Aus dem Robert Koch-Institut
eingereicht über das Institut für Mikrobiologie und Tierseuchen des
Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Epidemiology and Ecology of Bacillus cereus biovar anthracis in Taï National Park, Côte d'Ivoire

Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin

vorgelegt von
Fee Zimmermann
Tierärztin aus Siegburg

Berlin 2018
Journal-Nr.: 4033

mbvberlin | mensch und buch verlag

Aus dem Robert Koch-Institut
eingereicht über das Institut für Mikrobiologie und Tierseuchen des
Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Epidemiology and Ecology of *Bacillus cereus* biovar *anthracis* in Taï National Park, Côte d'Ivoire

---

Inaugural-Dissertation
zur Erlangung des Grades eines
Doktors der Veterinärmedizin
an der
Freien Universität Berlin

vorgelegt von
Fee Zimmermann
Tierärztin aus Siegburg

Berlin 2018

Journal-Nr.: 4033

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Jürgen Zentek
Erster Gutachter:	Prof. Dr. Lothar H. Wieler
Zweiter Gutachter:	Prof. Dr. Peter-Henning Clausen
Dritter Gutachter:	PD Dr. Wolfgang Beyer

Deskriptoren (nach CAB-Thesaurus):
bacillus cereus anthracis, phylogeny, serology, diptera, cadaver (MeSH), monte carlo method, real time pcr, cote d'ivoire

Tag der Promotion: 12.04.2018

Bibliografische Information der *Deutschen Nationalbibliothek*
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <https://dnb.de> abrufbar.

ISBN: 978-3-86387-946-4
Zugl.: Berlin, Freie Univ., Diss., 2018
Dissertation, Freie Universität Berlin
D188

 Choriner Str. 85 - 10119 Berlin
verlag@menschundbuch.de – www.menschundbuch.de

Ça va aller - on est ensemble

Teile dieser Dissertation sind veröffentlicht in:

Zimmermann F., Köhler S.M., Nowak K., Dupke S., Barduhn A., Düx A., Lang A., De Nys H.M., Gogarten J.F., Grunow R., Couacy-Hymann E., Wittig R.M., Klee S.R,, Leendertz F.H. (2017). Low antibody prevalence against *Bacillus cereus* biovar *anthracis* in Taï National Park, Côte d'Ivoire, indicates high rate of lethal infections in wildlife. *PLoS Neglected Tropical Diseases,* Sep 21;11(9):e0005960

Hoffmann C. & Zimmermann F.[*], Biek R., Kuehl H., Nowak K., Mundry R., Agbor A., Angedakin S., Arandjelovic M., Blankenburg A., Brazolla G., Corogenes K., Couacy-Hymann E., Deschner T., Dieguez P., Dierks K., Düx A., Dupke S., Eshuis H., Formenty P., Ginath Yuh Y., Goedmakers A., Gogarten J.F., Granjon A., McGraw S., Grunow R., Hart J., Jones S., Junker J., Kiang J., Langergraber K., Lapuente J., Lee K., Leendertz S.A., Léguillon F., Leinert V., Löhrich T., Marrocoli S., Mätz-Rensing K., Meier A., Merkel K., Metzger S., Murai M., De Nys H.M., Sachse A., Schenk S., van Schijndel J., Thiesen U., Ton E., Wu D., Wieler L.H., Boesch C., Klee S.R., Wittig, R.M., Calvignac-Spencer S., Leendertz F.H. (2017). Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. *Nature,* 548(7665),82-86.

[*] Equal contributions

Table of contents

Figures

Supplementary Figure S1: Maximum Likelihood tree based on chromosomal sequences of *Bcbva* isolates from TNP (Côte d'Ivoire) and Grebo (Liberia) - detailed version with sampling dates.

Tables

Abbreviations

ABAS	Ausschuss für Biologische Arbeitsstoffe
AFLP	Amplified fragment length polymorphism
ATP	Adenosine triphosphate
B. anthracis	*Bacillus anthracis*
Bcbva	*Bacillus cereus* biovar *anthracis*
B. cereus	*Bacillus cereus*
BEAST	Bayesian evolutionary analysis sampling trees
BF	Bayes factor
BMCMC	Bayesian markov chain monte carlo
B. mycoides	*Bacillus mycoides*
bp	Base pair
B. pseudomycoides	*Bacillus pseudomycoides*
B. thuringiensis	*Bacillus thuringiensis*
B. weihenstephanensis	*Bacillus weihenstephanensis*
CAR	Central African Republic
cAMP	Cyclic adenosine monophosphate
canSNP	Canonical single nucleotide polymorphism
CDC	Centers for Disease Control and Prevention
C.diana	*Cercopithecus diana*
CFU	Colony forming unit
CI	Confidence interval
C. maxwelli	*Cephalophus maxwelli*
C. polykomos	*Colobus polykomos*
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of the Congo
EF	Edema factor
ELISA	Enzyme-linked immunosorbent assays
ENA	European Nucleotide Archive
ENP	Etosha National Park
FU Berlin	Freie Universität Berlin
GATK	Genome analysis toolkit
GPS	Global positioning system

HIV	Human immunodeficiency virus
HPD	Highest posterior density
HRP	Horseradish peroxidase
KNP	Kruger National Park
LF	Lethal factor
P. badius	*Procolobus badius*
P.t. verus	*Pan troglodytes verus*
MAPKK	Mitogen-activated protein kinase kinase
MCC	Maximum clade credibility
ML	Maximum likelihood
MLST	Multilocus sequence typing
MRCA	Most recent common ancestor
MRM	Multiple regression on distance matrices
MLVA	Multi-locus-variable-number-tandem-repeat-analysis
MPI-EVA	Max Planck Institute for Evolutionary Anthropology
NNI	Nearest-neighbor-interchange
OD	Optical density
OIE	World Organization for Animal Health
PA	Protective antigen
PBS	Phosphate-buffered saline
(q) PCR	(quantitative) Polymerase chain reaction
PhD	Doctor of Philosophy
PHRANA	Progressive hierarchical resolving assays using nucleic acids
PPE	Personal protective equipment
PVDF	Polyvinylidene difluoride
R^2	coefficient of determination
RKI	Robert Koch Institute
RNA	Ribonucleic acid
RRW	Relaxed random walk
SNP	Single nucleotide polymorphism
SNR	Single nucleotide repeat
SeNP	Serengeti National Park
SPR	Subtree-pruning-regrafting
TBS	Tris-buffered saline

TCP	Taï Chimpanzee Project
TMP	Taï Monkey Project
tMRCA	Time to the most recent common ancestor
TNF-α	tumor necrosis factor α
TNP	Taï National Park
TRBA	Technischen Regeln für Biologische Arbeitsstoffe
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
USA	United States of America
VNTR	Variable number tandem repeat
ybp	Years before present
WAHIS	World animal health information system
WB	Western blot
WGS	Whole genome sequencing
WHO	World Health Organization
WNA	Western North America

1 Introduction

From October 2001 until June 2002 six chimpanzees (*Pan troglodytes verus*) in Taï National Park (TNP), Côte d'Ivoire, died of anthrax-like disease (*1*). This was the first observation of anthrax in wild living non-human primates and in a rainforest ecosystem. In depth investigations revealed that the causative agent of the outbreak differed from classical *Bacillus anthracis* (*2*, *5*). While carrying the *B. anthracis* virulence plasmids pXO1 and pXO2, the chromosomal background of the pathogen was genetically closer to atypically virulent *B. cereus* and *B. thuringiensis* strains. In the following years the subsequently called *Bacillus cereus* biovar *anthracis (Bcbva)* was detected in rainforest habitats in Cameroon, Central African Republic (CAR), Democratic Republic of Congo (DRC) (*2*, *6*) and Liberia (*4*), indicating a broad geographical distribution and the importance of this pathogen for the sub-Saharan region (Fig. 1).

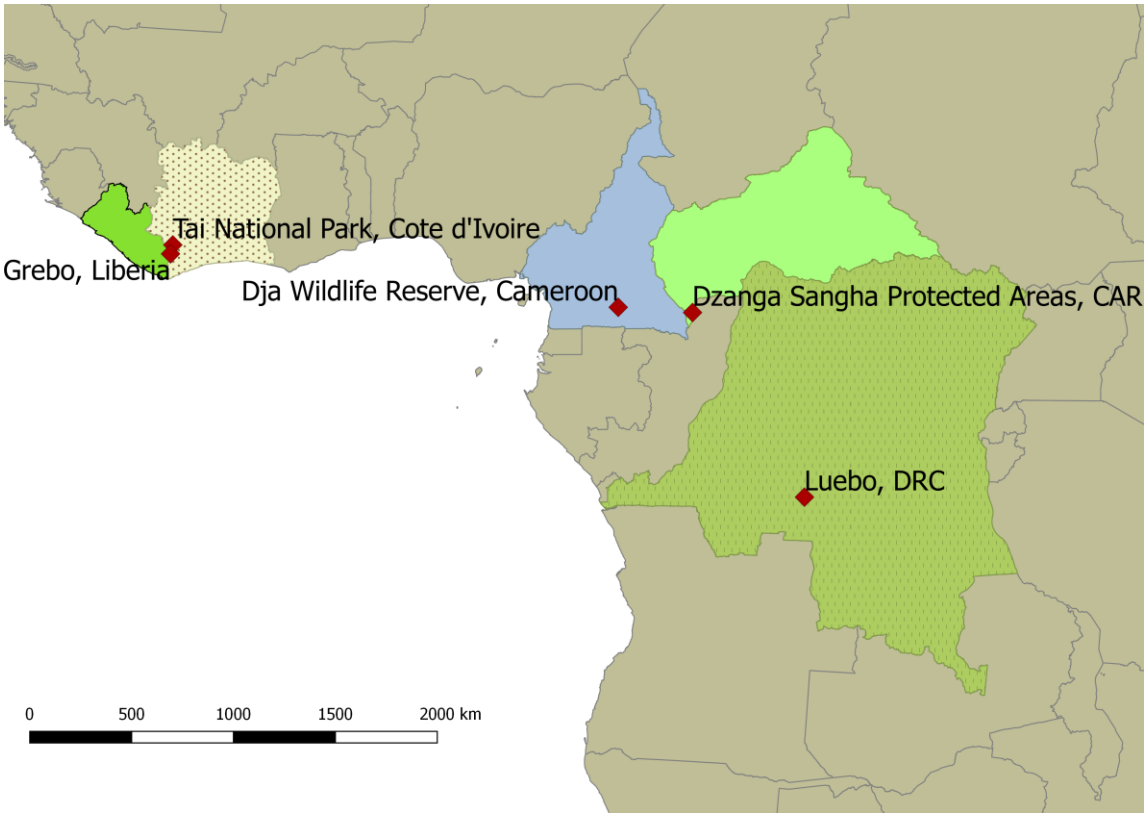

Figure 1. Illustration of sites where anthrax cases caused by *Bacillus cereus* biovar *anthracis* (*Bcbva)* are documented .

Long-term monitoring of carcasses in TNP has revealed continuous wildlife mortality due to *Bcbva* affecting a broad range of species. Out of 205 necropsies in ten different species 81 were tested positive for *Bcbva* (40 %; Table S1) (*4*). This is an exceptional anthrax burden, also compared to other African national parks where anthrax is endemic. For comparison, during systematic investigations of mortality in Etosha National Park, Namibia, anthrax was responsible for 13 % of wildlife deaths (*7*). Systematic molecular analysis of carrion flies confirmed the continuous occurrence of

Bcbva in TNP and molecular analysis of bones identified a *Bcbva* case in a chimpanzee already back in 1991 (*4*).

Bcbva does not only pose a threat to the endangered wildlife in several UNESCO World Heritage Sites. Due to the close phylogenetic relatedness of humans and great apes, diseases that are deadly for one always pose a potential threat to the other. The emergence of this novel pathogen requires research in a One Health framework combining multi-disciplinary efforts and a broad array of methods.

The present work uses two complementary approaches to investigate the epidemiology and ecology of *Bcbva* in TNP:

The core of the study is the analysis of 178 full *Bcbva* genomes derived from necropsy and environmental samples covering a period from 1996 to 2014. Whole genome SNP analysis in a Bayesian statistical framework was performed to gain insights into the disease dynamics at the interface of microbiology and epidemiology. With this approach information was derived about the diversity, activity, transmission pathways, spatial patterns and origin of *Bcbva*. Such a detailed phylodynamic study is not only unprecedented for *Bcbva*, but also in *B.anthracis* research.

Results based on molecular epidemiology were complemented by a classical veterinary approach. We conducted a comprehensive serological study in TNP, testing for anthrax antibodies in five species commonly affected by *Bcbva*.

Bcbva in Africa was the topic of two parallel PhD projects during the past four years. Detailed information about the distribution of *Bcbva* in sub-Saharan Africa and the use of carrion flies as *Bcbva* monitoring tools on small and large scale can be found in the PhD thesis of Constanze Hoffmann.

Although our studies yielded important insights into the lifestyle of *Bcbva*, many questions remain to be answered. Most importantly, investigations are needed to determine whether the disease is restricted to rainforest areas or if human populations around the forests are affected by *Bcbva*. Therefore, the disease prevalence in humans and associated risk factors need to be addressed in the future.

2 Objectives

The phylodynamic and serological studies of this work aimed at illuminating the following:

1. The role of Bcbva in the Taï National Park ecosystem

The diversity of *Bcbva* strains from Taï National Park (TNP) using whole genome SNP analysis of isolates derived from different hosts and sample materials was investigated as an indicator for *Bcbva* activity in the area; to put these results into context the *Bcbva in vitro* mutation rate was determined experimentally. The sequencing data was further used to reconstruct the pathogen evolution and population development in the area in a Bayesian statistical framework. The resulting data was compared to available *Bcbva* sequence data from sub-Saharan Africa. Serological testing of TNP wildlife species susceptible to *Bcbva* aimed at investigating the likelihood of survival following an infection with *Bcbva* in wildlife.

2. Comparative Ecology of *Bcbva* and other spore forming bacteria

Further analyses of the *Bcbva* whole genome sequencing data and mutation rate were performed to compare the lifestyle of so far cryptic *Bcbva* with what is known for other spore forming bacteria - especially *B. anthracis* - with a focus on activity patterns, within host diversity, geographical dynamics and transmission pathways.

3 Background

3.1 Study location: Taï National Park, Côte d'Ivoire

TNP is an evergreen rainforest situated in the south-west of Côte d'Ivoire near the Liberian border (0°15' – 6°07'N, 7°25' – 7°54'W) (Fig. 2). The climate in TNP is sub-equatorial with two rainy seasons (major: August-October, minor: March-June) and a total average annual rainfall of 1800 mm (*8*). The Upper Guinea Forest belt used to stretch from Ghana to Sierra Leone and today TNP represents its largest remain covering an area of 3300 km^2 plus a 200 km^2 buffer zone.

Figure 2. Location of Taï National Park (*figure by C. Hoffmann, published in* (*4*)).

TNP harbors almost 1000 species of vertebrates and was awarded UNESCO Natural World Heritage Site status in 1982 (*8*, *9*). The unique fauna of the park includes the rare pygmy hippopotamus *(Cheoropsis liberiensis),* the bushpig *(Potamochoerus porcus),* the giant forest hog *(Hylocheorus meinertzhageni),* seven species of duikers *(Cephalophus jentinki, C. sylvicultur, C. ogilbyi,* C. *dorsalis, C. zebra, C. maxwelli and Philantomba monticola),* the honey badger *(Mellivora capensis),* the long-nosed mongoose *(Mungos obscurus*), the giant pangolin (*Manis gigantea)*, the leopard *(Panthera pardus),* the golden cat (*Profelis aurata),* the pardine genet *(Genetta pardina),* and the civet cat *(Viverra civetta).* Besides chimpanzees (*Pan troglodytes verus*), ten species of monkeys live in the area: three colobus species *(Procolobus badius, Colobus polykomos, and C. verus),* four cercopithecoids species *(Cercopithecus diana, C. petaurista, C. campbelli and C. nictitans),* the sooty mangabey *(Cercocebus atys),* the dwarf galago *(Galago dernidovii),* and the Bosman's potto *(Perodicticus potto).*

Long-term behavioral study projects on chimpanzees and monkeys have existed since 1979 and 1989, respectively, and three habituated chimpanzee communities, famous for their unique nut-cracking techniques, are constantly observed by primatologists from

the Max Planck Institute for Evolutionary Anthropology Leipzig (MPI-EVA, Leipzig, Germany) organized in the Taï Chimpanzee project (TCP) (*8*). Habituation and observation of different monkey species is done by the Taï Monkey Project (TMP), run by the University of Neuchâtel and the Ohio State University.

Human pressure on the ecosystem has been increasing during the last 40 years and numerous villages are lining the edges of the park today (*10*, *11*). The population density in the Taï "sous-préfecture" adjoining our research site has increased from 8 inhabitants per km^2 in 1971 to 135 inhabitants per km^2 in 1991 (*8*) and since this survey the population of Côte d'Ivoire has increased by another 50 % (*11*).

Besides continuous anthrax cases, cases of ebola and monkeypox occurred in TNP in 1994 and 2012, respectively (*12*, *13*). In 2001 constant veterinary surveillance of wildlife has been implemented by a collaboration of Robert Koch Institute (RKI, Berlin, Germany) and Max Planck Institute for Evolutionary Anthropology (MPI-EVA, Leipzig, Germany). A necropsy is performed on every detected wildlife carcass and samples are routinely tested for *Bcbva* and Ebola virus (Table S1). Due to the sometimes difficult political situation only limited veterinary records are available for the years 2003 and 2010.

3.2 Bacillus cereus group

Bcbva clusters within the *Bacillus cereus* group (*Bacillus cereus sensu lato*), which is part of the Bacillaceae family. This group of Gram-positive, rod-shaped, endospore forming bacteria contains six different bacterial species, namely *B. weihenstephanensis, B. mycoides, B. pseudomycoides, B. thuringiensis, B. cereus* and *B. anthracis* (*14*). While these species are closely related and share highly conserved chromosomal sequences (*15*, *16*), they differ significantly concerning their phenotypes and most importantly their pathogenicity. *B. weihenstephanensis, B. mycoides* and *B. pseudomycoides* are considered to be apathogenic to weakly pathogenic. *B. thuringiensis* is commonly used as a biological pesticide based on its production of intracellular protein crystals toxic for a broad range of insect larvae (*17*). *B. cereus* and *B. anthracis* cause food poisoning and deadly septicemia respectively and will be described in detail below. Most of the relevant genetic information accounting for the different pathotypes is located on plasmid level, which has led to debates in the past about the criteria for taxonomic assignment of the different species (*16*, *18*).

3.2.1 *Bacillus cereus*

B. cereus is an aerobic or facultatively anaerobic, motile bacterium that can be found ubiquitously in the environment in soil and water (*19*), but also commensal in the gut of invertebrae (*20*). It is a well-known agent of food poisoning and hard to eliminate from food processing sites because of its production of adhesive endospores (*19*). Vomiting and diarrhea are caused by the plasmid encoded toxin cereulid and a combination of several chromosomally encoded enterotoxins respectively (*21*, *22*). A number of other virulence factors are encoded chromosomally and untypically pathogenic *B. cereus* have been reported to cause local and systemic infections in both immunologically compromised and competent patients (*19*, *23*).

3.2.2 *Bacillus anthracis*

B. anthracis causes anthrax, a disease characterized by fatal septicemia in a broad range of mammals, including humans. *B. anthracis* is aerobic or facultatively anaerobic and it forms stable endospores that can stay infective in soil for decades determining the dynamics of the disease (*24*). After entering a suitable host, spores germinate and re-establish vegetative growth. At death of the host high numbers of bacteria can be released in the environment and bacteria transform back into the endospore form upon contact with oxygen (*24*). The vegetative form of *B. anthracis* is less stable and it is subject to debates if the *B. anthracis* lifecycle can be maintained outside of a mammalian host (*24*, *25*). Anthrax is generally considered to be a point-source infection and no direct transmission between hosts occurs, but fly-borne spread has been discussed (*26*, *27*) (see 3.2.2.4).

In most cases, humans contract the disease by contact with contaminated animal carcasses or products, but in more recent times novel sources of infection have been reported. During the last 16 years contaminated heroin has led to an increase of human anthrax cases in Europe affecting drug users in Great Britain, Norway, Germany, Denmark and France. Injectional Anthrax presents with an unusual pathology, which is described in more detail below (*28–32*) (see 3.2.2.2). In 1979 in Sverdlovsk (former Soviet Union) 64 people died of inhalational anthrax after an accident in a factory producing *B. anthracis* spores for biological warfare (*33*). *B. anthracis* was also used in bioterrorist attacks, the most infamous example being the anthrax letter attacks in the USA in 2001, where 22 people contracted inhalational anthrax and five people succumbed to the disease (*34*, *35*). The potential use of *B. anthracis* for biological warfare and bioterrorism has led to its classification as category A agent by the Centers for Disease Control and Prevention (CDC, USA). This category includes agents that pose a high risk to the public, because they can easily be spread, result in high death rates and can potentially cause panic in the population. In Germany *B. anthracis* is classified as a risk group 3 pathogen by the german *Ausschuss für Biologische Arbeitsstoffe (ABAS)* (based on the *Biostoffverordnung* and stated in the *Technische Regeln für Biologische Arbeitsstoffe (TRBA)*); this risk group contains agents which cause severe disease in humans and spread easily, but are treatable.

3.2.2.1 Worldwide distribution

Formation of anthrax endospores ceases at temperatures lower than 9-12 °C (*26*), which explains why *B. anthracis* is mostly endemic in warm countries. Endemic zones are described in sub-Saharan Africa, central Asia, countries of the former Soviet Union, China, India, Australia, the USA and parts of South America. Sporadic cases occur in southern Europe and less frequently in northern Europe (*26*, *36*). The last (not imported) cases in Germany were reported from cows in Saxony-Anhalt in summer 2012 and spring 2014 (*37*, *38*).

However in developing countries, underreporting of the disease due to lack of experience, veterinary personnel and interest from governments is a severe problem and the worldwide distribution of *B. anthracis* can only be approximated (*36*). We know from personal communication by our colleagues in Côte d'Ivoire that regular anthrax outbreaks occur in the north of the country. This is not adequately reported by ProMED or in the World Organisation for Animal Health (OIE) database (WAHIS Interface).

3.2.2.2 Clinical manifestation

Cutaneous anthrax represents over 95 % of human anthrax cases (*26*). As the bacteria are not invasive, they enter the skin through small cuts or abrasions (*39*). They cause a localized necrotic lesion (eschar), which is often accompanied by marked edema (*26*). The incubation period lasts usually from 2-6 days, but incubation periods of up to 3 weeks have been described (*26*). Cutaneous anthrax is self-limiting and stays localized in the skin in most cases, so that 80-90 % of lesions resolve without complications even without treatment. However, antibiotic treatment is recommended, as malignant edema followed by symptoms of shock or anthrax meningitis can complicate the course of the disease (*40*). The importance of cutaneous anthrax for animals is unknown and cutaneous anthrax in animals is likely underreported. In a report about an anthrax outbreak in India with suspected transmission by biting flies 90 % of affected cattle showed the cutaneous form of the disease (*41*) (see also 3.2.2.4).

While symptoms of gastrointestinal anthrax in humans appear 2-5 days after ingestion of *B.anthracis* spores (*40*), the incubation period in livestock appears to vary more (*36*). Therefore the OIE suggests trade restrictions of 20 days after occurrence of anthrax (*39*). It is assumed that breaches in the mucosal lining facilitate manifestation of the disease, but it is unclear to date, where exactly the germination of endospores occurs. Ulceration of the gastrointestinal mucosa leads to bloody diarrhea and vomiting, which are accompanied by high fever (*40*). If not treated with antibiotics, gastrointestinal anthrax is usually fatal with death caused by intestinal perforation or anthrax toxemia (*40*).

Inhalational anthrax occurs rarely, as endospores must be aerosolized for inhalation. Cases were described in goat-hair mills, in people using animal hides for drum building and in a bioterrorism/biological warfare context (*40*, *42*). For animals, inhalation during grazing has been discussed (*43*) (see also 3.2.2.4). Although spores enter the host through the bronchi, pulmonary infection is not always observed in inhalational anthrax cases (*40*). Alveolar macrophages are generally thought to incorporate the spores and transport them to mediastinal and peribronchial lymphnodes, the starting point of mostly fatal bacteremia and toxaemia (*40*), but other models of dissemination are discussed (see 3.2.2.3) (*44*). Incubation usually lasts ten days, but incubation periods of up to six weeks have been described, indicating possible persistence of spores in the lungs (*40*).

Anthrax meningitis is a rare, but severe complication of all forms of anthrax disease. Infection of the meninges leads to a hemorrhagic meningitis, which is almost always fatal, even under antibiotic treatment (*26*, *40*).

The three classic types of anthrax infection (cutaneous, gastrointestinal and inhalational) have recently been complemented by the injectional form of the disease, observed in heroin users (*28*).

This newly described form of the disease presents with severe soft tissue infections around the injection sites, but no eschars typical for cutaneous anthrax (*29*, *31*). Large proportions of patients showed non-specific gastrointestinal and neurological symptoms (*29*). Mortality rates are significantly higher for injectional anthrax than for cutaneous anthrax, due to the frequently systemic course of the disease (*29*).

3.2.2.3 Pathogenesis and virulence factors

The non-invasive *B.anthracis* endospores enter the host via inhalation, ingestion, injection or skin abrasions. For inhalational anthrax it is widely assumed that spores are taken up by alveolar macrophages and dendritic cells and transported to regional lymphnodes (*45*, *46*). According to the most commonly known and cited "Trojan horse" model spores germinate in the macrophages during this transport. The vegetative bacilli spread through the lymphatic system and are finally released into the blood stream, where they reach concentrations of 10^7 to 10^8 bacteria per ml blood in the final stage of the disease (*40*, *45*). More recently, Weiner & Glomski (2012) have suggested the "Jailbreak" model, in which spores can also germinate directly at the site of entry if the integrity of the epithelial barrier is compromised. Bacteria colonize the entry site first, further damage the epithelial barrier by production of their exotoxins and subsequently travel to the regional draining lymphnode by lymphatic flow rather than phagocytic transport (*44*). Due to the potential of aerosolized *B.anthracis* spores as biological weapon, most studies focus on the pathogenesis of inhalational anthrax. The development of gastrointestinal anthrax is poorly understood, but spores have been shown to be captured and to proliferate in Peyer's patches (*47*).

Immune response against vegetative bacilli is impaired by the *B.anthracis* poly-γ-D-glutamic acid capsule encoded on plasmid pXO2, harboring the capsule gene operon *capBCADE* (*24*, *48*, *49*).

The anthrax exotoxins are encoded within the plasmid pXO1 pathogenicity island by the genes *pagA, lef* and *cya* (*24*). They are responsible for the production of protective antigen (PA), lethal factor (LF) and edema factor (EF) respectively (*24*). PA (size=83kDa) can bind to at least two cellular ANTXR receptors and mediates the entry of EF and LF into the host cells (*24*, *50*). After binding to the receptor the PA monomer is split into two subunits (PA20 and PA63). PA20 dissociates and EF and LF bind competitively to the oligomerized PA63 heptamer, with one PA 63 sub-unit binding a maximum of three LF/EF components (*24*, *51*). Intradermal application of PA + EF (the edema toxin) alone causes edema in the skin (*52*), while intravenous injection of PA + LF (lethal toxin) has been shown to be lethal (*53*).

EF (size=89kDa) is a calcium- and calmodulin-dependent adenylate cyclase that transforms intracellular adenosine triphosphate (ATP) into cyclic adenosine monophosphate (cAMP). Massive accumulation of cAMP in the cell disturbs homeostasis and is responsible for the severe edema observed in cutaneous anthrax (*50*). Additionally a neutrophil impairing effect was demonstrated for EF in *in vitro* experiments (*54*) and EF was shown to increase the expression of ANTXR receptors, promoting toxin intake into the cells (*55*).

LF (size=90kDa) stimulates macrophages to produce tumor necrosis factor α (TNF α) and interleukin-1-β (IL1-β) which are at least partly responsible for the sudden deaths following systemic anthrax infections (*24*, *50*). It is not clear yet, how exactly LF interacts with macrophages, but it was shown that LF is a zinc-dependent metalloprotease that inactivates mitogen-activated protein kinase kinases (MAPKKs) *in vitro* (*24*, *50*).

Expression of virulence factors depends on environmental factors and is generally favoured by temperatures > 37 °C, carbon dioxide concentrations > 5 % and presence of different serum factors (*24*, *56*, *57*), with capsule formation independent of temperature

(*58*). While regulation of toxins and capsule is mediated by the global regulator AtxA, encoded on pXO1, the capsule has an additional regulator AcpA, encoded on pXO2 (*59*). Elimination of plasmid pXO2 is enough to reduce *B. anthracis* virulence so much that it can be used for vaccination in animals. This has been done successfully for decades by the Sterne veterinary spore vaccine (*60*).

3.2.2.4 Ecology and infection of animals

Herbivores are known to be especially susceptible to *B. anthracis,* while deaths in carnivores are rarely reported (*26*, *36*). Grazers ingest endospores when feeding on contaminated soils (*61*) and oral microlesions caused by stones or thorny vegetation are suspected to enhance disease development as spores alone are not considered to be invasive (*36*). In a field experiment in Etosha National Park it was shown that *B. anthracis* spores can promote grass growth and subsequently attract potential hosts (*61*, *62*). Gastrointestinal anthrax is the most common form of anthrax in animals, although inhalational anthrax cases in cows have been reported (*43*). These inhalational anthrax cases suggest that inhalation of spores while grazing in dry and dusty areas may be possible.

However many anthrax scenarios in animals cannot be explained by grazing alone and a number of the affected species do not graze at all. Extensive research in this field has resulted in a number of alternative suggestions for routes of transmission and spread of the disease (*36*):

Animal experiments from the beginning of the 20th century already confirmed that anthrax transmission by biting-flies was possible *in vitro*. Haemophagic flies have since been suspected to cause anthrax transmission between animals in west Texas and India (*27*, *41*, *63*). These reports are mainly based on the observation, that anthrax incidence was correlated with high biting fly densities and the fact that cases spread fast over large distances. Fasanella et al. (2013) reported a case of human cutaneous anthrax in a farmer stung by a horse-fly (family *Tabanidae*) (*64*).

Involvement of blowflies (*Calliphoridae*) in anthrax transmission has first been discussed in Kruger National Park (KNP), where anthrax epidemics coincided with a large increase in the blowfly population (*65*). It was found that after feeding on a carcass the flies deposited fecal and discarded regurgitation droplets on the nearby vegetation in a height of 1-3 m. This is the feeding height of browsers like the kudu (*Tragelaphus strepsiceros*), the primary host of anthrax in KNP (*65*). Similar observations were made for white-tailed deer in North America (*27*) and during an anthrax outbreak in west Texas *B. anthracis* DNA was recovered from vegetation surrounding carcasses (*66*). In *in vitro* studies Fasanella et al. (2010) showed that after feeding on contaminated carcasses house fly (*Musca domestica*) droplets contained *B. anthracis* spores (*67*) and in an experimental study by von Terzi et al. (2014) living bacteria were cultured from carrion flies up to nine days after ingestion of vegetative bacilli (*68*). Insect mediated transmission of *B. anthracis* is also the most likely explanation for the anthrax outbreak in the Russian tundra in 2016 with thousands of affected reindeers and cases in humans (*69*).

Plenty evidence exists for the important role biting and non-biting flies may play for *B. anthracis* transmission, but final proof during a natural anthrax outbreak is missing for both (*63*).

During mass outbreaks in hippos (*Hippopotamus amphibius*) spread of the disease among the hippo population due to carnivory was described (*70*).

De Vos et al. (1990) suspected spread of anthrax by vultures (family *Accipitridae*) in KNP. They observed that vultures cleaned themselves in waterholes after feeding on carcasses, partly even regurgitated food into the water and might thus contaminate these important drinking sites with *B. anthracis* spores (*65*). Experimentally *B. anthracis* could be found in vulture feces after feeding them with *B. anthracis* spores, but not after feeding them with vegetative bacilli. Therefore vultures might also prevent anthrax spread by elimination of carcasses before formation of spores (*65*).

3.2.3 *Bacillus cereus biovar anthracis*

Bcbva was first described in Taï National Park (TNP) in 2001 and 2002, when six chimpanzees died of anthrax-like disease (*1*). All individuals had been in apparently good physical condition. The deaths of three individuals were observed by primatologists and it was reported that they had died within a couple of hours after the onset of disease symptoms. Pathological and histological examination revealed haemorrhages in nearly all inner organs and Gram-positive, rod-shaped bacteria were present intra- and extravascularly.

The causative agent of the disease carried the two *B. anthracis* virulence plasmids pXO1 and pXO2, but had a higher similarity with non-*B. anthracis* members of the *B. cereus* group at chromosomal level (*2*, *5*). It lacked the *B. anthracis* specific prophage regions A, C, D and E and in phylogenetic analysis based on multilocus sequence typing (MLST) its closest relatives were two atypically virulent members of the *B. cereus* group (*B. thuringiensis* serovar konkukian strain 97-27 and *B. cereus* strain E33L).

At bacteriological level it was also apparent that this novel pathogen could neither be classified as *B. anthracis*, nor as *B. cereus* (*5*). Like *B. anthracis*, it lacked beta-hemolytic and phospholipase C activity, produced a capsule and was sensitive to Penicillin. But it also shared properties with *B. cereus*, most importantly resistance to the gamma phage and motility. It was assumed that the lack of hemolytic and phospholipase C activity was caused by a frameshift mutation in the gene for the regulator PlcR, in contrast to the nonsense mutation observed in *B. anthracis* (*5*).

In 2010 the full genome sequence of *Bcbva* was first published (an isolate from the deceased TNP chimpanzee Léo) (*71*). It confirmed that the pathogen, though carrying the *B. anthracis* virulence plasmids, was not part of the monophyletic *B. anthracis* clade. It was suggested that it evolved, when an independent *B. cereus* strain acquired the two *B. anthracis* plasmids, which led to the name *Bacillus cereus* biovar *anthracis* (*71*). The genome of *Bcbva* consists of four replicons: the chromosome (5196054 bp), plasmid pXO1 (181907 bp), plasmid pXO2 (94469 bp) and the small plasmid pCI-14 (14219 bp), which is not present in all isolates. The plasmids pXO1 and pXO2 show 99 % to 100 % identity with the *B. anthracis* plasmids in comparative sequence analysis. Six unique genomic islands, all larger than 12 kbp, were identified within the chromosome of *Bcbva* CI. Interestingly Island IV is integrated in the gene for sporulation factor SigK (*71*). This island is also routinely used today as a marker in PCR assays for the differentiation of *Bcbva* from classical *B. anthracis.*

In 2004, 2006, 2012 and 2013 *Bcbva* cases were reported from Dja Faunal Reserve (Cameroon), Dzanga-Sangha Complex of Protected Areas (CAR) and Luebo (DRC) in chimpanzees (*Pan troglodytes verus*), gorillas (*Gorilla gorilla*), an elephant (*Loxodonta africana*) and a domestic goat (*Capra ageagrus hircus*), suggesting a broad sub-saharan distribution of the pathogen and a broad host range (*2*, *6*). Recently, a new site of *Bcbva* occurrence was detected in Grebo, Liberia, by systematic analysis of fly and bone samples (*4*) and isolates from Grebo are included in the phylogenetic analyses in this work.

