 16S amplicons were deep-sequenced, adapting the amplicon preparation protocol provided by Illumina (Supplementary Information). We used a custom pipeline to determine taxonomic assessment of each read to the genus and order level as described in the Supplementary Information (Supplementary Tables 5, 11, 12). Sequences assigned to domestic animals were regarded as contamination, as it was shown that even stringent anti-contamination procedures do not prevent the amplification of human and domestic animal sequences present in the environment and reagents[32].

Details on blow fly analyses and results are in the Supplementary Information.

Bones. Bones were collected in TNP and 12 Pan African Programme sites in nine countries (Supplementary Fig. 3 and Supplementary Tables 1, 7). Bones were transported and stored at ambient temperature. DNA was extracted using a silica-based method[33,34] (Suppementary Information). Bone extracts were tested by real-time PCR as described for necropsy samples (Supplementary Tables 1, 7). Powder from PCR-positive bones was also used for bacterial culture attempts after homogenization in sterile NaCl (Supplementary Table 7). We processed the homogenates as described above for necropsy samples with one native aliquot and one heat-treated aliquot. Details on bone analyses and additional results are in the Supplementary Information.

Whole-genome sequencing of Bcbva isolates and SNP calling. Supplementary Table 8 contains a complete list of all Bcbva isolates sequenced (Supplementary Fig. 6). Isolate preparation and extraction is described in the Supplementary Methods. Libraries for whole-genome sequencing were prepared with the Nextera XT DNA Library Preparation Kit (Illumina). Libraries were pooled and sequenced on the HiSeq 1500 platform (Illumina) in rapid run mode using either v1 (2 × 150 bp) or v2 (2 × 250 bp) chemistry.

Illumina adapters were removed using scythe version 0.993 (ref. 35) and trimmed with sickle version 1.33 (ref. 36) applying a quality threshold of 25. Quality trimmed reads were aligned to the reference genome (Bcbva strain CI, accession numbers CP001746–1749) with the BWA-MEM algorithm implemented in bwa version 0.7.12-r1039 (ref. 37). For conversion to .bam format, sorting, deduplication and indexing of aligned reads, we used the picard tools version 1.136 (ref. 31) software package applying the commands SortSam, MarkDuplicates and BuildBamIndex. Subsequent variant calling was performed using the Genome Analysis Toolkit (GATK) version 3.4 (refs 38–40). We realigned bam files with the tools RealignerTargetCreator and IndelRealigner. Variants were called with UnifiedGenotyper with a minimum phred-scaled confidence threshold of 30 for SNPs to be called. Hard filtering of SNP sites was done with the VariantFiltration

command using recommended filter settings. With the SelectVariants command, only SNP sites that passed the filter were selected for further processing. SelectVariants was also used to exclude all SNPs with a coverage of $<5\times$, a minor allele frequency of >0.1 and a GATK Genotype Quality value of <99. Final consensus sequences were composed with FastaAlternateReferenceMaker. We assessed the coverage of all samples with the GATK tools DepthOfCoverage and CoveredByNSamplesSites. Details and further analysis of whole-genome sequencing of Bcbva isolates and SNP calling can be found in the Supplementary Information.

Phylogenetic analyses. 126 genome sequences (one isolate per mammal or fly) from TNP and Grebo National Forest (Supplementary Table 8) were aligned and stripped of non-variant sites with Geneious Pro version 8.1.3 (Biomatters ltd.)[41]. The resulting alignments of variable sites were 298, 18 and 11 bp long for the chromosome, pXO1 and pXO2, respectively. Given the low number of variable sites in pXO1 and pXO2, we only performed phylogenetic analyses on the chromosome alignment. jModelTest version 2.1.4 (ref. 42) was used for determination of the best nucleotide substitution model in a maximum likelihood framework, resulting in the choice of TVMef[43].

Maximum likelihood analysis was performed with PhyML version 20131022 (ref/ 44) using a combination of subtree-pruning–regrafting and nearest-neighbour-interchange tree search algorithms. Branch support was estimated using non-parametric bootstrapping with 100 pseudo-replicates. The tree was rooted using the heuristic residual mean-squared function in TempEst version 1.5 (ref. 45), placing the root at the position resulting in the most clock-like structure of the data (Fig. 3).

We also performed phylogenetic analyses using the Bayesian Markov chain Monte Carlo sampling approach implemented in BEAST version 1.8.2 (ref. 46), specifying a constant population coalescent tree prior and assuming an uncorrelated log-normal relaxed molecular clock[47] (Supplementary Information). The maximum clade credibility tree derived from this analysis was very similar to the maximum likelihood tree (Fig. 3).

Another dataset was assembled to compare Bcbva from TNP to other strains from sub-Saharan Africa. It included the chromosomal sequences from a representative TNP genome, the sequences from Grebo National Forest, as well as previously published genomes determined from isolates derived from Bcbva cases in the Central African Republic and Cameroon[15,17] (Extended Data Fig. 6). The alignment was compiled as described above and contained 1,016 variable positions. Model selection with jModelTest version 2.1.4 (ref. 42) selected a TPM1 (ref. 48) nucleotide substitution model. We performed maximum likelihood analyses as described above.

Statistical analyses. No statistical methods were used to predetermine sample size. To test the effect of season on the probability of a carcass or fly, respectively, being anthrax-positive, we used a generalized linear mixed model (GLMM)[49] with binomial error structure and logit link function[50]. As predictors we included the species (monkeys, chimpanzees, duikers, others, blow flies), season and their interaction. 'Season' was modelled by first turning the sampling date into a circular variable and including its sine and cosine into the model. As random intercept effects we included trap-ID (that is, GPS location) and the combination of sampling date and GPS location, the latter accounting for potential non-independence of flies sampled on the same day from the same trap. We further included random slopes of season within the trap-ID[51,52]. To test the effect of season we compared the full model with a null model lacking the fixed effects of season and its interaction with species[53], using a likelihood ratio test[54]. Sample size for this model was 1,803 samples (carcasses and flies), collected at 352 locations and 328 combinations of sampling date and location including necropsy samples and flies.

In a second model we tested whether the probability of a fly to be tested anthrax-positive was influenced by season and the amount of mammalian DNA within the fly. We used a GLMM[49] with binomial error structure and logit link function[50]. Into this we included the amount of mammalian DNA found within the fly (determined by real-time PCR) and season as fixed effects. We modelled the season by first turning the sampling date into a circular variable and then including it sine and cosine into the model. Because the amount of mammalian DNA within the fly was highly skewed, we log-transformed the data before fitting the model. As random effects (random intercepts), we included the ID of the trap and the date of sampling. To avoid overconfident estimates we included random slopes of the amount of mammal DNA within trap-ID and trapping date[51,52]. As an overall test of the effects of the amount of mammal DNA and season, we compared the full model with a null model lacking these effects[53] using a likelihood ratio test[54]. We also used likelihood ratio tests to test for the individual predictors (comparing the full model with a respective reduced model lacking the predictor to be tested[51]). We fitted the model in R[55] using the function glmer of the R package lme4 (version 1.1–10)[56]. To estimate model stability we excluded levels of the random effects one at a time, which did not indicate influential levels that existed. The sample size for this model was a total of 474 flies caught on 43 days in 33 traps. (Extended Data Fig. 4, Supplementary Table 10 and Supplementary Information).

