doi:10.1038/nature23309

Supplementary Material

S1 – Study sites

a) Small scale: Taï National Park, Côte d'Ivoire

Taï National Park (TNP), Côte d'Ivoire (Fig. S1), covers an area of 3,300 km^2 and an additional 200 km^2 buffer zone, representing the largest remaining tropical rain forest block in West Africa. TNP harbors almost 1,000 species of vertebrates and eleven species of primates, including chimpanzees (*Pan troglodytes verus*). Long-term behavioral study projects on chimpanzees and monkeys started in 1979 and 1989 respectively and three habituated chimpanzee communities are observed by primatologists. TNP has two annual rainfall peaks in June and September and a distinct dry period from December to February. Human pressure on the ecosystem has been increasing during the last 40 years and numerous villages line the edges of the park today[23,27,61]. Since 2001 this long-term study site includes a veterinary program responsible for outbreak investigations in wildlife, though veterinary surveillance could not be maintained during 2003 and 2010 due to political instability. We defined the research area as the habitat ranges of the three habituated chimpanzee groups plus a 500 m buffer zone (103 km^2) (Fig. S2).

b) Large scale: West- and Central Africa

Samples belonging to the large scale data set were collected at 16 different sites in eleven sub-Saharan countries stretching from Senegal to Uganda (Fig. S3, Table S1). These sites represent different types of great ape habitats including moist rain forest and savanna. Most sites (14 out of 16) were temporary research sites of the *Pan African Programme* (www. panafrican.eva.mpg.de) led by the Max-Planck-Institute for Evolutionary Anthropology, Leipzig, Germany. Within this program ecological, social, demographic and behavioral data about 35 to 40 chimpanzee populations is systematically collected. *Bacillus cereus* biovar *anthracis* (*Bcbva*) has not been previously described at *Pan African Programme* sites. Additional samples were obtained from two sites where *Bcbva* is known to occur: Dja Faunal Reserve (DJR), Cameroon[18] and Dzanga-Sangha Protected Areas (DSPA), Central African Republic[17].

S2 – Necropsies

a) Sampling

Continuous carcass monitoring has been performed in TNP since 2001, following an anthrax outbreak observed in 2001[14] as part of a veterinary program implemented in collaboration with the Robert Koch-Institute, Berlin, Germany. A veterinarian permanently on site performs necropsies on every carcass reported by researchers working in the forest. Tissue samples of all inner organs are taken, as far as the state of carcass decomposition allows. Necropsies follow a standardized protocol, including use of full PPE (personal protective equipment) due to the occurrence of anthrax, Ebola and

monkeypox in the study area[62,63]. After every necropsy the carcass site is decontaminated according to World Health Organization (WHO) guidelines{Turnbull, 1998 #9;Turnbull, 2008 #10;Turnbull, 2008 #10;Turnbull, 2008 #10;Turnbull, 2008 #10;Turnbull, 2008 #10;Turnbull, 2008 #10}, involving burial of the carcass and incineration or disinfection with 10% - 30% formalin of all contaminated materials. In the field, samples are stored in liquid nitrogen and formalin. Frozen samples are transported on dry ice and stored at -80°C on arrival. In total, 173 necropsies were performed by the TNP veterinary program between 1998 and 2015. In addition, we received tissue samples from 31 carcasses sampled by the WHO in TNP between 1996 and 2000 (Extended Data Fig. S1, Extended Data Fig. S2, Table S2).

b) DNA Extraction

DNA was extracted from various tissues per animal (liver, spleen and lung if available) following the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany); extracts were subsequently quantified using Nanodrop (Thermo Scientific, Waltham, MA, USA) and stored at -20°C.

c) Real-time PCR assay for *Bcbva* detection

All DNA extracts were tested for anthrax in duplicate with real-time PCR. The full anthrax assay used includes three real-time PCRs each of them targeting one of the following three 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[17] (Table S3). The chromosomal *Island IV* marker is specific for *Bcbva* and thus can be used to discriminate *Bcbva* from *Bacillus anthracis*[15]. Samples were first tested for *pag*. Only samples positive in duplicate for *pag* were tested for *capB* and *Island IV*. Each PCR mix contained 0.2 mM dNTP (with dUTP replacing dTTP), 4 mM $MgCl_2$, 0.3 µM of each primer, 0.1 µM of the corresponding probe, 0.5 U Platinum® Taq Polymerase (Invitrogen, Carlsbad, CA, USA) and PCR water. PCRs were carried out in a total volume of 25 µl and seeded with 200 ng DNA or 5 µl of DNA extract if DNA concentration was below 40 ng/µl. Cycling conditions were: 600 sec at 95°C and 45 cycles of 15 sec at 95°C and 34 sec at 60°C. PCR runs were performed in a Stratagene qPCR MX3000 cycler (Stratagene, Lajolla, CA, USA) and fluorescence signals were evaluated with the software MXPRO (www.genomics.agilent.com). 81 out of 204 necropsy samples were tested positive for *Bcbva* by real-time PCR (Extended Data Fig. S1, Extended Data Fig. S2, Table S2).

d) Bacterial culture

Culture under BSL3 conditions was attempted for all PCR positive necropsy samples collected until the end of 2013 (June 2014 for duikers) when suitable material was available. One native and one heat- treated (65°C for 30 min, to assess presence of spores) aliquot were plated onto the following agar plates: Columbia blood agar (Oxoid, Wesel, Germany), blood-trimethoprim agar (1.6 mg trimethoprim, 6.4 mg sulfamethoxazole, 20 mg polymyxin B per liter agar medium) and Cereus Ident agar

(Heipha Diagnostica, Eppelheim, Germany) with the chromogenic substrate 5-bromo-4-chloro-3-indoxyl-myoinositol-1-phosphate[30]. Cultures were incubated at 37°C and monitored daily. Morphologically suspicious colonies were sub-cultured and tested in real-time PCR as described above. *Bcbva* could be cultured from native and heat-treated samples indicating the presence of heat-resistant spores. *Bcbva* isolates were retrieved from 55 of 60 samples (Table S2). Isolates were frozen in Microbank tubes (Mast Diagnostica, Reinfeld, Germany) at -80°C.

e) Histology

Histopathology was performed on a subset of 15 PCR positive necropsy samples (Table S2). The most consistent histopathologic finding was per-acute to acute anthrax related pneumonia characterized by mild lymphohistiocytic infiltrates and intraalveolar eosinophilic and proteinaceous or fibrinous material. Numerous bacilli were found intravascular and intraalveolar. Multifocal alveolar and peribronchiolar hemorrhages were present in all animals. Lymph node changes consisted of sinus histiocytosis, cortical hemorrages and edema especially in the mediastinal, tracheobronchiolar and mesentery lymph nodes. Huge amounts of bacilli were demonstrable in the sinusoids. Within the abdominal cavity the spleen was the organ most affected, 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.

S3 – Blow-flies

a) Sampling

Flies were caught on the ground and in the canopy using custom-made traps. Fly traps on the ground included a pyramidal net attached to nearby branches and a plastic bowl containing the bait (Fig. S4). Fly traps up to 30 m in the canopy consisted of a net hanging over a plastic disk equipped with a closing mechanism that could be operated from the ground (Fig. S5). Either meat or commercially available bait based on animal proteins (Unkonventionelle Produkte Feldner, Waldsee, Germany) was used as attractant. Contact of flies and bait was prevented by covering the bait containing plastic bowl with a net. Trapping duration varied from 20 to 60 min depending on catch success, with a maximum of 20 flies collected per trap. Flies were euthanized with ether and collected in 50 ml Falcon tubes (Carl Roth, Karlsruhe, Germany). Flies were stored at -20°C in 2 ml Cryotubes (Carl Roth) containing up to ten flies or at ambient temperature on silica in 50 ml Falcon tube containing up to 20 flies (Table S4). At a small scale, in TNP, 726 flies were randomly collected within the research area on the ground and in the canopy in the years 2008, 2009, 2012 and 2013 (referred to as „random flies“). Another 908 flies were collected over 19 days in May and June 2014 according to a 2x2 km grid system covering the research area and 225 km^2 surrounding the research area (referred to as “snapshot flies”) (Extended Data Fig. S7, Additional Data table S1). At a larger scale, 784 flies were collected at eight sites within five sub-Saharan countries (*Pan African Programme*) in the years 2012 to 2014 (Table S4). At all eight sites *Bcbva* had not been described

before. Another 305 flies were analyzed from two sub-Saharan sites where *Bcbva* cases had been detected before: 105 flies caught in DJR, Cameroon and 200 flies caught in DSPA, Central African Republic[17,18] (Table S4). In total, 2,723 flies caught following these sampling schemes were analyzed (Table S4).

b) DNA Extraction

DNA extraction was performed on individual flies using the GeneMATRIX Stool DNA Purification Kit (Roboklon, Berlin, Germany). Each fly was cut into small pieces using sterilized scissors before being homogenized using a Fast Prep® (MP Biomedicals, Santa Ana, CA, USA) and subsequently manufactures instructions were followed. DNA concentration of extracts was measured using a Nanodrop (Thermo Scientific, Waltham, MA, USA).

c) Real-time PCR assay for *Bcbva* detection

All flies were tested by real-time PCR as described for necropsy samples (Table S4, Additional Data table S1).

d) Bacterial culture

A subset of 50 flies containing high *pag* copy numbers was chosen for bacterial culture (Table S4, Additional Data table S1). Half of the fly mush remaining after DNA extraction was plated directly onto the same culture media described for 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* could be retrieved from native and heat-treated samples, indicating the presence of heat-resistant spores in flies. Culture was possible for 41 samples, and successfully cultured flies contained an average of 387 copies of *pag* per *5* µl of DNA extract. An on-site study of 204 flies from TNP used direct culture from flies without preceding PCR testing. Flies were homogenized and the complete material plated directly onto Cereus Ident agar. Suspicious colonies were sub-cultured on blood-trimethoprim agar and tested in real-time PCR. This approach retrieved an additional 21 *Bcbva* isolates (Table S4, Additional Data table S1).

e) Fly meal analysis

Flies are extremely efficient in detecting recently deceased mammals without having an apparent preference towards certain species. To get a hint whether certain mammals were preferentially affected by *Bcbva*, we tested whether there were differences in the fly meal composition of anthrax positive and negative flies. We first screened a subset of 750 flies from TNP for mammalian DNA using a real time PCR targeting a 130 bp fragment of mammalian 16S mitochondrial DNA. The assay is described in detail by Calvignac-Spencer et al.[19]. We chose a subset of mammal and anthrax positive (n=28) and the according number of mammal positive but anthrax negative flies (n=29) from the same trap (Additional Data table S1).

To dissect fly meal composition we used a metabarcoding approach, whereby 16S amplicons were deep-sequenced. We adapted the amplicon preparation protocol provided

by Illumina (https://support.illumina.com/downloads/16s_metagenomic_sequencing_library_preparation.html) to our needs. Instead of a two-step we followed a three-step PCR protocol to reduce amplification biases[64]. For each fly, two amplicons were generated in two independent experiments. The first step PCRs were carried out in a total volume of 25 µl and were seeded with 200 ng of DNA (DNA concentration > 40 ng/µl) or 5 µl DNA (DNA concentration < 40 ng/µl). Each PCR-Mix contained 0.2 mM dNTP (with dUTP replacing dTTP), 4 mM MgCl2, 0.2 µM of each 16Smam primer[65], 1 µM of blocking primer[19], 0.3 U Amperase® uracil N-glycolsylase (Invitrogen), 1.25U Platinum® Taq Polymerase (Invitrogen), 2.5 µl 10x PCR Buffer (Invitrogen) and PCR water. Cycling condition were: 45°C 7 min, 95°C 15 min, 42 cycles of 95°C 30 sec, 64 °C 30 sec, 72°C 1 min, followed by elongation at 72°C for 10 min. PCR products were visualized on an agarose gel. Only PCR products that yielded a clear band were included in downstream analyses. In the second PCR, Illumina specific overhang adapters needed for the following indexing PCR were added. Each PCR amplification contained 25 µl of 1:50 diluted PCR-product from the precedent PCR to minimize induced biases [64], 0.2 mM dNTP, 4 mM $MgCl_2$, 0.2 µM of each fusion primer (16Smam primer plus Illumina specific adapter sequence), 1µM of each blocking primer, 1.25 U Platinum® Taq Polymerase (Invitrogen), 10x PCR Buffer (Invitrogen) and PCR water in a total volume of 50 µl. Cycling conditions were: 95°C 5 min, 5 cycles of 95°C 30 sec, 64°C 30 sec, 72°C 1 min, followed by elongation at 72°C for 10 min. PCR products were visualized on agarose gel. Only PCR products that yielded a clear band were included for downstream analysis. At the third limited-cycle amplification step, multiplexing indices enable pooling of samples and Illumina sequencing adapters were added to each sample. Each PCR mix contained 5 µl of cleaned PCR product from the precedent PCR, 25 µl KAPA HiFiTM Hot Start ReadyMix (PeqLab, Erlangen, Germany), 10 µM of each Nextera Index Primer (Illumina) in a total volume of 50 µl. Cycling conditions were 95°C 3 min, 15 cycles of 95°C 30 sec, 55°C 30 sec, 72°C 30 sec followed by elongation at 72°C for 5 min. Two clean-ups were performed using the Agencourt® AMPure® XP PCR1 Purification system (Beckman Coulter, CA, USA). First, before Index PCR using 25 µl PCR product, 45 µl Ampure® beads eluting in 25 µl elution buffer and second, after Index PCR using 50 µl PCR product, 90 µl Ampure® beads and eluting in 25 µl elution buffer. The dual-indexed amplicon libraries were quantified using Quant-iT ™ dsDNA Broad-Range Assay Kit (Invitrogen). These were subsequently pooled in equimolar concentrations (4 nM) and sequenced using the Illumina MiSeq sequencer with MiSeq Reagent Kit v2 (Illumina) with 2x250 cycles and 25 Mio reads. All Illumina specific oligonucleotide sequences including indexes and adapters are available at the Illumina website (https://support.illumina.com/downloads/illumina-customer-sequence-letter.html). We obtained ~3.4 Mio raw reads and an average of 59,000 raw sequences per fly.