Multi-locus-variable-number-tandem-repeat-analysis (MLVA) using eight markers was performed for *Bcbva* strains from Côte d'Ivoire and Cameroon and revealed that *Bcbva* did not nest within the B. *anthracis* A or B lineages, commonly present in Africa (*72*, *73*). Isolates from Côte d'Ivoire and Cameroon also differ from each other, but together they form the new "forest anthrax" cluster F (*72*). This was later confirmed by whole genome SNP analysis of isolates from Côte d'Ivoire, Cameroon, CAR and DRC, which showed that all *Bcbva* strains cluster together in a monophyletic clade within the *B. cereus* group, but outside of the *B. anthracis* clade (*6*). Their closest relative in this analysis was *B. cereus* strain ISP3191 isolated from spice in Belgium (*74*). All *Bcbva* strains generally share similar microbiological properties, but isolates from Cameroon, CAR and DRC are penicillin resistant and the isolate from DRC lacks motility due to a pre-mature stop codon in the *fliP* gene of the flagella gene cluster (*5*, *6*).

The virulence of *Bcbva* was recently shown to be comparable to that of *B. anthracis* in small animal models (*3*). In addition to the *B. anthracis* polyglutamate capsule encoded on pXO2, *Bcbva* expresses a second capsule encoded by the *hasACB* operon on pXO1 and consisting of hyaluronic acid. Expression of the second capsule is regulated by the global transcription regulator AtxA, which also controls anthrax toxins and polyglutamate capsule in classical *B. anthracis* (*24*). It is important to note that unlike classical *B. anthracis*, the virulence of *Bcbva* is only slightly reduced when cured of plasmid pXO2 (*3*).

Bcbva was classified as a risk group 3 agent by the German ABAS and laboratory work with *Bcbva* needs to be performed in BSL 3 facilities.

3.2.4 *Other untypically pathogenic members of the B. cereus group*

Untypically pathogenic non-*B. anthracis* members of the *B. cereus* group have caused severe anthrax-like disease based on plasmid encoded virulence factors at several occasions in the past.

B. cereus G9241 was isolated from the sputum and blood of a patient with inhalational anthrax-like symptoms in the US in 1994 (*75*). It carried a plasmid homologous to pXO1 expressing the *B.anthracis* toxins and combined this with a polysaccharide capsule encoded on a previously unidentified plasmid. Additionally, it expressed a hyaluronic acid capsule encoded on pXO1 as described for *Bcbva* (*76*). Further cases of disease caused by *B.cereus* strains carrying pXO1 homologues and different types of capsules have been reported in the US since, for example in 2003 fatal pneumonia was caused by *B. cereus* 03BB102, a *B. cereus* isolate carrying a plasmid that harbored both the anthrax toxin genes and the capsule synthesis genes *cap A, B* and *C* (*77*, *78*).

Recently bovine anthrax cases have been published from Koza, Cameroon, where the causative pathogen JF3964 combined genetic traits of *B. cereus* with the two *B. anthracis* virulence plasmids, confirmed by southern blot analysis (*79*) (Fig. 3).

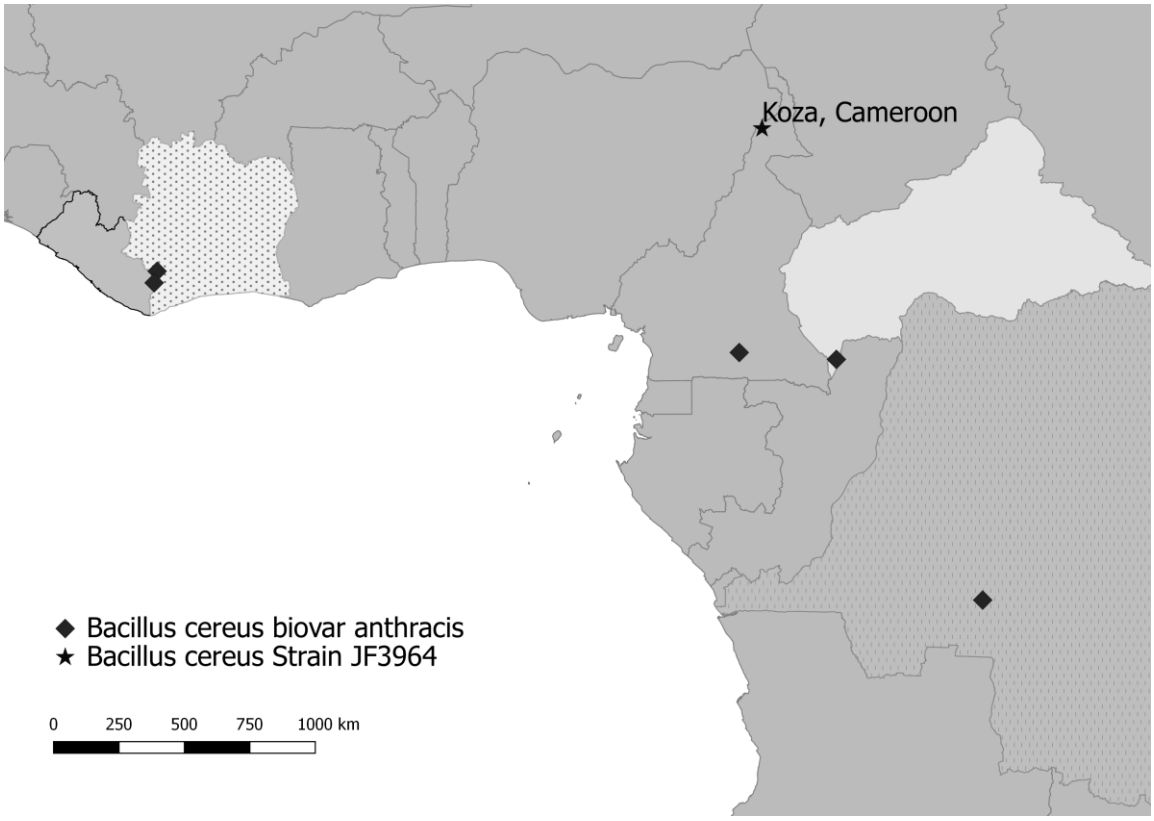

Figure 3. Illustration of sites where anthrax cases caused by *Bacillus cereus* biovar *anthracis* (diamonds) and *B. cereus* strain JF3964 (star) are documented. Strain JF3964 was reported from cattle in Koza, Cameroon.

In phylogenetic analysis based on MLVA results, the isolate JF3964 clusters with *Bcbva* and its microbiological traits are also in accordance with *Bcbva* strains from Cameroon. Unfortunately, no other information or full genome sequence is available for strain JF3964. The occurrence of both *B. anthracis* virulence plamids in a non-*B. anthracis* background has only been observed in *Bcbva* and isolate JF3964 to date.

Horizontal gene transfer of plasmids or single plasmid-encoded genes is not uncommon within the *B. cereus* group (*80*). This further challenges the strict taxonomic grouping of the different *B. cereus* group species and could be of importance for the emergence of *Bcbva*.

3.3 *B. anthracis* in African wildlife

Anthrax, caused by *B. anthracis*, is enzootic in several large African national parks (*27*). Numerous studies of the disease have been performed throughout Africa and underline the importance of investigating the pathogen in close relation with its ecosystem. Unlike TNP and other tropical rainforests, where the presence of *Bcbva* was described, all of the parks with frequent *B. anthracis* occurrence represent savannas. The typical anthrax

scenario in wildlife in savanna regions is characterized by sporadic fatalities throughout the year with seasonal epidemics affecting different species in waves (*27*). Over 99 % of anthrax fatalities in savanna habitats are recorded in ungulates (*7*, *81*). Anthrax is thought to be density dependent and it is considered a natural culling mechanism that is symbiotic with its ecosystem (*27*). It can become a threat for wildlife if the sensitive balance of the ecosystems is disturbed, i.e. by poaching or fencing in of parts of the habitat (*65*). Detailed anthrax surveillance data is available from Etosha National Park (ENP, Namibia), Kruger National Park (KNP, South Africa) and Serengeti National Park (SeNP, Tanzania). However, cases are rarely diagnosed by molecular methods or bacterial culture. If animals show typical pathology or cases co-occur with known anthrax cases, diagnosis of anthrax is often based on circumstantial evidence. Most studies rely on microscopic detection of rod-shaped Gram-positive bacilli. Epidemiological scenarios vary widely between parks and the ecology and epidemiology of anthrax in wildlife is still poorly understood (*27*, *82*).

3.3.1 *Anthrax in Etosha National Park, Namibia*

ENP is a 22 000 km^2 conservation area in northern Namibia consisting mainly of savanna woodland and saline desert. The park has a savanna desert climate with a distinct wet season from November until March (*83*).

Anthrax was first described here in 1964 (*84*) and distinct enzootic areas can be differentiated from non-enzootic areas (*85*). Sporadic anthrax cases occur year-round interspersed by occasional small-scale outbreaks. Larger epizootics are rare and occur usually only in one species at a time (*85*). Species mostly affected by anthrax in ENP are Burchell's zebra (*Equus quagga*), blue wildebeest (*Conochaetes taurinus*), springbok (*Antidorcas marsupialis*) and elephant (*Loxodonta africana*) (*85*). A seasonal anthrax peak was observed for plains ungulates at the end of the rainy season and elephants at the end of the dry season (*85*).

The cause of this pattern is still unknown, but it has been speculated that host immune status and behaviour may play an important role (*85*, *86*). This is favoured by the observation that during several major outbreaks in elephants no cases were observed in plains ungulates, although they were drinking from waterholes near elephant carcasses (*85*).

ENP wildlife mortatlity records have been stored since 1975 for over 6000 animals. Anthrax was confirmed as cause of death in 13 % of carcasses and suspected in another 11 % (*7*). It plays a major role for wildlife mortatilty in ENP, also compared to other infectious diseases, like rabies (2 %) (*7*). There is generally no anthrax control program in place, except for rare aerial dart gun vaccination of threatened species (*7*).

3.3.2 *Anthrax in Kruger National Park, South Africa*

KNP covers an area of 20.000 km^2 in north-eastern South Africa. The climate here is subtropical with a wet season from October until May and vegetation mainly consisting of savanna woodlands (*65*).

Isolation of *B. anthracis* from a radio-carbon dated bone allowed dating back of anthrax occurrence in KNP to the middle of the 18^{th} century (*65*). Regular outbreaks have been

recorded since 1954 and occur almost exclusively in the northern half and centre of the park with a hotspot in the Fever Tree Depression at the northern tip (*65*, *87*). The primary hosts for Anthrax infection in KNP are kudu (*Tragelaphus strepsiceros*) and impala (*Aepyceros melampus*) and infections follow a strict seasonal pattern with a distinct peak from July until August (*65*). Unlike for ungulates in ENP, anthrax mortality in KNP peaks in the dry season indicating a different ecology of the disease in this habitat. Large anthrax epidemics occur periodically every ± 10 years usually following a dry spell after several years of above average rainfall (*65*).

It has been suggested that for the anthrax transmission cycle in KNP concentration of anthrax spores in water holes during the dry period and subsequent spread of the disease by blowflies may play a role (*65*, *88*). Involvement of vultures in the distribution of anthrax spores into water holes has also been discussed (*65*).

Although anthrax is a part of the KNP ecosystem, it poses a threat to endangered species, like the roan antelope (*Hippotragus equinus*), because of the unnatural fenced-in setting of the park today (*65*).

3.3.3 *Anthrax in the Serengeti, Tanzania*

SeNP in the north of Tanzania consists of 15 000 km^2 of grassland plains, savanna, riverine forests and woodlands. It is famous for the largest terrestrial mammal migration in the world with 1.5 million western white-bearded wildebeests (*Conochaetes taurinus mearnsi*) and 250 000 zebras (*Equus quagga*) following the seasonal growth of grass in the plains. The climate of SeNP is moderate with a short rainy season from November until December and a long rainy season from March until May (www.eoearth.org).

Veterinary records from 1996 until 2009 revealed annual sporadic cases of anthrax in a broad range of ungulates (*81*). Four large spatially localized species-specific outbreaks have been described, affecting impalas (*Aepyceros melampus*), zebras & wildebeest or buffalos (*Syncerus caffer*) at a time (*81*, *82*). Large epidemics always followed heavy rains or long droughts, determining extreme weather conditions as a risk factor for anthrax in SeNP (*82*). Anthrax cases were heterogeneously distributed throughout the park indicating a correlation between anthrax prevalence and soil alkalinity (*82*).

3.3.4 *Anthrax in other African national parks*

Anthrax outbreaks have been observed in a number of other national parks, but less detailed data is available here.

Mass outbreaks of anthrax were described in the hippopotamus (*Hippopotamus amphibius*) populations of Luangwa River Valley in Zambia and the Queen Elizabeth National Park in Uganda (*89*, *90*). Carnivory seems to play an important role for the spread of the disease within the hippopotamus populations (*70*).

During the first described anthrax outbreak in Zimbabwe in 2004 in Malilangwe wildlife reserve almost all kudus (*Tragelaphus strepsiceros*) were exterminated, along with significant numbers of deaths in nyalas (*Tragelaphus angasii*), bushbucks (*Tragelaphus scriptus*), waterbucks (*Kobus ellipsiprymnus*), roan antelopes (*Hippotragus equinus*) and buffalos (*Syncerus caffer*) (*91*).

An outbreak affecting red hartebeest (*Alcelaphus buselaphus),* wildebeest (*Connochaetes taurinus*), zebra (*Equus quagga*), springbok (*Antedorcas marsupialis*) and eland (*Tragelaphus oryx*) was reported from Jwana Game reserve, Botswana, in 2004-2005 (*92*).

Further an anthrax outbreak in equids in the Wamba area of souther Samburu, Kenya, during a drought killed Burchell's zebras (*Equus quagga*), donkeys (*Equus asinus)* and a significant number of threatened Grevy's zebras (*Equus grevyi*) between December 2005 and March 2006 (*93*).

While reports about large epidemics are being published, it is likely that sporadic cases or smaller outbreaks in wildlife occur unnoticed.

3.4 *B. anthracis* genotyping and evolution

3.4.1 *B. anthracis genotyping*

B. anthracis is phenotypically and genotypically monomorphic and it has been impossible for a long time to differentiate *B. anthracis* isolates originating from different sources (*36*). Biochemical, serological and phage typing methods did not reveal strain differences and attempts to distinguish *B. anthracis* isolates with common molecular tools were not very fruitful, i.e. studies using amplified fragment length polymorphisms (AFLP) or ribotyping (*94*, *95*).

The search for areas of diversity within the *B. anthracis* genome resulted in a first success in the history of *B. anthracis* typing in 1996: the identification of the first variable number tandem repeat (VNTR) locus *vrrA* (*96*, *97*). VNTR loci are likely the result of slipped strand mispairing and they are prone to relatively fast mutation compared to the rest of the genome (*98*). Keim et al. published seven additional tandem repeat loci in 2000 (five chromosomal loci and one locus for each plasmid) and established the idea of the multi-locus-variable-number-tandem-repeat-analysis (MLVA) (*99*). The 8-marker MLVA was used to compile the first *B. anthracis* genotype database, including 426 isolates representing 89 different genotypes. This database was the basis for several *B. anthracis* typing projects, one of them being the investigation of the bioterrorist anthrax letter attacks in the USA in 2001 (*35*). In the following years further MLVA loci were identified by Van Ert et al. (*73*) and Lista et al. (*100*) resulting in the comprehensive 31-marker MLVA.

In addition to MLVA, Keim et al. suggested in 2004 to use a hierarchical approach to *B. anthracis* strain typing, the so called "Progressive Hierarchical Resolving Assays using Nucleic Acids" (PHRANA) (*98*). They proposed definition of phylogenetic groups by canonical Single Nucleotide Polymorphisms (canSNPs), followed by MLVA and sub-typing of resulting groups with Single Nucleotide Repeats (SNRs). This system should account for the fact that different genetic markers are appropriate on different scales. CanSNPs are SNPs at key phylogenetic positions that are considered to be highly stable. Whereas they are suitable to identify the deep phylogeny of strains, they are not discriminatory enough to sub-type populations. SNRs are a special case of VNTRs, considered to be mutational hot-spots, suitable for sub-typing but not stable enough to define phylogenies.

Van Ert et al. used a combination of canSNP analysis and 15-marker MLVA to derive the global distribution of *B. anthracis* genotypes from 1000 isolates in 2007 (*73*). All strains could be classified into the three main lineages A, B and C and represented 221 MLVA subtypes. The A lineage has a broad worldwide distribution, while B and C lineages are rather rare with the B lineage only present in Europe, North America and South Africa and the C lineage only detected in the USA to date (*73*, *101*). As mentioned above and in line with its divergent chromosomal clade *Bcbva* strains from Cameroon and Côte d'Ivoire did not cluster with the A or B branch, but formed their own independent "Forest Anthrax" cluster F (*72*).

While all these typing methods are functional and represent significant advances in understanding the epidemiology of *B. anthracis*, they still lack the discriminatory power that can be achieved by sequencing of the whole *B. anthracis* genome. In the past years whole-genome sequencing has become widely available and this allowed us to use SNPs as stable phylogenetic markers for the present study and still receive a sufficient resolution. An increasingly high number of *B. anthracis* genotyping studies by whole-genome-SNP analysis are being published (*32*, *102*, *103*) and this approach will likely replace the other *B. anthracis* typing systems in the near future.

3.4.2 *B. anthracis genotyping in African national parks*

The only two existing detailed African *B. anthracis* genotyping studies investigated isolates from ENP and KNP using PHRANA and MLVA based approaches, respectively.

Genetic diversity of 384 *B. anthracis* strains from ENP has been assessed by 31-marker-MLVA in combination with canSNPs and SNRs following the PHRANA system suggested by Keim et al. (2004) (*98*, *104*). 24 different MLVA genotypes were identified within the boundaries of ENP, all belonging to the *B.anthracis* A branch. The isolates represented a closely related clonal group differing by only one or two MLVA markers in most cases. Three prominent MLVA genotypes accounted for 80 % of isolates causing long lasting and recurrent outbreaks, while numerous different genotypes caused sporadic outbreaks. One of the three prominent strains was hypothesized to be the ancestor of all other genotypes in ENP. It was observed several times that two to five different MLVA genotypes could cause simultaneous outbreaks.

In KNP genetic diversity was investigated using 8-marker MLVA on 98 *B. anthracis* isolates (*87*). In total, four genotypes were identified with two prominent genotypes representing 97% of all isolates. Interestingly, the two most common genotypes in KNP clustered within the distinct *B. anthracis* A and B branches respectively, which could either be explained by ancient divergent evolution or introduction of divergent *B. anthracis* strains into the area. Isolates of the A branch occurred in the north and the center of the park, while isolates of the B branch were only found in the north. Simultaneous outbreaks caused by A and B strains were observed, contradicting the original idea that outbreaks always start with one distinct infection event and subsequently affect the whole park in waves.

No detailed *B. anthracis* genotyping has been done for isolates from other African national parks and so far no comparative whole genome SNP studies have been performed on the available isolates from ENP and KNP. Such analyses could provide

valuable insights into the epidemiology and ecology of anthrax in African wildlife in the future.

3.4.3 *B. anthracis mutation rate in vitro*

The *in vitro* per site mutation rate of *B. anthracis* was determined by Vogler et al. in 2002 by measuring the spontaneous development of rifampicin resistance with a Luria-Delbrück fluctuation test (*105*). The magnitude of the resulting rate 5.2 x 10^{-10} mutations/site/generation was confirmed by Agren et al. in 2014 (*102*). They combined a sequential passaging approach (n=3) with whole genome sequencing and determined the number of mutations introduced after ten days of passaging, which resulted in a rate of 8.3 x 10^{-10} mutations/site/generation. In the present study we present a more refined version of this setup resulting in an estimate for the *in vitro* mutation rate of *Bcbva*.

3.4.4 *B. anthracis evolution*

Due to the limited discriminatory power of *B. anthracis* typing tools it has been challenging to determine real-life evolutionary rates or the time to the most recent common ancestor (tMRCA) of *B. anthracis* strains in the past.

Van Ert et al. (2007) made a tentative attempt to derive age estimates from the whole-genome SNP data available for seven major *B. anthracis* lineages and estimated the separation of the C branch from the A/B branch to have happened 12 000- 25 000 years before present (ybp) (*73*). The worldwide spread of the A lineage was assumed to have occurred in the mid-Holocene (3000-6000 ybp) with continuous reintroductions into North America during colonial and industrial times. However, these estimates are based on the *in vitro* mutation rate estimates by Vogler et al. (*105*) and a very crude assumption for infection/death cycles.

In 2009, Kenefic et al. examined a selection of 2850 SNPs in 128 various *B.anthracis* isolates to investigate the origin of the Western North America (WNA) clade belonging to the A lineage (*106*). They concluded that *B. anthracis* must have been introduced to North America from Asia around 13 000 ybp via the Bering Street and spread southwards from there. This study also contains highly speculative elements and is mainly based on the relatively large evolutionary separation between the WNA clade and its nearest Old World relatives in the transeurasian clade (106 SNPs). The assumption for infection/death cycles used by Van Ert et al. (2007) was adapted according to empirical data, which led to the increase in the estimate for the time since radiation of the A lineage.

This work has been challenged by Vergnaud et al. (2016) who performed whole-genome SNP analysis on 122 French *B. anthracis* isolates and 45 *B. anthracis* isolates of worldwide origin (*103*). They showed that *B. anthracis* WNA is nested within the French *B. anthracis* phylogeny and postulated that it had been introduced to North America from Europe at the beginning of the 17th century, when French settlers populated the Canadian coast. They argued that the length of the WNA branch was not due to long persistence of *B. anthracis* in North America, but to accelerated expansion and therefore evolution in the naive North Amercian habitat.

Time-scaled phylogenetic tree reconstruction using Bayesian evolutionary analysis represents a less speculative approach to derive information about evolutionary rates and spread of pathogens. This is the method chosen in the present investigation of *Bcbva* evolution and it offers unique insights into the lifestyle of *Bcbva* in TNP. The concept of Bayesian evolutionary analysis and its challenges, when working with monomorphic bacteria, are described in detail below.

Leading scientists in the field have attempted comparable analyses for *B. anthracis*, but have not yet detected sufficient diversity in a *B. anthracis* outbreak scenario to derive a temporal signal (*P.Keim, personal communication*).

3.5 Bayesian evolutionary analysis of monomorphic bacteria

The concept of measurably evolving pathogen populations and related analytical approaches have been introduced by Drummond et al. in 2003 (*107*). It is based on the assumption that serially sampled pathogen sequences measurably accumulate *de novo* mutations over time, so that deductions about their rate of evolution can be made (*108*). This has successfully been applied to fast-evolving RNA viruses in the past, i.e. for timing the ancestors of HIV-1 and HIV-2 (*109*, *110*).

Monomorphic bacteria like *B. anthracis* or *Mycobacterium tuberculosis* have very low per-site mutation rates and it has long been assumed that they were too evolutionary stable to be included in this type of analysis (*108*). However, mutation rates and genome sizes are usually inversely related and the relatively large bacterial genome sizes in combination with the high resolution of whole genome sequencing results in a comparable amount of genetic information as found in fast-evolving, but small, RNA virus genomes (*108*). Therefore the rapidly developing field of Bayesian evolutionary analysis can also offer new insights into the population dynamics and disease epidemiology of monomorphic bacteria (*108*, *111*).

The use of molecular clock models in phylogenetic analyses, adapting root-to-tip distances to sampling dates, allows us to estimate the evolutionary rate and the tMRCA of serially sampled pathogen sequences in a Bayesian framework as implemented for example in the program BEAST (Bayesian Evolutionary Analysis by Sampling Trees) (*112*, *113*). This can be combined with coalescent based approaches revealing the demographic histories of pathogen populations (*114*). If further information is available, i.e. about the geographic origin of samples, analyses can be adapted to make implications about the geographic origin and spread of a disease outbreak as explored in the field of phylogeography (*115*, *116*). In recent studies information about the pathogen host was added to the models to reconstruct the likely directionality of transmission between different species (*117*).

Although these developments are very promising and methods are constantly improving, a number of challenges still need to be overcome, when deriving epidemiological information from sequence data (*108*):

For these approaches observable genetic variation must occur in the same time period, in which epidemiological relevant processes happen (*108*). Serially collected samples must cover adequately long time periods to yield information and extensive sampling is required to recover sufficient genetic diversity. If necessary, resolution can be increased

by including further markers in the analysis, like insertions and deletions (indels) or plasmids, but more research is needed to fully exploit their use (*108*).

Another important issue is the time-dependency effect of age estimates (*108*, *118*). If the age of pathogen ancestors is derived from serially sampled data and the sampling period does not cover a significant part of the tMRCA, age estimates will systematically be underestimated. The exact reasons for this are still unclear, but mutational saturation at variable sites over long time periods and the registration of slightly deleterious mutations when sampling over short intervals are supposed to play a role for the dependency of age estimates from sampling intervals (*108*).

When working with bacteria it is important to test and account for homologous recombination, which plays a more important role for the genomic diversification of many bacterial species than *de novo* mutation (*108*). Recombination is not useful for the reconstruction of evolution in a Bayesian framework and could potentially even conceal an existing temporal signal (*108*). However, *B. anthracis* and *Mycobacterium tuberculosis* are thought to be clonal and recombination does not occur to a significant degree in these species (*119*, *120*).

Lastly; *B.anthracis* genotyping studies using MLVA and SNP typing methods have reported infection of the same hosts with different genotypes, which presents another challenge that needs to be accounted for (*102*, *121*). If analyses are based on one isolate per host, this could potentially obscure epidemiological transmission links (*108*).

Despite these potential pitfalls evolutionary analysis of bacterial sequence data in a Bayesian framework is a powerful tool to gain insights into disease epidemiology and ecology of bacteria. Care must be taken during the set-up of the analyses and the interpretation of the data.

3.6 Serological surveillance of anthrax in wildlife

3.6.1 *Assays for anthrax antibody detection in wildlife*

Anthrax is considered a disease with extremely high mortality and therefore anthrax surveillance was often done by carcass monitoring alone in the past (*122*). However, carcass detection is heavily biased by the detection probability of affected species and only covers infections with a lethal outcome (*123*). Serological approaches offer an alternative means of understanding disease ecology: in combination with carcass monitoring data, insights can be made about the susceptibility of different species to the disease.

Due to the chromosomal similarity between *B. anthracis* and *B. cereus* most serological tests for anthrax antibody detection are based on plasmid encoded antigens (*39*). WHO generally recommends enzyme-linked immunosorbent assays (ELISAs) measuring antibodies against PA and LF (*26*) as an inexpensive, high throughput surveillance tool.

Most anthrax ELISAs are based on the competitive inhibition indirect ELISA described by Turnbull et al. in 1986 (*124*). Many variations of the Turnbull ELISA exist today, but adapting the protocol for testing of wildlife samples still presents a challenge. Due to the lack of species specific antibodies ‘‘double sandwich’’ competitive inhibition indirect ELISAs were developed (*125*). Here antisera to the immunoglobulins of the

species studied raised in rabbits are used and detected by commercially available anti-rabbit antibodies. Alternatively, commercially available conjugated antisera of the closest relatives of the wildlife species are used in a ‘‘single sandwich’’ competitive inhibition indirect ELISA (*126*, *127*). Due to the close relationship of humans and primates, ELISA tests developed for the use with human secondary antibodies can be used for testing of primates (*128–130*).The use of non-species specific secondary antibodies potentially leads to a loss of sensitivity and specificity, but validation of existing test protocols is often impaired by lack of positive controls (*122*).

The ELISA cut-off value is usually determined as the mean of negative controls plus one to three times their standard deviation. However, for wildlife, negative controls are often only available from animals in captivity or from related species (*122*). In many cases, no negative controls are available at all and background absorbance values have been used as control values (*131*). In most assays determination of end-point titres is not quantitative and cut-off rules and therefore results can vary significantly (*86*).

Another method used in studies to detect antibodies against PA in wildlife is the commercial QuickELISA Anthrax-PA Kit immunoassay (Immunetics Inc., Boston, MA, USA) (*81*, *132*). It is based on two recombinant PA conjugates, one bound to streptavidin and one bound to horse radish peroxidase (HRP). The rPA conjugates are incubated with the tested serum and in the case of a positive sample, the two rPA conjugates bind at two different sites of the multivalent antibody. Binding to the microtiter plate is mediated by binding of the streptavidin bound rPA to the biotin-coated microtiter plate, while a colour change is caused by the reaction of HRP with chromogenic peroxidase substrate. The kit specific cut-off value is determined to detect ≈300 ng/ml of PA-specific antibody. This non-species specific assay would be suitable for testing and comparing various wildlife species and has been used for the investigation of anthrax exposure in wildlife in the Serengeti (*81*). However, the kit was developed for use with human samples under contract from the CDC to provide a standard anthrax test for public health laboratory use and its use with wildlife samples has not been formally validated yet.

3.6.2 *Anthrax antibody prevalence in African national parks*

It has been shown in animal experiments that herbivores are generally very susceptible to anthrax disease, while carnivores are resistant up to very high infectious doses (*26*).

In African national parks this is mirrored by high levels of seropositivity in most carnivore species (63-100 %) in combination with low fatality rates suggesting regular non-fatal exposure (*81*, *122*). Cheetahs (*Acinonyx jubatus*) present an exception with an antibody prevalence of only 4 % in a study from Botswana and several reports about fatalities (*81*, *92*). Lions (*Panthera leo*) and hyenas (Fam. Hyaenidae) are likely regularly exposed through consumption of infected prey or scavenging (*81*) and generally no deaths in these species are observed (*7*, *65*, *82*). However, high anthrax mortality in lions in KNP has been reported following a period of low anthrax incidence (*133*). This suggests that more frequent exposure may trigger a protective immune response in carnivores (*81*, *133*).

Studies examining spatial distribution of anthrax reported a correlation between seropositivity in carnivores and the anthrax burden in their ranges (*81*, *133*). Similar

observations were made for seropositivity in vultures and it has been suggested to use carnivore or vulture sero-surveillance as a tool to map anthrax risk zones (*127*).

On the other hand, relatively low seroprevalences (0-46 %) in combination with high mortality rates in herbivores suggest higher susceptibility to the disease (*81*, *122*). However, Turnbull et al. (1992) found anti-PA titers in cows in the UK (27 % of untreated, symptomless individuals in a herd where several cows had died of anthrax previously) and in three springbok (*Antidorcas marsupialis*) and a giraffe (*Giraffa camelopardalis)* in ENP, presenting a first hint for possible non-lethal anthrax infection in herbivores (*133*). Anthrax antibodies were also found in wildebeest (*Conochaetes taurinus mearnsi*) (19 %) and buffalo (*Syncerus caffer*) (46 %) in the Serengeti underlining that not all herbivores succumb to infection. Significant seropositivity was also observed in zebra (*Equus quagga*) in ENP and Canadian wood bison *(Bison bison athabascae)* (*86*, *134*). Cizauskas et al. (2014) postulated that sublethal infection could temporarily trigger a protective immune response in herbivores and cause the seasonal variation in anthrax incidence in ENP (*86*). It is still problematic that the classical anthrax assays do not differentiate between anthrax infection and subsequent recovery and low dose exposure with seroconversion (*122*, *127*).

Today it is clear that anthrax is not an “all or nothing” disease and a bigger effort is needed for systematic anthrax sero-surveillance in wildlife (*122*). Linking serological data with mortality records could offer important insights into the dynamics of the disease and the use of vultures and carnivores as sentinel species could be an efficient tool for anthrax risk mapping. Lembo et al. (2011) extended this approach and correlated anthrax seropositivity in dogs with anthrax incidence in villages surrounding the Serengeti, a monitoring approach that could be useful for anthrax surveillance throughout sub-saharan Africa (*81*).

4 Materials

4.1 Chemicals, media and buffers/gels

4.1.1 *Chemicals*

Ammonium persulfate solution (APS) 25 % (w/v)	Carl Roth GmbH, Karlsruhe
Acrylamid-/Bisaycrylamidsolution 30%	Carl Roth GmbH, Karlsruhe
Ampure Beads	Beckman Coulter, Brea, USA
Deoxynucleoside triphosphate (dNTP)	Invitrogen, Carlsbad, USA
Ethanol (RNase-free)	Carl Roth GmbH, Karlsruhe
Fly attractant	Unkonventionelle Produkte Feldner, Waldsee, Germany
Glycin	Carl Roth GmbH, Karlsruhe
H_2SO_4 (2M)	Carl Roth GmbH, Karlsruhe
Magnesium chloride (PCR)	Invitrogen, Carlsbad, USA
Methanol	Carl Roth GmbH, Karlsruhe
Phosphate buffered saline	Biochrom GmbH, Berlin
Primers and probes	TIB Molbiol, Berlin
	Invitrogen, Carlsbad, USA
SDS	Carl Roth GmbH, Karlsruhe
Skimmed milk powder	Carl Roth GmbH, Karlsruhe
TEMED	Carl Roth GmbH, Karlsruhe
TMB SeramunBlau fast	Seramun Diagnostics GmbH, Heidesee, Germany
TMB Seramun prec	Seramun Diagnostics GmbH, Heidesee, Germany
Tris	Carl Roth GmbH, Karlsruhe
Tris-HCl	Sigma-Aldrich, Taufkirchen
Tween 20	Carl Roth GmbH, Karlsruhe
Water (nuclease free)	Biochrom GmbH, Berlin

4.1.2 *Media*

Media were either produced at the RKI medium centre or purchased commercially. All in-house produced media were autoclaved at 121 °C for 20 min and stored at 4 °C for up to four weeks.