To evaluate the reproducibility of fly meal identification for each fly, we correlated the proportion of sequence counts per amplicon (two amplicons per fly) that was assigned to different mammalian genera using a Spearman correlation. To test whether there were differences in fly meal composition of anthrax-positive and anthrax-negative flies, we tested whether detection of a given mammal taxon in a fly sample was associated with anthrax positivity. We used GLMMs[49] applied separately for each mammal genus identified in the flies. The response was whether the fly was anthrax-positive and the key predictor with fixed effect was presence of mammalian DNA. We considered mammalian DNA to be present, when DNA was detected in at least one of the two amplicons per fly. We included only those mammal genera in the model that were detected in at least five of all generated amplicons (two per fly). In addition to mammal presence, we included trap-ID and the factor sampling date as random effects (random intercepts)[51,52]. Models were fitted with binomial error structure and logit link function[50]. The sample size for all models was 57 flies, caught in 22 different traps on 13 days. To test whether mammal presence had an impact on anthrax positivity, we dropped mammal presence from the model[53] and compared the models using a likelihood ratio test[54]. Model stability was assessed as above. We fitted models at two different taxonomic resolutions: one with taxonomic assignment at genus level and the other at order level. GLMMs were fit in R[55] using the function glmer of the R package lme4 version 1.1-10 (ref. 56).

To evaluate geographic distribution of Bcbva in TNP, we checked whether, owing to higher mammal density[23], Bcbva was more likely to occur inside the research area. To test this hypothesis we analysed 908 flies from 83 different traps (Extended Data Fig. 7 and Supplementary Tables 4, 11). Of these, 21 traps were located within the research area and 62 traps in the adjoining forest belt. In total, 8 out of 21 traps within the research area and 8 out of 62 outside the research area were anthrax positive. We compared the two groups using Fisher's exact test (Supplementary Information).

To learn more about the spatial dynamics of Bcbva in TNP, we investigated the correlation between genetic and geographic distances. To correct for genetic and spatial autocorrelation, we excluded strains from the dataset that originated from the same fly catching point (in a 1-km^2 radius) on the same day or from the same followed-up outbreak in mammals. Only one strain was kept per outbreak or fly-catching point, the selection criterion being high average coverage of the genome (Supplementary Table 8 and Supplementary Information). Geographic distances (in km) were derived from GPS data using GeographicDistanceMatrixGenerator version 1.2.3 (ref. 57). Genetic distances were approximated using the relative distances drawn from a maximum likelihood tree built in PhyML version 20131022 (ref. 44) with the R package ape[58] using the cophenetic function. Multiple regression on distance matrices (MRM) as implemented in the R ecodist package[59] using 1,000 permutations and Spearman's correlations was performed on genetic and geographic distance matrices. To examine variation within genetic lineages, we binned our data by genetic distance (bin size = relative genetic distance of 0.03, approximately 2.5 SNPs) and focused on groups with low genetic distance (maximum relative genetic distance <0.5) and their mean geographic distance (Extended Data Fig. 9). Homogeneity of variance between groups was assured with the Fligner–Killeen test ($P=0.07$; $P>0.05$, the pre-specified significance level).

To evaluate the impact Bcbva could have on the TNP chimpanzee population, we conducted a simulation (Supplementary Information). We first defined a series of population parameters for the simulation[27]. We simulated the survival prospects of chimpanzee communities of a given size, with individuals reproducing at certain regular intervals after maturation, having a maximum age and an annual survival probability. Because most of these parameters are associated with considerable uncertainty and because we wanted to assess to what extent the simulation results depend on the particular parameters chosen, we parameterized the simulations as follows: initial community size: 20–80 individuals (increment: 10); inter-birth interval: 4–7 years (increment: 1); interval after death of infant: 1 year; maximum age: 40–50 years (increment: 2); age of first reproduction of males and females: 10 and 14 years, respectively. Because per capita annual survival probability without the influence of anthrax is unknown (mortality cases due to anthrax may not be detected in all cases, in particular before necropsies were made systematically), we simulated per capita annual survival probabilities from 0.93 to 0.99 (increment: 0.03). In addition, we made survival probability density dependent, as this is a common characteristic observed in many species including chimpanzees[60]. For this we introduced a logistic function $(1/(1+\exp(-(20-0.08\times \text{community size}))))$ that increased or reduced mortality rate as a function of chimpanzee community size. At the beginning of each simulation run we generated a community of the

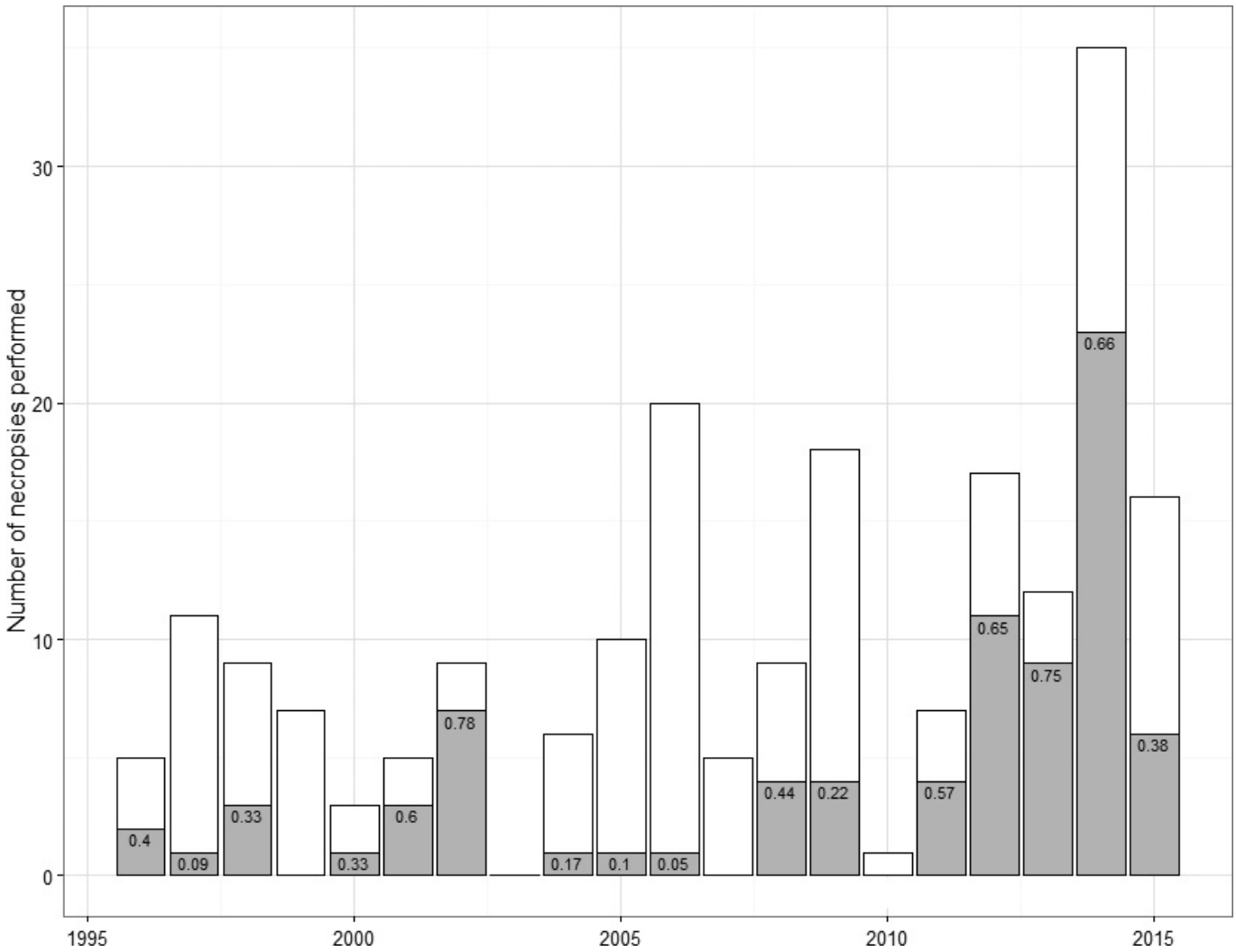
simulated size by randomly allocating sexes (proportion of females: 0.7) and ages (uniformly distributed between 10 and the simulated maximum age) to individuals. To avoid stochastic effects of the initially generated community, we let the simulation run for 50 time steps (that is, years) without anthrax presence before the evaluated time period began.