For quality trimming, cleaning and analyzing the NGS data we used the following bioinformatic pipeline: In a first step paired-end reads were merged with the program *illuminapairedend* of the software package *OBITools* v1.1.18[66] setting the minimum alignment score to 40. Primers were removed using the program *Cutadapt* v1.2.1[67]. Subsequent quality trimming was conducted utilizing the program *Trimmomatic* v0.35[68] setting the quality score to 30 over a sliding window of four bases (window size). For de-

replication of identical sequences and filtering for identical sequences that occurred at least ten times within the sample we used the *obiuniq* and *obigrep* command of the software package *OBITools.* Quality trimming and cleaning reduced the average number of raw reads per fly by 14%, to an average number of 51,000 sequences per fly.

For taxonomic assignment a reference database was built performing an *in silico* PCR on all mammalian and vertebrate sequences available at *Genbank* (http://www.ncbi.nlm.nih.gov/genbank/) using the program *ecoPCR* v0.2[69,70]. This database contained the reference sequences itself and a unique *taxid* linking each sequence to a taxonomy database including broader taxonomic information. Assignment of obtained sequences was done using the *ecotag* program of the software package *OBITools. Ecotag* implements the global alignment algorithm Needleman-Wunsch to find the most similar sequence to the query sequence in the reference database with a minimum identity level of 0.95 (primary reference sequence). The query sequence is assigned to the most common recent ancestor of the primary reference sequence and the secondary reference sequence that is the most similar reference sequence to the primary reference.

For each amplicon we obtained the counts of sequences taxonomically assigned at genus and order level (Additional Data table S2 and S3). Sequences assigned to domestic animals were regarded as contamination. In total, 21 genera belonging to six mammalian orders were identified (Table S5). It was striking that 61% of all flies tested (35/57) contained rodent DNA whereas only 4 % of the necropsies (9/204) were performed on rodents (Table S2). This finding underlines the potential of flies to give a more unbiased picture of mammal diversity and maybe mammal mortalities in the forest. Six of the 17 genera found in anthrax positive flies corresponded to wild mammal genera tested positive for *Bcbva* in necropsies (n=8). The additional eleven genera detected in anthrax positive flies suggest the host range *Bcbva* exploits is larger than previously estimated. Chimpanzees, a species majorly affected by *Bcbva*, were not detected in flies, though this is due to a technical reason; throughout the process of amplification, human blocking primers were used that also block the amplification of chimpanzee sequences. Genera detected in anthrax positive flies (n=17) and in anthrax negative flies (n=15) largely overlapped (12 genera detected in both groups) (Table S5). To test whether there were differences in fly meal composition between anthrax positive and anthrax negative flies we used General Linear Mixed Models (Supplementary Materials part S8).

Raw reads of 16S amplicons are available in the European Nucleotide Archive (ENA) under the project accession number PRJEB14554, sample accession numbers ERS1217219 to ERS1217336 and the alignment is available on request.

f) Fly species determination

Using flies as a monitoring tool, we detected *Bcbva* in TNP and Grebo, Liberia. However, flies collected in DJR, Cameroon (n=105) and DSPA, Central African Republic (n=200) were all negative for *Bcbva*, though *Bcbva* has been detected before at these sites. This could be due to various reasons: insufficient sampling effort, no anthrax activity at the time of sampling (not the case for DSPA) or differences in detectability. We suggest that variation in fly species composition may also influence *Bcbva* detectability. To identify fly species, we opted for a species delimitation approach relying

on phylogenetic information, the Generalized Mixed Yule Coalescent Model (GMYC)[71,72]. GMYC models the transition points between inter-species and intra-species branching assuming intra-species branching occurs at a higher rate than inter-species branching. GMYC identifies branching rate transition points and thus requires an ultrametric phylogeny. A single gene locus was used to infer phylogenetic information: a frequently used 710 bp region of the mitochondrial cytochrome c oxidase unit (COI)[73].

Sanger sequencing of the COI fragment was performed for 419 flies from TNP, 105 flies from DJR and 126 flies from DSPA. PCR amplification was carried out in a total volume of 25 µl containing 0.2 mM dNTP (with dUTP replacing dTTP), 4 mM $MgCl_2$, 0.2 µM of each primer (LCO1490, LCO2198)[73] 0.3 U Amperase® uracil N-glycolsylase (Invitrogen), 1.25U Platinum® Taq Polymerase (Invitrogen), 2.5 µl 10x PCR Buffer (Invitrogen) and PCR water. Each reaction was seeded with 200 ng DNA extract or 5 µl if DNA concentration was below 40 ng/µl. Cycling conditions were: 45°C 7 min, 95°C 15 min, 35 cycles of 95°C 30 sec, 50 °C 30 sec, 72°C 1 min, followed by elongation at 72°C for 10 min. PCR products were visualized on agarose gels and cleaned using ExoSAP-IT® (Affymetrix, Santa Clara, CA, USA). Sanger-sequencing was performed in both directions on the ABI PRISM® 3100 Genetic Analyzer (Applied Biosystems, Waltham, MA, USA) using amplifying primers and the BigDye Terminator v3.1 cycle kit (Thermofischer, Waltham, MA, USA). All chromatograms were evaluated using *Geneious Pro* v 8.1.3 (Biomatters Ltd., Auckland, New Zealand)[41]. Primers were truncated and sequences manually checked for quality. Sequences containing ambiguous bases were excluded from the dataset, leaving 606 sequences (TNP n=398, DSPA n=107, DJR n=101). De-replication of the dataset was performed using the clustering program *cd-hit-est*[74] (http://weizhongli-lab.org/cd-hit/servers.php) setting the identity cut-off to 1. This resulted in 316 clusters, each represented by one sequence used for further analysis. By translating unique sequences to protein sequences, one sequence was identified as a nuclear mitochondrial DNA segments (NUMTs) and excluded from down-stream analysis. Three sequences encoding *Wolbachia* sp. and one sequence belonging to a distant phylogenetic group were removed, leaving 601 sequences (TNP n=395, DSPA n=106, DJR n=101) and 312 unique clusters. Remaining unique sequences were aligned at the amino acid level using the *MUSCLE* algorithm[75] implemented in *SeaView* v4[76] and back translated to nucleotide sequence. We used the Bayesian information criterion as implemented in *jModelTest* v2.1.4[42,77] to select the best-fit model of nucleotide substitution (General Time Reversible (GTR)[78] with a gamma distribution and proportion of invariable sites[79]. A calibrated ultrametric tree for GMYC analysis was produced using *BEAST* v1.7.4[46] under the model of nucleotide substitution mentioned above, a relaxed clock (lognormal){Drummond, 2010 #82} and a coalescent model assuming constant population size. Four independent analyses were run over 50,000,000 generations and sampled each 1,000 steps. The convergence of the four runs was checked using *Tracer* v 1.6[80] and the burn-in fraction estimated accordingly (5,000,000). The four runs were pooled using *LogCombiner* v1.7.4 and the maximum clade credibility tree extracted using *TreeAnnotator* v1.7.4. The single-threshold GMYC method was implemented using the R package *splits* v1.0-19 (SPecies LImits by Threshold Statistics, available from https://r-forge.r-project.org/projects/splits/) using R v3.3.0[55]. Forty-five maximum likelihood entities (95% confidence interval: 42-56) were identified at the determined between/within species branching threshold. Representative sequences for each entity

were compared to the NCBI non-redundant database using BLAST[81]. Assignments were done at family level in case a clear barcoding gap was identified. Family level assignment was possible in 24 cases and revealed six different dipteran families (Calliphoridae, Glossinidae, Sarcophagidae, Staphylinidae, Syrphidae, Tachinidae) represented with one to eleven GMYC species (Table S6). In the other 21 cases an assignment based on BLAST results was not possible. This could be due either to erroneous and insufficient data in the database or potential nuclear mitochondrial DNA segments that inflate the number of species identified as described elsewhere[82]. To compare fly species composition between sites, all available sequences were compared to the 24 assigned GMYC entities using BLAST. At all three sites, Calliphoridae constituted the majority of flies (TNP 88% (n=395), DJR 82% (n=101), DSPA 48% (n=106)). However, the composition of Calliphoridae species varied across sites: in TNP 67% of the flies are represented by one Calliphoridae species. This species only represented 7% of flies in DJR and 22% in DSPA. Species belonging to other dipteran families also differ among sites (Extended Data Fig. S10). Varying fly species composition likely influences *Bcbva* detectability.

S4 – Bones

a) Sampling

Bones were collected in TNP as well as on a large scale at 12 *Pan African Programme* sites in nine countries (Fig. S3, Table S1). Bones collected in TNP were collected since 1989 in collaboration with the Taï Chimpanzee project and the Taï Monkey Project (Table S7). Bones were transported and stored at ambient temperature.

b) DNA Extraction

DNA was extracted using a silica-based method[33,34]. For each bone, 100-150 mg of fine powder drilled at low speed were extracted using a 48 h digestion in 5 ml of extraction buffer (0.5 M EDTA, 0.5% N-lauryl-Sarcosyl, 1 mg/mL Proteinase K at pH 8.0). The supernatant, recovered after spinning the solution at 2000 rpm for 5 min, was transferred into 20 ml of buffer (GuSCN 5 M, Tris 50 mM, NaCl 25 mM, EDTA 20 mM, TritonX-100 1x) with 80 µL of a fresh silica pellet[34]. The final pH was adjusted to 4.0–5.0 using 37% HCl and pH paper. DNA binding to silica surfaces was performed for 3 h at room temperature with agitation. After centrifugation at 2000 rpm for 2 min the supernatant (except for ~1 ml) was removed. The silica pellet was re-suspended, transferred to a 1.5-mL tube and centrifuged at 12,000 rpm for 2 min. The supernatant was removed and washed twice with 1 mL 80% ethanol, centrifuged again at 12,000 rpm for 2 min. The silica pellet was dried for 20 min at room temperature in a laminar flow-hood and re-suspended in 90 µl of nuclease free water. The suspension was incubated for 20-30 min at 37°C and centrifuged at 12,000 rpm for 2 min. The supernatant was transferred to a new tube and stored at -20°C.

c) Real Time PCR assay for *Bcbva* detection

All bones were tested by real-time PCR as described for necropsy samples (Table S1, Table S7).

d) Bacterial culture

Powder from PCR positive bones was homogenized in sterile NaCl and processed as described above for necropsy samples with one native aliquot and one heat- treated aliquot. *Bcbva* could be cultured from native and heat-treated samples indicating the presence of heat-resistant spores. *Bcbva* isolates were retrieved from 7 of 26 PCR positive bones (Table S7).

S5 – Soil

a) Sampling

Soil sampling was conducted in parallel with the fly snapshot sampling (see S3a). At every fly sampling location, one spoon of superficial soil was transferred into a 50 ml Falcon tube (Carl Roth) and stored at ambient temperature in the field until arrival at the Robert Koch-Institute, where samples were stored at 4°C. An evenly spatially distributed subset of 35 of 82 samples was analyzed.

b) Extraction

DNA extraction from soil was performed using the PowerMax® Soil DNA Isolation Kit (MO BIO Laboratories, CA, USA). Each extraction was seeded with 6 g of soil and followed the manufacturer´s instructions. No inhibition was detected with a dedicated in-house real-time PCR assay (*12*).

c) Real Time PCR assay for *Bcbva* detection

All soil DNA extracts were tested by real-time PCR as described for necropsy samples. *Bcbva* was not detected in soil samples.