Luria Bertani (LB)-medium
10 g Bacto-Trypton
5 g Bacto-yeast extract
10 g NaCl
ad 1 l $H_2O_{bidest.}$; pH 7.0

Columbia blood agar
Commercially available at Oxoid, Wesel, Germany
Or in-house production:
39 g Columbia-blood-agar-base
23 g Pepton
1 g Starch
5 g NaCl
10 g Agar
ad 1 l $H_2O_{bidest.}$; pH 7.0; autoclave, leave cool to 50-55 °C, then add 50 ml off freshly defibrinated sheep blood

Blood-trimethoprim agar
Blood- agar as described above, plus (per liter agar medium)
1.6 mg trimethoprim
6.4 mg sulfamethoxazole
20 mg polymyxin B
Cereus Ident agar
Commercially available at Heipha Diagnostika, Eppelheim, Germany

4.1.3 *Buffers and gels*

4.1.3.1 Buffers

PCR Buffer

10x Run-buffer	Invitrogen, Carlsbad, USA

Electrophoresis buffer
25 mM Tris
192 mM Glycin
0.1 % (w/v) SDS
pH: 7.5

Blotting buffer
50 mM Tris
39 mM Glycin
0.037 % (w/v) SDS
20 % Methanol
pH: 7.5

ELISA washing buffer

Phosphate-buffered saline with 0.02 % Tween-20 (PBS-T), pH: 7.2

ELISA blocking buffer

5 % skimmed milk powder in PBS-T

ELISA dilution buffer

Same as blocking buffer

Western Blot washing buffer

Tris-buffered saline with 0.05 % Tween-20 (TBS-T), pH: 7.6

Western Blot blocking buffer

10 % skimmed milk powder in TBS-T

Western Blot dilution buffer

3 % skimmed milk powder in TBS-T

4.1.3.2 SDS-PAGE

Collection gel (ca 5 %) for 2-3 gels:
2.9 ml H_2O
1.25 µl 0.5 M Tris-HCL, pH 6.8
0.85 µl 30 % Acrylamid/Bisacrylamid solution
50 µl 10 % SDS
10 µl TEMED
15 µl 25 % (w/v) Ammonium persulfate solution

Separation gel (12.5 %) for 2-3 gels:

5.1 ml H_2O
3.8 ml 1.5 M Tris-HCL, pH 8.8
6 ml 30% Acrylamid/Bisacrylamid solution
150 µl 10 % SDS
20 µl TEMED
20 µl 25 % (w/v) Ammonium persulfate solution

4.2 Enzymes

Platinum® *Taq* DNA- Polymerase	Invitrogen, Carlsbad, USA
Proteinase K	Carl Roth GmbH, Karlsruhe

4.3 Antibodies and antigens

goat anti-human HRP labelled conjugate	Dianova, Hamburg
rabbit anti-goat HRP labelled conjugate	Dianova, Hamburg
Protective Antigen (PA)	Quadratech Diagnostics, Surrey, UK
Lethal Factor (LF)	Quadratech Diagnostics, Surrey, UK

4.4 Kits

DNeasy Blood and Tissue Kit	Qiagen, Hilden
GenMatrix stool DNA purification Kit	Roboklon, Berlin
Nextera XT DNA preparation kit	Illumina, San Diego, USA
Nextera XT Index Kit (96 indexes, 384 samples)	Illumina, San Diego, USA

4.5 Technical equipment

2100 Bioanalyzer	Agilent Technologies, Santa Clara, USA
ABI 7500 Cycler	Applied Biosystems, Foster City, USA
Biological Safety cabinet HeraSafe	Thermo Fisher Scientific, Waltham, USA
Centrifuges Labofuge 400e	Heraeus, Hanau
Centrifuge Heraeus Sepatech	Heraeus, Hanau
Centrifuge 5402 PA	Eppendorf AG, Hamburg
Compass	Globetrotter, Hamburg
Dissecting set (Scissors and forceps)	Carl Roth GmbH, Karlsruhe
Electrophoresis chamber	Biometra, Göttingen
Faceshield	Carl Roth GmbH, Karlsruhe
FastPrep®, cell disruptor	MP Biomedicals, Santa Ana, USA
GPS 62s	Garmin, Schaffhausen, Switzerland
HiSeq 1500	Illumina, San Diego, USA
Mini-Rocker-Shaker	SIA Biosan, Riga, Latvia

Nanodrop	Thermo Fisher Scientific, Waltham, USA
Pipettes	Eppendorf AG, Hamburg
Power supply gel electrophoresis	Biometra, Göttingen
Qubit® 2.0 Fluorometer	Invitrogen, Carlsbad, USA
Scalpel	Carl Roth GmbH, Karlsruhe
Semi-Dry Blotter	Phase, Lübeck
Sling Shot	Sherrill Tree, Greensboro, USA
Tecan Reader	Tecan Group Ltd., Männedorf, Switzerland
Table centrifuge	Eppendorf AG, Hamburg
Table-top scale	Sartorius, Göttingen
Stratagene Mx3000P Cycler	Thermo Fisher Scientific, Waltham, USA
Vortexer (Labdancer)	Carl Roth GmbH, Karlsruhe

4.6 Consumables

Cryotubes (1.2 and 2 ml)	Carl Roth GmbH, Karlsruhe
Cell culture flasks	Nunc, Wiesbaden
Centrifugal filters (0,22 µm)	Merck Millipore, USA
Ceramic beads (Ø 1.4 mm)	Precellys, USA
EDTA tubes	Carl Roth GmbH, Karlsruhe
Falcon tubes (15 and 50 ml)	TPP, Swiss
FFP3 masks	Carl Roth GmbH, Karlsruhe
Maxisorp 96 well plates	Sigma-Aldrich, Taufkirchen
Mikrobank Tubes	Mast Diagnostica, Rheinfeld, Germany
Parafilm	Carl Roth GmbH, Karlsruhe
PVDF Membrane	Merck, Darmstadt
Reaction tubes (0.5 ml, 1.5 ml, 2 ml)	Carl Roth GmbH, Karlsruhe
Silica gel beads	Carl Roth GmbH, Karlsruhe
Tyvek Suits	Carl Roth GmbH, Karlsruhe
Whatman filter paper 1,3 mm	A. Hartenstein GmbH, Würzburg

4.7 Disinfection

Necropsy equipment

Formalin 30-10% (contains ~ 11 - 4 % Formaldehyde), depending on grade of contamination. Bought locally in pharmacies in Côte d'Ivoire.

Laboratory use

Work with spores:

Wofasteril E400 (Peracetic acid)	Kesla AG, Bitterfeld

(freshly diluted 1:100 in water and stored at 4 °C for a maximum of 4 days)

Other:

5 % Lysoform d	Carl Roth GmbH, Karlsruhe
Terralin liquid	Schülke&Mayer GmbH, Norderstedt

4.8 Software

BEAST software package *v. 1.8.2* (*112*, *135*)

BEAGLE *v.2.* (*136*)

bwa version *0.7.12-r1039* (*137*)

FigTree *v. 1.4.0* (*135*)

Geneious Pro *v. 8.1.3* (*138*)

Genome Analysis Toolkit (GATK) *version 3.4* (*139*, *140*)

GeographicDistanceMatrixGenerator *v1.2.3* (*141*)

Google Earth *http://earth.google.com*

iTOL *v. 3.2* (*142*)

jModelTest *v. 2.1.4* (*143*)

LogCombiner *v.1.8.2* (*135*)

Microsoft Office Microsoft, USA

MxPro 4.0 Stratagene, USA

Path-O-Gen *v1.2* (*144*)

PhyML *v. 20131022* (*145*)

picard *tools v. 1.136* (*146*)

QGis 2.14.3 (*147*)

R *v.3.3.0* (*148*)

R package ape (*149*)

R package ecodist (*150*)

R package TipDatingBeast (*151*)

Seaview 4 (*152*)

Scythe *v. 0.993* (*153*)

Sickle *v. 1.33* (*154*)

Spread *v. 1.0.6* (*155*)

Tracer v. *1.6* (*156*)

Tree Annotator (*135*)

5 Methods

5.1 Overview sample origin

Necropsies on deceased wildlife from 1998–2015: A. Blankenburg, H. De Nys, A. Düx, P. Formenty, F. Leendertz, S. A. Leendertz, T. Löhrich, S. Metzger, K. Nowak, S. Schenk, F. Zimmermann

Collection of blowflies on the ground and in the canopy: F. Zimmermann

Collection of duiker serum samples: K. Nowak and F.Zimmermann

Bcbva qPCR positive bones and further Bcbva qPCR positive fly samples were received from K. Nowak and C. Hoffmann

Monkey serum and chimpanzee whole blood samples were received from H. De Nys, A. Düx, A. Lang, F. Leendertz and K. Nowak

All wildlife samples were collected with permission of the Ivorian Office of National Parks (permits Nr. 048/MESRS/DGRSIT/KCS/TM and 90/MESRS/DGRSIT/mo) and have been exported under permission of the according CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora). The study was approved by the Centre Suisse de Recherche Scientifique en Côte d'Ivoire and the Laboratoire National de la Pathologie Animale, Bingerville, Côte d'Ivoire.

5.2 Sample Collection

5.2.1 *Necropsies*

Continuous carcass monitoring has been performed at our long-term study site in TNP. After the first detected anthrax outbreak in 2001 (*1*), a perennial veterinary program was implemented as a collaboration of RKI and the primatology research program of the MPI. A veterinarian is permanently on site and performs a necropsy on every carcass reported by researchers working in the forest (Fig. 4). Tissue samples of all inner organs are taken, as far as the state of decomposition of the carcass allows. Necropsies follow a standardized protocol, including the use of full personal protective equipment (PPE) due to the known occurrence of anthrax, Ebola and monkeypox in the study area (*13*, *157*). After every necropsy the carcass site is decontaminated following World Health Organization (WHO) guidelines (*26*, *158*), involving burial of the carcass and incineration or disinfection with 10 % - 30 % formalin of all contaminated materials. Samples are collected in 2 ml cryotubes (Carl Roth) for freezing or in 50 ml Falcon tubes (TPP) for conservation in formalin. In the field samples are stored in liquid nitrogen or at ambient temperature in 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 from 1998 to 2015. In addition to the samples collected by RKI/ MPI veterinarians, we received tissue samples from 31 carcasses sampled by the WHO in TNP in the period from 1996 to 2000 (Table S1).

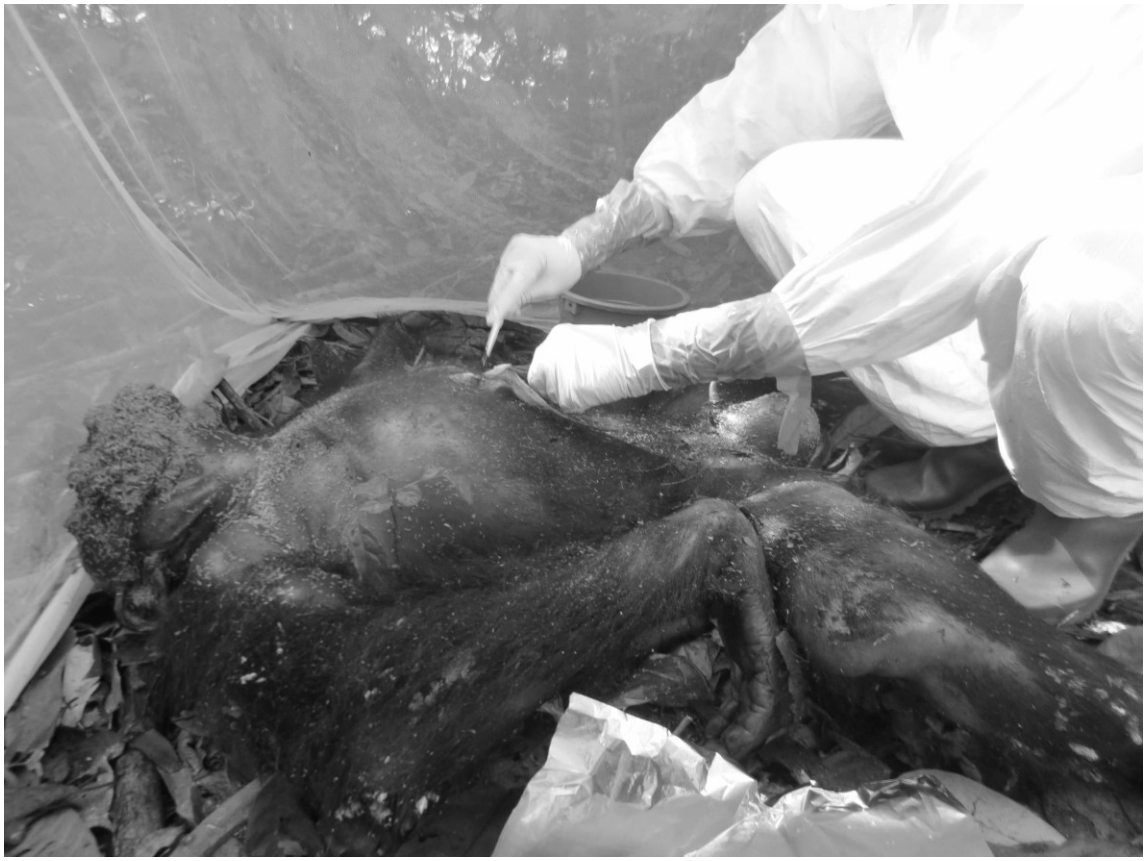

Figure 4. Necropsy on a *Bcbva* positive chimpanzee in TNP in 2013.

5.2.2 *Fly Trapping*

Carrion flies (Calliphoridae) were caught on the ground and up to 30 m high in the canopy using custom-made fly traps. Fly traps on the ground consisted of a pyramidal mosquito net hanging over a plastic container filled with bait. For canopy traps mosquito nets and bait container were hanging from a plastic disc and traps could be closed from the ground using a rope attached to the net (Fig. 5). To install traps in the canopy a big shot slingshot was used (Sherrill Tree).

Figure 5. Installation of a canopy fly trap with a big shot slingshot.

Flies were either baited with meat or commercially available fly attractant consisting of different animal proteins (Unkonventionelle Produkte Feldner). Plastic containers containing bait were covered with a net to avoid contamination of flies. Traps were operated for 20- 60 min depending on capture success and a maximum of 20 flies was collected per trap. Ether was used to euthanize flies and flies were stored at -20 °C in 2 ml cryotubes (Carl Roth) or at ambient temperature in 50 ml Falcon tubes (TPP) containing silica beads (Carl Roth).

In the course of this study a total of 103 flies from the canopy were caught and analyzed (Table S2) and 204 flies caught on the ground in spring 2013 were tested in direct culture in the field without preceding PCR testing (marked as "fly direct" in Table S4).

Additionally 1531 "ground flies" from TNP caught following different sampling schemes were tested in qPCR during the PhD project of C. Hoffmann (623 flies were randomly collected within the research area in the years 2008, 2009, 2012 and 2013. Another 908 flies were caught in a "fly snapshot" following a 2x2 km grid system within 19 days in May and June 2014). 43 flies from this study that were tested positive in qPCR and contained high *pagA* copy numbers were included in bacterial culture and genomic studies of this work (Table S3). Further, two qPCR positive flies from Grebo, Liberia, were received from C. Hoffmann and included in the present study (Table S3).

5.3 Nucleic acid extraction

The DNeasy Blood and Tissue Kit (Qiagen) was used for extraction of DNA from necropsy samples. Following the manufacturer's instructions DNA was extracted from several tissues per animal (liver, spleen and lung, if available).

For flies the GeneMATRIX Stool DNA Purification Kit (Roboklon) was used and homogenization steps were added before running the standard protocol to account for the tough chitin exoskeleton of flies. First, flies were cut into pieces with sterilized scissors and further homogenization was achieved using a Fast Prep® machine (MP Biomedicals).

Quantification of DNA extracts was done using a Nanodrop spectrophotometer (Thermo Scientific) and aliquots were stored at -20 °C.

5.4 Real-time PCR assay for *Bcbva* detection

For all sample types testing for *Bcbva* was done with three qPCR assays targeting two markers on both plamids and a chromosomal marker (Table 1). The presence of plasmid pXO1 is tested for with gene marker *pagA* (gene for protective antigen) and plasmid pXO2 is targeted with gene marker *capB* (gene for capsule formation). While *pagA* and *capB* can be found in *B. anthracis* and *Bcbva*, the chromosomal Island IV is specific for *Bcbva* and can be targeted to differentiate the two pathogens. All samples were first tested for *pagA* and samples positive for *pagA* in duplicate were tested for *capB* and Island IV for confirmation.

The 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 hydrolysis probe (Table 1), 0.5 U Platinum® Taq Polymerase (Invitrogen) and PCR water. 200 ng of template DNA or 5

µl of DNA extract (if DNA concentration was below 40 ng/µl) were added and the total volume of each reaction was 25 µl.

Cycling conditions were: 600 sec at 95 °C and 45 cycles of 15 sec at 95 °C and 30 sec at 60 °C, using a Stratagene qPCR MX3000 cycler (Stratagene). The MXPRO software was used for evaluation of fluorescence signals.

Table 1. Primer and probe sequences and annealing temperatures for *Bcbva* qPCR. *PagA* and *capB* are markers for both *B.anthracis* and *Bcbva*. Island IV is a chromosomal genetic island specific for *Bcbva*.

Primer name	Sequence (5′-3′)	Quencher	T_a (°C)	Amplicon size (bp)	Reference
pagA - for	CggATCAAgTATATgggAATATAgCAA		60	204	Ellerbrok et al. 2002
pagA - rev	CCggTTTAgTCgTTTCTAATggAT		60		Ellerbrok et al. 2002
pagA - TM	TEX-CTCgAACTggAgTgAAgTgTTACCgCAAAT	BBQ			Ellerbrok et al. 2002
capB - for	gggAAAACAACTggTACATCTgC		60	144	Antonation et al. 2016
capB - rev	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

5.5 Bacterial culture

Bacterial culture was done for most qPCR positive necropsy samples and a subset of qPCR positive fly samples. Further, *Bcbva* positive bone powder samples dating back to 1991 were received from K.Nowak and used for bacterial culture.

5.5.1 *Necropsy samples*

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 (Table S1). Material used for culture had been stored in liquid nitrogen in the field and subsequently at -80 °C. 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), blood-trimethoprim agar and Cereus Ident agar (Heipha Diagnostica)(*5*). Cultures were incubated at 37 °C and monitored daily. Morphologically suspicious colonies were sub-cultured and tested in real-time PCR as mentioned above. Isolates were frozen in Microbank tubes (Mast Diagnostica) at -80°C.

5.5.2 *Flies*

A subset of 52 flies containing high *pagA* copy numbers was chosen for bacterial culture (Table S2 and S3). Flies had either been stored at ambient temperature on silica beads or in liquid nitrogen/ at -80 °C. 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 then plated.

Further, in an on-site study for flies from TNP direct culture from flies without preceding PCR testing was used as a diagnostic tool. Here, 204 flies were homogenized and the complete material was plated directly onto Cereus Ident agar and incubated at ambient temperature in the TNP field laboratory. Suspicious colonies were sub-cultured on blood- trimethoprim agar and tested in real-time PCR (marked in Table S4 as “fly direct”). All field work was performed in a glove box and all agar plates were burnt after soaking in 10 % Formaldehyde over night.

5.5.3 *Bones*

In TNP bones have been collected by the Taï Chimpanzee Project (TCP) and Taï Monkey Project (TMP) since 1989. Bone powder was produced by drilling deep into bones with a sterile drill to minimize contamination and possibly gain bone marrow material. DNA extraction and qPCR of 75 bones done by K. Nowak dated the occurrence of *Bcbva* back to 1991. Culture was attempted for all PCR positive bones. Powder from qPCR positive bones was homogenized in sterile NaCl and then processed as described above for necropsy samples with one native aliquot and one heat-treated aliquot. Bones had been stored at ambient temperature since collection.

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

5.6.1 *DNA extraction*

Table S4 contains a complete list of all *Bcbva* isolates sequenced for this study. 5 ml of Luria- Bertani (LB) liquid medium were inoculated with one *Bcbva* colony from Columbia blood agar (Oxoid) and incubated overnight at 37 °C. To prevent sporulation, a fresh 5 ml LB liquid culture was inoculated with 200 µl of the overnight culture the following morning. After incubation for four hours, 3 ml of culture were centrifuged and DNA was extracted from the pellet using the DNeasy Blood and Tissue kit (Qiagen). Eluates were centrifuged at 12 000 g for 4 min using centrifugal filter units (0.22 µm pore size, Merck Millipore) to ensure sterility. 1/10 of the resulting DNA eluate was incubated in LB medium and on Columbia blood agar for seven days testing for sterility before DNA was taken out of the BSL3 laboratory. DNA concentrations were determined using the Qubit® 2.0 Fluorometer (Invitrogen).

5.6.2 *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) and normalized them for sequencing. Libraries were pooled and sequenced on the HiSeq 1500 platform (Illumina) in rapid run mode using either v1 (2 x 150bp) or v2 (2 x 250bp) 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.

Raw reads for the *Bcbva* isolates from the *in vitro* mutation rate experiment are available in the European Nucleotide Archive (ENA) under the project accession number PRJEB14597, sample accession numbers ERS1222218 to ERS1222233.

5.6.3 *Bioinformatic pipeline*

Illumina adapters were removed using scythe v0.993 (*153*) and raw reads were trimmed with sickle v1.33 (*154*) applying a quality threshold of 25. The quality trimmed reads were aligned to the reference genome (*Bcbva* strain CI, Accession numbers CP001746-CP001749) with the BWA-MEM algorithm as implemented in bwa v0.7.12-r1039 (*137*). For the conversion to bam format, sorting, deduplication and indexing of aligned reads, the picard tools 1.136 (*146*) software package was used applying the commands *SortSam, MarkDuplicates* and *BuildBamIndex*. Subsequent variant calling was done with the Genome Analysis Toolkit (GATK) v3.4 (*139*, *140*). In a first step, bam files were realigned with the tools *RealignerTargetCreator* and *IndelRealigner*. Variants were called with the *UnifiedGenotyper* and a minimum phred scaled confidence threshold of 30 for SNPs to be called. Hard filtering of the resulting SNP sites was done with the *VariantFiltration* command using the 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 < 5 x, a minor allele frequency of > 0.1 and a GATK Genotype Quality value < 99. Final consensus sequences were composed with the *FastaAlternateReferenceMaker*. Coverages of all samples were assessed with the GATK tools *DepthOfCoverage* and *CoveredByNSamplesSites*.

5.6.4 *Within-host diversity testing*

Several bacterial colonies from the same host (mammal or fly) were sequenced for seven mammals and eight flies to test for within-host diversity. For four mammals, colonies originated from separate organs.

5.7 *Bcbva in vitro* mutation rate analysis and latency estimate

The *in vitro* mutation rate of *Bcbva* was determined in a sequential passaging experiment, where 15 subcultures derived from the same starting culture were passaged 24 times and full genomes of starting culture and subcultures were sequenced to detect newly introduced SNPs. As a prerequisite for this study, the generation time of *Bcbva* was examined and the maximum concentration of bacteria reached in LB liquid medium in the stationary phase was determined:

For the determination of *Bcbva* generation time two liquid Luria-Bertani (LB) cultures were inoculated with *Bcbva* spores in concentrations of 10^1 and 10^2 CFU/ml, respectively. The number of CFUs was identified every 2 hours for a total of 10 hours and all time points falling into the exponential growth phase were used to derive an estimate for the generation time of 29.1 min (95 % CI: 21.8 – 43.5).

To ensure that the *Bcbva* cultures persisted in the stationary phase at the time of passaging, two liquid cultures were inoculated with concentrations of 10^3 CFU/ml and their concentrations (in CFU) were measured every two hours between 14 and 24 hours after inoculation. Concentrations stayed on the same level and did not start declining. Finally the exact bacterial concentration in the stationary phase was determined by inoculating 12 liquid LB cultures with a concentration of 10^3 CFU/ml *Bcbva* and determining CFU numbers after 24 hours. Under the conditions in this experiment, the bacteria grew very homogenously and reached a concentration of 2.4 x 10^8 CFU/ml in average.

In the actual passaging experiment one *Bcbva* colony from a Columbia blood agar plate was used to inoculate 15 ml of LB medium and this was left to incubate over night at 37 °C reaching a concentration of 2.4 x 10^8 CFU/ml. The DNA of the starting culture was extracted and the full genome was determined as described above. 15 x 15ml LB medium were inoculated with the starting culture at a concentration of 2.4 x 10^3 CFU/ml. Cultures were incubated at 37 °C and left to grow for 24 hours reaching the stationary phase with a concentration of 2.4 x 10^8 CFU/ml in average as determined above (equaling 16.6 generations) and then diluted again to a concentration of 2.4 x 10^3 CFU/ml. This was repeated for 24 passages in total, equaling 5976 generations of continuous bacterial growth for the 15 cultures over the 24 passages. After these passages, 10 µl of each liquid culture were plated onto Columbia blood agar (Oxoid) and incubated overnight. To avoid sequencing of mixed cultures, DNA extraction and

whole-genome sequencing were performed for one colony per culture, resulting in 15 full genomes to be compared to the genome of the starting culture. To the number of generations from the sequential passage (5976, see above) a number of generations was added accounting for the additional passaging steps during DNA extraction of the 15 isolates. This was approximated using *Bcbva* generation time (determined to be 29.1 min, see above) and resulted in a total of 6624 generations of bacterial growth.

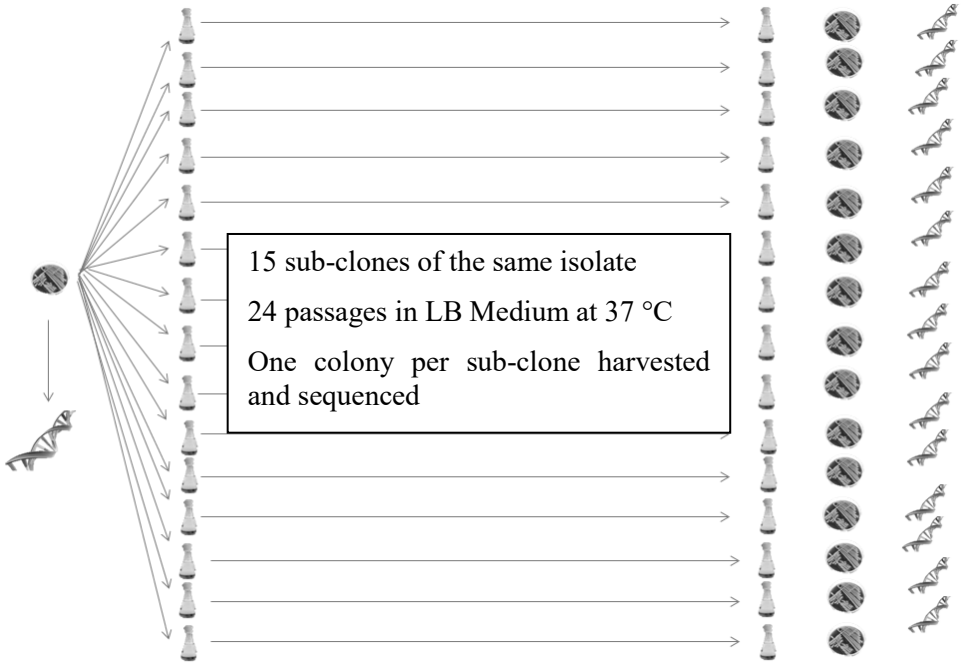

Figure 6. Illustration of *in vitro* mutation rate experiment. The full genome sequence of the starting culture was determined and then 15 sub-clones were passaged 24 times with one passage (with growth from 2.4 x 10^3 CFU/ml to 2.4 x 10^8 CFU/ml) representing in average 16.6 generations. In total bacteria grew over 5976 generations and one colony of each subculture was sequenced afterwards to compare the numbers of SNPs before and after passaging.

5.8 Phylogenetic and statistical analyses

5.8.1 *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 S4) were aligned and stripped of non-variant sites with *Geneious Pro* v8.1.3 (Biomatters ltd.) (*138*). The resulting alignments of variant sites were 298, 18 and 11 bp long for the chromosome, pXO1 and pXO2, respectively. Analyses were performed on chromosome and plasmid level, but as sufficient genetic information for further inference was only found on the chromosome, in depth analyses focus on the chromosome. *jModelTest* v2.1.4 (*143*) was used for determination of the best nucleotide substitution model in a Maximum Likelihood framework, resulting in the choice of TVMef for chromosomes from TNP with no proportion of invariant sites or rate variation among sites.

5.8.2 *Maximum Likelihood analysis*

Maximum Likelihood analysis was performed with PhyML v20131022 (*145*) implementing the selected substitution model and a combination of subtree-pruning-regrafting (SPR) and nearest-neighbor-interchange (NNI) algorithms. Branch support was estimated using non-parametric bootstrapping with 100 replicates. The tree was rooted using the "best-fit" option in Path-O-Gen v1.2 (*144*), placing the root at the position resulting in the most clock-like structure of the data.

5.8.3 *Bayesian Analyses*

5.8.3.1 Time-scaled phylogenetic tree reconstruction using BEAST

Bayesian Markov Chain Monte Carlo (BMCMC) sampling as implemented in the *BEAST* software package v1.8.2 (*135*) was used to estimate a time-scaled phylogenetic tree from TNP and Grebo sequences. Exact tip dates could be assigned for 121 out of 126 samples. If exact sampling day or month were unknown, the middle of the month/year was set as date. The TVMef nucleotide substitution model was applied as determined above with no invariant sites or rate variation among sites. As the alignment contained only variable sites, the number of constant sites was specified in the xml file to adjust clock rate estimates. Several combinations of different clock models and coalescent tree priors were tested and Bayes factor (BF) comparison of their log marginal likelihoods estimated by stepping stone sampling was done following Kass & Raftery's criteria (*159–161*). For each combination two independent MCMC chains were run for 100 Mio states sampling every 10000th state, excluding a 10 % burn-in. *Tracer* v1.6. (*156*) was used to verify convergence of chains and a sufficient effective sampling size (> 200) and runs were combined with *LogCombiner* v1.8.2. (*135*).

To ensure that detected signals originated in the data, analyses were repeated with randomized dates (n=20) as suggested by Duchêne et al. (2015) (*162*). Dates were reassigned to sequences using the *RandomDates* function of the *TipDatingBeast* package v1.0-2 (*151*) in R v3.3.0 (*148*).

Tree figures were produced with FigTree v1.4.0 (http://tree.bio.ed.ac.uk/software/figtree/).

5.8.3.2 Ancestral state reconstruction of host transmission

To investigate the vector potential of flies for *Bcbva,* Bayesian ancestral state reconstruction was performed, using discrete phylogenetic diffusion models, as implemented in BEAST v1.8.2 (*135*) in combination with the BEAGLE v2.1 library (*136*) for accelerated likelihood evaluation. Four transmission scenarios were tested, represented by different models following Mather et al. (2013) (*117*): symmetrical transmission, asymmetrical transmission, unidirectional transmission from mammals to flies and vice versa. For this analysis, a discrete trait was added to the dataset classifying the hosts either as "Fly" (n= 65) or "Mammal" (n=61). For the two unidirectional models respective rates were set to zero through their associated indicators.

5.8.3.3 Continuous phylogeography

The phylogeographic history of *Bcbva* in TNP was investigated in a Bayesian framework using four different continuous diffusion models. To correct for genetic and spatial autocorrelation, strains 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 were excluded from the data set. Only one strain was kept per outbreak or fly catching point, the selection criterion being high average coverage of the genome (Table S4). The resulting alignment contained 77 strains and 225 variant sites. Bayesian analysis using BEAST v1.8.2 was set up as described above with a constant population tree prior and an uncorrelated lognormal relaxed clock. GPS data was added as continuous trait and a 0.0001 jitter was applied. A Brownian diffusion model assuming homogenous dispersal was tested as well as three Relaxed Random Walk (RRW) models allowing for branch specific variation in dispersal rates (*116*).

5.8.3.4 Time-scaled phylogenetic tree reconstruction on plasmid level

We attempted to also derive temporal information from the plasmids pXO1 and pXO2. Alignments of variant sites were 18 and 11 bp long for pXO1 and pXO2, respectively. Best- fitting nucleotide substitution models determined by jModelTest v2.1.4 (*143*) were K80 (*163*) and JC (*164*) and Bayesian analyses were run in BEAST v1.8.2 as described above.

5.8.3.5 Time-scaled phylogenetic tree reconstruction for sub-Saharan Africa

The data set used to date *Bcbva* presence in sub-Saharan Africa included the chromosomal sequences of *Bcbva* cases from Central African Republic, Cameroon and Côte d'Ivoire (*2*, *6*) and chromosomal sequences of the two new isolates from Liberia (flies) (Table 2). The alignment was compiled as described above and contained 1016 variable positions. Model selection with jModelTest v2.1.4 (*143*) resulted in the use of TPM1 (*165*) as nucleotide substitution model. We specified a constant population coalescent tree prior, assumed an uncorrelated lognormal relaxed molecular clock (*113*) and used the evolutionary rate derived from TNP as prior.

Table 2. *Bcbva* strains from sub-Saharan Africa used for Bayesian inference of the sub-Saharan tMRCA.

Strain	Country	Species	Accession number
CI	Côte d'Ivoire	*Pan troglodytes*	SAMN03610233
CAM	Cameroon	*Pan troglodytes*	SAMN03610234
A-363/2	Central African Republic	*Loxodonta africana*	SAMN03610235
A-364/1	Central African Republic	*Gorilla gorilla*	SAMN03610236
505-4	Liberia	Calliphoridae	This work
505-6	Liberia	Calliphoridae	This work

5.8.4 *Correlation of genetic and geographic distances*

To learn more about the spatial dynamics of *Bcbva* in TNP, the correlation between genetic and geographic distances was investigated. The cleaned data set described for continuous phylogeographical analyses (see 5.8.3.3) was used to correct for genetic and spatial autocorrelation. Geographic distances (in km) were derived from GPS data using GeographicDistanceMatrixGenerator v1.2.3. (*141*). Genetic distances were approximated using the relative distances drawn from a Maximum Likelihood Tree built in PhyML v20131022 (*145*) with the *R* package ape (*149*) using the *cophenetic* function. Multiple regression on distance matrices (MRM) as implemented in the *R* ecodist package (*150*) using 1000 permutations and Spearman correlation was performed on genetic and geographic distance matrices.

To examine variation within genetic lineages, data was binned by genetic distance (bin size= relative genetic distance of 0.03, approx. 2.5 SNPs) with a focus on groups with low genetic distance (max relative genetic distance < 0.5) and their mean geographic distance. Homogeneity of variance between groups was assured with the Fligner Killeen test ($p=0.07$; > 0.05 as requested).

5.9 Serological analyses

5.9.1 *Serum collection*

5.9.1.1 Duikers

Sampling of Maxwell's duikers was performed in 2013 and 2014 (Table S6). A total of nine duikers were trapped using the night-time net capture technique described by Newing (*166*), which is based on the animals freezing when stunned with a strong flashlight. Following capture with the net capture technique, animals were initially anesthetized using medetomidine (0.1 mg/kg), ketamine (4.0 mg/kg) and midazolam (0.1 mg/kg). Because the level of anesthesia reached with this combination of

anesthetics was rather deep for our needs, dosages were adapted to 0.07 mg/kg medetomidine and 2.5 mg/kg ketamine with slightly increased midazolam (0.17 mg/kg). This combination yielded appropriate anesthetic depth for our sampling needs. Induction took 5-10 min and animals were anesthetized for 30-50 min. Samples were collected following a standardized protocol, including drawing blood and taking nasal, oral, genital, anal and cutaneous swabs. Blood samples were collected in EDTA coated tubes (Carl Roth). After antagonization with atipamezole, animals fully recovered within 20-90 minutes. Blood samples were centrifuged at 3000 rpm for 10 min upon return to our forest laboratory, separated into plasma, buffy coat and erythrocytes and subsequently stored in liquid nitrogen. Samples were transported on dry ice and conserved at -80 °C for long-term storage. A large well-trained team is needed for this method for capturing duikers and further optimization might be possible (e.g., combining dazzling with a flashlight with GPS marked darts for anesthesia).

5.9.1.2 Chimpanzees and Monkeys

One chimpanzee serum sample taken under anesthesia during surgery in a chimpanzee was available from 2009 (*167*). A further eight whole blood samples were obtained during necropsies of freshly deceased animals that died in outbreaks of respiratory disease in 2004, 2006 and 2009 (*168*) (Table S6).