We estimated the risk of annual anthrax outbreak probability, dependence on community size and the number of individuals affected as $\exp(-1.83+0.039\times\text{community size})$ from a Poisson regression (null, full model comparison, $\chi^2=7.89$, d.f. $=1$, $P<0.01$). We simulated both an anthrax and a non-anthrax scenario for 150 time steps (that is, years) with 100 replications each and for each possible combination of the simulated parameters. Communities were considered to be extinct, when no reproducing females were present.

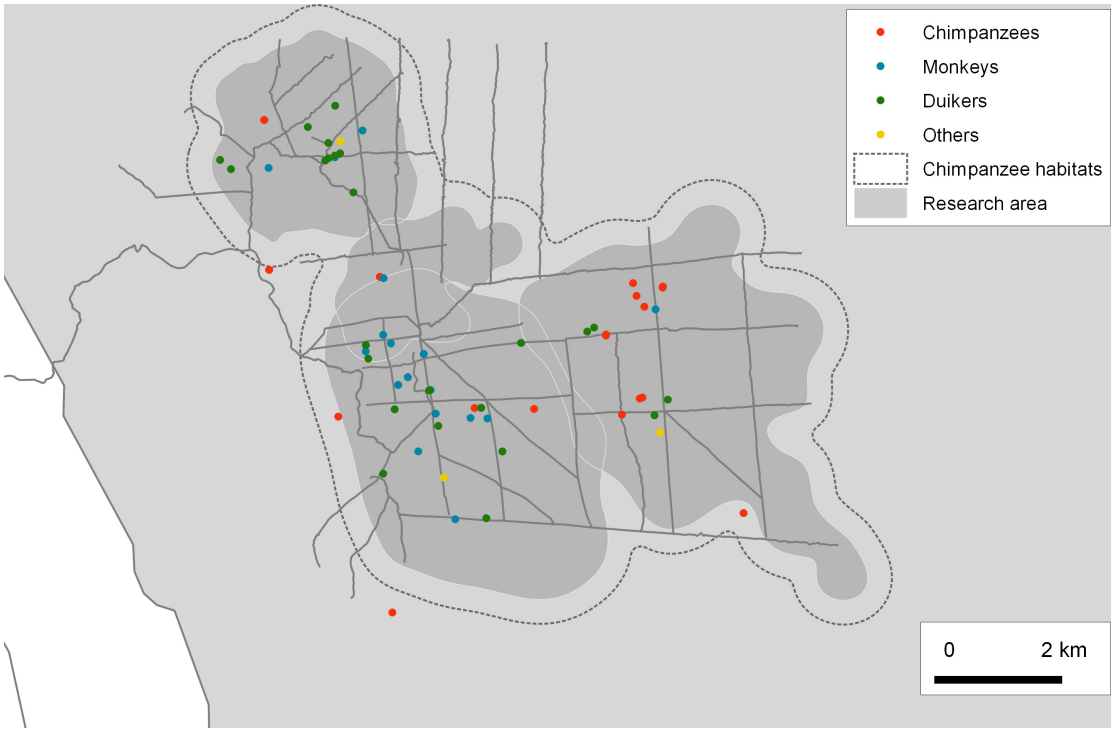
All R scripts are available upon request. Details on statistical analyses and additional results are in Supplementary Information.

Data availability. Raw reads of 16S amplicons are available in the European Nucleotide Archive (ENA) under project accession number PRJEB14554, sample accession numbers ERS1217219–ERS1217336. Raw reads for all 178 Bcbva isolates from TNP and Grebo National Forest are available in the ENA under project accession number PRJEB14616, sample accession numbers ERS1222903–ERS1223080. Variable position alignments are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.v8bn7.