S6 – Whole-genome sequencing of *Bcbva* isolates and SNP calling

a) DNA extraction

Table S8 contains a complete list of all *Bcbva* isolates sequenced for this study (Fig. S6). 5 ml of Luria-Bertani (LB) liquid medium were inoculated with one *Bcbva* colony from Columbia blood agar (Oxoid, Wesel, Germany) and incubated overnight at 37°C. To prevent the presence of spores, the following morning a fresh 5 ml LB liquid culture was inoculated with 200 µl of the overnight culture. After incubation for 4 hours, 3 ml of culture were centrifuged and DNA was extracted from the pellet using the DNeasy Blood and Tissue kit (Qiagen). We incubated 1/10 of the resulting DNA eluate in LB medium and on Columbia blood agar for 7 days, testing for sterility before DNA left the BSL3

laboratory. DNA concentrations were determined using the Qubit® 2.0 Fluorometer (Invitrogen).

b) Library preparation and sequencing

Libraries for whole-genome sequencing were prepared with the Nextera XT DNA Sample Preparation Kit (Illumina) and the Nextera XT Index Kit (Illumina) using 1 ng of input DNA. AMpure XP beads (Beckman Coulter) were used for PCR clean-up. We validated libraries with a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, US) and normalized them for sequencing. Libraries were pooled and sequenced on the HiSeq 1500 platform (Illumina) in rapid run mode using either v1 (2x150bp) or v2 (2x250bp) chemistry.

Raw reads for all 178 *Bcbva* isolates from TNP and Grebo are available in the European Nucleotide Archive (ENA) under the project accession number PRJEB14616, sample accession numbers ERS1222903 to ERS1223080.

c) Bioinformatic analysis

Illumina adapters were removed using scythe v0.993[35] and raw reads trimmed with *sickle* v1.33[36] applying a quality threshold of 25. Quality trimmed reads were aligned to the reference genome (*Bcbva* strain CI, Accession numbers CP001746- CP001749) with the BWA-MEM algorithm implemented in *bwa* v0.7.12-r1039[37]. For conversion to bam format, sorting, deduplication and indexing of aligned reads, we used the *picard tools* 1.136[43] software package applying the commands *SortSam*, *MarkDuplicates* and *BuildBamIndex*. Subsequent variant calling was done using the *Genome Analysis Toolkit* (GATK) v3.4 [38-40]. In a first step we realigned the bam files with the tools *RealignerTargetCreator* and *IndelRealigner*. Variants were called with *UnifiedGenotyper* and a minimum phred scaled confidence threshold of 30 for SNPs to be called. Hard filtering of resulting 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 < 5x, a minor allele frequency of > 0.1 and a GATK Genotype Quality value < 99. Final consensus sequences were composed with the *FastaAlternateReferenceMaker*. We assessed coverages of all samples with the GATK tools *DepthOfCoverage* and *CoveredByNSamplesSites*.

d) Within-host diversity testing

Several bacterial colonies from the same host (mammal or fly) were sequenced in a number of cases to examine within- host diversity (Table S9). For mammals, colonies originated partly from separate organs. We detected a maximum difference of two SNPs within necropsy and fly isolates on chromosome level. No within-host heterogeneity was found on pXO1; pXO2 sequences differed by one SNP in one case. One fly represented an exception, carrying two distinct strains differing by 42 SNPs on chromosome level and 2 SNPs for pXO2 (Isolates 600-1 and-2). This was also the only case where colonies differed morphologically, with one especially mucous colony. For each host, the isolate with the highest sequencing coverage and smallest number of autapomorphic SNPs was

included in further analyses, with the exception of fly isolates 600-1 and -2, where both strains were analyzed.

S7 – Phylogenetic analyses

a) Nucleotide substitution model selection

126 genome sequences (one isolate per mammal/fly, for fly 600 both divergent isolates were included) from TNP and neighboring Grebo (Liberia) (Table S8) were aligned and stripped of non-variant sites with *Geneious Pro* v8.1.3 (Biomatters ltd.)[41].The resulting alignments of variant sites were 298, 18 and 11 bp long for the chromosome, pXO1 and pXO2 respectively. As sufficient genetic information for further inference was only found on the chromosome, we focused on the chromosomal analyses. *jModelTest* v2.1.4[42] was used for determination of the best nucleotide substitution model in a Maximum Likelihood framework, resulting in the choice of TVMef[43] for chromosomes from TNP, with no proportion of invariant sites or rate variation among sites.
A second data set used to investigate *Bcbva* presence in sub-Saharan Africa included the chromosomal sequences of *Bcbva* cases from Central African Republic, Cameroon and Côte d'Ivoire[15,17] and chromosomal sequences of the two new isolates from Grebo (Liberia). The alignment was compiled as described above and contained 1016 variable positions. Model selection with *jModelTest* v2.1.4[42] resulted in the use of TPM1[48] as nucleotide substitution model with no invariant sites or rate variation among sites.

b) Maximum Likelihood analysis

Maximum Likelihood analyses were performed with *PhyML* v20131022[44] implementing the selected substitution models and a combination of subtree-pruning-regrafting (SPR) and nearest-neighbor-interchange (NNI) algorithms. Branch support was estimated using non-parametric bootstrapping with 100 pseudo-replicates. The trees were rooted using the heuristic residual mean squared function in *TempEst* v 1.5[45] (Fig. 3, Extended Data Fig. S6).

c) BMCMC reconstruction using BEAST

We also used Bayesian Markov Chain Monte Carlo (BMCMC) sampling as implemented in the *BEAST* software package v1.8.2[46] to confirm the structure of our phylogenetic trees. Nucleotide substitution models were applied as determined above with no invariant sites or rate variation among sites. We specified a constant population coalescent tree prior and assumed an uncorrelated lognormal relaxed molecular clock[47] (Extended Data Fig. S5).

S8 – Statistical analyses

a) Effect of season on anthrax positivity of carcasses

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, chimps, duikers, others, blowflies), season and their interaction. 'Season' was modelled by first turning the sampling date into a circular variable and then including its sine and cosine into the model. As random intercept effects we included trap id (i.e., 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 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]. We fitted the model in R[55] using the function *glmer* of the R package *lme4* v1.1-10[56]. The sample size for this model was 1803 samples (carcasses and flies), collected at 352 sites and 328 combinations of sampling date and location including necropsy samples and flies. The full null model comparison did not reveal a significant effect of season on the probability of a carcass or fly being anthrax positive (likelihood ratio test, $\chi 2=5.96$, df=10, P=0.819).

b) Effect of season and the amount of mammalian DNA within the fly on anthrax positivity of flies

To test whether the probability of a fly to be tested anthrax positive was influenced by season and the amount of mammalian DNA within in the fly we used a Generalized Linear Mixed Model (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 with real time PCR described above) and season as fixed effects. 'Season' we modeled by first turning the sampling date into a circular variable and then including it sine and cosine into the model. Since the amount of mammalian DNA within the fly was highly skewed, we log transformed it 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 to exist. The total sample size for this model was a total of 474 flies caught on 43 days in 33 traps. Overall, the full model was clearly superior to the null model (likelihood ratio test, $\chi 2=19.17$, df=3, P<0.001). More specifically, the probability of a fly to be positive clearly increased with the amount of mammal DNA found in it (estimate+SE=0.36+0.10, $\chi 2=9.44$, df=1, P=0.002) (Extended Data Fig. S3). Furthermore, we found seasonal variation in the probability of *pag gene* positivity ($\chi 2=6.9$, df=2, P=0.032) with positive flies being particularly common in February and March (Fig. S14, Table S10).

c) Fly meal analysis

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 Spearmen correlation. The relative sequence counts found in the two amplicons were clearly correlated. In fact, the average correlation 0.67 (n=56) and even the smallest correlation of 0.33 were clearly positive. Moreover, 53 out of 56 correlations were clearly significant ($p > 0.05$).

To test whether there were differences in fly meal composition of anthrax positive and anthrax negative flies, we tested whether the detection of a given mammal taxon in a fly sample was associated with anthrax positivity. We used GLMMs[49] which we applied separately for each mammal identified in the flies. The response in the model was whether the fly was anthrax positive and the key predictor with fixed effect was mammal presence. We considered a mammal to be present when it was detected in at least one of the two amplicons per fly. We included only those mammals in the analyses 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 into the model (random intercepts)[51,52]. The models where 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 a significant impact on anthrax positivity we dropped mammal presence from the model[53] and compared the two models using a likelihood ration test[54]. We run the model in R[55] using the function *glmer* of the R package *lme4* v1.1-10[56]. Model stability was estimated as indicated above. We fitted the models at two different taxonomic resolutions: First, with taxonomic assignment at genus level and second at order level. At genus level there was no recognizable impact of mammal presence on anthrax positivity. None of the full null model comparisons revealed significance ($p<0.05$). At lower taxonomic resolution (order level), one of six full null model comparisons revealed significance. For carnivores the full model was superior to the null model (estimate+SE=2.11+1.11, $\chi 2=5.09$, df=1, p=0.024). However, because of an extremely wide confidence interval (0.36, 153.27) these values have very limited interpretation.

d) Snapshot geographic correlations

To evaluate geographic distribution of *Bcbva* in TNP we checked whether, due to higher mammal density[23], *Bcbva* was more likely to occur inside the research area. To test this hypothesis we analyzed 908 flies from 83 different traps (Extended Data Fig. S7, Additional Data table S1). 21 of these traps were located within the research area and 62 in the adjoining forest belt. 8/21 traps within the research area were anthrax positive and 8/62 outside the research area. We compared the two groups using Fisher's Exacts Test which revealed a significant difference between groups (p=0.02). Hence, it was more likely to detect *Bcbva* within the research area compared to the adjoining forest. At time of submission, adequate mammal density data was not available to test for the correlation between mammal density and anthrax positivity in flies. However, we assume that the data shown here give a good indication.

e) Correlation genetic and geographic distance

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 data set 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 (Table S8). Geographic distances (in km) were derived from GPS data using *GeographicDistanceMatrixGenerator* v1.2.3[57]. Genetic distances were approximated using the relative distances drawn from a Maximum Likelihood Tree built in *PhyML v20131022*[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 1000 permutations and Spearman correlation was performed on genetic and geographic distance matrices. It revealed a significant correlation between genetic and geographic distances (p=0.006), however only a small fraction of geographic variation was explained by genetic variation (R^2=1.3%). This is likely due to different genetic lineages co-circulating in the same area leading to spatially close isolates with large genetic differences. To illustrate this, during our 19 day fly snapshot in June 2014 we retrieved 18 *Bcbva* isolates (Table S8, Additional Data table S1). Of the 18 isolates, 13 differed by > 2 SNPs implying at least 13 contemporaneous *Bcbva* positive carcasses (based on our within-host diversity data, see S6d). Deaths were caused by multiple independent transmission chains and *Bcbva* isolates formed geographically co-circulating clusters differing by up to 48 SNPSs (Extended Data Fig. S8). To examine variation within genetic lineages, we binned our data by genetic distance (bin size= relative genetic distance of 0.03, approx. 2.5 SNPs) and focused on groups with low genetic distance (max relative genetic distance < 0.5) and their mean geographic distance (Extended Data Fig. S9). Homogeneity of variance between groups was assured with the Fligner Killeen test (p=0.07; > 0.05 as requested). We found a significant correlation between mean genetic distances of groups and their respective mean geographic distances ($p=4x10^{-5}$, R^2=0.72) underlining spatially restricted transmission of *Bcbva*.

f) Simulation of chimpanzee population dynamics

To evaluate the impact *Bcbva* could have on the TNP chimpanzee population we conducted a simulation. 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. Since most of these parameters are associated with considerable uncertainty and since 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 to 80 individuals (increment: 10); inter-birth interval: 4 to 7 years (increment: 1); interval after death of infant: 1 year; maximum age: 40 to 50years (increment: 2); age of first reproduction of males and females: 10 and 14 years, respectively. Since 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). Additionally, 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*community size)))) that increased or reduced mortality rate as a function of chimpanzee community size. In principle, it seems plausible that also fecundity is density dependent. However, we could not identify such pattern in our data (N=23 years; community size range: 19 to 76; Spearman's rho=-0.028, P=0.9). Additionally, we are not aware of studies addressing this question more in depth for chimpanzees. Since a previous study across a number of chimpanzee research sites showed evidence for density dependence of chimpanzee infant mortality rates (Fig. 9 in Kuehl et al. 2008) we felt it being justified to base our simulation on a density dependent survival probability and not also on or instead on fecundity. 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 (i.e., 'years') without anthrax presence before the evaluated time period began.