Samples of red colobus monkeys, black-and-white colobus monkeys and sooty mangabeys were collected by RKI veterinarians (see above) between 2006 and 2015, resulting in a total of 41 monkey samples (Table S6).

5.9.2 *Test approach*

No standardized approaches are available to investigate anthrax seroprevalence in wildlife. Thus, samples were tested for antibodies against the anthrax protective antigen (PA, List Biological Laboratories Inc) using an in-house ELISA and an in-house Western Blot and for antibodies against anthrax lethal factor (LF, List Biological Laboratories Inc) using an in-house Western Blot. Assays are described in detail below.

Testing for PA and LF does not allow for a discrimination between classical *B. anthracis* and *Bcbva*, as both pathogens produce the typical anthrax toxins (*5*). However, in our extensive carcass monitoring all 81 anthrax cases in TNP were exclusively caused by *Bcbva* (*4*). Therefore, in the TNP ecosystem antibodies generated against PA and LF likely originate from exposure to *Bcbva*.

Human positive and negative controls were used for all monkey and chimpanzee assays as no species-specific controls were available. Negative controls were selected from a set of available human sera of unvaccinated donors, which were unreactive in Western Blot against PA and LF. Two of these sera that were representative of the range for PA-negative human sera in PA-ELISA were chosen as negative controls and included on each ELISA test-plate under the same conditions as the samples for inter-plate comparison. A positive control serum from a PA-vaccinated human donor was included as an 8-step $\log_2$ serial dilution curve (starting concentration: 1 in 4000) with repetitious reactivity and accurate results on every test-plate, yielding a titre of 512000.

For duiker assays a negative control was available from a red duiker (*Cephalophus natalensis*) from a zoo (*kindly provided by Berlin zoo*), which was unreactive in PA-

and LF- Western Blot. A pool of goats vaccinated with the *B. anthracis* Sterne spore live vaccine (*60*) (*kindly provided by W. Beyer*) was used as a positive control in the same fashion as stated for the positive human control (starting concentration: 1 in 1000), yielding a titre of 128000. No specific conjugated antibodies were available for any of the species tested. For primate samples, polyvalent goat anti-human horseradish peroxidase (HRP) labelled conjugate (H+L) (Dianova) was used, as described previously (*128*, *169*) . For duikers we tested the reactivity of duiker serum with different commercially available conjugates from the Bovidae family (sheep, cow, goat) in a comparative dot blot approach with logarithmic duiker serum dilutions starting at 1:10. We found that rabbit anti-goat HRP conjugate (Dianova) was the most suitable commercially available conjugate for duiker samples.

5.9.3 *PA- ELISA*

PA ELISA testing was performed by S. Köhler, but the method is described here for the sake of completeness.

PA- ELISA was performed as described by Hahn et al. (2004), with slight modifications (*129*, *130*). Briefly, each well of high-binding microtiter plates (Carl Roth) was coated with 0.1 µg of recombinant PA (Quadratech Diagnostics) in PBS at 4 °C over night. Wells were washed with phosphate-buffered saline containing 0.02 % (v/v) Tween 20 (Carl Roth) (PBS-T) and blocked with 5 % skimmed milk powder (Carl Roth) in PBS-Tween. Samples and negative controls were diluted 1:500 in blocking solution and incubated in duplicate together with the positive controls for 2 hours at room temperature. Secondary antibodies were used in a concentration of 1:10 000 and 1:4000 (previously evaluated and adjusted) for humanoid and duiker assays respectively. Plates were developed in the dark with 100 µl of TMB SeramunBlau fast (Seramun Diagnostics GmbH) substrate per well for 10 min and stopped with 100 µl H_2SO_4 (2M) (Carl Roth). Absorbance was measured at 450 nm (Reference wavelength 620nm) using a Tecan Sunrise 96-well-reader (Tecan Group Ltd.). The mean of the negative controls plus two times their standard deviation was set as cut-off value for each plate (Table S6).

5.9.4 *Western Blot*

Western Blot testing was performed by S. Dupke and F. Zimmermann.

For the PA- and LF-Western Blot assay, 380 ng of purified recombinant PA or recombinant LF (Quadratech Diagnostics) diluted in 125 µl of phosphate-buffered saline (PBS) were blotted onto a Immobilion-P PVDF-Membrane (Merck) after running in a 12,5 % SDS page. After blocking of the membrane with blocking buffer containing 10 % milk powder in tris-buffered saline with 0.05 % Tween (TBS-T), 5mm stripes were cut from the membrane and samples and controls were added in a dilution of 1:1000 in TBS-T and 3% powdered milk (dilution buffer). Samples were incubated at room temperature for one hour. Goat anti-human HRP conjugate (Dianova) was added to primate samples and human controls in a 1:10 000 dilution in dilution buffer (1:8000 for LF Western Blot). For duiker samples and goat/duiker controls, rabbit anti goat HRP (Dianova) was diluted 1:4000 in dilution buffer (1:8000 for LF Western Blot). The conjugate was left to incubate for one hour. Reactions were detected with precipitating

peroxidase substrate TMB SeramunBlau prec (Seramun Diagnostics GmbH) after 10 min of incubation.

6 Results

6.1 Necropsy samples

6.1.1 *PCR*

Out of the 204 carcasses sampled and tested between 1996 and 2015, 81 were found to be positive for *Bcbva* in qPCR (Table S1). A sample was defined as positive if it was positive for *pagA* in duplicate. All *pagA* positive samples were confirmed in qPCR for *capA* and *IslandIV*, therefore representing *Bcbva* and not *B. anthracis*.

6.1.2 *Culture*

Culture was attempted for all qPCR positive necropsy samples collected until the end of 2013 if appropriate material was available (for duiker necropsy samples until June 2014). *Bcbva* isolates could be retrieved from 55 out of 60 samples (Table S1). The five samples that failed culture were all aged samples with reduced sample quality. *Bcbva* could be cultured from native and heat-treated samples indicating the presence of heat-resistant spores. Culture was possible from various organs and tissues, including spleen, lung, liver, muscle and skin (Table S4) For carcasses in an advanced state of decomposition, maggots were found to be a suitable source for *Bcbva* culture.

6.2 Flies

6.2.1 *PCR*

Of the 103 flies caught in the canopy in spring 2014, 12 were tested positive for *Bcbva* in qPCR (11.7 %, Table S2, positivity defined as above for necropsy samples). In C. Hoffmann's study in randomly caught ground flies between 2008 and 2013, *Bcbva* prevalence was found to be 4.4 % (n= 1531, pos= 68) and a subset of those was included in the culture/genomic analyses in this work (Table S3).

6.2.2 *Culture*

Culture was attempted for 52 qPCR positive carrion flies in total (5 canopy flies, 45 TNP ground flies, 2 Grebo flies). Isolates were retrieved from 43 carrion flies in total (5 canopy flies, 36 TNP ground flies, 2 Grebo flies) (Table S2 and Table S3). Successfully cultured flies contained in average 387 copies of *pagA* per *5* µl of DNA extract). Culture of *Bcbva* from flies caught in forest strata up to 25 m high demonstrated the vector potential of carrion flies regarding infection of arboreal monkeys. *Bcbva* could be retrieved from native and heat-treated samples indicating the presence of heat-resistant spores in or on flies (preparation of flies did not allow to differentiate location of spores).

From April to June 2013 direct culture in the field was attempted from 204 flies during a pilot study. With this direct culture approach, an additional 21 *Bcbva* isolates were retrieved (marked in Table S4 as “fly direct”).

6.3 Bones

Bcbva isolates could be retrieved from 7 out of 26 PCR positive bones (positivity defiend as above for necropsy samples), the oldest one from a chimpanzee bone found in 1996 (Table 3).

Table 3. PCR positive bones received from K. Nowak, their results for culture and their isolate names in Whole Genome Sequencing (if applicable)

Species	Year	Culture	WGS
Pan troglodytes verus	1991	negative	
Pan troglodytes verus	1992	negative	
Pan troglodytes verus	1994	negative	
Pan troglodytes verus	1994	negative	
Pan troglodytes verus	1994	negative	
Pan troglodytes verus	1996	positive	529
Pan troglodytes verus	1998	negative	
Pan troglodytes verus	1999	negative	
Cercopithecus campbelli	2000	negative	
Pan troglodytes verus	2000	negative	
Pan troglodytes verus	2001	positive	503-1
Pan troglodytes verus	2002	negative	
Procolobus badius	2002	negative	
Colobus polykomos	2003	positive	530
Cercocebus atys	2004	positive	531
Pan troglodytes verus	2004	negative	
Pan troglodytes verus	2004	negative	
Pan troglodytes verus	2004	negative	
Pan troglodytes verus	2009	Not done, isolate available from tissue 140-1	
Pan troglodytes verus	2009	Not done, isolate available from tissue 92-4	
Cephalophus sp.	2013	positive	552
Cephalophus sp.	2013	positive	553
Cephalophus sp.	2013	positive	554
Cephalophus sp.	unknown	negative	
Pan troglodytes verus	unknown	negative	
Pan troglodytes verus	unknown	negative	

6.4 *Bcbva* Diversity

6.4.1 *Total Bcbva diversity in the Taï region*

In total 178 *Bcbva* isolates from TNP and Grebo were sequenced in the course of this study, including isolates from necropsy samples, bones and flies (Table S4). Some of the isolates originated in the same host to test for within-host diversity, but 126 of the isolates represent independent hosts, covering a period from 1996 to 2014. Average genome coverage of samples ranged from 18 to 394 x. Whole genome sequencing with this deep coverage was chosen to allow for the use of SNPs as stable phylogenetic marker and receive the highest possible resolution. As *B. anthracis* is know to be a very monomorphic bacterium, it was an interesting finding to detect relatively high diversity in *Bcbva* strains from the Taï region. Analyzing the 126 chromosomal sequences from separate hosts (mammals and flies) a total of 298 SNPs was detected on the 5196054 bp long *Bcbva* chromosome (in our analyses > 99% of the whole *Bcbva* genome were sequenced with a coverage of at least 5x). Of these SNPs 176 were parsimony-informative; a parsimony-informative SNP being a SNP that occurs in more than one isolate and therefore contains phylogenetic information. The maximum distance between any two chromosomes was 69 SNPs (median distance: 26 SNPs). These isolates originated in flies caught in two consecutive years only 6 km apart. Due to the smaller size of the two plasmids, less SNPs were located here, but significant diversity was still detectable. 18 (7 parsimony informative) and 11 (6 parsimony informative) SNPs were determined on pXO1 and on pXO2, respectively.

A limited number of samples was handled at a time and the work space was decontaminated with bleach and UV light between batches to avoid cross-contamination and mix-up of samples. Excluding all SNPs with a minor allele frequency of < 0.1 in our bioinformatics pipeline would also exclude artifacts due to slight cross-contaminations.

All alignments of variable sites are deposited in the Dryad database:
Doi: 10.5061/dryad.v8bn7

6.4.2 *Within-host diversity*

Within-host diversity of *Bcbva* isolates was tested on a subset of seven carcasses and eight flies. If available, isolates from different organs were chosen for carcasses to detect the highest existing amount of diversity and two to six isolates were sequenced per host (Table 4). A maximum difference of two SNPs between isolates for necropsies as well as flies was detected on chromosomal level (mean 0.35 SNPs). 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 had differed morphologically with one especially mucous colony. For each host, the isolate with the highest sequencing coverage and smallest number of autapomorphic SNPs (=SNPs that exclusively occur in one isolate and none of the other isolates) was included in further analyses with the exception of fly isolates 600-1 and -2, where both strains were analyzed.

Table 4. Within-host diversity. 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/pXO1/pXO2)					
				460	529				
Pan troglodytes 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	507-5	507-6
Cercopithecus diana 06.04.2012	**507-1**	**liver**	**135**	**-**	**0**	**0**	**0**	**0**	**0**
	507-2	liver	53	0	-	0	0	0	0
	507-3	liver	52	0	0	-	0	0	0
	507-4	spleen	107	0	0	0	-	0	0
	507-5	spleen	40	0	0	0	0	-	0
	507-6	spleen	61	0	0	0	0	0	-
				508-1	508-2	508-3	508-4	508-5	508-6
Cephalophus sp. 02.12.2013	508-1	liver	50	-	0	0	0	0	1/0/0
	508-2	liver	51	0	-	0	0	0	1/0/0
	508-3	liver	67	0	0	-	0	0	1/0/0
	508-4	spleen	36	0	0	0	-	0	1/0/0
	508-5	**spleen**	**89**	**0**	**0**	**0**	**0**	**-**	**1/0/0**
	508-6	spleen	53	1/0/0	1/0/0	1/0/0	1/0/0	1/0/0	-
				509-1	509-2	509-3	509-4	509-5	509-6
Cephalophus sp. 29.12.2013	**509-1**	**liver**	**174**	**-**	**0**	**1/0/0**	**0**	**0**	**0**
	509-2	liver	113	0	-	1/0/0	0	0	0
	509-3	liver	62	1/0/0	1/0/0	-	1/0/0	1/0/0	1/0/0
	509-4	spleen	63	0	0	1/0/0	-	0	0
	509-5	spleen	86	0	0	1/0/0	0	-	0
	509-6	spleen	17	0	0	1/0/0	0	0	-
				564-1	564-2	564-3	564-4	564-5	564-6
Pan troglodytes 15.10.2001	**564-1**	**muscle**	**31**	**-**	**0**	**0**	**0**	**0**	**0**
	564-2	muscle	31	0	-	0	0	0	0

	564-3	muscle	19	0	0	-	0	0	0
	564-4	muscle	24	0	0	0	-	0	0
	564-5	muscle	19	0	0	0	0	-	0
	564-6	muscle	25	0	0	0	0	0	-
				565-1	565-2	565-3	565-4	565-5	565-6
Cercopithecus petaurista 02.01.2012	565-1	muscle	26	-	0	1/0/0	0	1/0/0	1/0/0
	565-2	muscle	31	0	-	1/0/0	0	1/0/0	1/0/0
	565-3	muscle	25	1/0/0	1/0/0	-	1/0/0	2/0/0	2/0/0
	565-4	**muscle**	**32**	**0**	**0**	**1/0/0**	**-**	**1/0/0**	**1/0/0**
	565-5	muscle	34	1/0/0	1/0/0	2/0/0	1/0/0	-	2/0/0
	565-6	muscle	31	1/0/0	1/0/0	2/0/0	1/0/0	2/0/0	-
				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	607-5	607-6
Cephalophus sp. 17.04.2014	607-1	spleen	5	-	0	0	0	0	0
	607-2	spleen	45	0	-	0	0	0	0
	607-3	**spleen**	**50**	**0**	**0**	**-**	**0**	**0**	**0**
	607-4	spleen	26	0	0	0	-	0	0
	607-5	spleen	23	0	0	0	0	-	0
	607-6	spleen	17	0	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	-

6.5 *Bcbva in vitro* mutation rate analysis and latency estimate

The *in vitro* mutation rate of *B. anthracis* was determined to be 5.2 x 10^{-10} mutations/site/generation by Vogler et al. (2002) based on a Luria-Delbrück fluctuation test examining the mutation rate of rifampicin resistance (*105*). To assess if *Bcbva* mutates similarly slowly a long-term sequential passage approach of 15 cultures was combined with whole- genome sequencing.

We aimed at especially high coverages for this experiment, resulting in average coverages for the *Bcbva* genomes between 64 x and 150 x (Table S4). 99.7 % of each genome were covered by at least 5 x (the limit defined in this study for a SNP to be reliable). Applying the bioinformatics pipeline for SNP calling mentioned above, a total of 13 SNPs could be identified on the chromosome and 1 SNP on plasmid pXO1. All SNPs were well supported with coverages ranging from 20 x to 122 x.

The size of the *Bcbva* genome is 5472430 bp (chromosome: 5196054 bp, pXO1: 181907 bp, pXO2: 94469 bp; pCI-14 was ignored for this study, because it does not occur in all strains). We used 99.7 % of those as basis of our further calculations, as we know that in average 99.7 % of the 15 outcome genomes were covered by 5x (used for calculations: total genome: 5456013 bp, chromosome: 5180466 bp, pXO1: 181361, PXO2: 94186 bp).

To the number of generations from the sequential passage (5976, see methods section 5.7) a number of generations was added accounting for the additional passaging steps during DNA extraction of the 15 isolates. This was approximated using *Bcbva* generation time (determined to be 29.1 min, see methods section 5.7) and resulted in a total of 6624 generations of bacterial growth.

g being the number of generations our isolates went through to produce the *n* SNPs detected by whole genome sequencing and knowing the size *s* of the *Bcbva* genome (in base pairs, bp), the *Bcbva in vitro* mutation rate μ is defined by the equation $\mu = n/(s*g)$ (μ unit: mutations/site/generation).

The resulting *Bcbva in vitro* mutation rate is therefore:

μ_{total} = 14 SNPs/(5456013 bp x 6624 generations) =

3.87 x 10^{-10} mutations/site/generation

The 95% confidence interval (CI), determined under the assumption of a Poisson distribution with the *poisson.test* function in R (*37*), ranged from 2.12 x 10^{-10} to 6.5 x 10^{-10}.

Rates are μ_{Chrom}= 3.79 x 10^{-10} (95% CI: 2.02 x 10^{-10} – 6.48 x 10^{-10}) and μ_{pXO1}= 8.32 x 10^{-10} (95 % CI: 2.11 x 10^{-11} – 4.64 x 10^{-9}) for the chromosome and pXO1, respectively.

Due to the small size of plasmid pXO1 with only one newly introduced SNP during passaging, the confidence interval here is large and no further conclusions can be drawn from μ_{pXO1}.

6.6 Phylogenetic and statistical analyses

6.6.1 *Maximum Likelihood Analysis*

For a first impression of the phylogeny of *Bcbva* strains in TNP, an evolutionary tree was calculated in a Maximum Likelihood framework. In a Maximum Likelihood analysis the most likely tree of all theoretically possible trees, assuming a previously determined model of nucleotide substitution (see 5.8.1), is the outcome of the analysis. Subsequent non-parametric bootstrapping is performed to estimate the support for the single branches: For the present tree 100 pseudo-replicates of the original dataset were analysed and the frequency of appearance of any given branch in this set of pseudo ML-trees was plotted above the branches representing the "bootstrap values".

The resulting tree visualizes *Bcbva* diversity in TNP and illustrates that strains do not cluster by sampling date or species, therefore giving a first indication that outbreaks are not caused by independent newly introduced sources and that infection cycles are species independent (Fig. 7 and Fig. S1). Further, it is interesting to see that the isolates from Grebo, Liberia, nest within the diversity of TNP, rather than clustering separately. This indicates transmission links between Côte d'Ivoire and Liberia.

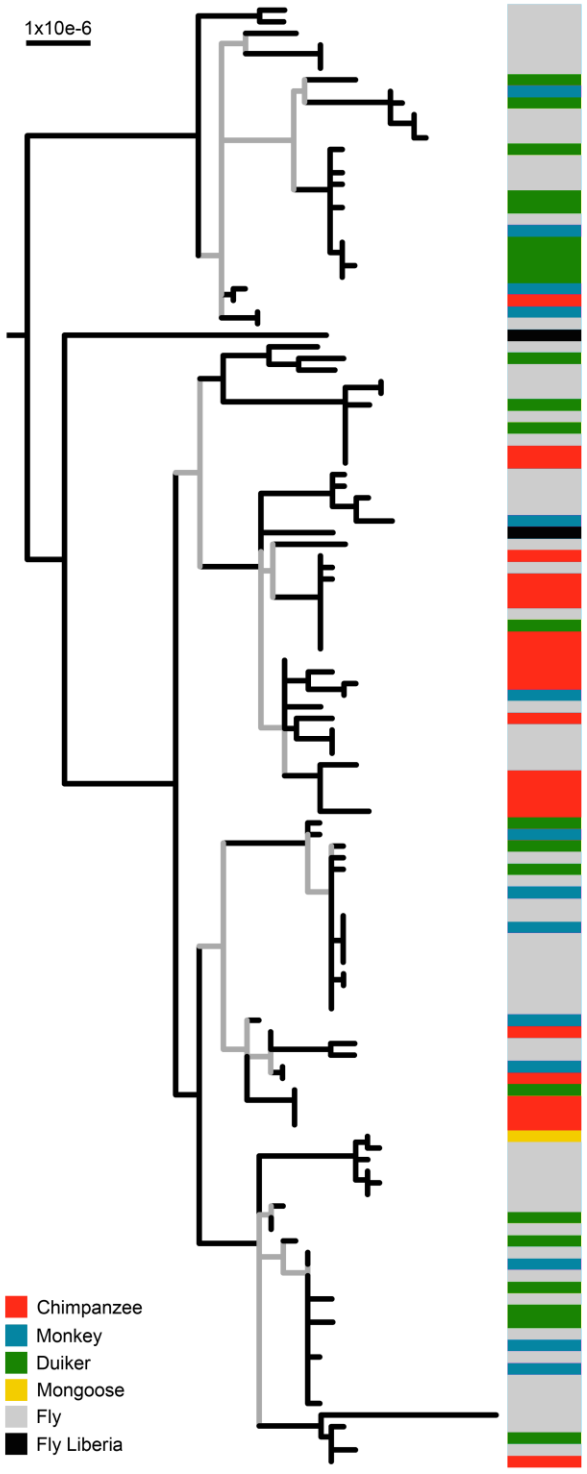

Figure 7. Maximum Likelihood tree based on chromosomal sequences of *Bcbva* isolates from TNP (Côte d'Ivoire, n=124) and Grebo (Liberia, n=2). One sequence per host (mammals/flies, two divergent isolates for fly 600) was included and the final alignment of variant sites measured 298bp.Internal branches with bootstrap values lower than 90 are colored in grey. The colored strip represents different host species. The tree was rooted using the heuristic residual mean squared function in *TempEst* v 1.5. The scale is given in substitutions/chromosome (*published in* (*4*)).

6.6.2 *Bayesian Analyses*

Phylogenies can also be calculated using a Bayesian Markov Chain Monte Carlo (BMCMC) approach. The phylogenetic tree resulting from a BMCMC analysis is called a Maximum Clade Credibility (MCC) tree. In contrast to Maximum Likelihood analyses, where the likelihood is the probability of the data given one specific tree, BMCMC analyses evaluate the probability distribution for all trees given the data. This conditional probability is referred to as the posterior probability in Bayesian statistics. The resulting MCC tree is therefore one sample tree from the distribution of trees that shows the posterior probability values for the whole tree distribution above its branches.

One advantage of this approach is that parameters like the sampling date can easily be incorporated in a Bayesian model. Based on the concept of measurable pathogen evolution the evolutionary rate of a pathogen and the time to its most recent common ancestor (tMRCA) can be determined in a Bayesian framework (*108*). If sufficient genetic information is available additional parameters like geographic sample origin and sample host can allow for conclusions on the geographic origin of a pathogen or the directionality of transmission between hosts. It is important to differentiate between mutation rate and evolutionary rate of a pathogen. While the mutation rate is determined experimentally during continuous growth in the laboratory (usual unit: mutations/site/generation), the evolutionary rate describes the amount of substitutions actually observed in the pathogen in its natural environment (usual unit: substitutions/site/year).

6.6.2.1 Time-scaled phylogenetic tree reconstruction using BEAST

The first model tested here contained only sequence data of the 126 isolates from independent hosts from TNP and Grebo and their respective sampling dates. Information on the molecular clock model and the development of the bacterial population (coalescent tree prior) is required as a priori information in BEAST. As this information was naturally not available, the model was run with all possible combinations of priors with the goal to compare and find the most appropriate combination in the end. For the molecular clock the options strict and uncorrelated lognormal relaxed clock (*113*) were tested. A strict molecular clock assumes that all lineages evolve at the same rate, while a relaxed clock allows for evolutionary rate heterogeneity between lineages. Options for coalescent tree priors included simple pathogen population growth models like exponential or logistic growth or a constant population size, but also more complex non-parametric tree priors that estimate the demographic history from the sequence data (Bayesian Skyline, GMRF Bayesian Skyride, Bayesian SkyGrid) (*170*).

After running all possible models, they can be compared by so called Bayes Factor comparison of marginal likelihoods. For every model a marginal likelihood is calculated, representing the probability of the data given the model type. Different approaches can be used to compute marginal likelihood values and we used the stepping stone method by Xie et al. (2011) that was shown to be superior to the harmonic mean method (*159*). Marginal likelihood values are usually processed in logarithmic scale to ensure numerical stability during the calculation process due to extremely small numbers. The Bayes Factor is the ratio of these likelihood values of two models and can test for evidence for model M_1 against Model M_0. Kass and Raftery (1995) defined

Bayes Factor values that represent a statistically significant difference between models, on the natural logarithmic scale (to the base *e*): A Bayes Factor between 0 and 1 is not worth more than a bare mention, values between 1 and 3 are considered as positive evidence against M_0 and values over 3 and 5 respectively show strong and very strong evidence against M_0 (*161*). So, if one looks for example at the difference between the log marginal likelihoods of the constant/relaxed clock model (M_1, log marginal likelihood -7208621,64) and the GMRF Bayesian Skyride/strict clock model (M_0, log marginal likelihood -7208665,62), M_1 is strongly preferred over M_0 with a BF of 44,02.

In the present analysis no coalescent tree prior was strongly preferred over another, except for the exclusion of the Bayesian Skyride model (Table 5). The most parsimonious scenario of a constant population size being as good an explanation for the data as more complex models, we therefore considered that the population size of *Bcbva* did not vary much over the phylogeny. This scenario was also in accordance with the pathogen population dynamics plot derived from the Bayesian SkyGrid Model (*171*) (Fig. 8). A relaxed clock model, which accounts for evolutionary rate heterogeneity, was preferred by Bayes Factor comparison (Table 5).

Table 5. Bayesian model comparison. Log marginal likelihood values and temporal estimates for different coalescent tree priors and molecular clock models.

Coalescent tree prior	Molecular clock prior	tMRCA (years)	Evolutionary rate (μ) (substitutions/site/year)	log marginal likelihood[1)]
Constant size	strict clock	129 (80-188)	3.77×10^{-8} (2.33-5.16)	-7208624,95
	uncorrelated lognormal relaxed clock	146 (70-237)	3.92×10^{-8} (2.32-5.52)	-7208621,64
Exponential growth	strict clock	123 (75-177)	3.84×10^{-8} (2.47-5.35)	-7208623,87
	uncorrelated lognormal relaxed clock	123 (65-198)	4.19×10^{-8} (2.55-5.92)	-7208623,67
Logistic growth	strict clock	121(74-179)	3.85×10^{-8} (2.42-5.31)	-7208624,7
	uncorrelated lognormal relaxed clock	122 (59-195)	4.12×10^{-8} (2.55-5.81)	-7208623,8
Bayesian Skyline	strict clock	133 (75-202)	3.56×10^{-8} (2.01-5.14)	-7208623,93
	uncorrelated lognormal relaxed clock	135 (68-222)	3.77×10^{-8} (2.21-5.53)	-7208622,09
GMRF Bayesian Skyride	strict clock	50 (36-67)	5.95×10^{-8} (4.2-7.64)	-7208665,62
	uncorrelated lognormal relaxed clock	31 (24-41)	7.55×10^{-8} (5.35-9.63)	-7208646,51
Bayesian SkyGrid	strict clock	121 (74-173)	3.9×10^{-8} (2.52-5.32)	-7208621,37
	uncorrelated lognormal relaxed clock	128 (60-208)	3.92×10^{-8} (2.26-5.67)	-7208619,32

1) Estimated by stepping stone sampling

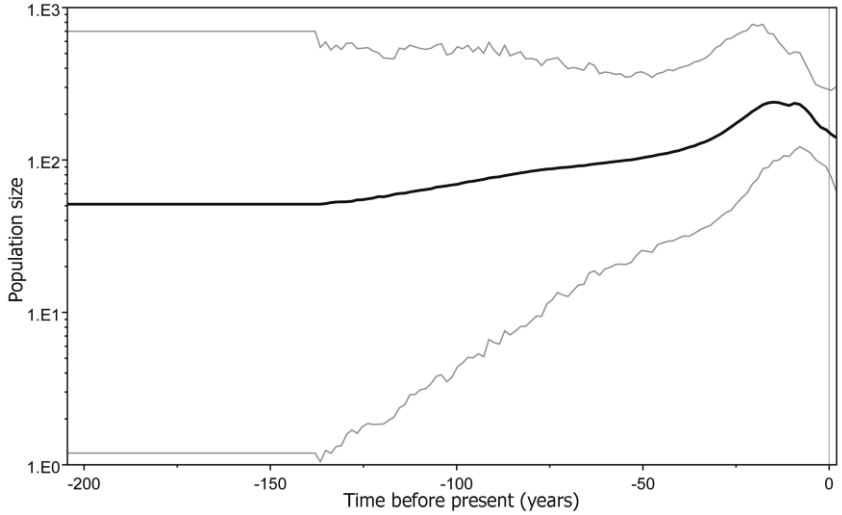

Figure 8. Sky Grid Plot TNP. Development of the *Bcbva* population in TNP over the last 128 years derived from the Bayesian SkyGrid Model for 126 full genomes from different hosts from TNP and neighboring Grebo (Liberia). Blue lines represent the 95 % confidence interval.

Similar rates for evolutionary rate and tMRCA were received for all relevant prior combinations with relatively small confidence intervals, a first indicator for reliable temporal estimates.

The resulting time to the most recent common ancestor (tMRCA) for the favored constant population/relaxed clock model was estimated to be 146 years (95 % CI: 70-237) with a mean clock rate of 3.92 x 10^{-8} substitutions/site/year (95 % CI: 2.32-5.52 x 10^{-8}) (Table 5). This equals 0.2 substitutions per year for a *Bcbva* chromosome.

The MCC tree resulting from this analysis is shown below (Fig. 9) and is largely in accordance with the ML tree, confirming that outbreaks are not species specific or caused by temporally independent sources.

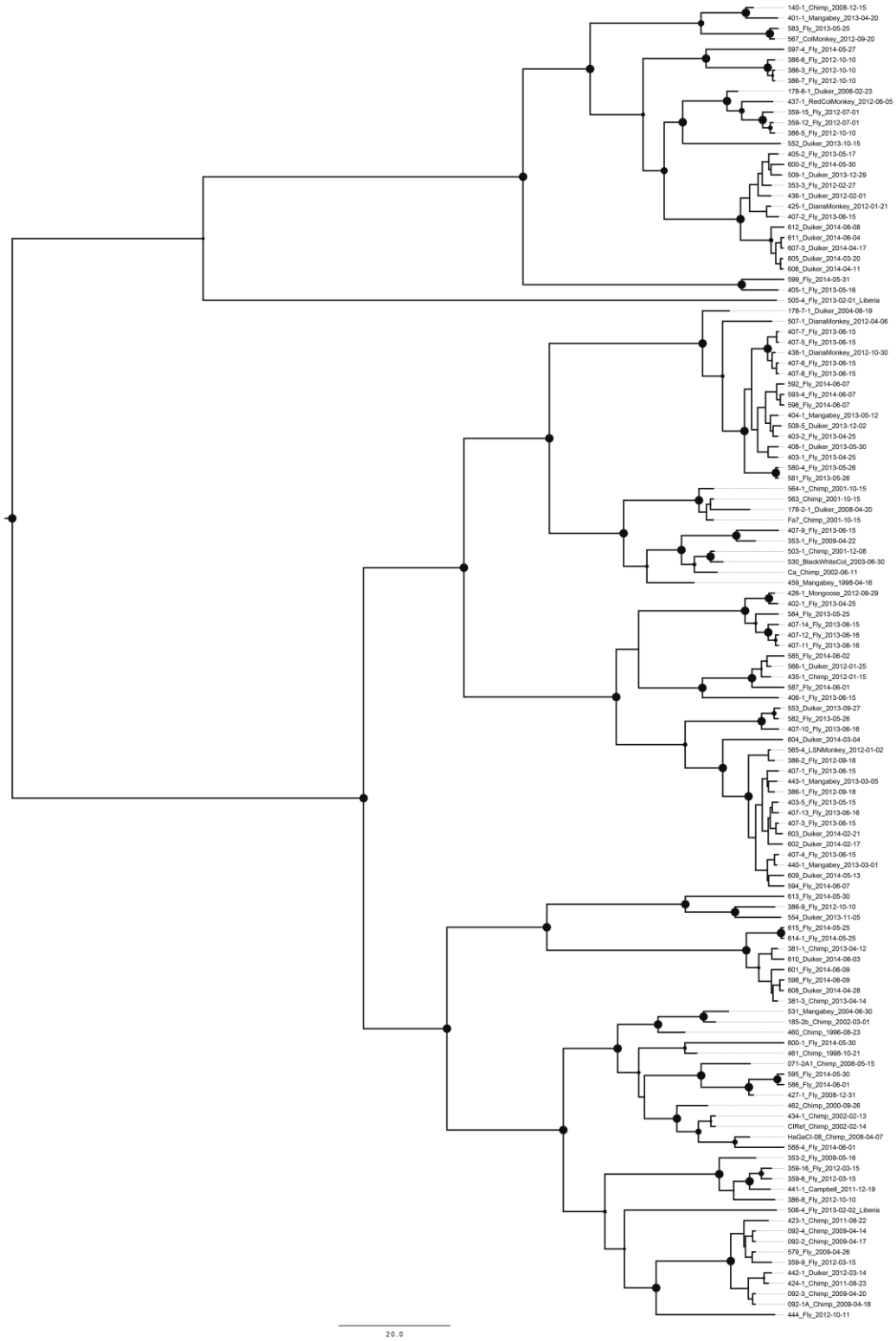

Figure 9. Maximum Clade Credibility tree based on chromosomal sequences of *Bcbva* isolates from TNP (Côte d'Ivoire) and Grebo (Liberia). Size of nodes represents posterior probability values. The location of the root received a posterior probability of 1 (*published in* (*4*)).

As this is the first time an evolutionary rate could ever be derived from temporally stamped anthrax sequences, a stringent randomization test with 20 date randomized replicates was performed to confirm the signal in the data.

As required to validate the signal none of the mean clock rate 95 % confidence intervals of the randomizations overlapped with the 95 % confidence interval of our data (Fig. 10).