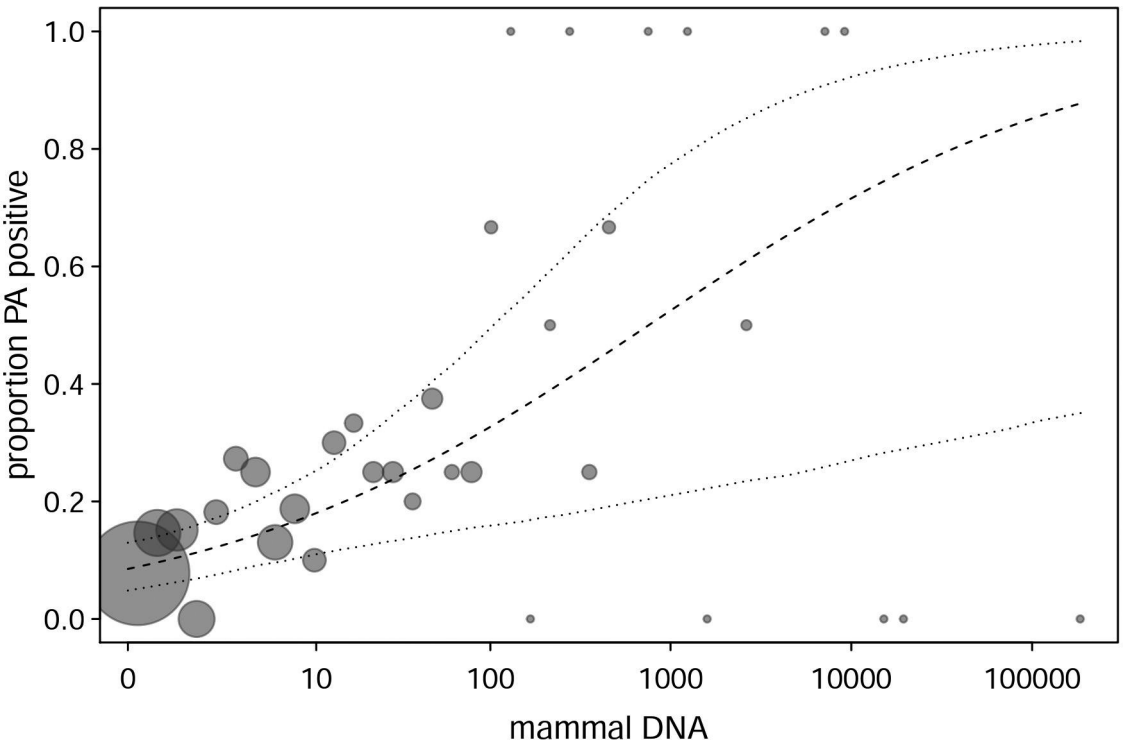
28. Turnbull, P. C. B. *Guidelines for the Surveillance and Control of Anthrax in Humans and Animals* 3rd edn (World Health Organization, Department of Communicable Diseases Surveillance and Response, 1998).
29. Ellerbrok, H. *et al.* Rapid and sensitive identification of pathogenic and apathogenic *Bacillus anthracis* by real-time PCR. *FEMS Microbiol. Lett.* **214,** 51–59 (2002).
30. Klee, S. R. *et al.* Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Cote d'Ivoire and Cameroon. *J. Bacteriol.* **188,** 5333–5344 (2006).
31. Panning, M. *et al.* Diagnostic reverse-transcription polymerase chain reaction kit for filoviruses based on the strain collections of all European biosafety level 4 laboratories. *J. Infect. Dis.* **196** (Suppl 2), S199–S204 (2007).
32. Leonard, J. A. *et al.* Animal DNA in PCR reagents plagues ancient DNA research. *J. Archaeol. Sci.* **34,** 1361–1366 (2007).
33. Gamba, C. *et al.* Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. *Mol. Ecol. Resour.* **16,** 459–469 (2016).
34. Rohland, N. & Hofreiter, M. Ancient DNA extraction from bones and teeth. *Nat. Protoc.* **2,** 1756–1762 (2007).
35. Buffalo, V. Scythe: a 3′-end adapter contaminant trimmer. https://github.com/vsbuffalo/scythe (2014).
36. Joshi, N. A. & Fass, J. N. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files. https://github.com/najoshi/sickle (2011).
37. Li, H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. Preprint at https://arxiv.org/abs/1303.3997 (2013).
38. Auwera, G. A. *et al.* From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* **43,** 11.10.1–11.10. 33 (2013).
39. DePristo, M. A. *et al.* A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat. Genet.* **43,** 491–498 (2011).
40. McKenna, A. *et al.* The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20,** 1297–1303 (2010).
41. Kearse, M. *et al.* Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28,** 1647–1649 (2012).
42. Darriba, D., Taboada, G. L., Doallo, R. & Posada, D. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* **9,** 772 (2012).
43. Posada, D. Using MODELTEST and PAUP* to select a model of nucleotide substitution. *Curr. Protoc. Bioinform.* **00,** 6.5.1–6.5.14 (2003).
44. Guindon, S. *et al.* New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59,** 307–321 (2010).
45. Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2,** vew007 (2016).
46. Drummond, A. J., Suchard, M. A., Xie, D. & Rambaut, A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29,** 1969–1973 (2012).
47. Drummond, A. J., Ho, S. Y., Phillips, M. J. & Rambaut, A. Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4,** e88 (2006).
48. Kimura, M. Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl Acad. Sci. USA* **78,** 454–458 (1981).
49. Baayen, R. H. *Analyzing Linguistic Data: A Practical Introduction to Statistics using R* (Cambridge Univ. Press, 2008).
50. McCullagh, P. & Nelder, J. A. *Generalized Linear Models* Vol. 37 (CRC, 1989).
51. Barr, D. J., Levy, R., Scheepers, C. & Tily, H. J. Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang.* **68,** 225–278 (2013).
52. Schielzeth, H. & Forstmeier, W. Conclusions beyond support: overconfident estimates in mixed models. *Behav. Ecol.* **20,** 416–420 (2009).
53. Forstmeier, W. & Schielzeth, H. Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol.* **65,** 47–55 (2011).
54. Dobson, A. J & Barnett, A. *An Introduction to Generalized Linear Models* (CRC, 2008).
55. R Core Team. *R: A Language and Environment for Statistical Computing* http://www.R-project.org/ (R Foundation for Statistical Computing, Vienna, Austria, 2013).
56. Bates, D., Mächler, M., Bolker, B. & Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67,** 1–48 (2015).
57. Ersts, P. *Geographic Distance Matrix Generator (version 1.2.3)*. http://biodiversityinformatics.amnh.org/open_source/gdmg (American Museum of Natural History, 2011).
58. Paradis, E., Claude, J. & Strimmer, K. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20,** 289–290 (2004).
59. Goslee, S. C. & Urban, D. L. The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22,** 1–19 (2007).
60. Kuehl, H. S., Elzner, C., Moebius, Y., Boesch, C. & Walsh, P. D. The price of play: self-organized infant mortality cycles in chimpanzees. *PLoS ONE* **3,** e2440 (2008).