We estimated the risk of annual anthrax outbreak probability, its dependence on community size and the number of individuals affected using a Poisson regression. This revealed the annual number individuals killed by an outbreak as clearly community size dependent (full-null model comparison, χ^2= 7.89, df=1, p <0.01) with the model result being number individuals killed = exp (-1.83 + 0.039 × community size) . We simulated both an anthrax and a non-anthrax scenarios for 150 time steps (i.e., '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. In anthrax scenarios we sampled for each year a random number from a Poisson distribution with a mean dependent on community size, according to the equation that resulted from the Poisson model just described. This number of individuals was then removed from the simulated community (whereby the individuals to be removed were randomly determined).

These simulations showed that under extremely low assumed annual per capita mortality rates, anthrax only slightly changed overall chimpanzee community survival rates. However, a decreasing background mortality rate that alone did not lead to an immediately reduced survival of communities, lead to a drastic decline in the number of surviving communities, when combined with additional mortalities due to anthrax (annual per capita survival of 0.96). This suggests that the combination of different causes of mortality (e.g., poaching, predation, respiratory infections, Ebola) has led to a situation in which anthrax induced mortality has moved overall mortality towards the tipping point beyond which the Taï chimpanzee communities are dramatically declining (Fig. S7 and S8).

All R scripts and alignments are available upon request.

References (61-83)

61 McGraw, W. S., Zuberbühler, K. & Noë, R. *Monkeys of the Taï Forest: an African primate community*. Vol. 51 (Cambridge University Press, 2007).

62 Formenty, P. *et al.* Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d'Ivoire. *The Journal of infectious diseases* **179 Suppl 1**, S120-126, doi:10.1086/514296 (1999).

63 Radonic, A. *et al.* Fatal monkeypox in wild-living sooty mangabey, Cote d'Ivoire, 2012. *Emerg Infect Dis* **20**, 1009-1011, doi:10.3201/eid2006.13-1329 (2014).

64 Berry, D., Ben Mahfoudh, K., Wagner, M. & Loy, A. Barcoded primers used in multiplex amplicon pyrosequencing bias amplification. *Appl Environ Microbiol* **77**, 7846-7849, doi:10.1128/AEM.05220-11 (2011).

65 Taylor, P. G. Reproducibility of ancient DNA sequences from extinct Pleistocene fauna. *Molecular Biology and Evolution* **13**, 283 - 285 (1996).

66 Boyer, F. *et al.* obitools: a unix-inspired software package for DNA metabarcoding. *Molecular ecology resources*, doi:10.1111/1755-0998.12428 (2015).

67 Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* **17**, pp. 10-12 (2011).

68 Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120, doi:10.1093/bioinformatics/btu170 (2014).

69 Bellemain, E. *et al.* ITS as an environmental DNA barcode for fungi: an in silico approach reveals potential PCR biases. *BMC Microbiol* **10**, 189, doi:10.1186/1471-2180-10-189 (2010).

70 Ficetola, G. F. *et al.* An in silico approach for the evaluation of DNA barcodes. *BMC Genomics* **11**, 434, doi:10.1186/1471-2164-11-434 (2010).

71 Pons, J. *et al.* Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst Biol* **55**, 595-609 (2006).

72 Fujisawa, T. & Barraclough, T. G. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Syst Biol* **62**, 707-724, doi:10.1093/sysbio/syt033 (2013).

73 Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular marine biology and biotechnology* **3**, 294-299 (1994).

74 Li, W. & Godzik, A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659, doi:10.1093/bioinformatics/btl158 (2006).

75 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).

76 Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* **27**, 221-224, doi:10.1093/molbev/msp259 (2010).

77 Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704 (2003).

78 Tavaré, S. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences* **17**, 57-86 (1986).

79 Lanave, C., Preparata, G., Saccone, C. & Serio, G. A new method for calculating evolutionary substitution rates. *J Mol Evol* **20**, 86-93 (1984).
80 Rambaut, A., Suchard, A., Xie, D. & Drummond, A. J. Tracer v1.6, Available from http://beast.bio.ed.ac.uk/Tracer. (2014).
81 Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *Journal of molecular biology* **215**, 403-410 (1990).
82 Song, H., Buhay, J. E., Whiting, M. F. & Crandall, K. A. Many species in one: DNA barcoding overestimates the number of species when nuclear mitochondrial pseudogenes are coamplified. *Proceedings of the National Academy of Sciences* **105**, 13486-13491 (2008).
83 QGIS Development Team. (Open Source Geospatial Foundation Project. , 2013).

Fig. S1.

Geographic location of Taï National Park. The Taï National Park (green) is located in the very western part of Côte d`Ivoire at the border to Liberia. It represents the largest remain of tropical rain forest in West Africa. All maps were created using QGIS [83].

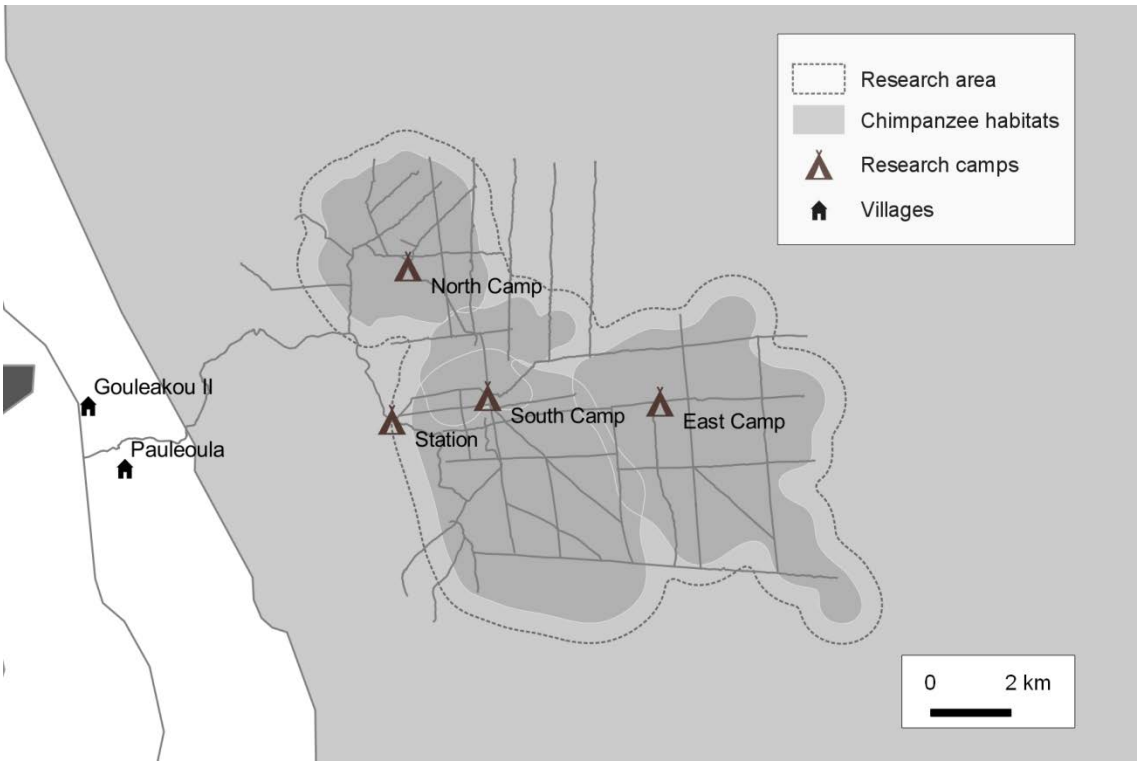

Fig. S2.

Research area of Taï National Park. The research area of Taï National Park (light grey) was defined as the home ranges of the habituated chimpanzee groups (dark grey) plus a 500 m buffer zone and is indicated with the dashed line. The research area covers 103 km^2.

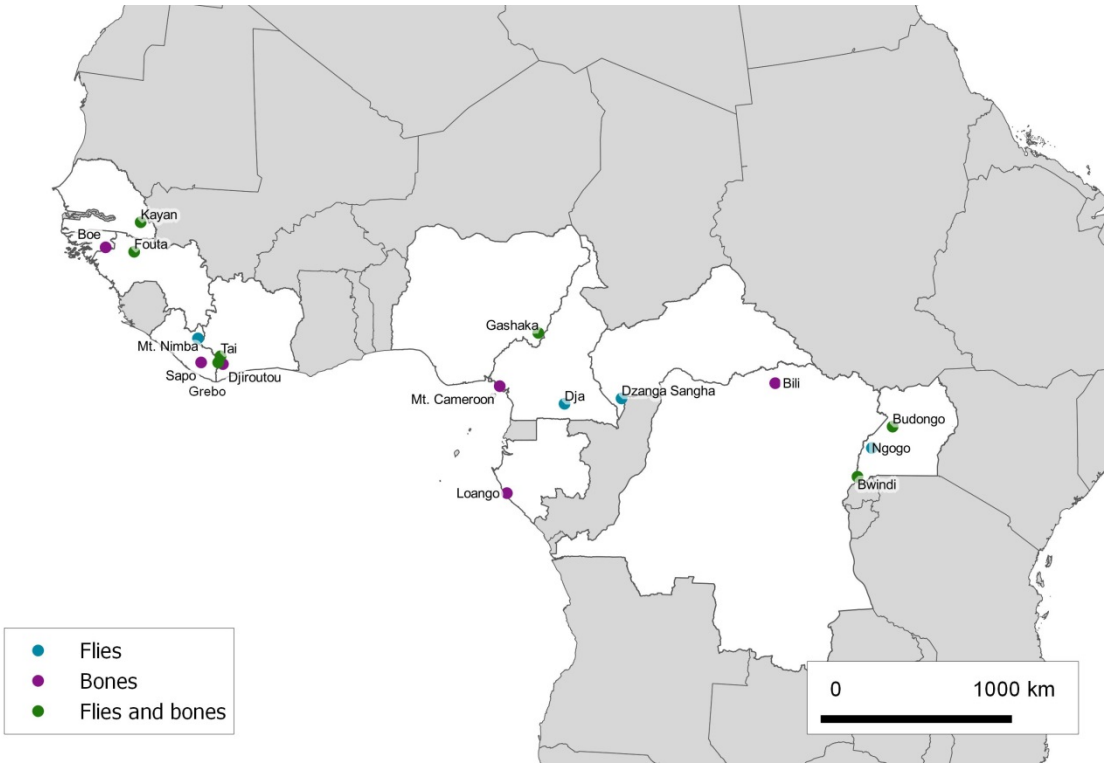

Fig. S3.

Sampling sites of flies and bones. At a large scale flies and bones were collected at 16 different sites in eleven sub-Saharan countries. At each site either flies or bones or both sample types were collected. All sites represent different types of chimpanzee habitats including moist rain forest as well as savanna.

Fig. S4.

Fly trap on the ground. Flies are caught using a pyramidal shaped net that is attached to nearby branches. Commercially available bait based on animal proteins is placed in a plastic bowl and covered with a net to prevent contact to the bait.

Fig. S5.

Fly trap in the canopy. Flies in the canopy (up to 30m high) were caught with the same bait as flies on the ground. The fly net can be closed by pulling the attached string from the ground.