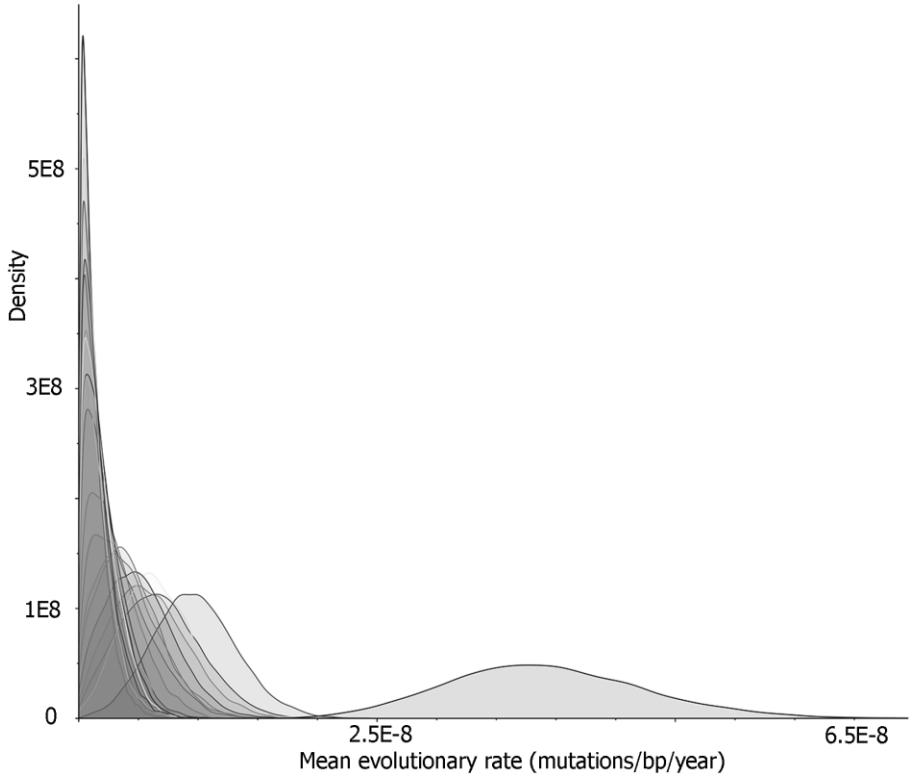

Figure 10. Date randomization chromosome. Posterior distribution of evolutionary rates estimated by Bayesian analysis from 126 chromosomal *Bcbva* sequences from TNP and Grebo (Liberia) (in grey) and from 20 date randomized replicates.

Comparing the chromosomal evolutionary rate determined during the Bayesian analysis of isolates from TNP to the mutation rate determined *in vitro,* an approximation could be made for the time *Bcbva* spends inactive in the environment in TNP:

The chromosomal mutation rate was determined to be 3.79×10^{-10} mutations/site/generation during the *in vitro* experiment. This is equivalent to 35.4 substitutions/chromosome/year, using the generation time of 29.1 min determined in this study (see 5.7).

However according to the Bayesian analysis in TNP *Bcbva* only evolves at a rate of 0.2 substitutions/chromosome/year, which would mean that in TNP the bacteria are only present in an actively replicating state for ~ 0.56 % of the time.

Further, to gain the observed number of yearly chromosomal substitutions in TNP, a *Bcbva* strain would go through 102 generations every year based on the calculated *in vitro* mutation rate. This is not a high number- but a lot compared to existing estimates for classical *B. anthracis* of 0.3 generations per year in Northern America (*106*)

6.6.2.2 Ancestral state reconstruction of host transmission

The existing Bayesian model can further be extended, i.e. by adding information about the host an isolate originated in. This approach was used by Mather et al. 2014 to investigate the amount and directionality of transmission for different *Salmonella typhimurium* hosts (*117*). We used their methodology to test for the directionality of transmission between mammals and flies, hoping to gain insight in the possible role of flies for disease transmission. "Host" was added as a discrete trait to the sequence date and four different transmission scenarios were calculated and compared. Log marginal likelihoods determined by stepping stone sampling (*159*, *160*) revealed evidence against a unidirectional transmission from flies to mammals (Table 6). Transmission from flies to mammals is not excluded, but does not seem to play a major role for the epidemic scenario in TNP.
This was also supported by the relative host transmission rates (an outcome from the asymmetrical model): the hypothetical transmission rate from mammals to flies was twice as high as the rate from flies to mammals (Table 7). Due to the balanced numbers of sequences from each host type it is unlikely that these results are influenced by sampling bias.

Table 6. Models ancestral state reconstruction. Log marginal likelihoods for the four different transmission scenarios tested in Bayesian phylogenetic analysis: bidirectional asymmetric, bidirectional symmetric, unidirectional mammal-to-fly and unidirectional fly-to-mammal.

	Log marginal likelihood [1)]
Bidirectional symmetric	-7208740.08
Unidirectional: mammal-to-fly	-7208741.48
Bidirectional asymmetric	-7208742.27
Unidirectional: fly-to-mammal	-7208752.9

[1)] Estimated by stepping stone sampling

Table 7. Rates ancestral state reconstruction. Relative host transition rates estimated from the asymmetric continuous-time Markov chain model.

	Fly	Mammal
Fly		0.64 (1.92×10^{-4} - 1.77)
Mammal	1.34 (0.03 - 3.31)	

6.6.2.3 Continuous phylogeography

Information about the sampling location can be added in the form of GPS data, as continuous trait. With this approach the geographical origin and spread of an outbreak can be approximated. The aim of this kind of analysis is to identify not just the time to,

but also the location of the most recent common ancestor of the tested isolates and to see how the disease spread from there. In this study we tested a Brownian diffusion model assuming homogenous dispersal, as well as three Relaxed Random Walk (RRW) models allowing for branch specific variation in dispersal rates (*116*). Formal model testing using Bayes factor comparison of log marginal likelihoods as described above did not favor any of the models, but excluded the Cauchy RRW model (Table 8). The three remaining models stated a mean rate of 0.43 to 0.45 km/year for *Bcbva* dispersal. However, visualization of the evolutionary tree in continuous space using *Spread* v1.0.6 (*155*) and the Google Earth software (http://earth.google.com) revealed that the 80 % highest posterior density (HPD) region (=the confidence interval) of the inferred location of the common ancestor coincided with the whole TNP research area, so that this analysis did not result in any useful information (Fig. 11).

Table 8. Models continuous phylogeography. Log marginal likelihood and parameter comparison for Brownian and different relaxed random walk (RRW) models

	dispersal rate (km/year)	Log marginal likelihood [1)]
Lognormal RRW	0.45	-7206753.25
Homogenous (strict Brownian)	0.45	-7206754.69
Gamma RRW	0.43	-7206758.15
Cauchy RRW	0.22	-7206778.86

[1)] Estimated by stepping stone sampling

Figure 11. Visualization continuous phylogeography. Screenshot of the video visualizing the continuous phylogeography of *Bcbva* in the TNP research area, based on Bayesian analysis in *BEAST* v1.8.2 assuming a constant population tree prior, an uncorrelated lognormal relaxed clock and a Brownian diffusion model. GPS data was added as continuous trait and a 0.0001 jitter was applied. Visualization was done using *Spread* v1.0.6 and the Google Earth software. The bright blue area represents the 80% highest posterior density (HPD) region of the inferred location of the common ancestor and this coincides with the whole TNP research area so that no information about the geographic origin of *Bcbva* in TNP can be derived (full video available on request).

6.6.2.4 Time-scaled phylogenetic tree reconstruction on plasmid level

For the sake of completeness it was also attempted to derive a temporal signal on plasmid level, although the two plasmids contained only limited genetic information (18 SNPs (7 parsimony informative) were determined on pXO1 and 11 (6 parsimony informative) on pXO2).

The BMCMC model converged for pXO1 and an evolutionary rate estimate could be derived (Table S6), but failed date randomization testing: the 95% confidence interval for the pXO1 evolutionary rate estimate overlapped with the estimates from randomized data in one out of ten runs (Fig. 12). The evolutionary rate estimated for pXO1 is therefore not valid. For pXO2 models did not even converge at the start and no information could be derived.

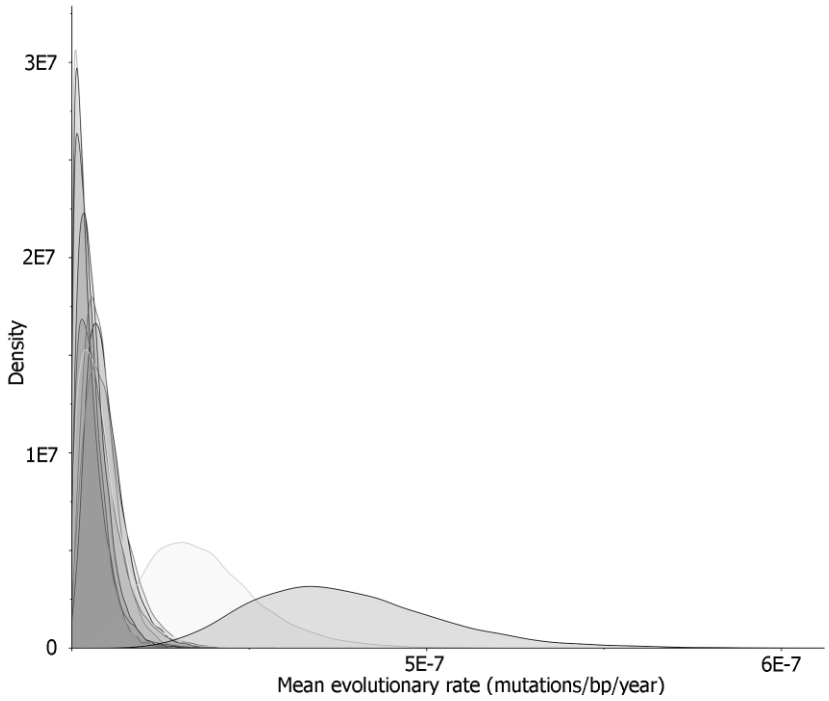

Figure 12. Date randomization pXO1. Posterior distributions of evolutionary rates estimated by Bayesian analysis from 126 *Bcbva* pXO1 sequences from TNP and Grebo (Liberia) (in grey) and from 10 date randomized replicates. For one of the replicates the 95% confidence interval of the evolutionary rate (in green) is overlapping with the 95 % confidence interval of the original data.

6.6.2.5 Time-scaled phylogenetic tree reconstruction for sub-Saharan Africa

Time-scaled phylogenetic tree reconstruction in a Bayesian framework was also attempted for strains from different sub-Saharan countries. As sampling depth here was not deep enough to derive an evolutionary rate from the data, the rate from TNP was used as prior in this analysis. Using the evolutionary rate derived from Bayesian

analyses in TNP for a conservative approximation, the tMRCA of *Bcbva* strains from sub-Saharan Africa strains was estimated to be at least 2157 years (95 % CI: 1129-3452), implying that *Bcbva* has been present and possibly active in sub-Saharan Africa for millennia (Fig. 13).

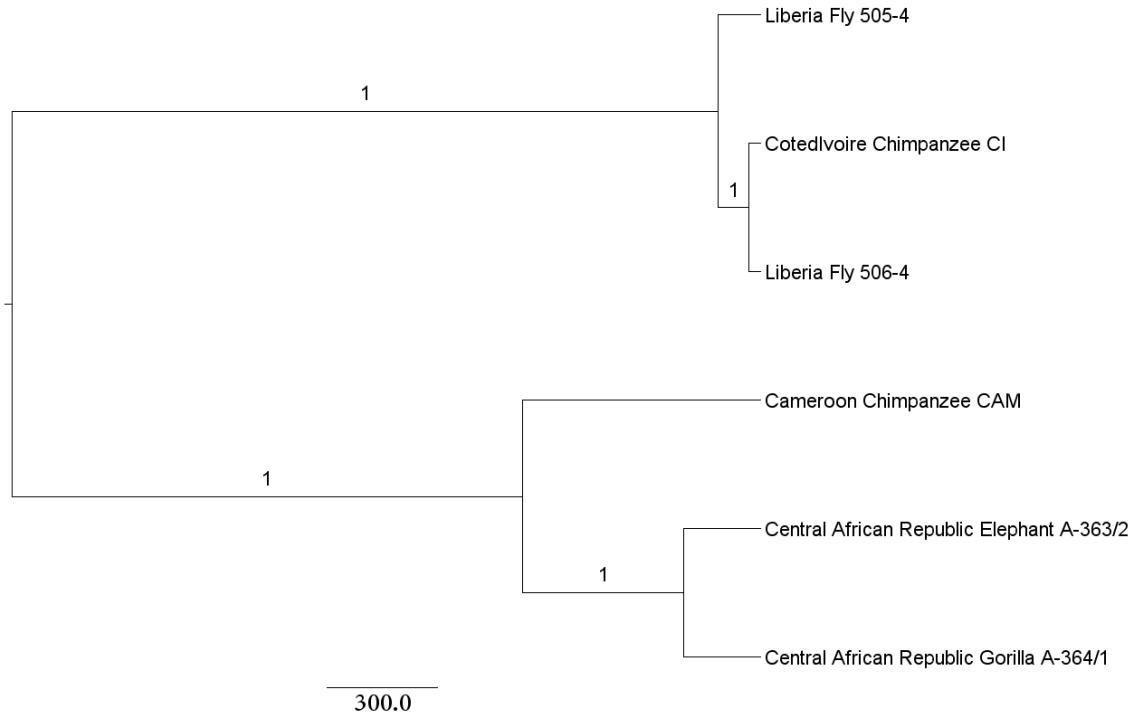

Figure 13. Maximum Clade Credibility tree for sub-Saharan *Bcbva* strains, including chromosomal sequences of *Bcbva* strains from Côte d'Ivoire, Cameroon, Central African Republic and Liberia. The evolutionary rate from TNP was used as a prior for this analysis. The alignment of variant sites measured 1016 bp. Nodes are labelled with their respective posterior probability values.

6.6.3 *Correlation of genetic and geographic distances*

During the 19 day fly snapshot in May and June 2014 (C. Hoffmann) 18 *Bcbva* isolates were retrieved from 17 carrion flies (Table S3, Table S4). Of the 18 isolates, 13 isolates differed by > 2 SNPs implying at least 13 contemporaneous *Bcbva* positive carcasses (based on within-host diversity data, see 6.4.2). Deaths were caused by multiple independent transmission chains and *Bcbva* isolates formed geographically co-circulating clusters differing by up to 48 SNPs (Fig. 14).

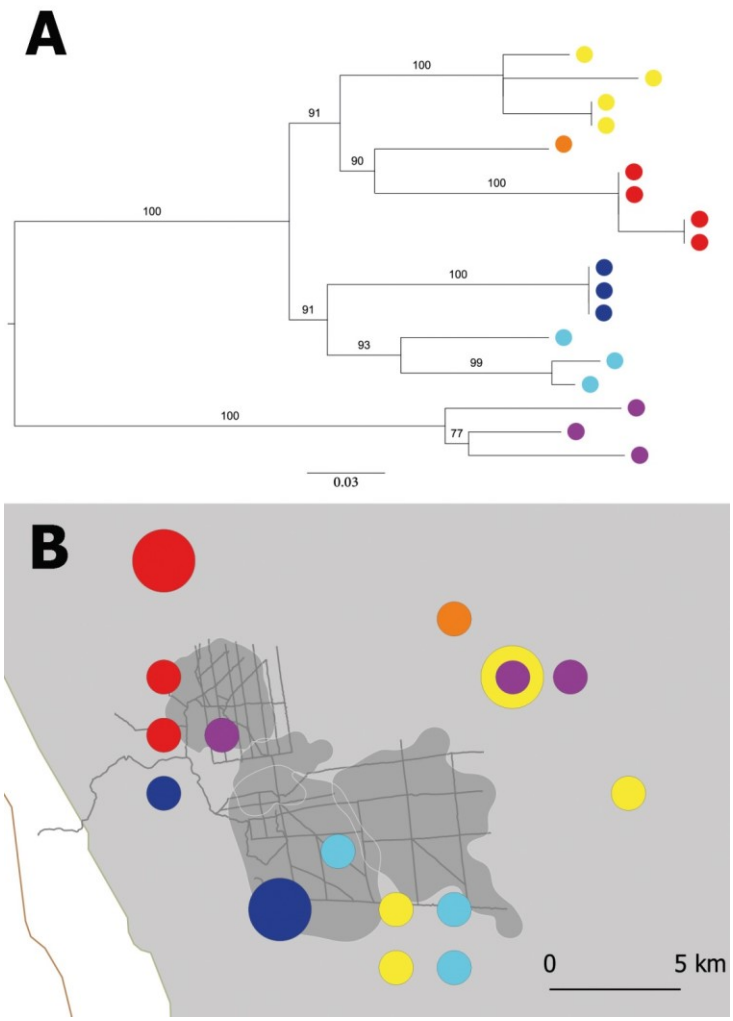

Figure 14. *Bcbva* strains active during fly snapshot. (A) Maximum Likelihood Tree based on chromosomal sequences of *Bcbva* isolates from the 19 day fly snapshot. The final alignment of variant sites measured 123bp. Bootstrap values are shown above all internal branches. The tree was rooted using the "best-fit" option in Path-O-Gen v1.2. (B) Geographic origin of *Bcbva* isolates collected during the fly snapshot. Colors correspond to ML tree. Big circles represent two isolates (*published in* (*4*)).

To learn more about the spatial dynamics of *Bcbva* in TNP the correlation between genetic and geographic distances of isolates was investigated. Multiple regression on distance matrices (MRM) testing 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 ($R^2=1.3$ %), possibly due to genetically different strains co-circulating in the same area.

In an examination of variation within genetic lineages we focused on groups of isolates with low genomic distances. A significant correlation between mean genetic distances of groups and their respective mean geographic distances was found here ($p=4x10^{-5}$, $R^2=0.72$) underlining spatially restricted transmission of *Bcbva* (Fig. 15).

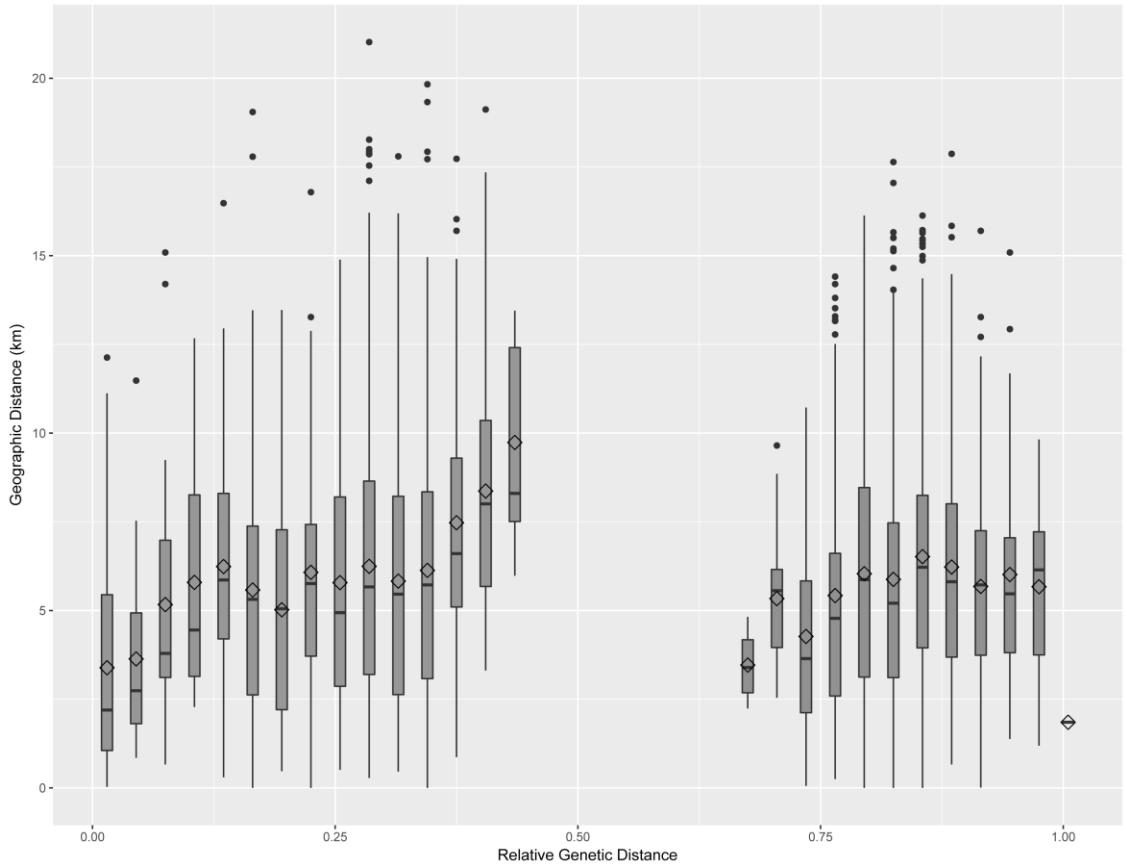

Figure 15. Genetic and geographic distances. *Bcbva* isolates from TNP were binned by relative genetic distance (bin size = 0.03, approx. 2.5 SNPs). To determine the relative genetic distance, the two most genetically distant isolates received a value of 1 and all other distances were scaled accordingly. Diamonds indicate the geographic distance means of the groups. To examine variation within genetic lineages, we only analyzed groups of isolates with low relative genetic distances (<0.5) and their mean geographic distances. Groups including isolates with larger relative genetic distances (>0.5) are shown for completeness, but were not included in the analysis. 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 ($p = 4 \times 10^{-5}$) (*published in* (*4*)).

6.7 Serological analyses

A total of 59 serum, plasma or whole blood samples from five different TNP wildlife species, mainly representing primates, were tested in a serological study. Despite the continuous occurrence of the disease in the research area, only a small number of samples contained anti-PA/LF antibodies.

For red colobus monkeys (n=15) and black-and-white colobus monkeys (n=10) one sample each was clearly positive in PA ELISA and could be confirmed in PA and LF Western Blot. One more red colobus sample reacted in PA Western Blot, but in none of the other assays.

None of the sooty mangabey (n=16) or chimpanzee (n=9) samples reacted in PA ELISA or PA/LF Western Blot.

For duikers (n=9) four samples were positive in PA ELISA, but could not be confirmed in PA/LF Western Blot.

Results are presented in detail in table 9 and table S6.

Table 9. Overview serological results for each tested species.

Results for PA ELISA, PA and LF Western Blot in each tested species. Observed seroprevalence includes all PA Western Blot positive animals. The 95 % confidence intervals for seroprevalence were approximated with the *prop.test* function in R (*148*).

Species	n	PA ELISA positive	PA WB positive	LF WB positive	Observed seroprevalence (95% CI)	positive control	negative control	conjugate
Western red colobus monkey (*Procolobus badius*)	15	1	2	1	13% (2-42%)	AVA vaccinated human	human	goat α human
Western black-and-white colobus monkey (*Colobus polykomos*)	10	1	1	1	10% (0-46%)	AVA vaccinated human	human	goat α human
Sooty mangabey (*Cercocebus atys*)	16	0	0	0	0% (0-24%)	AVA vaccinated human	human	goat α human
Chimpanzee (P*an troglodytes verus*)	9	0	0	0	0% (0-37%)	AVA vaccinated human	human	goat α human
Maxwell's duiker (*Cephalophus maxwelli*)	9	4	0	0	0% (0-37%)	Sterne vaccinated goat	Red forest duiker (*Cephalophus natalensis*) from zoo	rabbit α goat

7 Discussion

7.1 The role of *Bcbva* in the Taï National Park ecosystem

Bcbva continuously causes wildlife mortality in TNP affecting a broad range of hosts, mainly primate species. Since the beginning of the TCP veterinary program in 2001 tissue samples from 81 necropsies were tested positive for *Bcbva.* Thus, *Bcbva* was associated with 40 % of all detected mammal fatalities (Table S1) and with the help of molecular analysis of bones the first documented *Bcbva* case could be dated back to 1991 (*4*). This suggests a major importance of the pathogen for the TNP fauna. The only other African national park with accessible long-term carcass monitoring data for anthrax negative and positive carcasses is Etosha National Park (ENP). Here, anthrax caused by *B. anthracis* was responsible for 13 % of detected carcasses (and suspected for another 11 %) between 1975 and 1990 (*7*). In contrast to TNP outbreaks here were clearly associated with season mainly affecting plains ungulates and elephants (*7*). While in African savanna national parks wave-like anthrax outbreaks primarily affecting one (ungulate) species at a time are observed, we found *Bcbva* positive carcasses of ten species in TNP with no obvious seasonal pattern. Certainly the difference in host range between classical anthrax and *Bcbva* is also due to the different composition of the fauna in tropical rainforests and savanna habitats. However, numerous primates populate the Serengeti National Park (SeNP) (*Cercopithecus patas, Cercopithecus aethiops, Papio Anubis, Colobus guereza)* and not a single anthrax fatality in a primate has been reported from here to date (*81*). TNP mortality data indicates that *Bcbva* dynamics in TNP differ distinctly from classical anthrax in terms of host range, seasonality and amount of mortality caused by the disease. The results of the present genetic and serological studies further confirm the outstanding dynamics of *Bcbva* in TNP.

The high genetic diversity of *Bcbva* in TNP reflects the hyper endemic nature of the pathogen in the area. In total 298 SNPs were detected, when chromosomal *Bcbva* sequences from TNP were compared. The maximum difference between isolates was 69 SNPs and the mean difference 26 SNPs. For comparison, in a whole genome sequencing study investigating *B. anthracis* diversity in France Vergnaud et al. (2016) reported a maximum difference of 20 SNPs between any two isolates from the French Alps collected between 1997 and 2009, using a pipeline for SNP detection comparable to ours (*103*). Not many comparable data sets are available yet, and we cannot compare the diversity in TNP to other more related ecosystems. *B. anthracis* enzootic areas in Africa may show a similar diversity, but unfortunately no whole genome sequencing data exists to date. Using the MLVA or PHRANA system on the TNP dataset would allow for a comparison with genetic diversity data from Etosha National Park (*98*, *104*) and Kruger National Park (*87*), but it is questionable if these are suitable tools for the in depth analysis of *Bcbva*. Both systems were specifically designed for a *B. anthracis* genetic background and genetically *Bcbva* clusters with other members of the *B. cereus* group (*6*, *71*). The MLST typing schemes routinely applied for *B. cereus,* however, are not discriminatory enough to type *Bcbva* strains on a fine scale, resulting in the same MLST type for different *Bcbva* strains from TNP (*5*) .

It was tested if *Bcbva* had a higher mutation rate than *B. anthracis*, which could potentially cause the observed high genetic diversity in *Bcbva*. Our detailed approach for the determination of the *in vitro* mutation rate combined sequential passaging of liquid cultures (15 cultures, 24 days) with whole genome sequencing of starting culture and resulting cultures. This approach is more robust than the set-ups chosen so far for determination of the *B. anthracis* mutation rate: While Vogler et al. (2002) based their estimate only on the gene for rifampicin resistance, Agren et al. (2014) conducted a sequential passaging/whole genome sequencing approach, but with a very limited sample size (3 cultures, 10 days) and non-empirical assumptions for growth generations (*102*, *105*). However, all approaches support the result that *in vitro* mutation rates of *B. anthracis* and *Bcbva* are rather low and have the same magnitude. Thus, the high genetic diversity in TNP is not a result of accelerated mutation of *Bcbva* compared to *B. anthracis*, but seems to be indeed a result of relatively high transmission activity of the pathogen in the area.

The *Bcbva* isolates from TNP investigated in this study cover a period from 1996 until 2014. It was found that over this time period and in this area *Bcbva* has measurably evolved. The time to the most recent common ancestor (tMRCA) of all isolates was 146 years (95 % HPD: 70-237) in a scenario with a relaxed molecular clock, revealing that the pathogen has been present in TNP for at least multiple decades. Bayes Factor comparison of demographic models did not reject a constant population size model, suggesting that the pathogen has not emerged recently and that local transmission of *Bcbva* in TNP has taken place during this period. *Bcbva* has not recently been introduced into TNP leading to numerous wildlife deaths in a naïve population- but on the contrary: our data indicate that *Bcbva* has been present and possibly active in TNP for a long time.

While a high number of *Bcbva* positive carcasses was found in the TNP study area, our carcass monitoring still only covered a small proportion of wildlife mortality due to low visibility and fast decomposition of carcasses in the rainforest environment. Whole Genome Sequencing of *Bcbva* isolates from flies indicated a large proportion of undetected anthrax cases. During the 19 day fly snapshot in May and June 2014, *Bcbva* was isolated from 17 carrion flies revealing 13 strains differing from any other strain by > 2 SNPs (2 to 48 SNPs distance, median: 25 SNPs). Since two SNPs appears to be the upper level of within-host diversity (see 6.4.2), these 13 isolates must originate from separate hosts, implying at least 13 contemporaneous *Bcbva* positive carcasses. In the same time period, despite the same sampling effort, only three *Bcbva* positive carcasses were discovered and their isolates all corresponded to one of the fly *Bcbva* lineages (≤ 2 SNPs difference). This shows that carcass monitoring underestimates wildlife mortality by at least an order of magnitude.

Pathology and histo-pathology of *Bcbva* cases are comparable to that of anthrax cases caused by *B. anthracis* and in small animal models *Bc*bv*a* exhibits a similar virulence to what has been observed for *B. anthracis* (*1*, *3*). Unfortunately, for rainforest fauna living in the wild nothing is known about the likelihood of survival following infection with *Bcbva*. Monkeys and chimpanzees were observed to die within hours of the onset of disease symptoms (*Taï Chimpanzee Project & Taï Monkey Project, personal communication*). The rapid lethality following the onset of symptoms could potentially be a product of the generally reduced expression of weakness in wild animals (*172*). We do not know the incubation time for *Bcbva*, though available evidence suggests *Bcbva* is

highly virulent (*3*). The results of the serological investigation in five different species (chimpanzee (*Pan troglodytes verus*), black-and-white colobus monkey (*Colobus polykomos*), red colobus monkey (*Procolobus badius*), sooty mangabey (*Cercocebus atys*) and Maxwell's duiker (*Cephalophus maxwelli*) in combination with the TNP necropsy data allowed us to investigate the likelihood of sub-lethal *Bcbva* infections in wildlife. While *Bcbva* is responsible for a substantial proportion of deaths for each species (chimpanzees: 31 out of 55 carcasses; black-and-white colobus monkeys: 1/5; red colobus monkeys: 3/3; sooty mangabey: 11/23; all duikers: 26/40), only two animals, a red colobus monkey and a black- and white colobus monkey, showed a strong antibody response to PA and LF and one red colobus monkey was positive in PA Western Blot, but none of the other assays (Table S6). Four duikers were tested positive in PA ELISA, but could not be confirmed by Western Blot (Table S6). While a species-specific negative control from a zoo was available for duikers, the artificial ELISA cut-off derived from the negative control was likely too low, because zoo samples may show a different background in ELISA than wildlife samples.

It has been shown previously that sublethal anthrax infections do not only occur in carnivores. Different degrees of anthrax seroprevalence have been reported for herbivores in African savannas (*81*, *86*, *133*). The finding of anthrax antibodies in a significant amount of wildebeest (*Conochaetes taurinus mearnsi*) (19 %) and buffalo (*Syncerus caffer*) (46 %) in the Serengeti showed that not all herbivores succumb to infection and that susceptibility can vary widely between herbivore species (*81*).

Three individuals of the arboreal monkey species tested in TNP had antibodies against PA and two of them showed high OD values in ELISA in the range of the positive control and were confirmed in LF Western Blot. These strong immune reactions might suggest that these animals recently survived a sublethal infection. One potential scenario for infection of arboreal monkeys could be cutaneous infection transmitted by a vector, e.g., biting flies, as observed previously in cattle and humans (*41*, *173*) or transmission into the canopy by carrion flies (see 7.2). A weakly positive sample, like observed in one red colobus monkey, could either be caused by an old sublethal infections or by contact to small amounts of *Bcbva* or its toxin components. Serological testing for antibodies against PA cannot differentiate between infection with subsequent recovery and low dose exposure with seroconversion. Seroconversion through intestinal absorption of PA and other toxin components was suggested for scavengers by Turnbull et al. (2008), but has not yet been described for herbivores (*127*). Despite this seroconversion through contact to toxin components must also be considered when investigating a newly detected pathogen in an ecosystem where transmission pathways are unknown to date (see 7.2).

Mangabeys, chimpanzees and duikers suffer the highest *Bcbva* related mortality in this ecosystem and are likely regularly exposed to soil containing anthrax spores. It is striking that in these three species not a single animal showed an immune response against anthrax that would suggest a recently overcome infection.

These results are based on non-validated in-house tests and due to the nature of the study only limited numbers of controls were available. While the sample set presented here may not appear extensive in comparison to datasets collected in savanna ecosystems, it represents the largest available collection of serum from TNP wildlife. The small sample sizes cause considerable uncertainty, when calculating species-specific seroprevalence (Table 9), but they do overall indicate low *Bcbva*

seroprevalence. The wide distribution and long-term presence of *Bcbva* in TNP suggests the observed low number of reactive animals is not likely due to low rates of exposure. *Bcbva* has been present in TNP for at least decades , yet the TNP primate and duiker species tested here do not appear to have developed significant immunity to this disease.

In other African national parks, *B. anthracis* is considered a natural culling mechanism that is intrinsic to its ecosystem (*63*). Based on the reconstruction of the Bayesian pathogen population dynamics, this also appears to be the case in TNP, despite the high anthrax burden and likely high odds to succumb to a *Bcbva* infection in wildlife, that we find here. However, we suggest that, like in other parks, the disease can become a substantial threat to the endangered wildlife population if anthropogenic pressure increases (*63*). The number of villages bordering the park has grown drastically over the last 40 years (*8*, *10*, *11*) . People are intruding deeper into the park in search of cultivatable land and bushmeat is a common source of protein in the area. The research area itself is protected by the presence of researchers and ecoguards, but the less protected areas of the park are at risk (*10*). Wildlife is forced into the small well protected parts, threatening the sensitive ecological balance. The only species for which detailed demographic data is available, are the habituated TNP chimpanzees. A simulation of their population development using models including and excluding anthrax-induced mortality (*analysis by Hjalmar Kühl, MPI-EVA)* (*4*) showed a sharp population decline if anthropogenic pressure and *Bcbva* pressure are combined. A background mortality of 3-5 % combined with the observed mortality due to anthrax could lead to the extirpation of the TNP chimpanzees within the next 150 years. Chimpanzees have a low reproductive rate and the situation may not be as dramatic for other species, but our results suggest that *Bcbva* will contribute significantly to the extirpation of endangered species in the future if the overall protection of the national park cannot be improved. Infectious diseases are an important factor for the survival or extinction of threatened species and are often still neglected in conservation efforts (*174*).

On a bigger scale we investigated how *Bcbva* evolution in TNP relates to *Bcbva* sequences throughout sub-Saharan Africa. It has long been assumed that *B. anthracis* has evolved clonally, since the acquisition of the two virulence plasmids pXO1 and pXO2, as the only anthrax causing agent of the *B. cereus* group. However, horizontal transfer of small mobile elements has frequently been observed in the *B. cereus* group in the past (*16*, *80*). The existence of several atypically pathogenic members of the *B. cereus* group carrying parts of, or complete, *B. anthracis* virulence plasmids questions the taxonomic structure of the group. Most importantly, *Bcbva* and *B. cereus* JF3964 (found in cattle in Cameroon) demonstrate that not only the *B. anthracis* clade members are capable of causing anthrax (*79*). Whole genome SNP analysis showed that all *Bcbva* isolates found in sub-Saharan Africa to date group in a monophyletic clade within the *B. cereus* group, clearly distinct from *B. anthracis* (*6*). Limited information is available on *B. cereus* JF3964, but it was shown to cluster with *Bcbva* from Côte d'Ivoire in MLVA analysis by Pilo et al. (2011) (*79*). This *B. cereus* could either be part of the *Bcbva* clade or even represent a third clade of anthrax-causing *B. cereus* carrying both virulence plasmids, which might occur much more frequently than generally assumed.