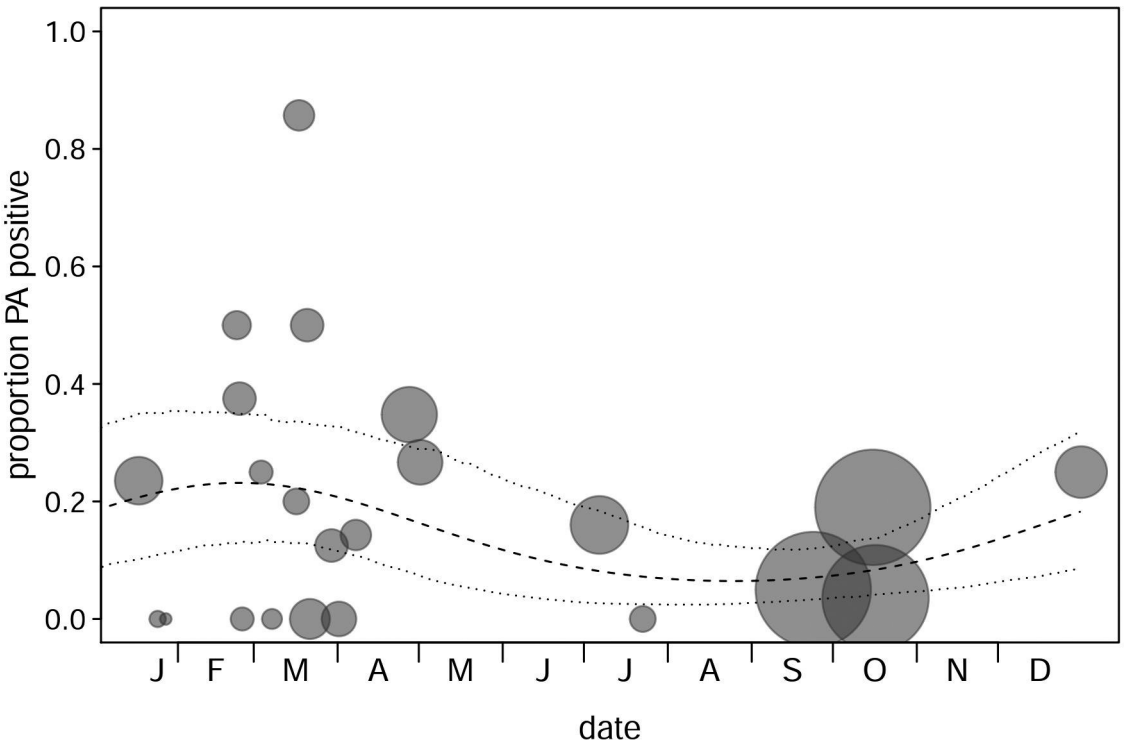
Extended Data Figure 1 | Necropsies performed since 1996. The total number of necropsies performed per year in TNP from 1996 to 2015. Grey bars indicate the number of Bcbva-positive necropsies and are annotated with the associated proportion. In the years 2003 and 2010 only limited veterinary service was available at TNP owing to political insecurity in the region.



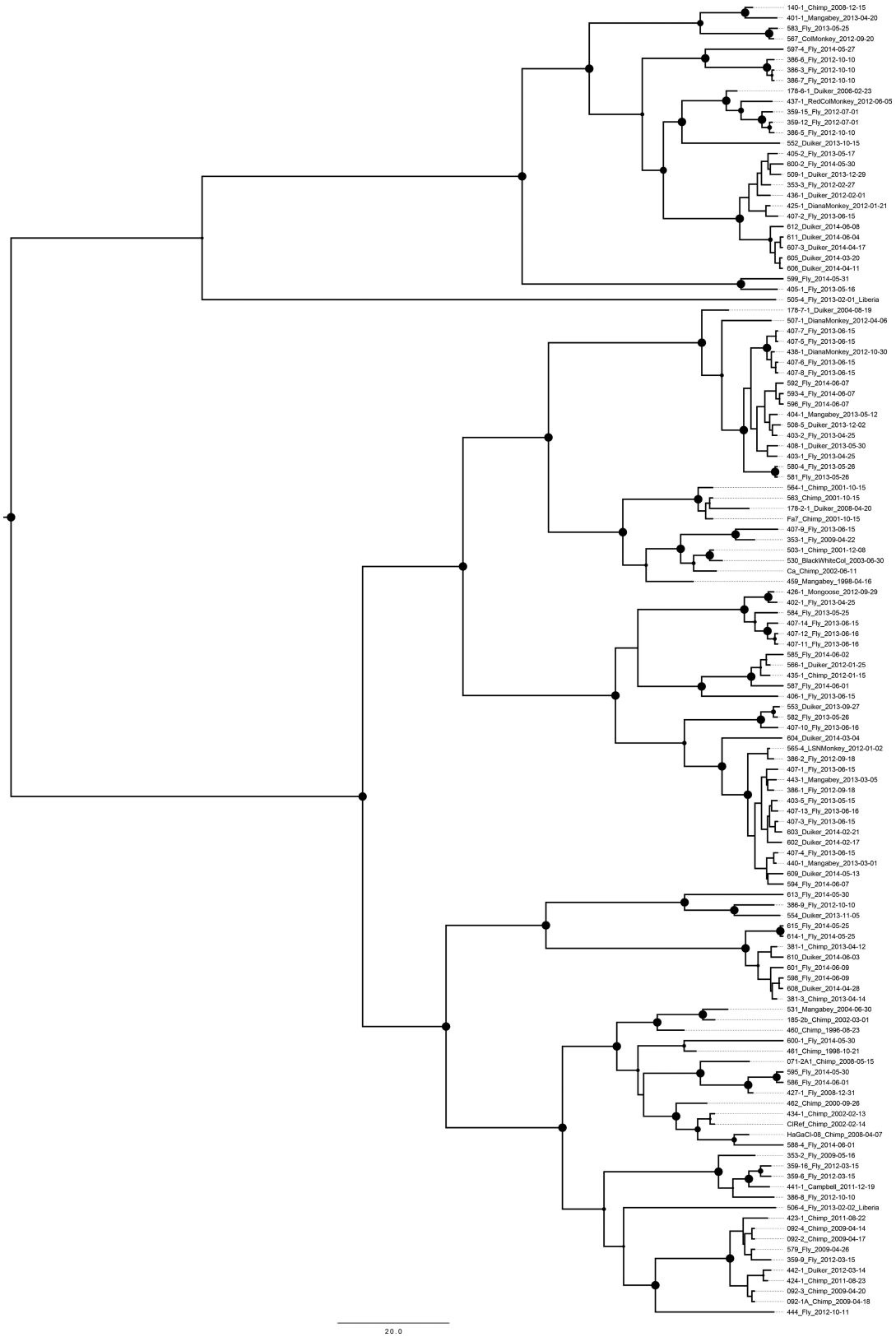
Extended Data Figure 2 | Geographic location of Bcbva-positive carcasses in TNP. Necropsies that tested Bcbva-positive in TNP since 2001. GPS data was available for 70 of all necropsies that tested positive ($n = 81$).