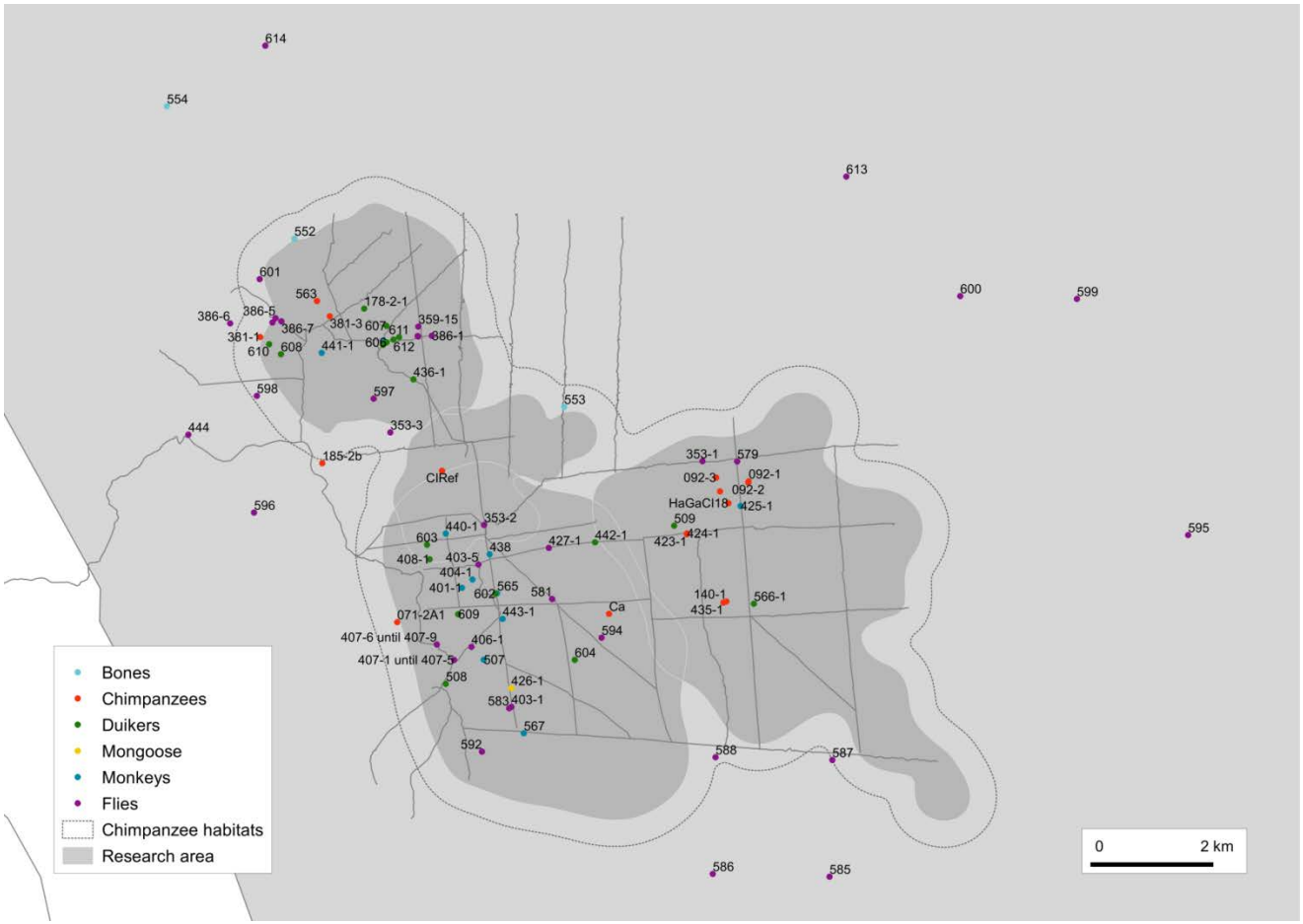

Fig. S6.

Geographic locations of *Bcbva* cases and *Bcbva* positive flies used for genetic analyses. GPS data is missing for the following isolates: 178-6-1, 178-7-1, 459, 460, 461, 462, 503-1, 530, 531. The two isolates from Grebo (Liberia) included in our study are not shown, but the location of Grebo (~40 km from TNP) is presented in Fig. S1.

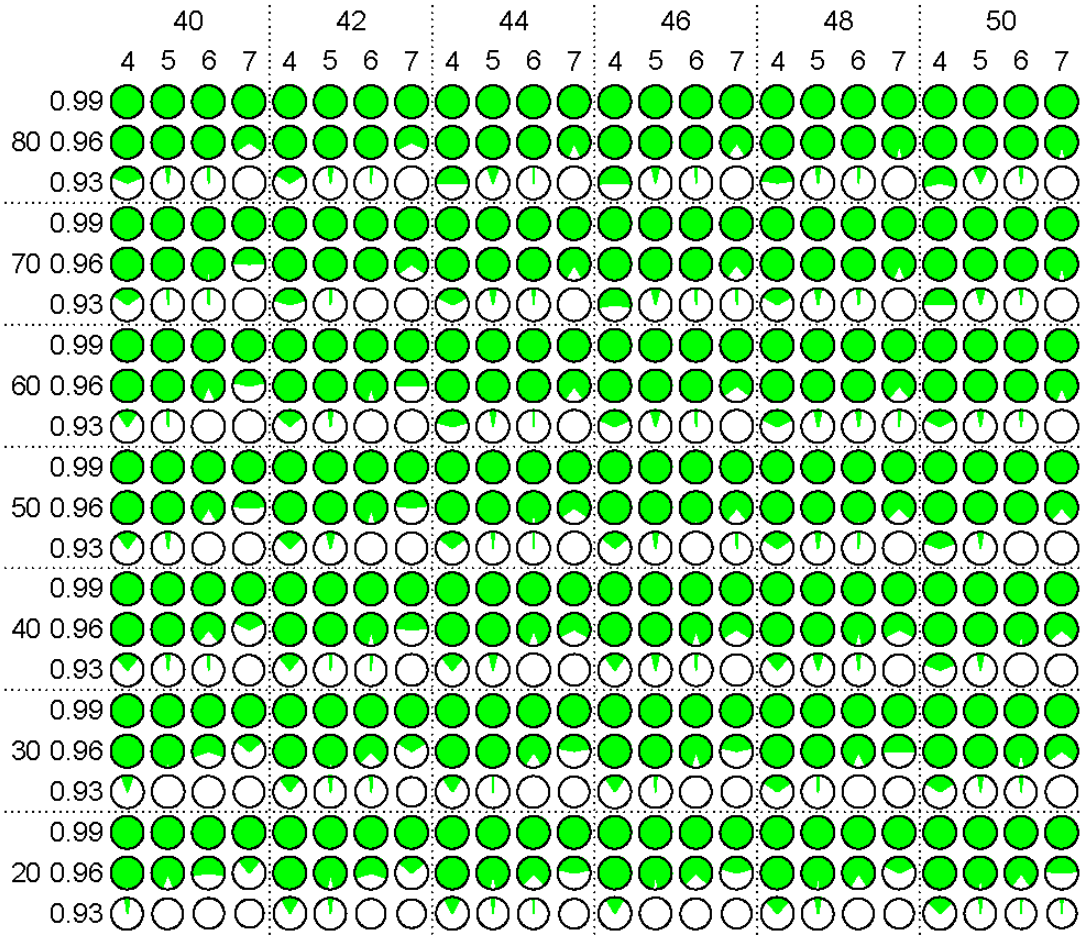

Fig. S7.

Proportions of simulated chimpanzee communities surviving 150 years without of the presence of anthrax. Shown are results for different community sizes (outer y-axis), per capita annual survival probabilities (inner y-axis), maximum ages upper x-axis, inter birth intervals (lower x-axis)). The area of the circles depicts the number of communities present at the beginning of the simulation (in a few simulations the simulated community did not survive the 50-years burn-in phase; maximum area corresponds to 100 simulations). The green area depicts the proportion of populations that survived for 150 time steps after the end of the burn-in phase.

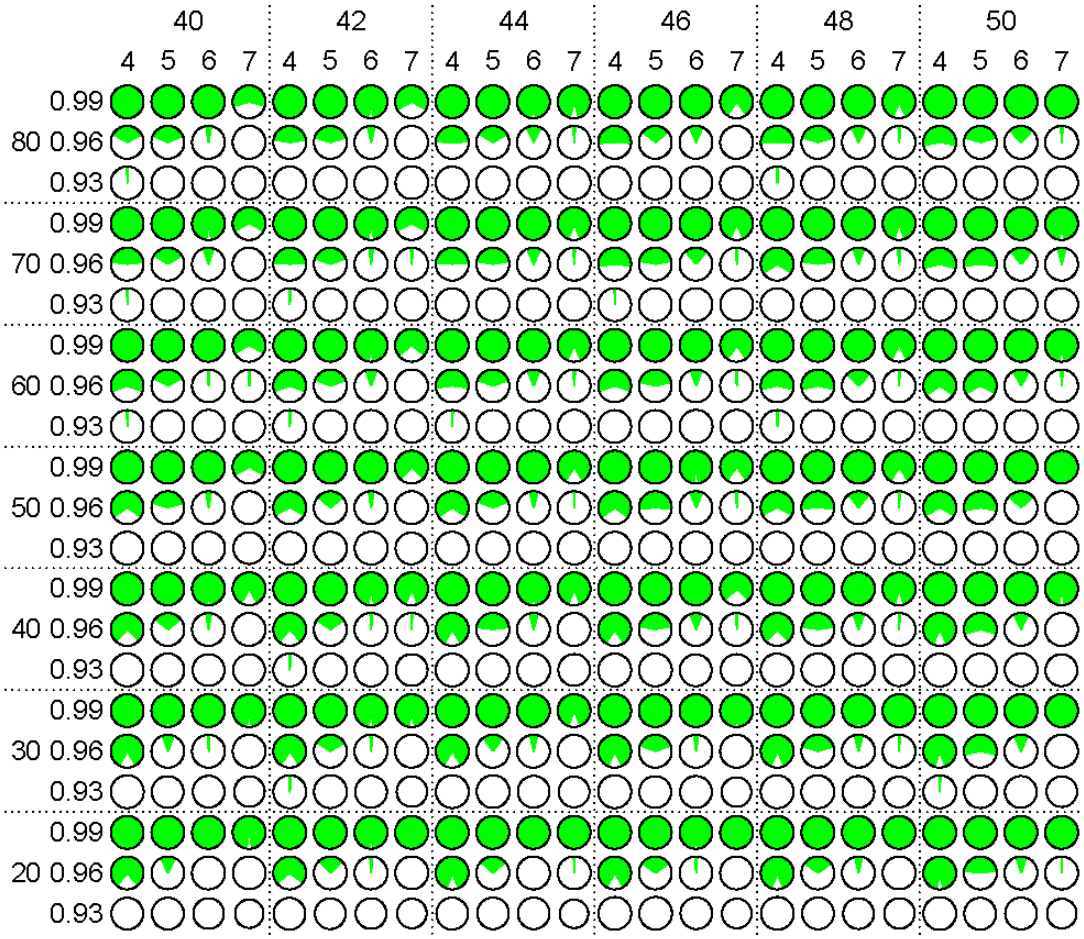

Fig. S8.

Proportions of simulated chimpanzee communities surviving 150 years with the presence of anthrax. Shown are results for different community sizes (outer y-axis), per capita annual survival probabilities (inner y-axis), maximum ages upper x-axis, inter birth intervals (lower x-axis). The area of the circles depicts the number of communities present at the beginning of the simulation (in a few simulations the simulated community did not survive the 50-years burn-in phase; maximum area corresponds to 100 simulations). The green area depicts the proportion of populations that survived for 150 time steps after the end of the burn-in phase.

Table S1.

Samples of the large scale data set.

Site	Country	Project	Flies	PCR-positive	Isolate	Bones	PCR-positive	Isolate
Dja Faunal Reserve	Cameroon	other	105	0		0	0	
Mount Cameroon	Cameroon	*Pan African Programme*	0	0		0	0	
Dzanga Sangha Protected Area	Central African Republic	other	200	0		0	0	
Djiroutou	Côte d´Ivoire	*Pan African Programme*	0	0		6	0	
Bili-Uele Forest	Democratic Republic Congo	*Pan African Programme*	0	0		25	0	
Loango National Park	Gabon	*Pan African Programme*	0	0		41	0	
Fouta Djallon	Guinea	*Pan African Programme*	105	0		9	0	
Boé	Guinea-Bisseau	*Pan African Programme*	0	0		7	0	
"nationwide"	Liberia	*Pan African Programme*	0	0		4	0	
Grebo National Forest	Liberia	*Pan African Programme*	105	2	2	8	1	0
Mount Nimba Strict Nature Reserve	Liberia	*Pan African Programme*	49	0		2	0	
Sapo National Park	Liberia	*Pan African Programme*	0	0		10	0	
Gashaka-Gumti National Park	Nigeria	*Pan African Programme*	105	0		7	0	
Kayan	Senegal	*Pan African Programme*	105	0		5	0	
Budongo Forest	Uganda	*Pan African Programme*	105	0		10	0	
Bwindi Impenetrable National Park	Uganda	*Pan African Programme*	105	0		2	0	
Ngogo in Kibale National Park	Uganda	*Pan African Programme*	105	0		0	0	
			Σ 1089			Σ 136		

Table S2.

Necropsies performed since 1996 in TNP. All samples were tested for *Bcbva* by PCR. PCR positive sample were cultured. For a subset of 15 cases diagnosis was confirmed with histo-pathology.