A conservative temporal estimate on the origin of *Bcbva* in sub-Saharan Africa was performed based on the determined evolutionary rate from TNP. The tMRCA of strains from Côte d'Ivoire, Cameroon, Central African Republic and Liberia was estimated to

be at least 2157 years in a Bayesian analysis (Fig. 13). TNP is an area of high *Bcbva* activity and *Bcbva* likely evolves faster here than in other sub-Saharan regions, where less cases occur. Therefore, and because of the time-dependency of evolutionary rate estimates (*108*), an estimate for sub-Saharan Africa based on the TNP evolutionary rate should be considered a minimum estimate of *Bcbva* occurrence in sub-Saharan Africa.

Due to the high diversity of anthrax-causing *B. cereus* group members in west- and central Africa, especially in Côte d'Ivoire, this region likely represents the geographic origin of anthrax. It will be of big interest to extend the strain collection in the area and to perform extensive Bayesian analyses on comprehensive *B. anthracis*, *B. cereus* and *Bcbva* datasets in the future to clarify the origin of anthrax disease and to investigate why *B. anthracis* has evolved into a globally active killer, while *Bcbva* has only been found in rain forest regions to date.

7.2 Comparative ecology of *Bcbva* and other spore forming bacteria

Although all members of the *B. cereus* group are known to produce spores, the way they interact with their environment can differ significantly. While some are thought to only germinate and replicate in mammalian hosts, others are known to multiply in different substrates (*20*). *B. anthracis* is well-known for spending a major part of its lifecycle dormant in the ground. The spore-forming nature is a major factor in shaping *B. anthracis* disease dynamics marked by distinct seasonal outbreaks. It has been subject to discussions if replication of the bacteria outside of the mammalian host can occur (*25*, *67*, *68*). While Braun et al. (2015) found hints to possible replication in soil, no evidence for replication in invertebrate guts was observed *in vitro* by von Terzi et al. (2014) (*25*, *68*). For other members of the *B. cereus* group replication in the gut of soil asscociated invertebrates (*20*), as saprophytes in the rhizosphere of grass plants (*175*) and in sterilized soil (*176*) have been described.

We hypothesized that an extra-mammalian lifecycle might play a role for *Bcbva*, due to its similarity to members of the *B. cereus* group known to replicate in soil and in invertebrate hosts. To analyse *Bcbva*'s activity in the environment we compared its evolutionary rate under natural conditions in TNP determined through Bayesian inference (see 6.6.2) to the mutation rate during permanent replication *in vitro* (see 6.5.). Assuming that no mutation occurs during the dormant phase of spores due to the dehydrated nature of spores and the lack of UV light on the forest ground, the quotient of the two rates can indicate *Bcbva* activity patterns (*177*).

It was found that *Bcbva* evolved ~ 200 times slower in TNP than during the *in vitro* experiment meaning that *Bcbva* is only present in TNP in an actively replicating stage for 0,56 % of its lifetime representing 102 generations per year. This finding indicates that replication in the environment is not an important factor in the *Bcbva* lifecycle. One could argue that we might find slightly deleterious mutations *in vitro* that would be eliminated *in vivo*, leading to an artificially increased *in vitro* mutation rate. It was shown in an experimental study on *E. coli* that (deleterious) mutations occur more frequently during passaging on agar plates than during growth in a well-mixed environment (*178*). On the other hand selective pressure in the mammalian host might be increased compared to an ideal *in vitro* scenario favouring mutations *in vivo,* so that

Bcbva would need even less time in an active state to achieve the observed mutations. This has for example been observed during an anthrax outbreak in cattle in Sweden, where antibiotic treatment favoured mutations causing antibiotic resistancy (*102*). While it is clear that this estimate of the *Bcbva* mutation rate can only be a rough approximation, the reliability of the estimate is favoured by the fact that TNP is a naïve environment, where animals are generally not treated with antibiotics which could increase mutation of bacteria *in vivo* (*102*). The *Bcbva* lifecycle appears to be comparable to that of *B. anthracis* with the majority of replication taking place in the mammalian host. No significant amount of replication seems to occur in invertebrate hosts, limiting their capacities as vectors of the disease.

Another valuable information, when comparing *in vitro* mutation rate and evolutionary rate wold be the actual *in vivo* mutation rate. This would allow for a more accurate comparison between our calculations and what we observe in TNP. In a different setting this has been done for *Clostridium difficile* by Eyre et al. (2013): They investigated patients and derived mutation rates and within-host diversity from whole genome sequences obtained from the first and last sample from 145 patients at a median interval of 51 days allowing them to estimate the actual speed of mutation within the mammalian host (*179*). This is of course difficult to do in a wildlife disease that likely ends lethal, but an approximation of the *in vivo* mutation rate of *Bcbva* in animal models would be valuable.

Significant within-host diversity has been described for *B. anthracis* and we were interested to see how *Bcbva* compares (*102*). To detect the highest possible amount of variation within one animal, *Bcbva* isolates derived from different organs were included in the study if available. It was found that for *Bcbva* in TNP two SNPs is the upper limit of variation found within any one host (mammal or fly) and that the mean within-host diversity is 0.35 SNPs (Table 4) – this is almost identical to the 0.3 SNPs mean within-host diversity detected in the above mentioned study on *C. difficile* diversity by Eyre et al. (*179*).The only exception were two isolates found in a carrion fly; here we detected two distinct *Bcbva* isolates differing by 42 SNPs (Fly 600). These isolates differed even morphologically with one colony being especially mucous and phylogenetically they clustered with separate strains, suggesting that this one carrion fly was contaminated while feeding on two separate anthrax carcasses (Fig. 14), another indicator for high *Bcbva* activity. For *B. anthracis,* infections with mixed genotypes have been reported in studies using MLVA and SNPs (*102*, *121*). Animals in ENP and in the Pyrenées (France) were found to be infected with several (highly related) MLVA genotypes. During an anthrax outbreak in cows in Sweden isolates from the same host were found to differ by up to five SNPs in whole genome SNP analysis. In both studies the observed diversity was attributed to heterogeneity in the infectious dose, which occurs in the so called “champs maudits” areas (French for “cursed field”), where in the past big amounts of animals repeatedly got infected and died, leading to a heterogeneous mixture of anthrax genotypes in the soil. The fact that *Bcbva* and *C. difficile* show a similarly low within-host diversity supports the assumptions that heterogenous bacteria within a host originate in the infectious dose, rather than developing during replication in the host. The lack of diversity in anthrax carcasses in TNP may be due to the fact that animals die evenly distributed over the research area, rather than in certain “champ maudit” areas. Another explanation would be that a different route of infection, other than grazing, could present a bottle neck against infection with heterogeneous infectious doses.

For *B. anthracis,* simultaneous outbreaks caused by different (MLVA) genotypes have been observed in ENP (*104*). However, according to the ENP within-host diversity study by Beyer et al. (2012) this could also be caused by above-mentioned heterogeneity in the infectious dose and must not necessarily suggest independent sources of infection (*121*). Several co-circulating *Bcbva* strains were detected during the short time period of the fly snapshot in TNP. Strains differed by up to 48 SNPs (median: 25), suggesting multiple contemporaneous sources of infection and transmission chains.
While correlation between genetic and geographic distances of isolates was significant in MRM testing (p=0.006), only a small fraction of the 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 (Fig. 14). For low genomic distances (≤ 35 SNPs), genomic and geographic distances of TNP isolates were positively correlated (Fig. 15), providing indication of spatially restricted transmission, typical for point-source infection likely originating in carcasses. The phylogeographic history of *Bcbva* in TNP was modelled in a Bayesian framework. Several continuous diffusion models were tested to infer dispersal patterns and the geographic origin of the TNP strains. This approach has successfully been applied to viral disease outbreaks in the past, for example for mapping the introduction and spread of West Nile virus lineage 2 in Europe (*180*). In our study, however, the 80 % highest posterior density (HPD) region (=confidence interval) of the inferred geographic location of the common ancestor covered the whole TNP research area (Fig. 11). This is likely due to the small geographic scale of the study in combination with the lifecycle of *Bcbva*, including latent phases in the ground, and prevents inference of the historic movements of *Bcbva* within TNP.
Bcbva causes anthrax in rain forest habitats differing from the savanna habitats typically known for anthrax infection and infects a broad range of unusual hosts. Thus, one important question in *Bcbva* research is to evaluate how wildlife in TNP gets infected with the pathogen. The limited diagnostic capacities of on-site pathology have so far not yet revealed specific organ lesions that would hint to a certain route of infection, but the absence of cutaneous lesions in the examined carcasses disagrees with cutaneous anthrax as cause of the fatalities. Histo-pathology of anthrax carcasses (*work by K. Mätz-Rensing, German primate Center*) showed Gram positive rod shaped bacteria in all tested organs as a result of anthrax bacterimea (*4*), but also no sign for the entry of *Bcbva* via one route or the other. However, the moist rain forest climate of TNP points against inhalational infection with *Bcbva*, as spores are not likely to get aerolized in the humid environment (*181*). The most likely scenario for anthrax infection in TNP is therefore ingestion via the oral route.

For some of the infected species (chimpanzee, mangabey, duiker, porcupine, mongoose) scenarios comparable to the classic route of anthrax infection via grazing can be imagined. All of these species feed, at least occasionally, on the ground and ingest soil while doing so. The chimpanzees in TNP are also known to hunt monkeys (*8*), which might present a possible route of infection, as sick and weak monkeys are especially prone to be caught. However, the vegetative form of the bacteria, present in a sick live monkey, would only be suitable for oral infection if lesions in the oral cavity are present, as vegetative bacilli are thought to be destroyed in the stomach (*127*). Duikers, on the other hand, have been observed to feed on carrion, usually containing fully developed spores, which could certainly present a source of oral infection for this species (*166*).

While different scenarios for the infection of the above mentioned species can be imagined, it is completely unclear, how the remaining, strictly arboreal, monkey species (colobus monkeys and diana monkeys) get in contact with *Bcbva*. As they never feed on the ground, infection with soil containing spores is not an option here (*9*). In the TNP *Bcbva* phylogeny strains infecting arboreal species did not cluster separately from strains infecting the species at least partially feeding on the ground, indicating transmission links between ground and canopy. Another indicator for the existence of alternative routes of infection is the fact that the two *Bcbva* sequences retrieved from carrion flies in Grebo (Liberia) nest within the diversity of *Bcbva* in TNP, with one of the Liberia branches splitting from the TNP branches as little as 30 years ago (Fig. 9). The geographic distance between the two sites could not have been bridged in this short time by point-source infection from carcasses alone. Further, as mentioned above, no significant *Bcbva* genetic diversity was discovered in carcasses indicating a possible bottle neck during transmission leading to homogeneity in the infectious dose. This could for example be due to infection through an (invertebrate) vector, rather than direct infection with contaminated soil. Non-biting flies, like carrion flies and house flies have been found to carry various bacteria, like *Salmonella* spp., *Shigella* spp. and also *Bacillus* spp (*182*). Their involvement in the transmission of anthrax has long been discussed (*63*). Experimental studies have shown that *B. anthracis* can persist in carrion flies for up to nine days and vegetation around anthrax carcasses has been found to be contaminated with fly droplets containing *B. anthracis* (*68*, *183*). The final proof of fly-mediated anthrax transmission in the field is still missin, however. Another example for invertebrate vector transmitted spore-forming bacteria, is the honey bee pathogen *Paenibacillus larvae*, that causes the lethal bee disease American foulbrood (*184*). *P. larvae* is known to be transmitted horizontally and vertically by adult bees who do not get infected themselves, but spread the spores in their larvae popuations and neighboring hives (*185*).

Bcbva DNA was detected in total in 5 % of randomly caught blowflies in TNP throughout the year (*this study and C. Hoffmann, PhD thesis*). It was further shown here that it is possible to culture living *Bcbva* from flies (n = 43).

To evaluate the capacity of carrion flies to transport *Bcbva* into the canopy, carrion flies were trapped in specially designed traps up to 25 m high in the canopy.

The finding of *Bcbva* positive flies in the canopy indicates that flies could be involved in the transmission of *Bcbva* to arboreal animals. However, the result of the above mentioned mutation rate-evolution rate comparison analysis speaks against replication of *Bcbva* in the environment (and flies), which would limit their vector capacities. Transmission rates in the Bayesian reconstruction of ancestral host stages also point against a substantial role of flies in the transmission of *Bcbva*.

It needs to be further investigated, if monkeys would get infected by ingesting flies directly or by feeding on leaves or fruits contaminated by fly feces/regurgitation droplets. Replication of *Bcbva* on certain fruits after contamination by flies could be necessary to reach a sufficient infectious dose. So far we can only state that carrion flies represent a useful *Bcbva* monitoring tool in TNP and that more research is necessary to clarify the role of carrion flies in the transmission of the pathogen. Basic entomological data about the carrion fly species present in TNP is missing, like information on their reproductive cycle, life span and the maxium distance and height they are able to fly. This data could partially be produced in the lab, but mark-recapture experiments would

be necessary in the field to understand the carrion fly moving patterns. In a second step carrion flies should be infected with *Bcbva in vitro* to tests for *Bcbva* impact on flies at different development stages, how long *Bcbva* can survive in flies, if the bacteria can germinate and replicate in the fly, and where exactly in the fly they are located. Experiments examining the potential of different TNP fruits as *Bcbva* culture medium would also be an important contribution to the discussion.

Generally, the search for vectors should be broadened. Vultures, suggested to spread *B. anthracis* in Kruger National Park (KNP), would be able to transport *Bcbva* spores over heights and distances, but they are not very common in TNP (*127*, *186*). More promising could be the testing of other ubiquitous invertebrate vectors, like termites and ants. Metagenomic studies in termites found them to harbor a broad range of bacteria in their gut, icluding bacteria of the genera *Bacillus* and *Paenibacillus* (*187*). Alternative scenarios like the transport of bacteria within plants, which has been suggested for other bacteria of the *B. cereus* group could also play a role in the cryptic lifestyle of *Bcbva* (*175*, *188*). Though few anthrax spores suffice to infect an animal intravenously, oral infective doses are very high experimentally and further factors like wounds or compromised host immune status likely play a role for the development of infection (*36*, *158*). This must be considered when new routes of infection are evaluated.

7.3 Measurable evolution of *Bcbva* in TNP- methodical aspects

A Bayesian approach was used linking pathogen sequence data with sampling date information to derive a rate for *Bcbva* evolution in TNP, further allowing us to estimate the time to the most recent common ancestor (tMRCA) of the *Bcbva* strains sampled in TNP. The field of phylodynamics that connects (full) genome sequencing data with epidemiological information has been developing fast over the last decade (*108*). Starting with rapidly evolving RNA viruses, today analyses are frequently performed on bacterial datasets. As genome size and evolutionary rate are usually inversely related, the amount of mutations that occur in the different species over a certain period has the same magnitude for most pathogens (*108*). For example, Hepatitis C virus evolves at a speed of 7.6 substitutions/genome/year (genome size: ~10000 bp, evolutionary rate: 0.79×10^{-3} substitutions/site/year) and *C. difficile* introduces 1.4 substitutions/genome/year (genome size: ~4300000 bp, evolutionary rate: 3.2×10^{-7} substitutions/site/year) (*177*). Our estimate of 0.2 substitutions/genome/year for *Bcbva* represents the lower end of that range and no estimate is available yet for *B. anthracis*. Some of the best studied examples for bacteria investigated in this field are *Escherichia coli*, *Salmonella typhimurium,* the clonal *Mycobacterium* spp. and spore forming *C. difficile* (*117*, *189–191*). For *C. difficile* detailed studies investigated evolutionary rates and patient-to-patient contact data in the UK and found that a large amount of patients gets infected with newly introduced *C. difficile* strains rather than strains that persist within hospitals (*177*, *179*). A similar discovery was made by Biek et al. (2012) in Ireland, when they traced contacts between dairy cow herds and compared them to the *Mycobacterium bovis* evolution in the area. Their findings demonstrated the significant importance of badgers in tuberculosis persistence (*190*). These examples underline how the field of phylodynamics can help us change our understanding of disease dynamics.

As described above successful inference of measurable evolution from pathogen sequences always depends on several conditions - most importantly an extensive serially collected sample set covering a sufficient period of time.

The present genome dataset from the continuous outbreak in TNP is unique in anthrax research regarding the number of samples available through time and space. The dataset contains whole *Bcbva* sequences originating in nine different mammal species and carrion flies covering a period from 1996 to 2014 (Table S4, Fig.16).

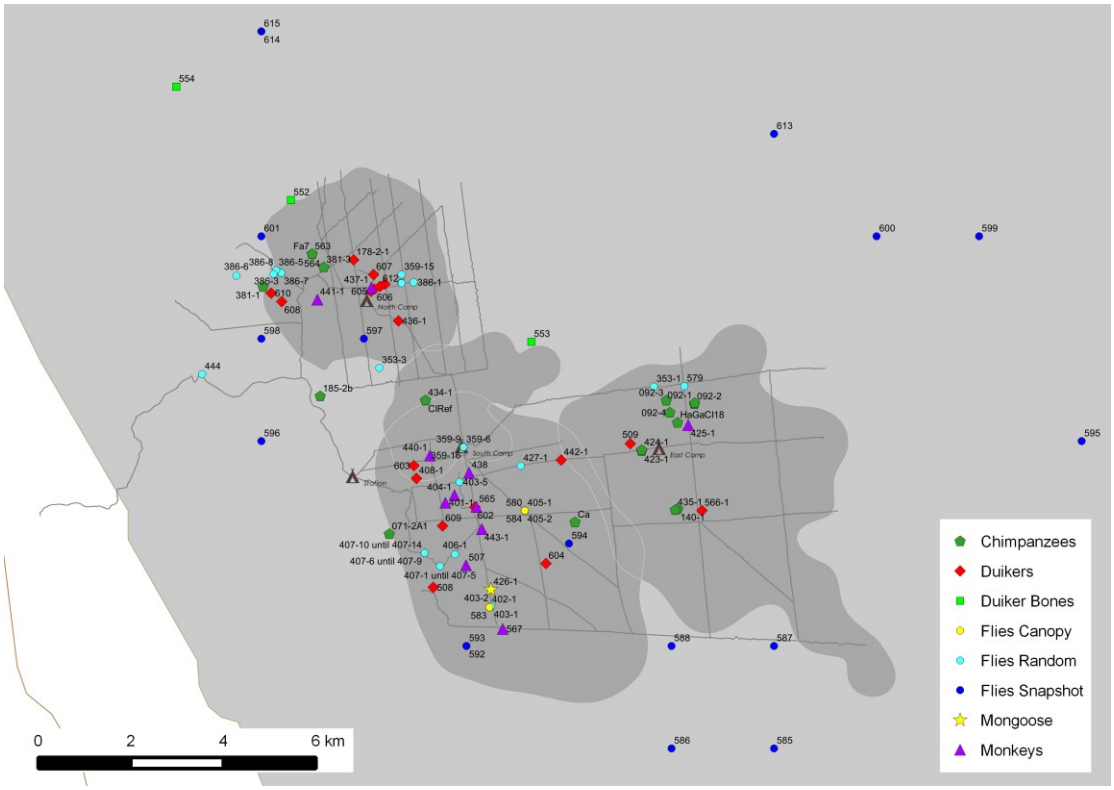

Figure 16. Geographic locations of *Bcbva* cases and *Bcbva* positive flies used for genetic analyses. The two isolates from Grebo (Liberia) included in our study are not shown, but the location of Grebo (~40 km from TNP) is presented in Figure 2. Numbers next to symbols represent the isolate numbers, under which the sequences can also be found in the European Nucleotide Archive (*published in* (*4*)).

However, even the most thorough sampling scheme can only be successful if sufficient genetic diversity occurs in the examined pathogen. The continuous activity of *Bcbva* in TNP has led to comparably high genetic diversity (see 6.4.1) enabling us to perform in depth analyses. The resulting estimates of evolutionary rate (3.92×10^{-8} substitutions/site/year) and tMRCA (146 years), do not only give unique insights into the ecology of *Bcbva* in TNP. On a bigger scale they allow for a new perspective on the *B.cereus* group.

Recombination could potentially obscure temporal signal, but *B. anthracis* is generally thought to be a highly clonal pathogen and the same is assumed for *Bcbva*. This is supported by the finding of the *Bcbva* specific genetic Island IV throughout *Bcbva* isolates from different sub-Saharan countries and the fact that all *Bcbva* isolates form a monophyletic clade within the *B. cereus* group (*6*).

In this study no significant within-host diversity was detected, with the exception of one carrion fly carrying two distinct *Bcbva* isolates (see 6.4.2). As we did not investigate direct epidemiological transmission links, within-host diversity would not affect the outcome of our analyses in any case.

The small 95 % confidence interval of the *Bcbva* evolutionary rate was a first indicator for a reliable estimate. To further test the robustness of our results against structural artefacts a randomization test was designed, in which sampling dates were randomly reassigned to the *Bcbva* sequences. Analyses were repeated with 20 date-randomized datasets. The 95 % confidence intervals for the evolutionary rates of real data and randomized data never overlapped, underlining the reliability of the information derived from our data. Date randomization tests can be evaluated using different levels of strictness. In this study the strictest possible criteria suggested by Duchêne et al. (2015) were applied and we are confident to have found robust estimates (*162*).

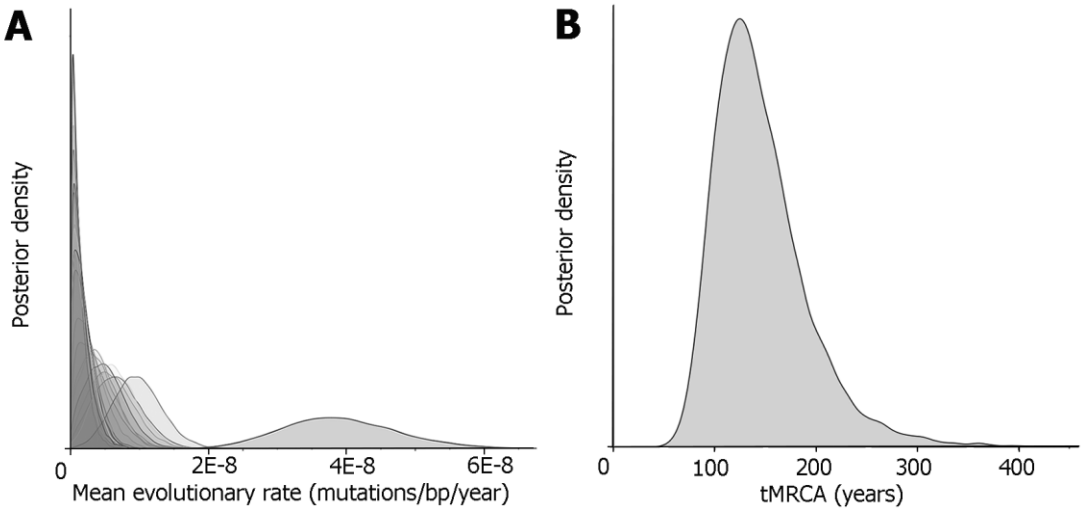

Figure 17. Overview results of genetic analyses of *Bcbva* isolates. (A) Posterior density distribution of evolutionary rates estimated by Bayesian analysis of chromosomal *Bcbva* sequences from TNP (Côte d'Ivoire) and neighboring Grebo (Liberia) (in grey) and from 20 date-randomized simulations (in colors). 95 % HPD intervals of real and simulated data do not overlap, validating the temporal signal in our data-set. (B) Posterior density distribution of the time to the most recent common ancestor of these *Bcbva* sequences (tMRCA).

The above mentioned analyses are all based on chromosomal sequences. Due to the smaller size of the plamids (pXO1: 181907 bp; pXO2: 94469 bp) compared to the chromosome (5196054 bp) the plasmids contain significantly less genetic information, preventing temporal estimates in this study.

Although we could not verify if pXO1 and pXO2 co-evolved with the chromosome in TNP, we propose that this is the most likely scenario. Antonation et al. (2016) (*6*) showed that, like the chromosomal sequences, plasmids of all sub-Saharan *Bcbva* strains cluster in a distinct monophyletic clade within the *B. cereus* group. Therefore the most parsimonious scenario for *Bcbva* evolution would be horizontal transfer of plasmids to a *Bcbva* ancestor in the past and subsequent co-evolution. Comprehensive *B. anthracis*, *B. cereus* and *Bcbva* isolate collections and phylodynamic analyses would be necessary to further illuminate the evolutionary history of the anthrax causing

members of the *B. cereus* group and to possibly predict the existence of further unusual anthrax causing agents.

8 Conclusions and Outlook

- *Bcbva* is a significant cause of wildlife mortality in the TNP ecosystem and has been present here for a long time, rather than presenting a newly introduced pathogen.
- Despite a low *in vitro* mutation rate comparable to that of classical *B. anthracis,* high genetic diversity was found in *Bcbva* isolates from TNP, indicating frequent *Bcbva* infections. Several genetically diverging *Bcbva* strains were found to circulate simultaneously in the TNP research area.
- While *Bcbva* prevalence is high throughout the TNP research area, serological testing of five frequently affected wildlife species only revealed a significant antibody response to the disease in two monkey individuals. Despite the limitations of our non-validated tests and small sample sizes, this is an indicator for high virulence of the pathogen in the tested species.
- Living *Bcbva* bacteria could be isolated from randomly caught carrion flies in TNP caught in all strata of the rainforest. Carrion flies could therefore present a source of infection for arboreal monkeys. However comparison of *Bcbva in vitro* mutation rate and *Bcbva* evolutionary rate points against replication of the pathogen outside of the mammalian host, limiting their vector capacities. The role of carrion flies as vectors of *Bcbva* needs to be further investigated and the search for vectors needs to be broadened. Evaluation of transmission pathways in a comparative approach with other sites of *Bcbva* occurrence (i.e. Dzanga-Sangha, Central African Republic) could help to identify relevant factors. *In vitro* experiments in flies should be performed using *Bcbva* and carrion fly species from TNP.
- A temporal signal for *Bcbva* evolution could be derived from the whole genome sequences from TNP. Combination of the TNP evolutionary rate with *Bcbva* sequence data from other sub-Saharan countries allowed for a conservative estimate of *Bcbva* presence in sub-Saharan Africa of ~2100 years. In the context of rapidly growing *B. anthracis* whole genome datasets our *Bcbva* data will be valuable input in the investigation of the origin of *B.anthracis.*
- Increased awareness for *Bcbva* is not only necessary to protect wildlife, and veterinarians working with wildlife, but also human populations in the area. On a bigger scale in the future the risk/ impact of anthrax, caused by *B.anthracis* and *Bcbva,* on humans and domestic animals in the sub-Saharan region needs to be re-assessed. This thesis was part of an ongoing DFG funded German-Ivorian project building capacities for the research on neglected diseases in Côte d'Ivoire. In the course of the project a study on anthrax seroprevalence in humans and domestic animals throughout Côte d'Ivoire is currently conducted and the development of assays for the serological differentiation between *B.anthracis* and *Bcbva* is in progress. Several meetings with decision makers and health services in Côte d'Ivoire have been held and the evaluation of risk factors for the human population will be part of the next funding phase of the project. We hope that this project can contribute to a better control and prevention of anthrax disease in sub-Saharan Africa in the future.

9 Abstract

Sudden death was observed in chimpanzees and gorillas in Côte d'Ivoire (Taï National Park) and Cameroon (Dja Reserve) in 2001 and 2004 and the causative agent was identified as a before unknown *Bacillus anthracis*-like bacterium designated *Bacillus cereus* biovar *anthracis* (*Bcbva*) (*1*, *2*). *Bcbva* combines the chromosomal background of *B. cereus* with the virulence plasmids of *B. anthracis* and was shown to be as virulent as *B. anthracis* in small animal models (*3*). As part of an ongoing long-term study in Taï National Park (TNP), autopsies on all available deceased wildlife were performed and environmental samples were collected. *Bcbva* positive samples dated back up to 1991 and *Bcbva* was identified as a continuous cause of death for a number of large mammal species (*4*).

The aim of the present study was to investigate the epidemiology and ecology of *Bcbva* in its natural rainforest habitat in TNP using a broad array of phylogenetic and serological methods.

Bacterial isolates could be retrieved from various sample types (tissue samples, bones, carrion flies) and 178 whole *Bcbva* genomes were derived covering a period from 1996 to 2014. SNP-based phylogenomic analysis in a Bayesian statistical framework revealed that *Bcbva* chromosomes had measurably evolved over this time span. This allowed us to determine that the median time to their most recent common ancestor (tMRCA) was 146 years (95 % highest posterior density (HPD): 70-237) and that 0.2 new mutations are introduced in average per *Bcbva* chromosome/year. Bayesian Markov Chain Monte Carlo (BMCMC) analyses were run under several demographic models. No model outperformed a constant population size model, suggesting *Bcbva* population size did not vary much over this period. To further inform the latter, *in vitro* experiments were performed determining the whole genome mutation rate of *Bcbva*, which was shown to be comparable to the *in vitro* mutation rate of *B. anthracis*. Comparison of *Bcbva* evolutionary rate and *in vitro* mutation rate approximated hypothetical *Bcbva* activity to be in average 102 generations/year, pointing against a significant extra mammalian lifecycle of *Bcbva.*The whole genome sequencing data provided evidence for multiple co-existing transmission chains that appear to be localized within different areas of the park. Genomic and geographic distances were correlated, particularly when genomic distances were low, reflecting the carcass-mediated dynamics of the disease. To reveal further possible transmission pathways, carrion flies were trapped in the rainforest canopy and we were able to isolate *Bcbva* from flies caught up to 25m high.

While we observe constant substantial pressure by *Bcbva* on the TNP fauna, only two monkeys were found to have significant amounts of antibodies against the disease in a serological study in five different wildlife species (n=59), indicating that *Bcbva* infections in these species are likely often lethal.

Combining data from TNP with sequence data from other sites of *Bcbva* occurrence revealed that *Bcbva* has been present in sub-Saharan Africa for millennia.

10 Zusammenfassung

Epidemiologie und Ökologie von *Bacillus cereus* biovar *anthracis* im Taï Nationalpark, Elfenbeinküste

Der plötzliche Tod von Schimpansen in Côte d'Ivoire (Taï National Park) und Gorillas in Kamerun (Dja Reservat) in den Jahren 2001 und 2004 wurde von einem unbekannten *Bacillus anthracis*-ähnlichen Erreger verursacht, der später als *Bacillus cereus* biovar *anthracis* (*Bcbva*) bezeichnet wurde (*1*, *2*). *Bcbva* kombiniert den chromosomalen Hintergrund von *Bacillus cereus* mit den Virulenzplasmiden von *Bacillus anthracis*. In Tierversuchen wurde gezeigt, dass die Virulenz von *B. cereus* und *B. anthracis* vergleichbar ist (*3*). Im Rahmen einer Langzeitstudie im Taï National Park (TNP) wurden alle gefundenen Wildtierkadaver autopsiert und Umweltproben gesammelt. *Bcbva* positive Proben reichen zurück bis 1991 und es konnte gezeigt werden, dass *Bcbva* kontinuierlich Todesfälle bei mehreren Säugetierarten verursacht (*4*).

Das Ziel dieser Dissertation war es, unter Anwendung phylogenetischer und serologischer Methoden die Epidemiologie und Ökologie von *Bcbva* im Taï Nationalpark, dem natürlichen Regenwaldhabitat des Erregers, zu untersuchen.

Bcbva konnte aus verschiedenen Probenmaterialien kultiviert werden (Gewebe, Knochen, Aasfliegen) und 178 *Bcbva* Vollgenome wurden sequenzanalysiert. Diese decken den Zeitraum von 1996 bis 2014 ab. SNP-basierte phylogenetische Analysen unter Anwendung bayesischer statistischer Methoden zeigten eine messbare Evolution des Erregers über diesen Zeitraum. Die Zeit bis zum letzten gemeinsamen Vorfahren der Isolate (tMRCA) beträgt 146 Jahre (95% Konfidenzintervall: 70-237 Jahre) und pro *Bcbva* Chromosom und Jahr enstehen im Durchschnitt 0.2 neue Mutationen. In bayesischen Monte-Carlo-Sampling-Verfahren (BMCMC) unter Annahme verschiedener demografischer Szenarien war kein Modell wahrscheinlicher als die Annahme einer konstanten Bakterienpopulation. Dies deutet darauf hin, dass die Größe der Bakterienpopulation über diesen Zeitraum nicht bedeutend variiert hat. Um diese Ergebnisse besser einordnen zu können, wurde die Mutationsrate des *Bcbva* Gesamtgenoms in *in vitro* Experimenten bestimmt. Diese ist vergleichbar mit der Mutationsrate von *B. anthracis*. Der Vergleich der Evolutions- und der Mutationsrate von *Bcbva* deutet darauf hin, dass der Erreger im TNP im Durschnitt pro Jahr nur 102 Generationen durchläuft. Dies spricht gegen eine bedeutende Vermehrung des Erregers außerhalb von Säugetierwirten. Weiterhin deckten die Sequenzdaten mehrere parallele Infektionsketten in unterschiedlichen Teilen des Parks auf. Für kleine genomische Distanzen sind genomische und geographische Distanz der Isolate korreliert, eine typische Folge der Ansteckung an eingegrenzten Quellen, z.B. Kadavern. Um weitere mögliche Übetragungswege aufzudecken, wurden Aasfliegen auch im Blätterdach des Regenwaldes gefangen. *Bcbva* konnte aus Fliegen aus einer Höhe von bis zu 25m kultiviert werden.

Trotz des großflächigen, konstanten Drucks von *Bcbva* auf die TNP Fauna zeigten in einer serologischen Studie in fünf Wildtierspezies (n=59) nur zwei Affen eine eindeutige Antikörperantwort gegen Milzbrand. Dies ist ein Hinweis dafür, dass *Bcbva*-Infektionen in den getesteten Spezies oft tödlich enden.

Abschließend führte eine Kombination unserer im TNP erhobenen Daten mit *Bcbva* Sequenzdaten aus anderen Ländern zu der Einschätzung, dass *Bcbva* schon seit mehreren tausend Jahren im Afrika südlich der Sahara existiert.

11 References

1. F. H. Leendertz *et al.*, Anthrax kills wild chimpanzees in a tropical rainforest. *Nature.* **430**, 451–452 (2004).

2. F. H. Leendertz *et al.*, Anthrax in Western and Central African great apes. *Am. J. Primatol.* **68**, 928–933 (2006).

3. C. Brézillon *et al.*, Capsules, toxins and AtxA as virulence factors of emerging *Bacillus cereus* biovar *anthracis*. *PLoS Negl. Trop. Dis.* **9**, e0003455 (2015).

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

5. S. R. Klee *et al.*, Characterization of *Bacillus anthracis*-like bacteria isolated from wild great apes from Côte d'Ivoire and Cameroon. *J. Bacteriol.* **188**, 5333–5344 (2006).

6. K. S. Antonation *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).

7. H. H. Berry, Surveillance and control of anthrax and rabies in wild herbivores and carnivores in Namibia. *Rev. Sci. Tech.* **12**, 137–146 (1993).