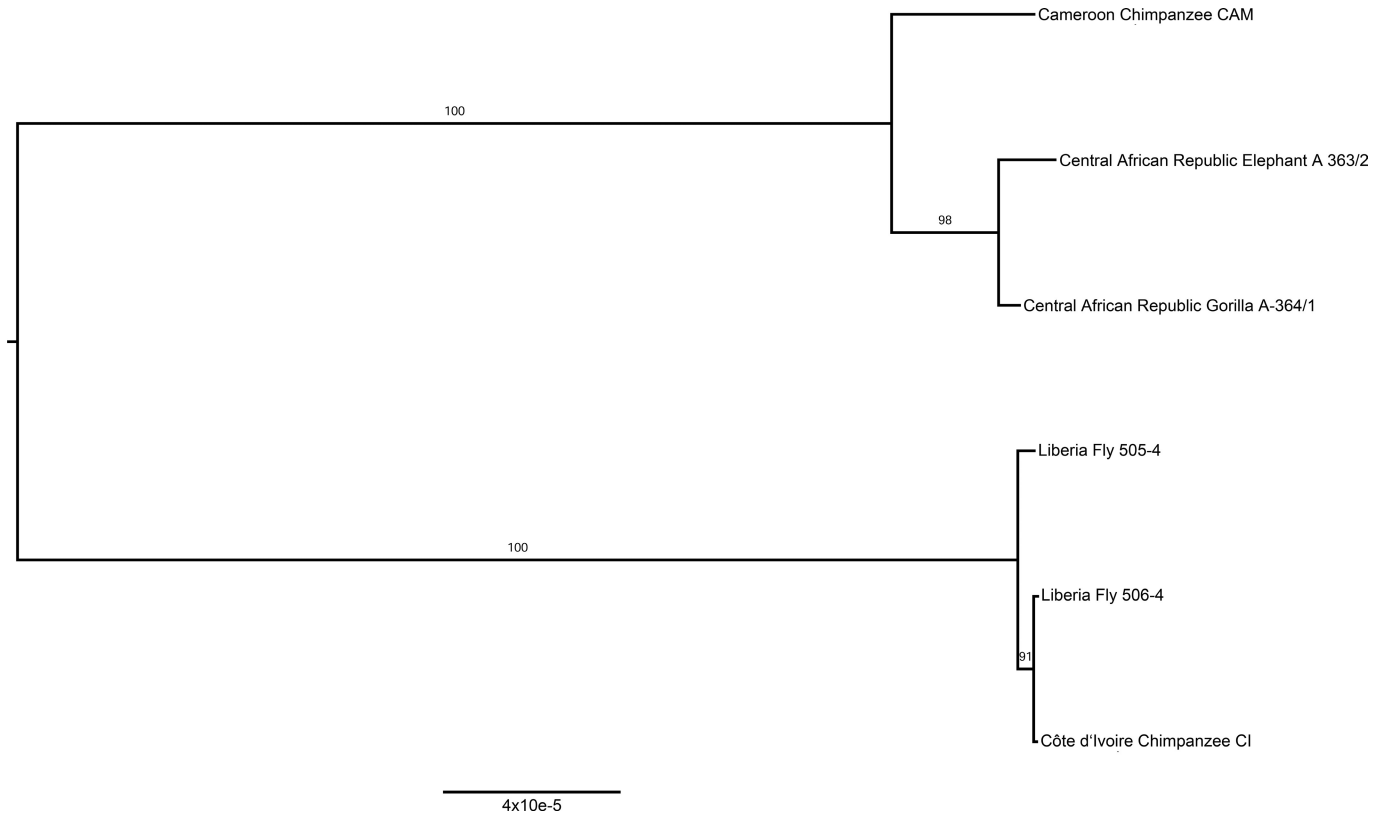
Extended Data Figure 3 | Effect of mammalian DNA content on anthrax positivity in flies. Shown is the probability of Bcbva positivity (PA) as a function of the amount of mammalian DNA (copies) found in a fly. The amount of mammal DNA was binned (bin width of 0.25) and the area of the points depicts the number of flies (range, 1–206) in the respective bins. The dashed line indicates the fitted model and the dotted lines the 95% confidence interval.



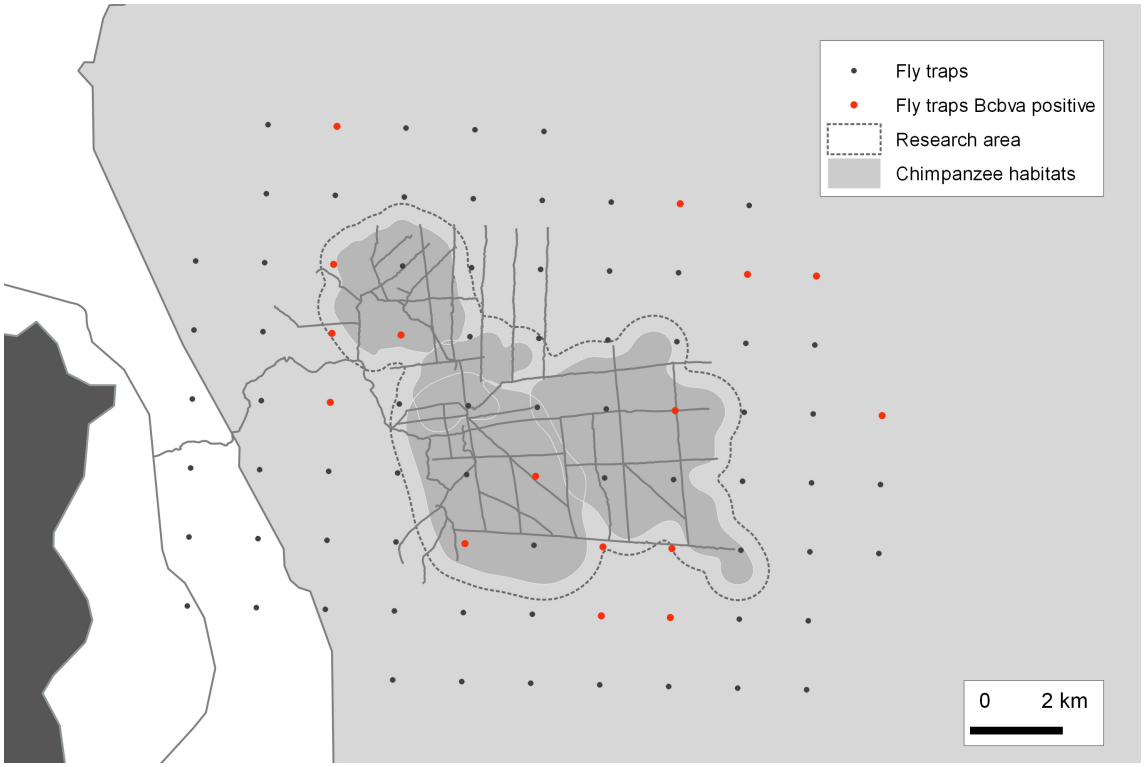
Extended Data Figure 4 | Effect of season on anthrax positivity in flies. The probability of Bcbva positivity (PA) over the course of a year (binned in 10-day periods) is shown. The area of the points depicts the number of flies in the respective 10-day period. The dashed line indicates the fitted model and the dotted lines the 95% confidence interval.