Species	Date of necropsy	PCR-positive	Culture	Histology[1)]	Longitude	Latitude	Other known cause of death
*Procolobus badius**	18.06.1996	x	0				
*Pan troglodytes verus**	23.08.1996	x	x				
*Cercopithecus petaurista**	09.10.1996						
*Procolobus badius**	18.10.1996						
*Procolobus badius**	30.10.1996						
*Procolobus badius**	12.06.1997						
*Cercocebus atys**	13.06.1997						
*Cercopithecus petaurista**	03.07.1997						
*Cercopithecus diana**	23.07.1997						
*Cercocebus atys**	25.07.1997	x	0				
*Cercopithecus petaurista**	31.07.1997						
*Procolobus badius**	14.08.1997						
*Procolobus badius**	26.08.1997						
*Procolobus verus**	07.09.1997						
*Procolobus badius**	22.10.1997						
*Procolobus badius**	26.10.1997						
*Cercopithecus diana**	04.02.1998						
*Cercocebus atys**	16.04.1998	x	x				
*Cercopithecus campbelli**	14.05.1998						
*Procolobus verus**	16.05.1998						
*Pan troglodytes verus**	16.08.1998	x	0				
*Procolobus badius**	26.08.1998						
*Procolobus badius**	09.10.1998						

*Pan troglodytes verus**	21.10.1998	x	x				
Pan troglodytes verus	15.12.1998						
*Procolobus badius**	09.02.1999						
*Procolobus badius**	11.02.1999						
Pan troglodytes verus	10.05.1999						Respiratory disease
Pan troglodytes verus	14.05.1999						Respiratory disease
Pan troglodytes verus	08.06.1999						Respiratory disease
*Cercopithecus campbelli**	21.06.1999						
*Colobus polykomos**	01.07.1999						
*Cercopithecus diana**	10.02.2000						
*Procolobus badius**	03.07.2000						
*Pan troglodytes verus**	26.09.2000	x	x				
undefined rodent	23.09.2001						
Pan troglodytes verus	15.10.2001	x	x		-7.350100	5.876500	
Pan troglodytes verus	15.10.2001	x	x		-7.350100	5.876500	
Pan troglodytes verus	15.10.2001	x	x		-7.350100	5.876500	
Pan troglodytes verus	11.12.2001						
Procolobus badius	05.02.2002	x	0	x			
Pan troglodytes verus	13.02.2002	x	x	x	-7.327980	5.847932	
Pan troglodytes verus	14.02.2002	x	x	x	-7.327980	5.847932	
Procolobus badius	14.02.2002				-7.327277	5.847704	
Pan troglodytes verus	01.03.2002	x	x		-7.348521	5.848744	
Pan troglodytes verus	01.03.2002	x	not cultured		-7.348521	5.848744	
Pan troglodytes verus	01.03.2002	x	not cultured		-7.348521	5.848744	
Procolobus badius	11.03.2002						
Pan troglodytes verus	11.06.2002	x	x		-7.298814	5.824126	
Procolobus badius	26.02.2004						

Pan troglodytes verus	10.03.2004					Respiratory disease
Pan troglodytes verus	10.03.2004					Respiratory disease
Procolobus badius	17.03.2004					
Pan troglodytes verus	19.03.2004					Respiratory disease
Cephalophus sp.	19.08.2004	x	x			
Pan troglodytes verus	02.03.2005					
Procolobus badius	13.04.2005					
Procolobus badius	22.05.2005					
undefined rodent	12.07.2005					
Pan troglodytes verus	13.07.2005	x	0			
Procolobus badius	19.07.2005					
Hexaprotodon liberiensis	20.07.2005					
Procolobus badius	13.08.2005					
Pan troglodytes verus	02.09.2005					
Perodicticus sp.	15.10.2005					
Soricidae	12.01.2006					
Perodicticus sp.	01.02.2006					
Pan troglodytes verus	01.02.2006					Respiratory disease
Pan troglodytes verus	01.02.2006					Respiratory disease
Pan troglodytes verus	09.02.2006					Respiratory disease
Procolobus badius	19.02.2006					
Cephalophus sp.	23.02.2006	x	x			
Potamocherus porcus	28.03.2006					
undefined rodent	23.04.2006					
Cercocebus atys	01.06.2006					
Procolobus badius	12.10.2006			-7.333232	5.879561	
Cephalophus sp.	14.10.2006			-7.347290	5.870976	
Cercopithecus diana	16.10.2006					

Procolobus badius	16.10.2006					
Soricidae	12.11.2006					
Cephalophus sp.	15.11.2006					
Cephalophus sp.	30.11.2006				-7.330046	5.839484
Cercocebus atys	06.12.2006					
Cephalophus sp.	10.12.2006				-7.321243	5.845181
Cephalophus sp.	07.01.2007					
Pan troglodytes verus	01.06.2007					
Cercocebus atys	05.11.2007					
Cephalophus sp.	09.12.2007					
Manidae	15.12.2007					
Pan troglodytes verus	07.04.2008	x	x	x	-7.278828	5.843579
Cephalophus sp.	10.04.2008				-7.284479	5.852811
Cephalophus sp.	20.04.2008	x	x		-7.341995	5.875391
Cercopithecus diana	20.04.2008					
Pan troglodytes verus	May 2008	x	x		-7.335046	5.821829
Cercocebus atys	29.09.2008					
Procolobus badius	26.10.2008				-7.329280	5.854302
Cercopithecus diana	01.12.2008					
Pan troglodytes verus	December 2008	x	x		-7.278780	5.826733
Perodicticus sp.	24.01.2009					
Manidae	05.02.2009					
Procolobus badius	30.03.2009					
Pan troglodytes verus	14.04.2009	x	x		-7.280321	5.845546
Pan troglodytes verus	17.04.2009	x	x	x	-7.275530	5.847195
Pan troglodytes verus	18.04.2009	x	x	x	-7.275466	5.847403
Pan troglodytes verus	20.04.2009	x	x	x	-7.281018	5.847899
Nandinia binotata	22.05.2009					

Nandinia binotata	01.07.2009						
Colobus polykomos	28.07.2009						
Pan troglodytes verus	06.08.2009				-7.314201	5.793401	
Pan troglodytes verus	01.12.2009						Respiratory disease
Pan troglodytes verus	02.12.2009						Respiratory disease
Pan troglodytes verus	06.12.2009						Respiratory disease
Pan troglodytes verus	06.12.2009						Respiratory disease
Pan troglodytes verus	07.12.2009				-7.329666	5.830711	Respiratory disease
Pan troglodytes verus	08.12.2009				-7.309060	5.824935	Respiratory disease
Pan troglodytes verus	11.12.2009						Respiratory disease
Colobus polykomos	11.03.2010						
Pan troglodytes verus	22.08.2011	x	x		-7.285815	5.838058	
Pan troglodytes verus	23.08.2011	x	x		-7.285806	5.838229	
Pan troglodytes verus	24.08.2011				-7.285706	5.838338	
Cercopithecus sp.	25.08.2011				-7.279716	5.821654	
Pan troglodytes verus	01.09.2011	x	not cultured		-7.324173	5.785753	
Cercopithecus diana	01.09.2011						
Cercopithecus campbelli	19.12.2011	x	x		-7.349070	5.867626	
Cercopithecus petaurista	02.01.2012	x	x		-7.318055	5.827177	
Pan troglodytes verus	15.01.2012	x	x	x	-7.279277	5.826527	
Cercopithecus diana	21.01.2012	x	x	x	-7.276743	5.843139	
Cephalophus sp.	25.01.2012	x	x		-7.274059	5.826447	
Cephalophus sp.	01.02.2012	x	x		-7.333262	5.863474	
undefined rodent	15.02.2012				-7.331164	5.864399	
undefined rodent	20.02.2012				-7.325428	5.855955	
Cephalophus sp.	14.03.2012	x	x	x	-7.304462	5.822751	
Cercocebus atys	14.03.2012				-7.301514	5.836306	
Cephalophus sp.	15.03.2012				-7.321115	5.815449	

Cephalophus sp.	18.03.2012				-7.306623	5.806508
Cercopithecus diana	06.04.2012	x	x		-7.320085	5.815754
Procolobus badius	05.06.2012	x	x	x	-7.338552	5.869983
Cercocebus atys	08.09.2012				-7.325691	5.815960
Colobus sp.	20.09.2012	x	x		-7.312917	5.803317
Herpestidae	29.09.2012	x	x		-7.315241	5.811010
Cercopithecus diana	30.10.2012	x	x	x	-7.319489	5.833828
Cercocebus atys	13.02.2013				-7.320745	5.839646
Cercocebus atys	01.03.2013	x	x	x	-7.327100	5.837196
Cercocebus atys	04.03.2013				-7.320618	5.840007
Cercocebus atys	05.03.2013	x	x		-7.317003	5.822825
Pan troglodytes verus	12.04.2013	x	x	x	-7.359682	5.870090
Pan troglodytes verus	14.04.2013	x	x	x	-7.347832	5.873935
Cercocebus atys	20.04.2013	x	x		-7.324103	5.827982
Cercocebus atys	12.05.2013	x	x		-7.322338	5.829451
Cephalophus sp.	30.05.2013	x	x		-7.329750	5.832719
Cercocebus atys	07.11.2013				-7.332019	5.871347
Cephalophus sp.	02.12.2013	x	x		-7.326472	5.811450
Cephalophus sp.	29.12.2013	x	x		-7.288041	5.839493
Procolobus badius	06.01.2014				-7.310206	5.834261
Pan troglodytes verus	08.01.2014	x	not cultured		-7.282501	5.823462
Colobus polykomos	09.01.2014	x	not cultured		-7.310540	5.822172
Manidae	12.01.2014				-7.331242	5.862474
Pan troglodytes verus	15.01.2014	x	not cultured		-7.307575	5.823502
Cephalophus sp.	17.02.2014	x	x		-7.318345	5.827097
Cephalophus sp.	21.02.2014	x	x		-7.330257	5.835207

Cercocebus atys	21.02.2014	x	not cultured	-7.330261	5.834050
Cercopithecus sp.	01.03.2014			-7.301679	5.808157
Potamocherus porcus	01.03.2014			-7.259130	5.798352
Cephalophus sp.	04.03.2014	x	x	-7.304491	5.816105
Cercopithecus diana	13.03.2014			-7.347766	5.865760
Cephalophus sp.	20.03.2014	x	x	-7.338608	5.869250
Cephalophus sp.	11.04.2014	x	x	-7.338011	5.869692
Cephalophus sp.	17.04.2014	x	x	-7.338120	5.872531
Cercocebus atys	19.04.2014			-7.348171	5.878638
Herpestidae	27.04.2014	x	not cultured	-7.275368	5.820312
Cephalophus sp.	28.04.2014	x	x	-7.356042	5.867231
Pan troglodytes verus	02.05.2014			-7.293111	5.796023
Hystricidae	06.05.2014	x	not cultured	-7.335862	5.872913
Cercocebus atys	10.05.2014	x	not cultured	-7.336855	5.869896
Cephalophus sp.	13.05.2014	x	x	-7.324640	5.823444
undefined rodent	21.05.2014			-7.338453	5.869829
undefined rodent	21.05.2014			-7.339480	5.867562
undefined rodent	21.05.2014			-7.339480	5.867562
Cephalophus sp.	03.06.2014	x	x	-7.358087	5.868892
Cephalophus sp.	04.06.2014	x	x	-7.335895	5.870653
Cephalophus sp.	08.06.2014	x	x	-7.336863	5.870204
Colobus polykomos	11.06.2014			-7.290994	5.824645
Cercocebus atys	29.06.2014	x	not cultured	-7.325633	5.835628
Pan troglodytes verus	22.07.2014	x	not cultured	-7.309839	5.824033

Pan troglodytes verus	29.10.2014	x	not cultured	-7.259531	5.805741
Cephalophus sp.	19.12.2014	x	not cultured	-7.289299	5.838710
Cephalophus sp.	24.12.2014			-7.304824	5.825347
Cephalophus sp.	29.12.2014	x	not cultured	-7.308611	5.824093
Cephalophus sp.	01.01.2015	x	not cultured	-7.276469	5.823480
Cercocebus atys	07.01.2015	x	not cultured	-7.298985	5.854464
Cephalophus sp.	07.01.2015			-7.307416	5.822154
Cephalophus sp.	08.01.2015	x	not cultured	-7.337016	5.879464
Cephalophus sp.	14.01.2015			-7.314343	5.818693
Potamocherus porcus	16.01.2015			-7.286380	5.836432
Cephalophus sp.	23.01.2015	x	not cultured	-7.307164	5.803652
Genetta sp.	05.02.2015			-7.332095	5.873363
Cercocebus atys	07.02.2015	x	not cultured	-7.331846	5.874990
Anomaluridae	09.02.2015			-7.324983	5.886978
Procolobus badius	12.03.2015			-7.286110	5.824395
Cercocebus atys	14.03.2015			-7.329481	5.868554
Anomaluridae	15.03.2015			-7.335016	5.868607
Cephalophus sp.	23.03.2015			-7.307957	5.786139
Cephalophus sp.	15.04.2015	x	not cultured	-7.316495	5.820553
	20.04.2015			-7.298468	5.852265
Σ 204		81			

* samples provided by World Health Organization
[1)] diagnosis confirmed in histo- pathology

Table S3.

Primers used in this study.