8. C. Boesch, H. Boesch-Achermann, *The chimpanzees of the Taï Forest: Behavioural ecology and evolution* (Oxford University Press, 2000).

9. S. McGraw, K. Zuberbühler, R. Noë, Eds., *Monkeys of the Taï Forest* (Cambridge University Press, 2007).

10. G. Campbell, H. Kuehl, A. Diarrassouba, P. K. N'goran, C. Boesch, Long-term research sites as refugia for threatened and over-harvested species. *Biol. Lett.* **7**, 723–726 (2011).

11. G. Campbell, H. Kuehl, P. N'Goran Kouamé, C. Boesch, Alarming decline of West African chimpanzees in Côte d'Ivoire. *Curr. Biol.* **18**, R903–R904 (2008).

12. B. Le Guenno *et al.*, Isolation and partial characterisation of a new strain of Ebola virus. *Lancet.* **345**, 1271–1274 (1995).

13. A. Radonic *et al.*, Fatal monkeypox in wild-living sooty mangabey, Côte d'Ivoire, 2012. *Emerg. Infect. Dis.* **20**, 1009–1011 (2014).

14. F. G. Priest, M. Barker, L. W. J. Baillie, E. C. Holmes, M. C. J. Maiden, Population structure and evolution of the *Bacillus cereus* group. *J. Bacteriol.* **186**, 7959–7970 (2004).

15. C. Ash, J. A. Farrow, M. Dorsch, E. Stackebrandt, M. D. Collins, Comparative analysis of *Bacillus anthracis, Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. *Int. J. Syst. Bacteriol.* **41**, 343–346 (1991).

16. A.-B. Kolstø, N. J. Tourasse, O. A. Økstad, What sets *Bacillus anthracis* apart from other *Bacillus*

species? *Annu. Rev. Microbiol.* **63**, 451–476 (2009).

17. A. Bravo, S. S. Gill, M. Soberón, Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Toxicon*. **49** (2007), pp. 423–435.

18. E. Helgason *et al.*, *Bacillus anthracis, Bacillus cereus*, and *Bacillus thuringiensis*--one species on the basis of genetic evidence. *Appl. Environ. Microbiol.* **66**, 2627–30 (2000).

19. E. J. Bottone, *Bacillus cereus*, a volatile human pathogen. *Clin. Microbiol. Rev.* **23**, 382–398 (2010).

20. G. B. Jensen, B. M. Hansen, J. Eilenberg, J. Mahillon, The hidden lifestyles of *Bacillus cereus* and relatives. *Environ. Microbiol.* **5** (2003), pp. 631–640.

21. M. Ehling-Schulz *et al.*, Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1. *BMC Microbiol.* **6**, 20 (2006).

22. L. P. Stenfors Arnesen, A. Fagerlund, P. E. Granum, From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol. Rev.* **32**, 579–606 (2008).

23. J. M. Miller *et al.*, Fulminating bacteremia and pneumonia due to *Bacillus cereus*. *J. Clin. Microbiol.* **35**, 504–507 (1997).

24. M. Mock, A. Fouet, Anthrax. *Annu. Rev. Microbiol.* **55**, 647–671 (2001).

25. P. Braun *et al.*, Microevolution of anthrax from a young ancestor (M.A.Y.A.) suggests a soil-borne life cycle of *Bacillus anthracis*. *PLoS One*. **10**, e0135346 (2015).

26. World Health Organization (WHO), World Organisation for Animal Health (OIE), Food and Agriculture Organization of the United Nations (FAO), *Anthrax in humans and animals* (Geneva, Fourth Edi., 2008).

27. M. E. Hugh-Jones, V. De Vos, Anthrax and wildlife. *Rev. Sci. Tech. L`Office Int. Des Epizoot.* **21**, 359–383 (2002).

28. S. H. Ringertz *et al.*, Injectional anthrax in a heroin skin-popper. *Lancet*. **356**, 1574–1575 (2000).

29. T. Berger, M. Kassirer, A. A. Aran, Injectional anthrax - New presentation of an old disease. *Eurosurveillance*. **19** (2014).

30. M. Hanczaruk *et al.*, Injectional anthrax in heroin users, Europe, 2000-2012. *Emerg. Infect. Dis.* **20**, 322–323 (2014).

31. R. Grunow *et al.*, Anthrax among heroin users in europe possibly caused by same *Bacillus anthracis* strain since 2000. *Eurosurveillance*. **18**, 1 (2013).

32. P. Keim *et al.*, Whole genome analysis of injectional anthrax identifies two disease clusters spanning more than 13 years. *EBioMedicine*. **2**, 1613–1618 (2015).

33. F. A. Abramova, L. M. Grinberg, O. V Yampolskaya, D. H. Walker, Pathology of inhalational anthrax in 42 cases from the Sverdlovsk outbreak of 1979. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 2291–2294 (1993).

References

34. J. A. Jernigan *et al.*, Bioterrorism-related inhalational anthrax: The first 10 cases reported in the United States. *Emerg. Infect. Dis.* **7**, 933–944 (2001).

35. A. R. Hoffmaster, C. C. Fitzgerald, E. Ribot, L. W. Mayer, T. Popovic, Molecular subtyping of *Bacillus anthracis* and the 2001 bioterrorism-associated anthrax outbreak, United States. *Emerg. Infect. Dis.* **8**, 1111–1116 (2002).

36. W. Beyer, P. C. B. Turnbull, Anthrax in animals. *Mol. Aspects Med.* **30**, 481–489 (2009).

37. M. Antwerpen, M. Elschner, W. Gaede, A. Schliephake, G. Grass, Genome sequence of *Bacillus anthracis* strain Stendal, isolated from an anthrax outbreak in cattle in Germany. *Am. Soc. Microiology.* **4**, 10–11 (2016).

38. Friedrich Loeffler Institute, Riems, Germany. *Anim. Dis. Report. Syst.*

39. OIE, *Terrestrial animal health code* (2013), vol. II.

40. T. Dixon, M. Meselson, J. Guillemin, P. Hanna, Anthrax. *New Engl. J. Med.* **341**, 815–826 (1999).

41. S. Krishna Rao, S. Mohiyudekn, Tabanus flies as transmitters of anthrax - a field experience. *Indian Vet. J.* **38**, 348–353 (1958).

42. C. K. Marston *et al.*, Molecular epidemiology of anthrax cases associated with recreational use of animal hides and yarn in the United States. *PLoS One.* **6**, e28274 (2011).

43. W. J. Bell, P. W. Laing, Pulmonary anthrax in cattle. *Vet. Rec.* **100**, 573–4 (1977).

44. Z. P. Weiner, I. J. Glomski, Updating perspectives on the initiation of *Bacillus anthracis* growth and dissemination through its host. *Infect. Immun.* **80**, 1626–1633 (2012).

45. C. Guidi-Rontani, The alveolar macrophage: the Trojan horse of *Bacillus anthracis*. *Trends Microbiol.* **10**, 405–409 (2002).

46. P. L. Goossens, J. N. Tournier, Crossing of the epithelial barriers by *Bacillus anthracis*: The Known and the Unknown. *Front. Microbiol.* **6**, 1–15 (2015).

47. I. J. Glomski, A. Piris-Gimenez, M. Huerre, M. Mock, P. L. Goossens, Primary involvement of pharynx and Peyer's patch in inhalational and intestinal anthrax. *PLoS Pathog.* **3**, 0699–0708 (2007).

48. A. Fouet, The surface of *Bacillus anthracis*. *Mol. Aspects Med.* **30**, 374–385 (2009).

49. T. Candela, M. Mock, A. Fouet, CapE, a 47-amino-acid peptide, is necessary for *Bacillus anthracis* polyglutamate capsule synthesis. *J. Bacteriol.* **187**, 7765–7772 (2005).

50. J. N. Tournier, A. Quesnel-Hellmann, A. Cleret, D. R. Vidal, Contribution of toxins to the pathogenesis of inhalational anthrax. *Cell. Microbiol.* **9**, 555–565 (2007).

51. K. Cunningham, D. B. Lacy, J. Mogridge, R. J. Collier, Mapping the lethal factor and edema factor binding sites on oligomeric anthrax protective antigen. *Proc. Natl. Acad. Sci. U. S. A.* **99**, 7049–7053 (2002).

52. J. L. Stanley, H. Smith, Purification of factor I and recognition of a third factor of the anthrax toxin. *J. Gen. Microbiol.* **26**, 49–63 (1961).

References

53. F. A. Beall, M. J. Taylor, C. B. Thorne, Rapid lethal effect in rats of a third component found upon fractionating the toxin of *Bacillus anthracis*. *J. Bacteriol.* **83**, 1274–1280 (1962).

54. J. O'Brien, A. Friedlander, T. Dreier, J. Ezzell, S. Leppla, Effects of anthrax toxin components on human neutrophils. *Infect. Immun.* **47**, 306–310 (1985).

55. F. J. Maldonado-Arocho, J. a. Fulcher, B. Lee, K. a. Bradley, Anthrax oedema toxin induces anthrax toxin receptor expression in monocyte-derived cells. *Mol. Microbiol.* **61**, 324–337 (2006).

56. J. C. Sirard, M. Mock, A. Fouet, The three *Bacillus anthracis* toxin genes are coordinately regulated by bicarbonate and temperature. *J. Bacteriol.* **176**, 5188–5192 (1994).

57. J. M. Bartkus, S. H. Leppla, Transcriptional regulation of the protective antigen gene of *Bacillus anthracis*. *Infect Immun.* **57**, 2295–2300 (1989).

58. A. Fouet, M. Mock, Differential influence of the two *Bacillus anthracis* plasmids on regulation of virulence gene expression. *Infect. Immun.* **64**, 4928–4932 (1996).

59. A. Fouet, AtxA, a Bacillus anthracis global virulence regulator. *Res. Microbiol.* **161**, 735–742 (2010).

60. P. C. B. Turnbull, Anthrax vaccines: past, present and future. *Vaccine.* **9**, 533–539 (1991).

61. W. C. Turner *et al.*, Fatal attraction: vegetation responses to nutrient inputs attract herbivores to infectious anthrax carcass sites. *Proc. Biol. Sci.* **281** (2014), doi:10.1098/rspb.2014.1785.

62. H. H. Ganz *et al.*, Interactions between *Bacillus anthracis* and plants may promote anthrax transmission. *PLoS Negl. Trop. Dis.* **8**, e2903 (2014).

63. M. Hugh-Jones, J. Blackburn, The ecology of *Bacillus anthracis*. *Mol. Aspects Med.* **30**, 356–67 (2009).

64. A. Fasanella *et al.*, Suspect vector transmission of human cutaneous anthrax during an animal outbreak in southern Italy. *Vector Borne Zoonotic Dis.* **13**, 769–71 (2013).

65. V. de Vos, The ecology of anthrax. *Salisbury Med Bull.* **68S**, 19–23 (1990).

66. J. K. Blackburn, M. Van Ert, J. C. Mullins, T. L. Hadfield, M. E. Hugh-Jones, The necrophagous fly anthrax transmission pathway: empirical and genetic evidence from wildlife epizootics. *Vector Borne Zoonotic Dis.* **14**, 576–83 (2014).

67. A. Fasanella *et al.*, Evaluation of the house fly *Musca domestica* as a mechanical vector for an anthrax. *PLoS One.* **5**, e12219 (2010).

68. B. von Terzi, P. C. B. Turnbull, S. E. Bellan, W. Beyer, Failure of Sterne- and Pasteur-like strains of *Bacillus anthracis* to replicate and survive in the urban bluebottle blow fly *Calliphora vicina* under laboratory conditions. *PLoS One.* **9**, e83860 (2014).

69. *http://www.promedmail.org/post/20160727.4373872*.

70. J. P. Dudley *et al.*, Carnivory in the common hippopotamus *Hippopotamus amphibius*: implications for the ecology and epidemiology of anthrax in African landscapes. *Mamm. Rev.* **46**, 191–203 (2015).

71. S. R. Klee *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).

References

72. F. H. Leendertz *et al.*, A new *Bacillus anthracis* found in wild chimpanzees and a gorilla from west and central Africa. *PLoS Pathog.* **2**, e8 (2006).

73. M. N. Van Ert *et al.*, Global genetic population structure of *Bacillus anthracis*. *PLoS One*. **2**, e461 (2007).

74. G. A. Van der Auwera, M. Feldgarden, R. Kolter, J. Mahillon, Whole-genome sequences of 94 environmental isolates of *Bacillus cereus* sensu lato. *Genome Announc.* **1**, e00380-13 (2013).

75. A. R. Hoffmaster *et al.*, Identification of anthrax toxin genes in a *Bacillus cereus* associated with an illness resembling inhalation anthrax. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 8449–54 (2004).

76. S.-Y. Oh, J. M. Budzik, G. Garufi, O. Schneewind, Two capsular polysaccharides enable *Bacillus cereus* G9241 to cause anthrax-like disease. *Mol. Microbiol.* **80**, 455–470 (2011).

77. A. R. Hoffmaster *et al.*, Characterization of *Bacillus cereus* isolates associated with fatal pneumonias: Strains are closely related to *Bacillus anthracis* and harbor B. anthracis virulence genes. *J. Clin. Microbiol.* **44**, 3352–3360 (2006).

78. C. K. Marston *et al.*, Anthrax toxin-expressing *Bacillus cereus* isolated from an anthrax-like eschar. *PLoS One*. **11**, e0156987 (2016).

79. P. Pilo *et al.*, Bovine *Bacillus anthracis* in Cameroon. *Appl. Environ. Microbiol.* **77**, 5818–21 (2011).

80. R. Patiño-Navarrete, V. Sanchis, Evolutionary processes and environmental factors underlying the genetic diversity and lifestyles of *Bacillus cereus* group bacteria. *Res. Microbiol.* **168**, 309–318 (2017).

81. T. Lembo *et al.*, Serologic surveillance of anthrax in the Serengeti ecosystem, Tanzania, 1996-2009. *Emerg. Infect. Dis.* **17**, 387–394 (2011).

82. K. Hampson *et al.*, Predictability of anthrax infection in the Serengeti, Tanzania. *J. Appl. Ecol.* **48**, 1333–1344 (2011).

83. C. A. Cizauskas, W. C. Turner, N. Pitts, W. M. Getz, Seasonal patterns of hormones, macroparasites, and microparasites in wild african ungulates: The interplay among stress, reproduction, and disease. *PLoS One*. **10**, 1–29 (2015).

84. H. Ebedes, Anthrax epizootics in Etosha National Park. *Madoqua*. **10**, 99–118 (1976).

85. P. M. Lindeque, P. C. Turnbull, Ecology and epidemiology of anthrax in the Etosha National Park, Namibia. *Onderstepoort J. Vet. Res.* **61**, 71–83 (1994).

86. C. A. Cizauskas, S. E. Bellan, W. C. Turner, R. E. Vance, W. M. Getz, Frequent and seasonally variable sublethal anthrax infections are accompanied by short-lived immunity in an endemic system. *J. Anim. Ecol.* **83**, 1078–1090 (2014).

87. K. L. Smith *et al.*, *Bacillus anthracis* diversity in Kruger National Park. *J Clin Microbiol.* **38**, 3780–3784 (2000).

88. L. E. O. Braak, V. de Vos, Feeding habits and flight-range of blow-flies (*Chrysomia* spp.) in relation to anthrax transmission in the Kruger National Park, South Africa. *Onderstepoort J. oaf Vet. Res.* **57**, 141–142 (1990).

References

89. M. M. Wafula, A. Patrick, T. Charles, Managing the 2004/05 anthrax outbreak in Queen Elizabeth and Lake Mburo National Parks, Uganda. *Afr. J. Ecol.* **46**, 24–31 (2008).

90. P. Turnbull *et al.*, Anthrax in wildlife in the Luangwa Valley, Zambia. *Vet. Rec.* **128**, 399–403 (1991).

91. S. B. Clegg, P. C. B. Turnbull, C. M. Foggin, P. M. Lindeque, Massive outbreak of anthrax in wildlife in the Malilangwe Wildlife Reserve, Zimbabwe. *Vet. Rec.* **160**, 113–118 (2007).

92. K. M. Good, A. Houser, L. Arntzen, P. C. B. Turnbull, Naturally acquired anthrax antibodies in a cheetah (*Acinonyx jubatus*) in Botswana. *J. Wildl. Dis.* **44**, 721–723 (2008).

93. P. K. Muoria *et al.*, Anthrax outbreak among Grevy's zebra (*Equus grevyi*) in Samburu, Kenya. *Afr. J. Ecol.* **45**, 483–489 (2007).

94. P. Keim *et al.*, Molecular evolution and diversity in *Bacillus anthracis* as detected by amplified fragment length polymorphism markers. *J. Bacteriol.* **179**, 818–824 (1997).

95. G. Patra, A. Fouet, J. Vaissaire, J. L. Guesdon, M. Mock, Variation in rRNA operon number as revealed by ribotyping of *Bacillus anthracis* strains. *Res. Microbiol.* **153**, 139–148 (2002).

96. G. L. Andersen, J. M. Simchock, K. H. Wilson, Identification of a region of genetic variability among *Bacillus anthracis* strains and related species. *J. Bacteriol.* **178**, 377–384 (1996).

97. P. J. Jackson *et al.*, Characterization of the variable-number tandem repeats in vrrA from different *Bacillus anthracis* isolates. *Appl. Environ. Microbiol.* **63**, 1400–1405 (1997).

98. P. Keim *et al.*, Anthrax molecular epidemiology and forensics: using the appropriate marker for different evolutionary scales. *Infect. Genet. Evol.* **4**, 205–13 (2004).

99. P. Keim *et al.*, Multiple-locus variable-number tandem repeat analysis reveals genetic relationships within *Bacillus anthracis*. *J. Bacteriol.* **182**, 2928–36 (2000).

100. F. Lista *et al.*, Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis. *BMC Microbiol.* **6**, 33 (2006).

101. D. Sue, C. K. Marston, A. R. Hoffmaster, P. P. Wilkins, Genetic diversity in a *Bacillus anthracis* historical collection (1954 to 1988). *J. Clin. Microbiol.* **45**, 1777–1782 (2007).

102. J. Ågren, M. Finn, B. Bengtsson, B. Segerman, Microevolution during an Anthrax outbreak leading to clonal heterogeneity and penicillin resistance. *PLoS One.* **9**, e89112 (2014).

103. G. Vergnaud *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).

104. W. Beyer *et al.*, Distribution and molecular evolution of *Bacillus anthracis* genotypes in Namibia. *PLoS Negl. Trop. Dis.* **6**, e1534 (2012).

105. A. J. Vogler *et al.*, Molecular analysis of rifampin resistance in *Bacillus anthracis* and *Bacillus cereus*. *Antimicrob. Agents Chemother.* **46**, 511–3 (2002).

106. L. J. Kenefic *et al.*, Pre-Columbian origins for North American anthrax. *PLoS One.* **4** (2009), doi:10.1371/journal.pone.0004813.

References

107. A. J. Drummond, O. G. Pybus, A. Rambaut, R. Forsberg, A. G. Rodrigo, Measurably evolving populations. *Trends Ecol. Evol.* **18**, 481–488 (2003).

108. R. Biek, O. G. Pybus, J. O. Lloyd-Smith, X. Didelot, Measurably evolving pathogens in the genomic era. *Trends Ecol. Evol.* **30**, 306–313 (2015).

109. B. Korber *et al.*, Timing the ancestor of the HIV-1 pandemic strains. *Science (80-.).* **288**, 1789–1796 (2000).

110. P. Lemey *et al.*, Tracing the origin and history of the HIV-2 epidemic. *Proc. Natl. Acad. Sci.* **100**, 6588–6592 (2003).

111. X. Didelot, R. Bowden, D. J. Wilson, T. E. a. Peto, D. W. Crook, Transforming clinical microbiology with bacterial genome sequencing. *Nat. Rev. Genet.* **13**, 601–612 (2012).

112. A. Drummond, A. Rambaut, BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

113. A. J. Drummond, S. Y. W. Ho, M. J. Phillips, A. Rambaut, Relaxed phylogenetics and dating with confidence. *PLoS Biol.* **4**, e88 (2006).

114. B. C. Emerson, E. Paradis, C. Thébaud, Revealing the demographic histories of species using DNA sequences. *Trends Ecol. Evol.* **16**, 707–716 (2001).

115. P. Lemey, A. Rambaut, A. J. Drummond, M. a Suchard, Bayesian phylogeography finds its roots. *PLoS Comput. Biol.* **5**, e1000520 (2009).

116. P. Lemey, A. Rambaut, J. J. Welch, M. a Suchard, Phylogeography takes a relaxed random walk in continuous space and time. *Mol. Biol. Evol.* **27**, 1877–85 (2010).

117. A. E. Mather *et al.*, Distinguishable epidemics within different hosts of the multidrug resistant zoonotic pathogen *Salmonella typhimurium* DT104. *Science (80-.).* **341**, 1514–1517 (2014).

118. I. Comas *et al.*, Out-of-Africa migration and Neolithic coexpansion of *Mycobacterium tuberculosis* with modern humans. *Nat. Genet.* **45**, 1176–1182 (2013).

119. P. Keim *et al.*, The genome and variation of *Bacillus anthracis*. *Mol. Aspects Med.* **30**, 397–405 (2009).

120. N. H. Smith, S. V. Gordon, R. de la Rua-Domenech, R. S. Clifton-Hadley, R. G. Hewinson, Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat. Rev. Microbiol.* **4**, 670–681 (2006).

121. W. Beyer, P. C. B. Turnbull, Co-infection of an animal with more than one genotype can occur in anthrax. *Lett. Appl. Microbiol.* **57**, 380–4 (2013).

122. K. H. Bagamian, K. A. Alexander, T. L. Hadfield, J. K. Blackburn, Ante- and postmortem diagnostic techniques for anthrax: Rethinking pathogen exposure and the geographic extent of the disease in wildlife. *J. Wildl. Dis.* **49**, 786–801 (2013).

123. S. E. Bellan, O. Gimenez, R. Choquet, W. M. Getz, A hierarchical distance sampling approach to estimating mortality rates from opportunistic carcass surveillance data. *Methods Ecol. Evol.* **4** (2013),

doi:10.1111/2041-210x.12021.

124. P. C. B. Turnbull, M. G. Broster, J. a. Carman, R. J. Manchee, J. Melling, Development of antibodies to protective antigen and lethal factor components of anthrax toxin in humans and guinea pigs and their relevance to protective immunity. *Infect. Immun.* **52**, 356–363 (1986).

125. P. C. B. Turnbull *et al.*, Further progress in understanding anthrax in the Etosha National Park, Namibia. *Madoqua.* **16**, 93–104 (1989).

126. P. C. B. Turnbull *et al.*, Vaccine-induced protection against anthrax in cheetah (*Acinonyx jubatus*) and black rhinoceros (*Diceros bicornis*). *Vaccine.* **22**, 3340–7 (2004).

127. P. C. B. Turnbull *et al.*, Naturally acquired antibodies to *Bacillus anthracis* protective antigen in vultures of southern Africa. *Onderstepoort J. Vet. Res.* **75**, 95–102 (2008).

128. S. Knauf, F. Dahlmann, E. K. Batamuzi, S. Frischmann, H. Liu, Validation of serological tests for the detection of antibodies against *Treponema pallidum* in nonhuman primates. *PLoS Negl. Trop. Dis.* **9**, 1–15 (2015).

129. R. Grunow *et al.*, Monitoring of ELISA-reactive antibodies against anthrax protective antigen (PA), lethal factor (LF), and toxin-neutralising antibodies in serum of individuals vaccinated against anthrax with the PA-based UK anthrax vaccine. *Vaccine.* **25**, 3679–83 (2007).

130. U. K. Hahn, M. Alex, C. P. Czerny, R. Böhm, W. Beyer, Protection of mice against challenge with *Bacillus anthracis* STI spores after DNA vaccination. *Int. J. Med. Microbiol.* **294**, 35–44 (2004).

131. R. Aloni-Grinstein *et al.*, Oral spore vaccine based on live attenuated nontoxinogenic *Bacillus anthracis* expressing recombinant mutant protective antigen. *Infect. Immun.* **73**, 4043–4053 (2005).

132. K. H. Bagamian *et al.*, Serological anthrax surveillance in wild boar (*Sus scrofa*) in Ukraine. *Vector Borne Zoonotic Dis.* **14**, 618–20 (2014).

133. P. C. Turnbull, M. Doganay, P. M. Lindeque, B. Aygen, J. McLaughlin, Serology and anthrax in humans, livestock and Etosha National Park wildlife. *Epidemiol. Infect.* **108**, 299–313 (1992).

134. D. J. New, Epidemiology of Anthrax outbreaks in wood bison (*Bison bison athabascae*) of the Mackenzie bison population. *PhD thesis* (2014).

135. A. J. Drummond, M. a Suchard, D. Xie, A. Rambaut, Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* **29**, 1969–73 (2012).

136. D. L. Ayres *et al.*, BEAGLE: An application programming interface and high-performance computing library for statistical phylogenetics. *Syst. Biol.* **61**, 170–173 (2012).

137. H. Li, Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv:1303.3997 [q-bio.GN].* (2013).

138. M. Kearse *et al.*, Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics.* **28**, 1647–1649 (2012).

139. M. A. DePristo *et al.*, A framework for variation discovery and genotyping using next- generation

DNA sequencing data. *Nat Genet.* **43**, 491–498 (2011).

140. G. A. Van der Auwera *et al.*, From FastQ data to high confidence variant calls: the Genonme Analysis Toolkit best practices pipeline. *Curr Protoc Bioinforma.* **11** (2014).

141. P. J. Ersts, Geographic Distance Matrix Generator [Software] Available at http://biodiversityinformatics.amnh.org/open_source/gdmg (2016).

142. I. Letunic, P. Bork, Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* (2016), doi:10.1093/nar/gkw290.

143. D. Darriba, G. L. Taboada, R. Doallo, D. Posada, jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods.* **9**, 772–772 (2012).

144. A. Rambaut, T. T. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2** (2016), doi:10.1093/ve/vew007.

145. S. Guindon *et al.*, New algorithms and methods to estimate maximum-likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321 (2010).

146. Picard Tools: A set of command line tools (in Java) for manipulating high-throughput sequencing (HTS) data and formats such as SAM/BAM/CRAM and VCF [Software]. Available at https://github.com/broadinstitute/picard.

147. QGIS Development Team, QGIS Geographic Information System. Open Source Geospatial Foundation Project. [Software]. Available at http://www.qgis.org/.

148. R Core Team, *R: A language and environment for statistical computing* (R Foundation for Statistical Computing, Vienna, Austria, 2015).

149. E. Paradis, J. Claude, K. Strimmer, APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics.* **20**, 289–290 (2004).

150. S. C. Goslee, D. L. Urban, The ecodist package for dissimilarity-based analysis of ecological data. *J. Stat. Softw.* **22**, 1–19 (2007).

151. A. Rieux, C. Khatchikian, TipDatingBeast: Using tip dates with phylogenetic trees in BEAST [Software]. R package version 1.0-2. Available at https://CRAN.R-project.org/package=TipDatingBeast (2016).

152. M. Gouy, S. Guindon, O. Gascuel, SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–4 (2010).

153. V. Buffalo, Scythe: A 3'-end adapter contaminant trimmer [Software]. Available at https://github.com/vsbuffalo/scythe.

154. N. Joshi, J. Fass, Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files [Software]. Available at https://github.com/najoshi/sickle (2011).

155. F. Bielejec, A. Rambaut, M. a. Suchard, P. Lemey, SPREAD: Spatial phylogenetic reconstruction of

evolutionary dynamics. *Bioinformatics*. **27**, 2910–2912 (2011).

156. A. Rambaut, M. Suchard, D. Xie, A. Drummond, Tracer [Software]. Available at http://beast.bio.ed.ac.uk/Tracer (2014).

157. P. Formenty *et al.*, Ebola virus outbreak among wild chimpanzees living in a rain forest of Cote d ' Ivoire. *J. Infect. Dis.* **179**, 120–126 (1999).

158. P. C. B. Turnbull, Guidelines for the surveillance and control of anthrax in humans and animals. *World Heal. Organ. Dep. Commun. Dis. Surveill. Response*. **3rd Editio** (1998).

159. W. Xie, P. O. Lewis, Y. Fan, L. Kuo, M. H. Chen, Improving marginal likelihood estimation for bayesian phylogenetic model selection. *Syst. Biol.* **60**, 150–160 (2011).

160. G. Baele *et al.*, Improving the accuracy of demographic and molecular clock model comparison while accommodating phylogenetic uncertainty. *Mol. Biol. Evol.* **29**, 2157–2167 (2012).

161. R. E. Kass, A. E. Raftery, Bayes Factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).

162. S. Duchêne, D. Duchêne, E. C. Holmes, S. Y. W. Ho, The performance of the date-randomisation test in phylogenetic analyses of time-structured virus data. *Mol. Biol. Evol.* **32** (2015), doi:10.1093/molbev/msv056.

163. M. Kimura, A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* **16**, 111–120 (1980).

164. T. H. Jukes, C. R. Cantor, Evolution of protein molecules. *Mamm. protein Metab.* **3**, 132 (1969).

165. M. Kimura, Estimation of evolutionary distances between homologous nucleotide sequences. *Proc. Natl. Acad. Sci.* **78**, 454–458 (1981).

166. H. S. Newing, Behavioural ecology of duikers (Cephalophus spp.) in forest and secondary growth, Tai, Côte d'Ivoire. *PhD thesis, Univ. Stirling* (1994).

167. S. Köndgen *et al.*, Pasteurella multocida Involved in Respiratory Disease of Wild Chimpanzees. *PLoS One*. **6**, e24236 (2011).

168. S. Köndgen, S. Schenk, G. Pauli, C. Boesch, F. H. Leendertz, Noninvasive monitoring of respiratory viruses in wild chimpanzees. *Ecohealth*. **7**, 332–41 (2010).

169. F. H. Leendertz *et al.*, Non-invasive testing reveals a high prevalence of simian T-lymphotropic virus type 1 antibodies in wild adult chimpanzees of the Taï National Park, Côte d'Ivoire. *J. Gen. Virol.* **85**, 3305–12 (2004).

170. S. Y. W. Ho, B. Shapiro, Skyline-plot methods for estimating demographic history from nucleotide sequences. *Mol. Ecol. Resour.* **11**, 423–434 (2011).

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

172. T. R. Gillespie, C. L. Nunn, F. H. Leendertz, Integrative approaches to the study of primate infectious disease: Implications for biodiversity conservation and global health. *Am. J. Phys. Anthropol.*

137, 53–69 (2008).

173. A. Fasanella *et al.*, Suspect vector transmission of human cutaneous anthrax during an animal outbreak in Southern Italy. *Vector Borne Zoonotic Dis.* **13**, 769–71 (2013).

174. R. D. E. MacPhee, A. D. Greenwood, Infectious disease, endangerment, and extinction. *Int. J. Evol. Biol.* **2013**, 1–9 (2013).

175. L. Halverson, M. Clayton, J. Handelsman, Population biology of *Bacillus cereus* UW85 in the rhizosphere of field-grown soybeans. *Soil Biol. Biochem.* **25**, 485–494 (1993).

176. S. Vilain, Y. Luo, M. B. Hildreth, V. S. Brozel, Analysis of the life cycle of the soil saprophyte *Bacillus cereus* in liquid soil extract and in soil. *Appl. Environ. Microbiol.* **72**, 4970–4977 (2006).

177. X. Didelot *et al.*, Microevolutionary analysis of *Clostridium difficile* genomes to investigate transmission. *Genome Biol.* **13**, R118 (2012).

178. L. Bosshard *et al.*, Accumulation of deleterious mutations during bacterial range expansions. *Genetics.* **in press** (2017).

179. D. W. Eyre *et al.*, Diverse sources of *C. difficile* infection identified on whole-genome sequencing. *N. Engl. J. Med.* **369**, 1195–1205 (2013).

180. G. Zehender *et al.*, Reconstructing the recent West Nile virus lineage 2 epidemic in Europe and Italy using discrete and continuous phylogeography. *PLoS One.* **12**, e0179679 (2017).

181. P. C. Turnbull, P. M. Lindeque, J. Le Roux, a M. Bennett, S. R. Parks, Airborne movement of anthrax spores from carcass sites in the Etosha National Park, Namibia. *J. Appl. Microbiol.* **84**, 667–76 (1998).

182. T. Chaiwong *et al.*, The blow fly, *Chrysomia megacephala*, and the house fly, *Musca domestica*, as mechanical vectors of pathogenic bacteria in Northeast Thailand. **31**, 336–346 (2014).

183. J. K. Blackburn, M. Van Ert, J. C. Mullins, T. L. Hadfield, M. E. Hugh-Jones, The necrophagous fly anthrax transmission pathway: empirical and genetic evidence from wildlife epizootics. *Vector Borne Zoonotic Dis.* **14**, 576–83 (2014).

184. J. Ebeling, H. Knispel, G. Hertlein, A. Fünfhaus, E. Genersch, Biology of *Paenibacillus larvae*, a deadly pathogen of honey bee larvae. *Appl. Microbiol. Biotechnol.* **100**, 7387–7395 (2016).

185. E. O. Jatulan, J. F. Rabajante, C. G. B. Banaay, A. C. Fajardo, E. C. Jose, A mathematical model of intra-colony spread of American foulbrood in European honeybees (*Apis mellifera*). *PLoS One.* **10**, 1–13 (2015).

186. J.M. Thiollay, The birds of Ivory Coast. *Malimbus.* **7**, 1–59 (1985).

187. M. Wenzel, I. Schönig, M. Berchtold, P. Kämpfer, H. König, Aerobic and facultatively anaerobic cellulolytic bacteria from the gut of the termite *Zootermopsis angusticollis*. *J. Appl. Microbiol.* **92**, 32–40 (2002).

188. E. Saile, T. M. Koehler, *Bacillus anthracis* multiplication, persistence, and genetic exchange in the

rhizosphere of grass plants. *Apllied Environ. Microbiol.* **72**, 3168–3174 (2006).

189. A. von Mentzer *et al.*, Identification of enterotoxigenic *Escherichia coli* (ETEC) clades with long-term global distribution. *Nat. Genet.* **46**, 1321–1326 (2014).

190. R. Biek *et al.*, Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. *PLoS Pathog.* **8**, e1003008 (2012).

191. D. W. Eyre *et al.*, Detection of mixed infection from bacterial whole genome sequence data allows assessment of its role in *Clostridium difficile* transmission. *PLoS Comput. Biol.* **9**, e1003059 (2013).

12 Appendix

12.1 Supplementary tables

Table S1. All necropsies conducted in TNP from 1996 until 4/2015 and results for *Bcbva* testing (table by C. Hoffmann and F. Zimmermann).