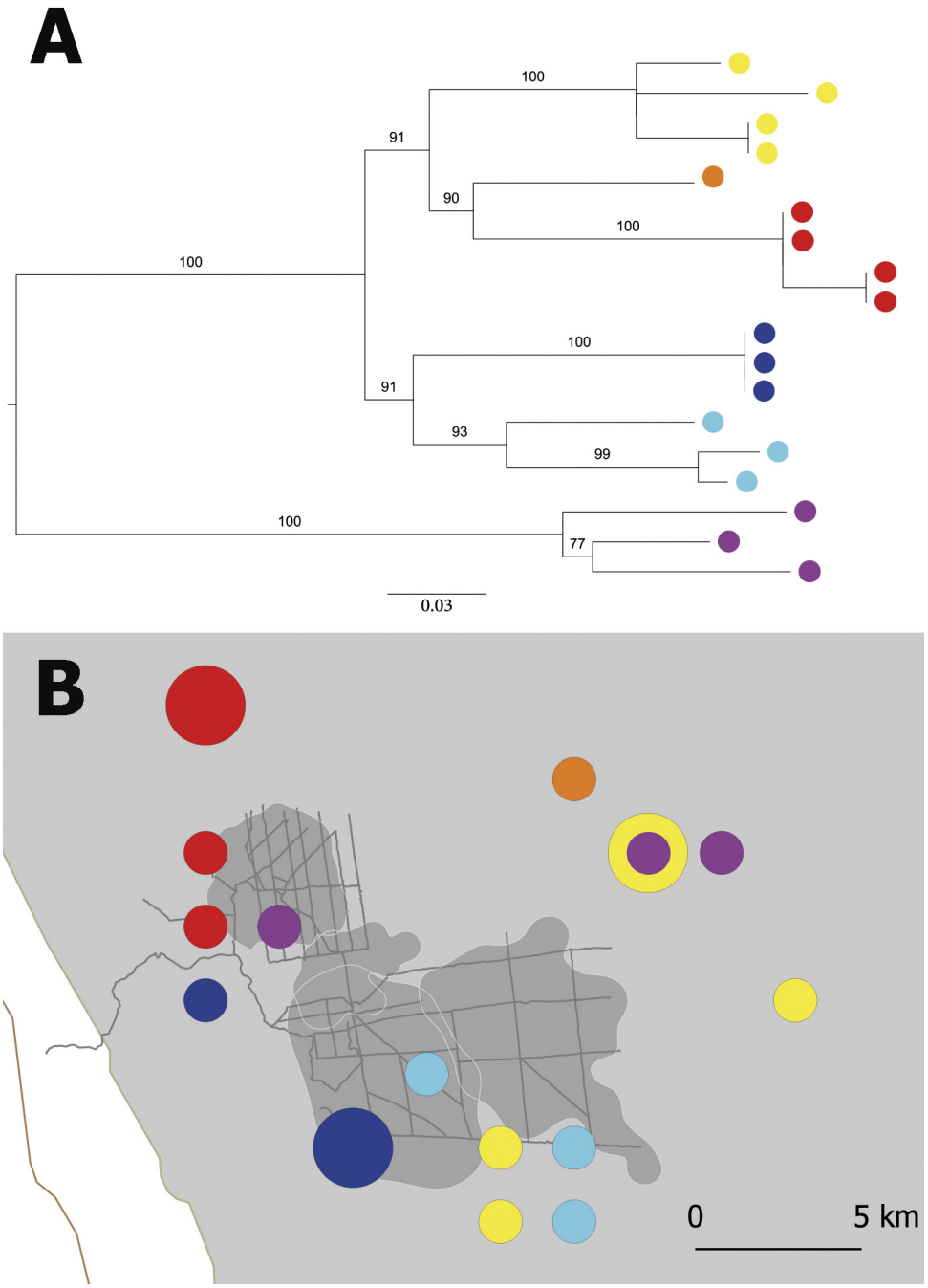
Extended Data Figure 5 | Maximum clade credibility tree based on chromosomal sequences of Bcbva isolates from TNP (Côte d'Ivoire, $n = 124$) and Grebo (Liberia, $n = 2$). One sequence per host (mammals or flies, two divergent isolates for fly 600) was included and the final alignment of variant sites measured 298 bp. The size of the nodes represents posterior probability values. The location of the root received a posterior probability of 1.



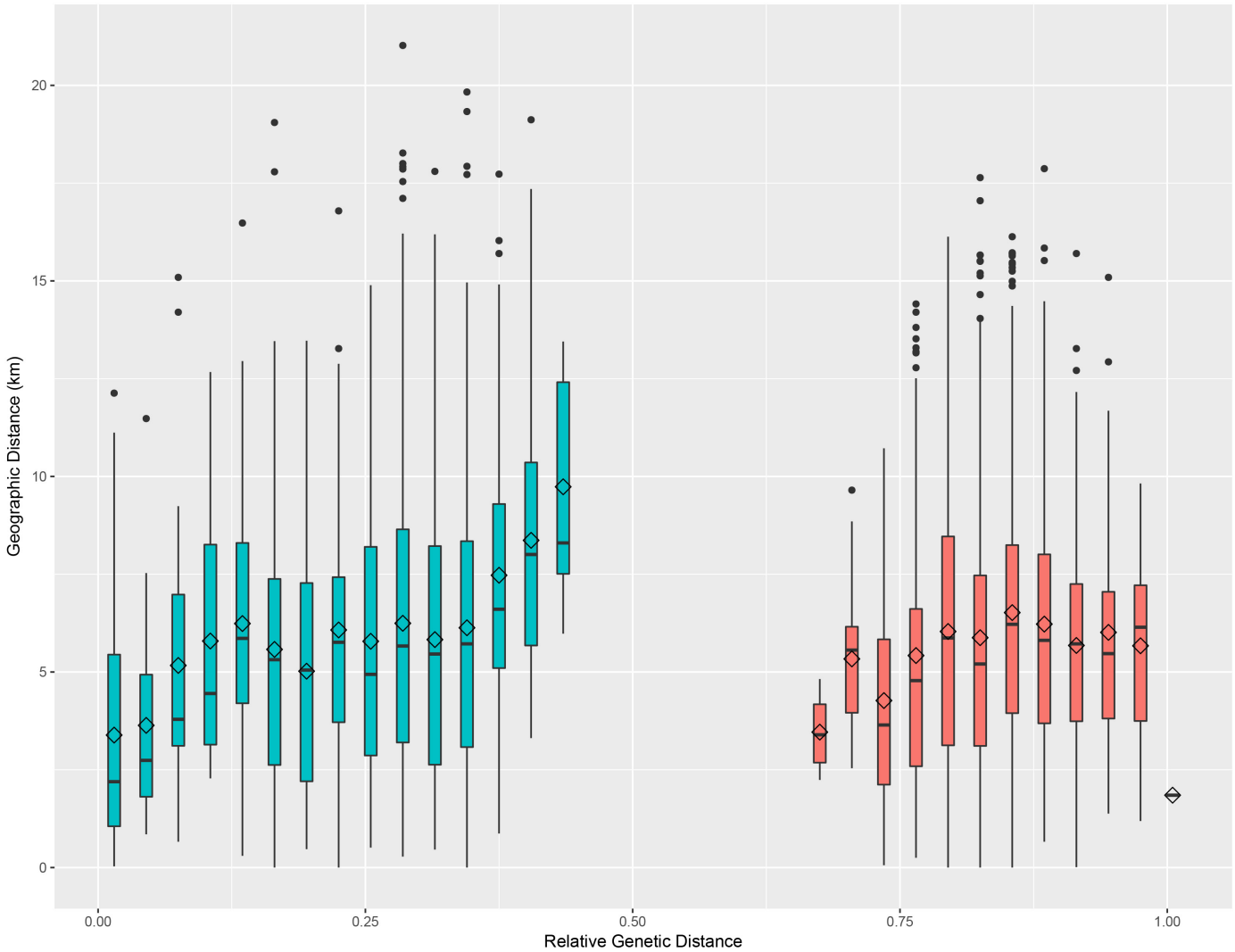
Extended Data Figure 6 | Maximum likelihood tree for sub-Saharan Bcbva strains. Maximum likelihood tree based on chromosomal sequences of Bcbva strains from Côte d'Ivoire, Cameroon, Central African Republic and Liberia. The alignment of variant sites measured 1,016 bp. Bootstrap values are shown above the branches and the scale bar indicates substitution per chromosomal site. The tree was rooted using TempEst version 1.5.