Primer name	Sequence (5´-3´)	Quencher	T_a (°C)	Fragment length (bp)	Reference
Bcbva detection					
PA F	CggATCAAgTATATgggAATATAgCAA		60	204	Ellerbrok et al. 2002
PA R	CCggTTTAgTCgTTTCTAATggAT		60		Ellerbrok et al. 2002
PA TM	TEX-CTCgAACTggAgTgAAgTgTTACCgCAAAT	BBQ			Ellerbrok et al. 2002
capB F	gggAAAACAACTggTACATCTgC		60	144	Antonation et al. 2016
capB R	AAgTgCTTCTgCTTCTAAATCAgC		60		Antonation et al. 2016
capB TM	6FAM-CCTCTTTAACTACCCTgCgTTgCTCACCg	TMR			Antonation et al. 2016
IslandIV-for	ggAgATATTAACAAgAgATggATTggA		60	140	Antonation et al. 2016
IslandIV-rev	CAgTAggCTTgTCTgCTCTAATAAAATT		60		Antonation et al. 2016
IslandIV-TM	6FAM-ACATgCCAgCgTTTTTTgCCTCTACACA	BHQ1			Antonation et al. 2016
Fly meal analysis					
16Smam1	CggTTggggTgACCTCggA		64	130	Taylor 1996
16Smam2	gCTgTTATCCCTAgggTAACT		64		Taylor 1996
16Smam_blkhum3	CggTTggggCgACCTCggAgCAgAACCC--spacerC3				Boessenkool et al. 2012
16Smam_blksus1	CggTTggggTgACCTCggAgTACAAAAAAC--spacerC3				Calvignac-Spencer et al. 2013

Table S4.

Overview of all flies analyzed in this study comprising flies from TNP as well as flies from the large scale data set. PCR results and culture outcome are included.

Small scale	Project	Site	Country	# analyzed	PCR positive	Isolates	Year
	random flies	Taï National Park	Côte d´Ivoire	726[1)]	62	24[2)]	2008/2009/2012/2013
	snapshot flies	Taï National Park	Côte d´Ivoire	908	18	18[4)] out of 18	2014
Large scale	**Project**	**Site**	**Country**	**# analyzed**	**PCR positive**	**Isolates**	**Year**
	known *Bcbva* site	Dja Faunal Reserve	Cameroon	105	0	0	2014
	known *Bcbva* site	Dzanga Sangha Protected Area	Central African Republic	200	0	0	2013
	Pan African Programme	Fouta Djallon	Guinea	105	0	0	2013
	Pan African Programme	Grebo National Forest	Liberia	105	2	2	2013
	Pan African Programme	Mount Nimba Nature Reserve	Liberia	49[3)]	0	0	2013/2014
	Pan African Programme	Gashaka-Gumti National Park	Nigeria	105	0	0	2013
	Pan African Programme	Kayan	Senegal	105	0	0	2013
	Pan African Programme	Budongo Forest	Uganda	105	0	0	2012/2013
	Pan African Programme	Bwindi Impenetrable National Park	Uganda	105	0	0	2012/2013
	Pan African Programme	Ngogo in Kibale National Park	Uganda	105	0	0	2013
				Σ 2723			

[1)] Including 103 canopy flies.
[2)] Culture was only attempted for a subset of 32 flies with high *pag* copy numbers. Additionally we retrieved 21 isolates from an on-site study testing culture as a diagnostic tool (without preceding PCR).

[3)] Not more flies available.

[4)] From one fly two isolates were obtained, for one fly culture was not possible.

Table S5.

Results fly meal analysis. Mammalian genera detected in flies and in *Bcbva* positive carcasses.

Order	Genus	number of mammal positive and *Bcbva* positive flies (n=28)	number of mammal positive and *Bcbva* negative flies (n=29)	number of *Bcbva* positive necropsies (n=81)
Artiodactyla	*Cephalophus*	23	21	26
Artiodactyla	*Hexaprotodon*	0	1	0
Artiodactyla	*Potamochoerus*	2	2	0
Carnivora	*Civettictis*	1	0	0
Carnivora	*Crossarchus*	2	0	0
Carnivora	*Genetta*	2	0	0
Carnivora	*Herpestes*	2	0	2
Carnivora	*Nandinia*	0	1	0
Chiroptera	*Hypsignathus*	2	3	0
Chiroptera	*Myonycteris*	4	2	0
Hyracoidea	*Dendrohyrax*	1	0	0
Primates	*Cercocebus*	3	5	11
Primates	*Cercopithecus*	12	17	5
Primates	*Colobus*	6	4	2
Primates	*Pan*	0*	0*	31
Primates	*Procolobus*	7	5	3
Rodentia	*Atherurus*	4	1	0
Rodentia	*Cricetomys*	2	1	0
Rodentia	*Hylomyscus*	10	11	0
Rodentia	*Hystrix*	0	0	1
Rodentia	*Praomys*	0	2	0
Rodentia	*Protoxerus*	1	3	0

*most probably due to use of human blocking primer

Table S6.

Fly species composition among *Bcbva* sites. Species were identified using the Generalized Mixed Yule Coalescent Model. Forty-five species were identified of which 24 were assignable at family level.

GMYC entity	family	TNP (n=325)	DJR (n=101)	DSPA (n=106)
7	Calliphoridae		1 %	8 %
12	Calliphoridae	17 %	58 %	10 %
13	Calliphoridae			2 %
14	Calliphoridae	2 %	10 %	1 %
15	Calliphoridae	67 %	7 %	22 %
16	Calliphoridae	2 %	2 %	3 %
27	Calliphoridae			1 %
28	Calliphoridae			1 %
37	Calliphoridae		2%	
39	Calliphoridae		1%	
40	Calliphoridae		1%	
8	Sarcophagidae	6 %	4 %	4 %
9	Sarcophagidae			12 %
10	Sarcophagidae	1 %	1 %	
11	Sarcophagidae	1 %		
30	Sarcophagidae			1 %
42	Glossinidae		1%	
44	Glossinidae		1%	
34	Staphylinidae			1 %
26	Syrphidae			1 %
31	Syrphidae			1 %
32	Syrphidae			1 %
25	Tachinidae			1 %
41	Tachinidae		1 %	
	not assignable	4 %	11 %	30 %

Table S7.

Mammal bones collected in TNP since 1989 including results from PCR and culture.

Year	Species	PCR-positive	Culture
1989	*Pan troglodytes verus*		
1991	*Pan troglodytes verus*	x	0
1992	*Pan troglodytes verus*	x	0
1992	*Pan troglodytes verus*		
1992/93	*Pan troglodytes verus*		
1992/93	*Pan troglodytes verus*		
1992/93	*Pan troglodytes verus*		
1993	*Pan troglodytes verus*		
1994	*Cercocebus atys*		
1994	*Colobus polykomos*		
1994	*Pan troglodytes verus*	x	0
1994	*Pan troglodytes verus*	x	0
1994	*Pan troglodytes verus*	x	0
1994	*Pan troglodytes verus*		
1994	*Pan troglodytes verus*		
1994	*Pan troglodytes verus*		
1994	*Pan troglodytes verus*		
1994	*Procolobus badius*		
1994	*Procolobus badius*		
1996	*Pan troglodytes verus*	x	x
1996	*Pan troglodytes verus*		
1997	*Procolobus badius*		
1998	*Pan troglodytes verus*	x	0
1999	*Pan troglodytes verus*	x	0
1999	*Pan troglodytes verus*		
1999	*Pan troglodytes verus*		
2000	*Cercocebus atys*		
2000	*Cercopithecus campbelli*	x	0
2000	*Cercopithecus diana*		
2000	*Pan troglodytes verus*	x	0
2001	*Pan troglodytes verus*	x	x
2001	*Procolobus badius*		
2002	*Cercopithecus diana*		
2002	*Pan troglodytes verus*	x	0
2002	*Pan troglodytes verus*		
2002	*Procolobus badius*	x	0
2002	*Procolobus badius*		

2002	*Procolobus badius*		
2003	*Colobus polykomos*	x	x
2003	*Procolobus badius*		
2004	*Cercocebus atys*	x	x
2004	*Pan troglodytes verus*	x	0
2004	*Pan troglodytes verus*		
2004	*Pan troglodytes verus*	x	1
2004	*Pan troglodytes verus*		
2004	*Pan troglodytes verus*	x	2
2004	*Procolobus badius*		
2006	*Cercocebus atys*		
2007	*Cercopithecus diana*		
2009	*Pan troglodytes verus*	x	isolate available from tissue
2009	*Pan troglodytes verus*	x	isolate available from tissue
2009	*Pan troglodytes verus*		
2009	*Pan troglodytes verus*		
2010	*Procolobus badius*		
2013	*Cephalophus* sp.	x	x
2013	*Cephalophus* sp.	x	x
2013	*Cephalophus* sp.	x	x
2013	*Cephalophus* sp.		
2013	*Cephalophus* sp.		
2013	*Cercocebus atys*		
2013	*Colobus* sp.		
2013	*Procolobus badius*		
2013	*Procolobus badius*		
2014	*Hippopotamus* sp.		
2014	*Pan troglodytes verus*		
unknown	*Cephalophus* sp.	x	0
unknown	*Cercocebus atys*		
unknown	*Cercopithecus* sp.		
unknown	*Colobus* sp.		
unknown	*Loxodonta africana*		
unknown	*Loxodonta africana*		
unknown	*Pan troglodytes verus*	x	0
unknown	*Pan troglodytes verus*	x	0
unknown	*Pan troglodytes verus*		
unknown	*Potamochoerus porcus*		
∑	75	26	7

Table S8.

All *Bcbva* isolates sequenced for this study. All isolates were sequenced on the Illumina HiSeq 1500 Platform in rapid run mode (except for 359-9 and 509-6: Illumina MiSeq).

Isolate ID	Sampling Date	Species	Material	Isolates	Chemistry	Average Coverage
071-2A1	May 2008	*Pan troglodytes verus*	skin/hair	1	v1 Rapid Run 2x150bp	319
092- 1A	18.04.2009	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	376
092-2	17.04.2009	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	311
092-3	20.04.2009	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	328
092-4	14.04.2009	*Pan troglodytes verus*	bone with meat	1	v1 Rapid Run 2x150bp	349
140-1	Dec 2008	*Pan troglodytes verus*	skin/maggots	1	v2 Rapid Run 2x250bp	108
178-2-1	20.04.2008	*Cephalophus* sp.	spleen	1	v1 Rapid Run 2x150bp	296
178-6-1	23.02.2006	*Cephalophus* sp.	bone with meat	1	v1 Rapid Run 2x150bp	352
178-7-1	19.08.2004	*Cephalophus* sp.	lung	1	v2 Rapid Run 2x250bp	90
185-2b	01.03.2002	*Pan troglodytes verus*	connective tissue	1	v1 Rapid Run 2x150bp	283
353-1	22.04.2009	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	291
353-2	16.05.2009	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	253
353-3	27.02.2012	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	343
359-6	15.03.2012	Calliphoridae	fly mush after DNA extraction	1	v3 2x 300bp	18
359-9	15.03.2012	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	394
359-12	01.07.2012	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	391
359-15	01.07.2012	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	303
359-16	15.03.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	105
381-1	12.04.2013	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	281
381-3	14.04.2013	*Pan troglodytes verus*	liver	1	v1 Rapid Run 2x150bp	257
386-1	18.09.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	58
386-2	18.09.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	110
386-3	10.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	70
386-5	10.10.2012	Calliphoridae	fly mush after	1	v2 Rapid Run	103

			DNA extraction		2x250bp	
386-6	10.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	94
386-7	10.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	104
386-8	10.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	104
386-9	10.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	83
401-1	20.04.2013	*Cercocebus atys*	spleen	1	v1 Rapid Run 2x150bp	324
402-1	25.04.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	79
403-1	25.04.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	28
403-2	25.04.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	54
403-5	15.05.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	34
404-1	12.05.2013	*Cercocebus atys*	liver	1	v2 Rapid Run 2x250bp	49
405-1	16.05.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	113
405-2	17.05.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	72
406-1	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	79
407-1	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	56
407-2	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	108
407-3	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	63
407-4	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	29
407-5	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	99
407-6	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	19
407-7	15.06.2013	Calliphoridae	fly direct	1	v1 Rapid Run 2x150bp	335
407-8	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	46
407-9	15.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	40
407-10	16.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	41
407-11	16.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	63
407-12	16.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	35
407-13	16.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	44
407-14	16.06.2013	Calliphoridae	fly direct	1	v2 Rapid Run 2x250bp	100