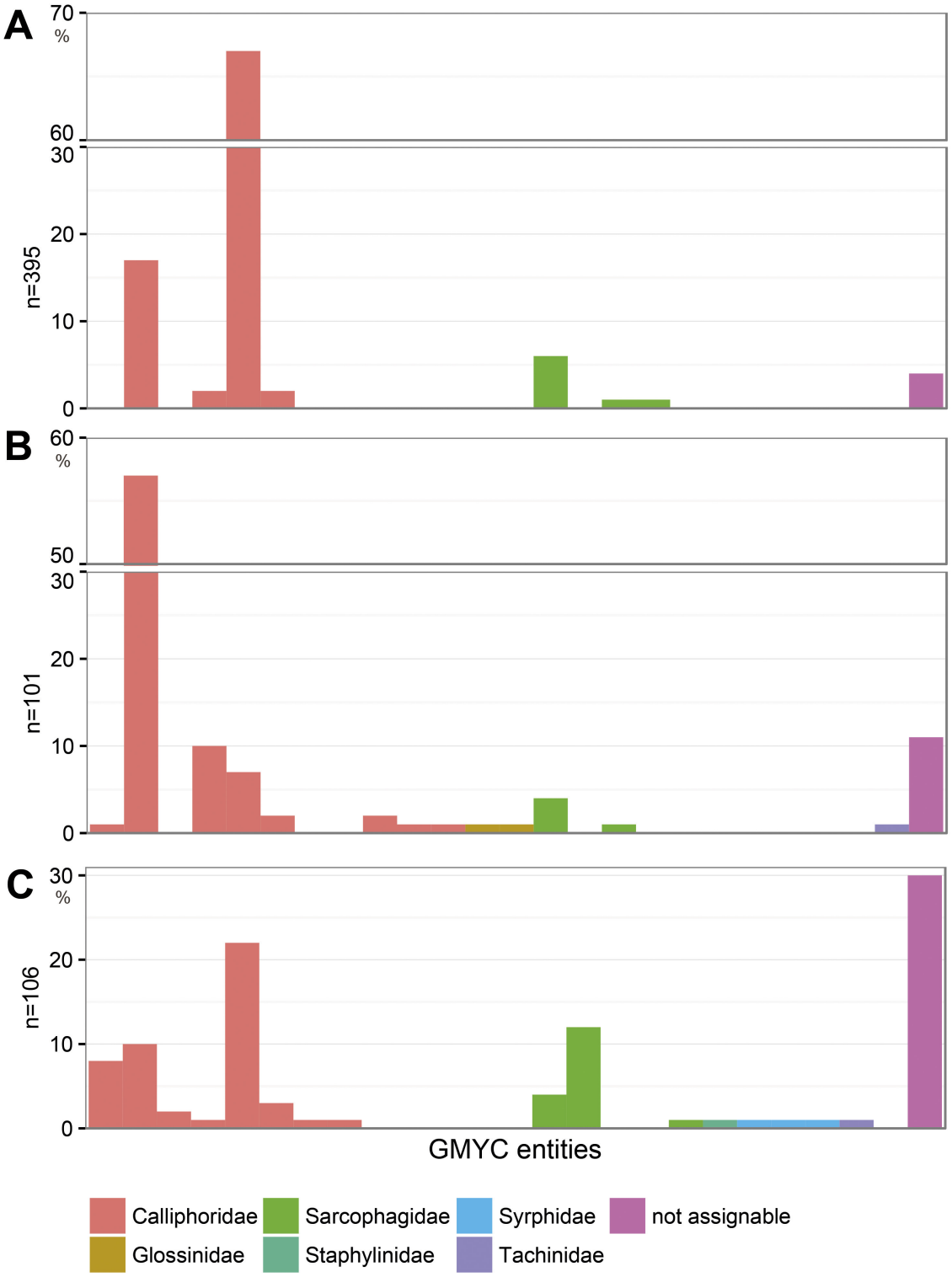
Extended Data Figure 7 | Fly snapshot sampling scheme. For the fly snapshot, flies were caught following a 2 × 2-km grid system within and outside the research area within 19 days. In total 908 snapshot flies were analysed.



Extended Data Figure 8 | Genetic and geographic distances of Bcbva isolates from the fly snapshot. a, Maximum likelihood tree based on chromosomal sequences of Bcbva isolates from the 19-day fly snapshot. Each dot represents one fly isolate. Colours were chosen to illustrate the distribution of genetically clustering isolates on the map presented in **b**. The final alignment of variant sites measured 123 bp. Bootstrap values are shown above all internal branches. The tree was rooted using the 'best-fit' option in Path-O-Gen version 1.2. The scale bar shows substitutions per site. **b**, Geographic origin of Bcbva isolates collected during the fly snapshot. Colours correspond to maximum likelihood tree in **a**. Large circles represent two isolates.



Extended Data Figure 9 | Box plot of genetic and mean geographic distances. Bcbva isolates from TNP were binned by relative genetic distance (bin size = 0.03, approximately 2.5 SNPs).The two most genetically distant isolates received a value of 1 and all other distances were scaled accordingly. Diamonds indicate the geographic distance means of the groups. To examine variation within genetic lineages, we analysed isolates with low genetic distance (maximum relative genetic distance <0.5, marked in blue) and their mean geographic distance. For low genomic distances, the linear regression of geographic distances on genetic distances has an R^2 of 0.72 and a slope coefficient that differs significantly from zero (Student's t-test, $P = 4 \times 10^{-5}$).



Extended Data Figure 10 | Fly species composition based on generalized mixed Yule-coalescent model (GMYC) analysis. a–c, Fly species composition for three sites with known Bcbva occurrence: TNP, Côte d'Ivoire (**a**); Dja Faunal Reserve, Cameroon (**b**); Dzanga-Sangha Protected Areas, Central African Republic (**c**). The proportions of flies per site (%) belonging to a single fly species identified with GMYC models are shown. Different colours indicate different taxonomic fly families.