408-1	30.05.2013	*Cephalophus* sp.	fat	1	v1 Rapid Run 2x150bp	326
423-1	22.08.2011	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	278
424-1	23.08.2011	*Pan troglodytes verus*	liver	1	v1 Rapid Run 2x150bp	317
425-1	21.01.2012	*Cercopithecus diana*	spleen	1	v1 Rapid Run 2x150bp	254
426-1	29.09.2012	Herpestidae	liver	1	v1 Rapid Run 2x150bp	307
427-1	31.12.2008	Calliphoridae	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	360
434-1	13.02.2002	*Pan troglodytes verus*	spleen	1	v2 Rapid Run 2x250bp	36
435-1	15.01.2012	*Pan troglodytes verus*	spleen	1	v2 Rapid Run 2x250bp	66
436-1	01.02.2012	*Cephalophus* sp.	lung	1	v2 Rapid Run 2x250bp	115
437-1	05.06.2012	*Procolobus badius*	spleen	1	v2 Rapid Run 2x250bp	115
438-1	30.10.2012	*Cercopithecus diana*	spleen	1	v2 Rapid Run 2x250bp	110
440-1	01.03.2013	*Cercocebus atys*	spleen	1	v2 Rapid Run 2x250bp	67
441-1	19.12.2011	*Cercopithecus campbelli*	muscle	1	v2 Rapid Run 2x250bp	132
442-1	14.03.2012	*Cephalophus* sp.	spleen	1	v2 Rapid Run 2x250bp	145
443-1	05.03.2013	*Cercocebus atys*	spleen	1	v2 Rapid Run 2x250bp	50
444	11.10.2012	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	31
459	16.04.1998	*Cercocebus atys*	liver	1	v2 Rapid Run 2x250bp	88
460&529	23.08.1996	*Pan troglodytes verus*	liver & bones	2	v2 Rapid Run 2x250bp	116/33
461	21.10.1998	*Pan troglodytes verus*	muscle	1	v2 Rapid Run 2x250bp	94
462	26.09.2000	*Pan troglodytes verus*	lung	1	v2 Rapid Run 2x250bp	87
503-1	08.12.2001	*Pan troglodytes verus*	bone	1	v2 Rapid Run 2x250bp	95
505 (1-4)	01.02.2013	Calliphoridae[1)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	68/53/48/108
506 (1-4)	02.02.2013	Calliphoridae[1)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	82/56/131/99
507 (1-6)	06.04.2012	*Cercopithecus diana*	liver & spleen	6	v2 Rapid Run 2x250bp	135/53/52/107/4 61
508 (1-6)	02.12.2013	*Cephalophus* sp.	liver & spleen	6	v2 Rapid Run 2x250bp	50/51/67/36/89/5
509 (1-6)	29.12.2013	*Cephalophus* sp.	liver & spleen	6	v2 Rapid Run 2x250bp/ v3 2x300bp (509-6)	174/113/62/63/8 17
530	2003	*Colobus polycomos*	bone	1	v2 Rapid Run 2x250bp	27
531	2004	*Cercocebus atys*	bone	1	v2 Rapid Run	29

					2x250bp	
552	October 2013	*Cephalophus* sp.	bone	1	v2 Rapid Run 2x250bp	28
553	27.09.2013	*Cephalophus* sp.	bone	1	v2 Rapid Run 2x250bp	31
554	05.11.2013	*Cephalophus* sp.	bone	1	v2 Rapid Run 2x250bp	39
563	15.10.2001	*Pan troglodytes verus*	muscle	1	v2 Rapid Run 2x250bp	30
564 (1-6)	15.10.2001	*Pan troglodytes verus*	muscle	6	v2 Rapid Run 2x250bp	31/31/19/24/19/2
565 (1-6)	02.01.2012	*Cercopithecus petaurista*	muscle	6	v2 Rapid Run 2x250bp	26/31/25/32/34/3
566	25.01.2012	*Cephalophus* sp.	muscle	1	v2 Rapid Run 2x250bp	32
567	20.09.2012	Colobus sp.	maggot	1	v2 Rapid Run 2x250bp	26
579	26.04.2009	Calliphoridae	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	18
580 (1-4)	26.05.2013	Calliphoridae[2)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	16/25/18/26
581	26.05.2013	Calliphoridae[2)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	20
582	26.05.2013	Calliphoridae[2)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	24
583	25.05.2013	Calliphoridae[2)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	35
584	25.05.2013	Calliphoridae[2)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	27
585	02.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	23
586	01.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	49
587	01.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	39
588 (1-4)	01.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	21/24/26/26
592	07.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	23
593 (1-4)	07.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	26/27/53/65
594	07.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	39
595	30.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	95
596	07.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	31
597 (1-4)	27.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	19/20/23/24/
598	09.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	19
599	31.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	23
600 (1-2)	30.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	2	v2 Rapid Run 2x250bp	24/ 44

601	09.06.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	19
602	17.02.2014	*Cephalophus* sp.	corpse swab	1	v2 Rapid Run 2x250bp	28
603	21.02.2014	*Cephalophus* sp.	spleen	1	v2 Rapid Run 2x250bp	22
604	04.03.2014	*Cephalophus* sp.	spleen	1	v2 Rapid Run 2x250bp	24
605	20.03.2014	*Cephalophus* sp.	spleen	1	v2 Rapid Run 2x250bp	28
606	11.04.2014	*Cephalophus* sp.	skin	1	v2 Rapid Run 2x250bp	28
607 (1-6)	17.04.2014	*Cephalophus* sp.	spleen	6	v2 Rapid Run 2x250bp	5/45/50/26/23/1
608	28.04.2014	*Cephalophus* sp.	spleen	1	v2 Rapid Run 2x250bp	29
609	13.05.2014	*Cephalophus* sp.	liver	1	v2 Rapid Run 2x250bp	28
610	03.06.2014	*Cephalophus* sp.	liver	1	v2 Rapid Run 2x250bp	22
611	04.06.2014	*Cephalophus* sp.	liver	1	v2 Rapid Run 2x250bp	24
612	08.06.2014	*Cephalophus* sp.	lung	1	v2 Rapid Run 2x250bp	26
613	30.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	42
614 (1-4)	25.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	69/26/22/29
615	25.05.2014	Calliphoridae[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	22
Ca	11.06.2002	*Pan troglodytes verus*	lung	1	v1 Rapid Run 2x150bp	349
CIRef	14.02.2002	*Pan troglodytes verus*	spleen	1	v2 Rapid Run 2x250bp	263
Fa7	15.10.2001	*Pan troglodytes verus*	spleen	1	v1 Rapid Run 2x150bp	321
HaGaCI8	07.04.2008	*Pan troglodytes verus*	spleen/liver	1	v1 Rapid Run 2x150bp	264

[1)] Samples from Grebo, Liberia

[2)] Flies caught in canopy

[3)] Flies caught during snapshot

Table S9.

Comparison of SNPs between different isolates retrieved from the same host. Differences on chromosomal and plasmid level are shown. For each host one isolate was included in further phylogenetic analyses (marked in bold), selection criteria being high average coverage and low number of autapomorphic SNPs.

Host	Isolate ID	Material	Average coverage	SNP Difference (Chromosome/pXC	
				460	529

Pan troglodytes verus 23.08.1996	**460**	**liver**	**116**	**-**	**0**		
	529	bone	33	0	-		
				505-1	505-2	505-3	505-4
Calliphoridae 01.02.2013	505-1	fly mush	68	-	1/0/0	0	0
	505-2	fly mush	53	1/0/0	-	1/0/0	1/0/0
	505-3	fly mush	48	0	1/0/0	-	0
	505-4	**fly mush**	**108**	**0**	**1/0/0**	**0**	**-**
				506-1	506-2	506-3	506-4
Calliphoridae 02.02.2013	506-1	fly mush	82	-	1/0/0	1/0/0	0
	506-2	fly mush	56	1/0/0	-	2/0/0	1/0/0
	506-3	fly mush	131	1/0/0	2/0/0	-	1/0/0
	506-4	**fly mush**	**99**	**0**	**1/0/0**	**1/0/0**	**-**
				507-1	507-2	507-3	507-4
Cercopithecus diana 06.04.2012	**507-1**	**liver**	**135**	**-**	**0**	**0**	**0**
	507-2	liver	53	0	-	0	0
	507-3	liver	52	0	0	-	0
	507-4	spleen	107	0	0	0	-
	507-5	spleen	40	0	0	0	0
	507-6	spleen	61	0	0	0	0
				508-1	508-2	508-3	508-4
Cephalophus sp. 02.12.2013	508-1	liver	50	-	0	0	0
	508-2	liver	51	0	-	0	0
	508-3	liver	67	0	0	-	0
	508-4	spleen	36	0	0	0	-
	508-5	**spleen**	**89**	**0**	**0**	**0**	**0**
	508-6	spleen	53	1/0/0	1/0/0	1/0/0	1/0/0
				509-1	509-2	509-3	509-4
Cephalophus sp. 29.12.2013	**509-1**	**liver**	**174**	**-**	**0**	**1/0/0**	**0**
	509-2	liver	113	0	-	1/0/0	0
	509-3	liver	62	1/0/0	1/0/0	-	1/0/0
	509-4	spleen	63	0	0	1/0/0	-
	509-5	spleen	86	0	0	1/0/0	0
	509-6	spleen	17	0	0	1/0/0	0
				564-1	564-2	564-3	564-4
Pan troglodytes verus 15.10.2001	**564-1**	**muscle**	**31**	**-**	**0**	**0**	**0**
	564-2	muscle	31	0	-	0	0
	564-3	muscle	19	0	0	-	0
	564-4	muscle	24	0	0	0	-
	564-5	muscle	19	0	0	0	0
	564-6	muscle	25	0	0	0	0
				565-1	565-2	565-3	565-4
Cercopithecus petaurista 02.01.2012	565-1	muscle	26	-	0	1/0/0	0
	565-2	muscle	31	0	-	1/0/0	0
	565-3	muscle	25	1/0/0	1/0/0	-	1/0/0
	565-4	**muscle**	**32**	**0**	**0**	**1/0/0**	**-**
	565-5	muscle	34	1/0/0	1/0/0	2/0/0	1/0/0

	565-6	muscle	31	1/0/0	1/0/0	2/0/0	1/0/0
				580-1	580-2	580-3	580-4
Calliphoridae 26.05.2013	580-1	fly mush	16	-	0	1/0/0	0/0/1
	580-2	fly mush	25	0	-	1/0/0	0/0/1
	580-3	fly mush	18	1/0/0	1/0/0	-	1/0/1
	580-4	**fly mush**	**26**	**0/0/1**	**0/0/1**	**1/0/1**	**-**
				588-1	588-2	588-3	588-4
Calliphoridae 01.06.2014	588-1	fly mush	21	-	0	2/0/0	0
	588-2	fly mush	24	0	-	2/0/0	0
	588-3	fly mush	26	2/0/0	2/0/0	-	2/0/0
	588-4	**fly mush**	**26**	**0**	**0**	**2/0/0**	**-**
				593-1	593-2	593-3	593-4
Calliphoridae 07.06.2014	593-1	fly mush	26	-	0	0	0
	593-2	fly mush	27	0	-	0	0
	593-3	fly mush	53	0	0	-	0
	593-4	**fly mush**	**65**	**0**	**0**	**0**	**-**
				597-1	597-2	597-3	597-4
Calliphoridae 27.05.2014	597-1	fly mush	19	-	1/0/0	1/0/0	1/0/0
	597-2	fly mush	20	1/0/0	-	0	0
	597-3	fly mush	23	1/0/0	0	-	0
	597-4	**fly mush**	**24**	**1/0/0**	**0**	**0**	**-**
				600-1	600-2		
Calliphoridae 30.05.2014	**600-1**	**fly mush**	**24**	**-**	**42/0/2**		
	600-2	**fly mush**	**44**	**42/0/2**	**-**		
				607-1	607-2	607-3	607-4
Cephalophus sp. 17.04.2014	607-1	spleen	5	-	0	0	0
	607-2	spleen	45	0	-	0	0
	607-3	**spleen**	**50**	**0**	**0**	**-**	**0**
	607-4	spleen	26	0	0	0	-
	607-5	spleen	23	0	0	0	0
	607-6	spleen	17	0	0	0	0
				614-1	614-2	614-3	614-4
Calliphoridae 25.05.2014	**614-1**	**fly mush**	**69**	**-**	**0**	**0**	**0**
	614-2	fly mush	26	0	-	0	0
	614-3	fly mush	22	0	0	-	0
	614-4	fly mush	29	0	0	0	-

Table S10.

Estimates of GLMM for testing the effect of season and the amount of mammalian DNA within the fly on anthrax positivity of flies.

Predictor	Estimate	SE	95% CI $_{lower}$	95% CI $_{upper}$	χ^2	P-value
Intercept	-2,37	0,27	-2.98	-1.90	[1]	[1]
mammal DNA	0,36	0,1	0.12	0.56	9.44	0,002
sin(date)	0,59	0,23	0.11	1.05	6.9 [2]	0.032[2]
cos(date)	0,45	0,45	-0.32	1.26		

[1] not shown because of not having a meaningful interpretation

[2] tested was the combined effect of sin(date) and cos(date)