Species	Date of necropsy	PCR positive	Culture	WGS	Longitude	Latitude	Other known cause of death
*Procolobus badius**	18.06.1996	x	negative (blood)				
*Pan troglodytes verus**	23.08.1996	x	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	negative (spleen)				
*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	x			
*Cercopithecus campbelli**	14.05.1998						

*Procolobus verus**	16.05.1998						
*Pan troglodytes verus**	16.08.1998	x	negative (lung)				
*Procolobus badius**	26.08.1998						
*Procolobus badius**	09.10.1998						
*Pan troglodytes verus**	21.10.1998	x	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	x			
undefined rodent	23.09.2001						
Pan troglodytes verus	15.10.2001	x	x	x	-7,350100	5,876500	
Pan troglodytes verus	15.10.2001	x	x	x	-7,350100	5,876500	
Pan troglodytes verus	15.10.2001	x	x	x	-7,350100	5,876500	
Pan troglodytes verus	11.12.2001						
Procolobus badius	05.02.2002	x	negative (bone with tissue)				
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	x	-7,348521	5,848744	
Pan troglodytes verus	01.03.2002	x	not done (hair)		-7,348521	5,848744	
Pan troglodytes verus	01.03.2002	x	not done (hair)		-7,348521	5,848744	
Procolobus badius	11.03.2002						
Pan troglodytes verus	11.06.2002	x	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	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	negative (bone)				
Procolobus badius	19.07.2005						
Hexaprotodon liberiensis	20.07.2005						
Procolobus badius	13.08.2005						
Pan troglodytes verus	02.09.2005						
Perodicticus	15.10.2005						
Soricidae	12.01.2006						
Perodicticus	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	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	x	-7,341995	5,875391
Cercopithecus diana	20.04.2008					
Pan troglodytes verus	May 2008	x	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	x	-7,278780	5,826733

Perodicticus	24.01.2009						
Manidae	05.02.2009						
Procolobus badius	30.03.2009						
Pan troglodytes verus	14.04.2009	x	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	x	-7,285815	5,838058	
Pan troglodytes verus	23.08.2011	x	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 done (tooth)		-7,324173	5,785753	
Cercopithecus diana	01.09.2011						
Cercopithecus campbelli	19.12.2011	x	x	x	-7,349070	5,867626	
Cercopithecus petaurista	02.01.2012	x	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	x	-7,274059	5,826447
Cephalophus sp.	01.02.2012	x	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	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	x	-7,312917	5,803317
Herpestidae	29.09.2012	x	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	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	x	-7,324103	5,827982
Cercocebus atys	12.05.2013	x	x	x	-7,322338	5,829451
Cephalophus sp.	30.05.2013	x	x	x	-7,329750	5,832719
Cercocebus atys	07.11.2013				-7,332019	5,871347
Cephalophus sp.	02.12.2013	x	x	x	-7,326472	5,811450
Cephalophus sp.	29.12.2013	x	x	x	-7,288041	5,839493

Procolobus badius	06.01.2014				-7,310206	5,834261
Pan troglodytes verus	08.01.2014	x	not done		-7,282501	5,823462
Colobus polykomos	09.01.2014	x	not done		-7,310540	5,822172
Manidae	12.01.2014				-7,331242	5,862474
Pan troglodytes verus	15.01.2014	x	not done		-7,307575	5,823502
Cephalophus sp.	17.02.2014	x	x	x	-7,318345	5,827097
Cephalophus sp.	21.02.2014	x	x	x	-7,330257	5,835207
Cercocebus atys	21.02.2014	x	not done		-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	x	-7,304491	5,816105
Cercopithecus diana	13.03.2014				-7,347766	5,865760
Cephalophus sp.	20.03.2014	x	x	x	-7,338608	5,869250
Cephalophus sp.	11.04.2014	x	x	x	-7,338011	5,869692
Cephalophus sp.	17.04.2014	x	x	x	-7,338120	5,872531
Cercocebus atys	19.04.2014				-7,348171	5,878638
Herpestidae	27.04.2014	x	not done		-7,275368	5,820312
Cephalophus sp.	28.04.2014	x	x	x	-7,356042	5,867231
Pan troglodytes verus	02.05.2014				-7,293111	5,796023
Hystricidae	06.05.2014	x	not done		-7,335862	5,872913
Cercocebus atys	10.05.2014	x	not done		-7,336855	5,869896
Cephalophus sp.	13.05.2014	x	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	x	-7,358087	5,868892
Cephalophus sp.	04.06.2014	x	x	x	-7,335895	5,870653
Cephalophus sp.	08.06.2014	x	x	x	-7,336863	5,870204

Colobus polykomos	11.06.2014			-7,290994	5,824645
Cercocebus atys	29.06.2014	x	not done	-7,325633	5,835628
Pan troglodytes verus	22.07.2014	x	not done	-7,309839	5,824033
Pan troglodytes verus	29.10.2014	x	not done	-7,259531	5,805741
Cephalophus sp.	19.12.2014	x	not done	-7,289299	5,838710
Cephalophus sp.	24.12.2014			-7,304824	5,825347
Cephalophus sp.	29.12.2014	x	not done	-7,308611	5,824093
Cephalophus sp.	01.01.2015	x	not done	-7,276469	5,823480
Cercocebus atys	07.01.2015	x	not done	-7,298985	5,854464
Cephalophus sp.	07.01.2015			-7,307416	5,822154
Cephalophus sp.	08.01.2015	x	not done	-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 done	-7,307164	5,803652
Genetta sp.	05.02.2015			-7,332095	5,873363
Cercocebus atys	07.02.2015	x	not done	-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 done	-7,316495	5,820553
Cephalophus sp.	20.04.2015			-7,298468	5,852265

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

Table S2. Canopy Flies.

PCR, culture and WGS for flies trapped in the canopy in TNP.

Date	PCR positive	Culture	WGS	Longitude	Latitiude
07.05.2013				-7,31552	5,80755
15.05.2013				-7,30861	5,82643
15.05.2013				-7,30861	5,82643
16.05.2013				-7,30861	5,82643
17.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
18.05.2013				-7,30861	5,82643
20.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,31552	5,80755
25.05.2013				-7,31552	5,80755
25.05.2013	x (22m)	x	583	-7,31552	5,80755
25.05.2013				-7,31552	5,80755
25.05.2013				-7,31552	5,80755
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013	x (25m)	not done		-7,30861	5,82643
25.05.2013				-7,30861	5,82643
25.05.2013	x (25m)	not done		-7,30861	5,82643
25.05.2013	x (25m)	x	584	-7,30861	5,82643
25.05.2013				-7,30861	5,82643

25.05.2013	x (25m)	not done		-7,30861	5,82643
26.05.2013				-7,31552	5,80755
26.05.2013				-7,31552	5,80755
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013	x (25m)	not done		-7,30861	5,82643
26.05.2013				-7,31552	5,80755
26.05.2013				-7,31552	5,80755
26.05.2013				-7,30861	5,82643
26.05.2013	x (25m)	not done		-7,30861	5,82643
26.05.2013	x (25m)	x	580/ 1-4	-7,30861	5,82643
26.05.2013	x (25m)	x	581	-7,30861	5,82643
26.05.2013	x (25m)	x	582	-7,30861	5,82643
26.05.2013	x (25m)	not done		-7,30861	5,82643
26.05.2013	x (25m)	not done		-7,30861	5,82643
26.05.2013				-7,30861	5,82643
26.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,30861	5,82643
27.05.2013				-7,31552	5,80755
27.05.2013				-7,31552	5,80755
27.05.2013				-7,31552	5,80755
27.05.2013				-7,31552	5,80755
27.05.2013				-7,31552	5,80755
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
27.05.2013				-7,3086	5,8264
16.06.2013				-7,32223	5,81790
17.06.2013				-7,32223	5,81790

17.06.2013	-7,32223	5,81790
17.06.2013	-7,3222	5,8179
17.06.2013	-7,3222	5,8179
18.06.2013	-7,32816	5,81817
18.06.2013	-7,32816	5,81817
18.06.2013	-7,32816	5,81817
18.06.2013	-7,32816	5,81817
18.06.2013	-7,32816	5,81817
18.06.2013	-7,32223	5,81790
18.06.2013	-7,32223	5,81790
18.06.2013	-7,32223	5,81790
18.06.2013	-7,32223	5,81790
18.06.2013	-7,3222	5,8179
18.06.2013	-7,3222	5,8179
18.06.2013	-7,3222	5,8179
18.06.2013	-7,3222	5,8179
18.06.2013	-7,3222	5,8179
19.06.2013	-7,32223	5,81790

Table S3. Ground flies.

Culture and WGS for PCR positive flies received from C. Hoffmann (randomly caught flies and Snapshot flies from TNP and flies from Grebo, Liberia).

Date	Project	Culture[1]	WGS	Origin	Longitude	Latitiude
31.12.2008	random	x	427-1	TNP	-7.30939	5.83517
22.04.2009	random	negative		TNP	-7.27751	5.85075
16.05.2009	random	negative		TNP	-7,32057	5,83883
16.05.2009	random	x	353-2	TNP	-7,32057	5,83883
26.04.2009	random	x	579	TNP	-7.27751	5.85075
04.03.2012	random	x	not done	TNP	-7,337	5,8543
15.03.2012	random	x	359-16	TNP	-7.32057	5.83883
01.07.2012	random	x	359-12	TNP	-7.33272	5.87092
15.03.2012	random	x	not done	TNP	-7.32057	5.83883
15.03.2012	random	x	359-6	TNP	-7.32057	5.83883
22.04.2009	random	negative		TNP	-7.27751	5.85075
22.04.2009	random	x	353-1	TNP	-7.28345	5.85065
01.07.2012	random	x	359-15	TNP	-7.33265	5.87253
18.09.2012	random	x	386-1	TNP	-7.3303	5.87101
18.09.2012	random	x	386-2	TNP	-7.33264	5.87084
10.10.2012	random	x	386-3	TNP	-7.3561	5.87286
10.10.2012	random	negative		TNP	-7.35712	5.87338
10.10.2012	random	x	386-5	TNP	-7.35712	5.87338
10.10.2012	random	x	386-6	TNP	-7.36487	5.8723
10.10.2012	random	x	386-7	TNP	-7.3561	5.87286

10.10.2012	random	x	386-8	TNP	-7.35712	5.87338
10.10.2012	random	x	386-9	TNP	-7.35764	5.87263
11.10.2012	random	x	444	TNP	-7.37155	5.85308
17.05.2013	random	negative		TNP	-7.34182	5.91426
09.03.2013	random	negative		TNP	-7.31604	5.88683
02.06.2014	Snapshot	x	585	TNP	-7,26	5,78
01.06.2014	Snapshot	x	586	TNP	-7,28	5,78
01.06.2014	Snapshot	x	587	TNP	-7,26	5,8
01.06.2014	Snapshot	x	588/ 1-4	TNP	-7,28	5,8
07.06.2014	Snapshot	x	592	TNP	-7,32	5,8
07.06.2014	Snapshot	x	593/ 1-4	TNP	-7,32	5,8
07.06.2014	Snapshot	x	594	TNP	-7,3	5,82
30.05.2014	Snapshot	x	595	TNP	-7,2	5,84
31.05.2014	Snapshot	negative		TNP	-7,26	5,84
07.06.2014	Snapshot	x	596	TNP	-7,36	5,84
27.05.2014	Snapshot	x	597/1-4	TNP	-7,34	5,86
09.06.2014	Snapshot	x	598	TNP	-7,36	5,86
31.05.2014	Snapshot	x	599	TNP	-7,22	5,88
30.05.2014	Snapshot	x	600/ 1-2	TNP	-7,24	5,88
09.06.2014	Snapshot	x	601	TNP	-7,36	5,88
30.05.2014	Snapshot	x	613	TNP	-7,26	5,9
25.05.2014	Snapshot	x	614/1-4	TNP	-7,36	5,92
25.05.2014	Snapshot	x	615	TNP	-7,36	5,92
01.02.2013	random	x	505/1-4	Grebo	NA	NA
02.02.2013	random	x	506/1-4	Grebo	NA	NA

An additional 21 Bcbva isolates were retrieved during a pilot study in TNP through direct culture of 204 flies (caught from April-June 2013) in the field without preceding PCR testing.

Table S4. All sequenced *Bcbva* strains from TNP and Grebo.

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).

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.

Raw reads for the *Bcbva* isolates from the *in vitro* mutation rate experiment are available in the European Nucleotide Archive (ENA) under the project accession number PRJEB14597, sample accession numbers ERS1222218 to ERS1222233.

Isolate ID	Sampling Date	Species	Material	Isolates	Chemistry	Average Coverage	Included in Phylogeography
071-2A1	May 2008	*Pan troglodytes verus*	skin/hair	1	v1 Rapid Run 2x150bp	319	x
092- 1A	18.04.2009	*Pan troglodytes verus*	Spleen	1	v1 Rapid Run 2x150bp	376	x
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	x
178-2-1	20.04.2008	*Cephalophus* sp.	Spleen	1	v1 Rapid Run 2x150bp	296	x
178-6-1	23.02.2006	*Cephalophus* sp.	bone with meat	1	v1 Rapid Run 2x150bp	352	no GPS data
178-7-1	19.08.2004	*Cephalophus* sp.	Lung	1	v2 Rapid Run 2x250bp	90	no GPS data
185-2b	01.03.2002	*Pan troglodytes verus*	connective tissue	1	v1 Rapid Run 2x150bp	283	x
353-1	22.04.2009	*Calliphoridae*	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	291	x
353-2	16.05.2009	*Calliphoridae*	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	253	x
353-3	27.02.2012	*Calliphoridae*	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	343	x
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	x
359-12	01.07.2012	*Calliphoridae*	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	391	x
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	x
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	x
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 DNA extraction	1	v2 Rapid Run 2x250bp	103	
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	x
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	x
402-1	25.04.2013	*Calliphoridae*	fly direct	1	v2 Rapid Run 2x250bp	79	x
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	x
404-1	12.05.2013	*Cercocebus atys*	Liver	1	v2 Rapid Run 2x250bp	49	x
405-1	16.05.2013	*Calliphoridae*	fly direct	1	v2 Rapid Run 2x250bp	113	x
405-2	17.05.2013	*Calliphoridae*	fly direct	1	v2 Rapid Run 2x250bp	72	x
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	x
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	x
408-1	30.05.2013	*Cephalophus* sp.	Fat	1	v1 Rapid Run 2x150bp	326	x
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	x
425-1	21.01.2012	*Cercopithecus diana*	Spleen	1	v1 Rapid Run 2x150bp	254	x
426-1	29.09.2012	*Herpestidae*	Liver	1	v1 Rapid Run 2x150bp	307	x
427-1	31.12.2008	*Calliphoridae*	fly mush after DNA extraction	1	v1 Rapid Run 2x150bp	360	x
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	x
436-1	01.02.2012	*Cephalophus* sp.	Lung	1	v2 Rapid Run 2x250bp	115	x
437-1	05.06.2012	*Procolobus badius*	Spleen	1	v2 Rapid Run 2x250bp	115	x
438-1	30.10.2012	*Cercopithecus diana*	Spleen	1	v2 Rapid Run 2x250bp	110	x
440-1	01.03.2013	*Cercocebus atys*	Spleen	1	v2 Rapid Run 2x250bp	67	x
441-1	19.12.2011	*Cercopithecus campbelli*	Muscle	1	v2 Rapid Run 2x250bp	132	x
442-1	14.03.2012	*Cephalophus* sp.	Spleen	1	v2 Rapid Run 2x250bp	145	x
443-1	05.03.2013	Cercocebus atys	Spleen	1	v2 Rapid Run 2x250bp	50	x
444	11.10.2012	*Calliphoridae*	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	31	x
459	16.04.1998	*Cercocebus atys*	Liver	1	v2 Rapid Run 2x250bp	88	no GPS data
460&529	23.08.1996	*Pan troglodytes verus*	liver & bones	2	v2 Rapid Run 2x250bp	116/33	no GPS data
461	21.10.1998	*Pan troglodytes verus*	Muscle	1	v2 Rapid Run 2x250bp	94	no GPS data
462	26.09.2000	*Pan troglodytes verus*	Lung	1	v2 Rapid Run 2x250bp	87	no GPS data
503-1	08.12.2001	*Pan troglodytes verus*	Bone	1	v2 Rapid Run 2x250bp	95	no GPS data

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	135/53/52/107/40/61	x
508 (1-6)	02.12.2013	*Cephalophus* sp.	liver & spleen	6	v2 Rapid Run 2x250bp	50/51/67/36/89/53	x
509 (1-6)	29.12.2013	*Cephalophus* sp.	liver & spleen	6	v2 Rapid Run 2x250bp/ v3 2x300bp (509-6)	174/113/62/63/86/17	x
530	2003	*Colobus polycomos*	Bone	1	v2 Rapid Run 2x250bp	27	no GPS data
531	2004	*Cercocebus atys*	Bone	1	v2 Rapid Run 2x250bp	29	no GPS data
552	October 2013	*Cephalophus* sp.	Bone	1	v2 Rapid Run 2x250bp	28	x
553	27.09.2013	*Cephalophus* sp.	Bone	1	v2 Rapid Run 2x250bp	31	x
554	05.11.2013	*Cephalophus* sp.	Bone	1	v2 Rapid Run 2x250bp	39	x
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	31/31/19/24/19/25	
565 (1-6)	02.01.2012	*Cercopithecus petaurista*	Muscle	6	v2 Rapid Run 2x250bp	26/31/25/32/34/31	x
566	25.01.2012	*Cephalophus* sp.	Muscle	1	v2 Rapid Run 2x250bp	32	x
567	20.09.2012	*Colobus* sp.	Maggot	1	v2 Rapid Run 2x250bp	26	x
579	26.04.2009	*Calliphoridae*	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	18	x
580 (1-4)	26.05.2013	*Calliphoridae*[2)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	16/25/18/26	x
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	x
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	x
586	01.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	49	x
587	01.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	39	x
588 (1-4)	01.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	21/24/26/26	x

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	x
594	07.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	39	x
595	30.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	95	x
596	07.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	31	x
597 (1-4)	27.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	19/20/23/24/	x
598	09.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	19	x
599	31.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	23	x
600 (1-2)	30.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	2	v2 Rapid Run 2x250bp	24/ 44	600-2
601	09.06.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	19	x
602	17.02.2014	*Cephalophus* sp.	corpse swab	1	v2 Rapid Run 2x250bp	28	x
603	21.02.2014	*Cephalophus* sp.	Spleen	1	v2 Rapid Run 2x250bp	22	x
604	04.03.2014	*Cephalophus* sp.	Spleen	1	v2 Rapid Run 2x250bp	24	x
605	20.03.2014	*Cephalophus* sp.	Spleen	1	v2 Rapid Run 2x250bp	28	x
606	11.04.2014	*Cephalophus* sp.	Skin	1	v2 Rapid Run 2x250bp	28	x
607 (1-6)	17.04.2014	*Cephalophus* sp.	Spleen	6	v2 Rapid Run 2x250bp	5/45/50/26/23/17	x
608	28.04.2014	*Cephalophus* sp.	Spleen	1	v2 Rapid Run 2x250bp	29	x
609	13.05.2014	*Cephalophus* sp.	Liver	1	v2 Rapid Run 2x250bp	28	x
610	03.06.2014	*Cephalophus* sp.	Liver	1	v2 Rapid Run 2x250bp	22	x
611	04.06.2014	*Cephalophus* sp.	Liver	1	v2 Rapid Run 2x250bp	24	x
612	08.06.2014	*Cephalophus* sp.	Lung	1	v2 Rapid Run 2x250bp	26	x
613	30.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	1	v2 Rapid Run 2x250bp	42	x
614 (1-4)	25.05.2014	*Calliphoridae*[3)]	fly mush after DNA extraction	4	v2 Rapid Run 2x250bp	69/26/22/29	x
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	x
CIRef	14.02.2002	*Pan troglodytes verus*	Spleen	1	v2 Rapid Run 2x250bp	263	x

Fa7	15.10.2001	*Pan troglodytes verus*	Spleen	1	v1 Rapid Run 2x150bp	321	x
HaGaCI8	07.04.2008	*Pan troglodytes verus*	spleen/liver	1	v1 Rapid Run 2x150bp	264	x

[1)] Samples from Grebo, Liberia

[2)]Caught in canopy

[3)] Caught during snapshot

Mutation Rate Experiment						
MR Start	21.01.2014	NA	NA	1	v2 Rapid Run 2x250bp	102
MR 1	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	95
MR 2	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	76
MR 3	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	68
MR 4	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	64
MR 5	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	104
MR 6	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	82
MR 7	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	80
MR 8	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	75
MR 9	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	150
MR 10	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	100
MR 11	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	103
MR 12	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	82
MR 13	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	126
MR 14	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	94
MR 15	18.02.2014	NA	NA	1	v2 Rapid Run 2x250bp	90

Table S5. Results for BMCMC reconstruction of temporal signal on plasmid level.

The model for pXO1 converged, but failed randomization testing.The model for pXO2 did not converge.

Plasmid	Molecular clock prior	tMRCA (years)	Evolutionary rate(μ) (substitutions/site/year)	log marginal likelihood[1)]	comment
pXO1	strict clock	39 (20-67)	3.9×10^{-7} (1.5 -6.5)	-252645.14	**Failed randomization test**
	uncorrelated lognormal relaxed clock	42 (20-76)	3.9×10^{-7} (1.5 -6.7)	-252648.63	Date randomization not done, as more conservative strict clock already failed
pXO2		NA	NA	NA	**No convergence of model**

[1)] estimated by stepping stone sampling

Table S6. Individual serological data (ELISA, WB).

Table S6 contains information on the origin of individual samples and results for Western Blot and ELISA testing. Given are the mean measured ELISA OD_{450nm} values for all samples tested (sorted by species) and their respective values for the assay internal negative control and standard deviation (see caption). Measurements for ELISA were conducted in a 1:500 dilution and the mean of the blanks was subtracted. To assess the quality of the measurements the coefficient of variation was calculated and readings were rerun if it was higher than 0.2. All Western Blot results refer to a 1:1000 sample dilution.

ID	species	date	sex	sample type	PA ELISA	ELISA (cut-off: mean NC + 2*SD)	PA Western Blot	LF Western Blot
R1	Western red colobus monkey	09.12.2006	m	Plasma	0,0123 [1]	neg	neg	neg
R2	Western red colobus monkey	09.01.2007	m	Plasma	0,44 [2]	**pos**	**pos**	**pos**
R3	Western red colobus monkey	09.01.2007	NA	Plasma	0,0038 [1]	neg	neg	neg
R4	Western red colobus monkey	12.01.2007	m	Plasma	0,0043 [3]	neg	neg	neg
R5	Western red colobus monkey	17.01.2007	m	Plasma	0,0023 [3]	neg	neg	neg
R6	Western red colobus monkey	20.01.2007	f	Plasma	0,0083 [1]	neg	neg	neg
R7	Western red colobus monkey	21.01.2007	f	Plasma	0,0028 [3]	neg	neg	neg
R8	Western red colobus monkey	28.01.2007	m	Plasma	0,006 [2]	neg	neg	neg
R9	Western red colobus monkey	31.01.2007	f	Plasma	0 [1]	neg	neg	neg
R10	Western red colobus monkey	01.02.2007	m	Plasma	0,0148 [1]	neg	neg	neg
R11	Western red colobus monkey	05.05.2012	m	Plasma	0,0018 [1]	neg	neg	neg
R12	Western red colobus monkey	07.05.2012	m	Plasma	0,0013 [3]	neg	neg	neg
R13	Western red colobus monkey	11.05.2012	f	Plasma	0,0028 [3]	neg	neg	neg
R14	Western red colobus monkey	18.05.2012	m	Plasma	0,0053 [3]	neg	neg	neg
R15	Western red colobus monkey	12.07.2012	f	Plasma	0,0333 [1]	neg	**pos**	neg
B1	Black-and-white colobus monkey	07.01.2007	m	Plasma	0,0043 [3]	neg	neg	neg

B2	Black-and-white colobus monkey	18.01.2007	f	Plasma	0,0153 [3]	neg	neg	neg
B3	Black-and-white colobus monkey	30.01.2007	m	Plasma	0,0048 [3]	neg	neg	neg
B4	Black-and-white colobus monkey	02.02.2007	f	Plasma	0,0108 [1]	neg	neg	neg
B5	Black-and-white colobus monkey	08.06.2007	f	Plasma	0,00188 [3]	neg	neg	neg
B6	Black-and-white colobus monkey	13.06.2007	f	Plasma	0,0203 [3]	neg	neg	neg
B7	Black-and-white colobus monkey	14.06.2007	m	Plasma	0,0063 [3]	neg	neg	neg
B8	Black-and-white colobus monkey	15.06.2007	f	Plasma	0,0018 [3]	neg	neg	neg
B9	Black-and-white colobus monkey	16.06.2007	m	Plasma	0,0028 [3]	neg	neg	neg
B10	Black-and-white colobus monkey	16.06.2007	m	Plasma	1,0328 [1]	**pos**	**pos**	**pos**
M1	Sooty mangabey	08.03.2010	m	Plasma	0,0415 [2]	neg	neg	neg
M2	Sooty mangabey	09.03.2010	m	Plasma	0,0533 [1]	neg	neg	neg
M3	Sooty mangabey	27.08.2013	f	Plasma	0,0905 [2]	neg	neg	neg
M4	Sooty mangabey	31.01.2014	m	EDTA blood	0,0218 [4]	neg	neg	neg
M5	Sooty mangabey	03.02.2014	m	EDTA blood	0,0098 [3]	neg	neg	neg
M6	Sooty mangabey	02.02.2014	m	EDTA blood	0,0033 [1]	neg	neg	neg
M7	Sooty mangabey	07.02.2014	f	Plasma	0,101 [2]	neg	neg	neg
M8	Sooty mangabey	31.05.2014	f	Plasma	0,0593 [3]	neg	neg	neg
M9	Sooty mangabey	31.05.2014	f	Plasma	0,0825 [2]	neg	neg	neg
M10	Sooty mangabey	03.06.2014	f	Plasma	0,0453 [3]	neg	neg	neg

M11	Sooty mangabey	04.06.2014	f	Plasma	0,0345 [2]	neg	neg	neg
M12	Sooty mangabey	07.06.2014	m	Plasma	0,0523 [4]	neg	neg	neg
M13	Sooty mangabey	12.06.2014	f	Plasma	0,0363 [1]	neg	neg	neg
M14	Sooty mangabey	26.11.2015	m	Plasma	0,0405 [2]	neg	neg	neg
M15	Sooty mangabey	28.11.2015	m	Plasma	0,058 [2]	neg	neg	neg
M16	Sooty mangabey	29.11.2015	m	Plasma	0,03075 [4]	neg	neg	neg
C1	Chimpanzee	03.05.2009	m	Serum	0,04 [2]	neg	neg	neg
C2	Chimpanzee	19.03.2004	f	necropsy blood	0,0358 [4]	neg	neg	neg
C3	Chimpanzee	Feb 06	f	necropsy blood	0,0613 [1]	neg	neg	neg
C4	Chimpanzee	09.02.2006	m	necropsy blood	0,0113 [4]	neg	neg	neg
C5	Chimpanzee	08.12.2009	f	necropsy blood	0,0023 [4]	neg	neg	neg
C6	Chimpanzee	07.12.2009	f	necropsy blood	0,032 [2]	neg	neg	neg
C7	Chimpanzee	07.12.2009	f	necropsy blood	0,0058 [4]	neg	neg	neg
C8	Chimpanzee	11.12.2009		necropsy blood	0,0148 [4]	neg	neg	neg
C9	Chimpanzee	02.12.2009	f	necropsy blood	0 [4]	neg	neg	neg
D1	Maxwell's duiker	05.06.2013	m	Plasma	0,021 [5]	neg	neg	neg
D2	Maxwell's duiker	06.06.2013	m	Plasma	0,0805 [5]	**pos**	neg	neg
D3	Maxwell's duiker	07.06.2013	f	Plasma	0,0545 [5]	**pos**	neg	neg
D4	Maxwell's duiker	08.06.2013	f	Plasma	0,0265 [5]	neg	neg	neg
D5	Maxwell's duiker	11.06.2013	m	Plasma	0,017 [5]	neg	neg	neg
D6	Maxwell's duiker	25.06.2014	f	Plasma	0,0565 [5]	**pos**	neg	neg
D7	Maxwell's duiker	30.06.2014	m	Plasma	0,065 [5]	**pos**	neg	neg
D8	Maxwell's duiker	03.07.2014	m	Plasma	0,0075 [5]	neg	neg	neg

D9	Maxwell's duiker	05.07.2014	f	Plasma	0,0215 [5]	neg	neg	neg

[1] Mean negative control for this plate: 0,068; standard deviation: 0,04

[2] Mean negative control for this plate: 0,0533 ; standard deviation: 0,0289

[3] Mean negative control for this plate: 0,0595; standard deviation: 0,0393

[4] Mean negative control for this plate: 0,0388; standard deviation: 0,0242

[5] Mean negative control for this plate: 0,015; standard deviation: 0,011

12.2 Supplementary Figures

Figure S1. Maximum Likelihood tree based on chromosomal sequences of *Bcbva* isolates from TNP (Côte d'Ivoire) and Grebo (Liberia) – detailed version with sampling dates. Bootstrap values are shown above all internal branches. The tree was rooted using the "best-fit" option in *Path-O-Gen* v1.2. The scale of the tree is given in substitutions per site (referring to the 298 bp long alignment of variable sites).

13 Ethics statement

All samples were collected with permission of the Ivorian Ministry of the Environment and Forests as well as the Ministry of Research and authorization of the directorship of Tai National Park (permits Nr. 048/MESRS/DGRSIT/KCS/TM and 90/MESRS/DGRSIT/mo). All wildlife samples have been exported under permission of the according CITES (Convention on International Trade in Endangered Species of Wild Fauna and Flora).

14 Published work

Research articles

Zimmermann F., Köhler S.M., Nowak K., Dupke S., Barduhn A., Düx A., Lang A., De Nys H.M., Gogarten J.F., Grunow R., Couacy-Hymann E., Wittig R.M., Klee S.R,, Leendertz F.H. (2017). Low antibody prevalence against *Bacillus cereus* biovar *anthracis* in Taï National Park, Côte d'Ivoire, indicates high rate of lethal infections in wildlife. *PLoS Neglected Tropical Diseases,* Sep 21;11(9):e0005960

Hoffmann C. & Zimmermann F.[*], Biek R., Kuehl H., Nowak K., Mundry R., Agbor A., Angedakin S., Arandjelovic M., Blankenburg A., Brazolla G., Corogenes K., Couacy-Hymann E., Deschner T., Dieguez P., Dierks K., Düx A., Dupke S., Eshuis H., Formenty P., Ginath Yuh Y., Goedmakers A., Gogarten J.F., Granjon A., McGraw S., Grunow R., Hart J., Jones S., Junker J., Kiang J., Langergraber K., Lapuente J., Lee K., Leendertz S.A., Léguillon F., Leinert V., Löhrich T., Marrocoli S., Mätz-Rensing K., Meier A., Merkel K., Metzger S., Murai M., De Nys H.M., Sachse A., Schenk S., van Schijndel J., Thiesen U., Ton E., Wu D., Wieler L.H., Boesch C., Klee S.R., Wittig, R.M., Calvignac-Spencer S., Leendertz F.H. (2017). Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. *Nature,* 548(7665),82-86.

[*] Equal contributions

Antonation K. S., Grützmacher K., Dupke S., Mabon P., **Zimmermann F.,** Lankester F., Peller T., Feistner A., Todd A., Herbinger I., DeNys H.M., Muyembe-Tamfun J., Karhemere S., Wittig R.M., Couacy-Hymann E., Grunow R., Calvignac-Spencer S., Corbett C.R., Klee S.R., Leendertz F.H. (2016). *Bacillus cereus* biovar *anthracis* causing anthrax in sub-Saharan Africa- Chromosomal monophyly and broad geographic distribution. *PLoS Neglected Tropical Diseases,* 10(9), 1-14.

Leendertz, F. H., Hoffmann C., **Zimmermann F.** (2016). Chapter „Anthrax". *International Encyclopedia of Primatology*

Saéz, A. M., Weiss, S., Nowak, K., Lapeyre, V., **Zimmermann F.**, Düx A., Kühl H., Kaba M., Regnaut S., Merkel K., Sachse A., Thiesen U., Villányi L., Boesch C., Dabrowski P.W., Radonic A., Nitsche A., Leendertz S.J.,Petterson S., Becker S., Krähling V., Couacy-Hymann E., Akoua-Koffi C., Weber N., Schaade L., Fahr J., Borchert M., Gogarten J.F., Calvignac-Spencer S. & Leendertz, F. H. (2015). Investigating the zoonotic origin of the West African Ebola epidemic. *EMBO Mol Med, 7*(1), 17–23.

Oral presentations

11/2016 **F. Zimmermann,** C. Hoffmann, R. Biek, K. Nowak, S. Dupke, E. Couacy-Hymann, R. Grunow, S. Klee, S. Calvignac-Spencer, F. Leendertz: Population dynamics of *Bacillus cereus* biovar *anthracis* in Tai National Park, Côte d'Ivoire. *The Biology of Anthrax Conference, St.Petersburg, USA.*

04/2016 **F. Zimmermann**, C. Hoffmann, K. Nowak, S. Dupke, E. Couacy-Hymann, R. Grunow, S. Calvignac- Spencer, S. Klee, F. Leendertz: Using full genome SNP Analysis as a tool to investigate the epidemiology and ecology of *Bacillus cereus* biovar *anthracis. Symposium Functional Molecular Infection Epidemiology, Berlin, Germany.*

11/ 2015 **F. Zimmermann**, S. Dupke, S. Klee, F. Leendertz : Collection et conservation des échantillons- Qu'est-ce qu'on fait avec un cas suspect d'Anthrax? *Workshop for the Detection and Prevention of Anthrax in Côte d'Ivoire, Bingerville, Côte d'Ivoire.*

10/ 2015 **F. Zimmermann**, C. Hoffmann, K. Nowak, S. Dupke, E. Couacy-Hymann, R. Grunow, S. Calvignac- Spencer, S. Klee, F. Leendertz: Using full genome SNP analysis as a tool to investigate the epidemiology and ecology of *Bacillus cereus* biovar *anthracis. Bacillus ACT (anthracis, cereus, thuringiensis) Conference, Delhi, India.*

Poster presentations

04/ 2016 **F.Zimmermann,** C. Hoffmann, K. Nowak, S. Dupke, E. Couacy-Hymann, R. Grunow, S. Calvignac-Spencer, S. R. Klee, F. Leendertz: A genomic approach to understanding the epidemiology and ecology of a new anthrax causing bacterium from tropical Africa. *Medical Biodefense Conference, Munich, Germany.*

10/2014 **F. Zimmermann,** S. Dupke, S. Klee, M. Stockhausen, C. Hoffmann, K. Merkel, T. Franz, E. Couacy-Hymann, S. Calvignac- Spencer, R. Grunow, F. Leendertz: A DFG-funded project to analyze epidemiology and ecology of *Bacillus cereus* biovar *anthracis* in Côte d'Ivoire- are blowflies a potential vector? *Medical Biodefense Conference, Munich, Germany.*

9/2013 **F. Zimmermann,** S. Dupke, S. Klee, E. Couacy-Hymann, F. Leendertz: How do the chimpanzees in Taï National Park, Côte d'Ivoire, get Anthrax? A field study to reveal transmission pathways. *National Symposium for Zoonoses Research, Berlin, Germany*

Awards & Stipends

11/2016	Travel stipend by the organizers of *The Biology of Anthrax Conference*, St.Petersburg, USA
04/2016	1st prize poster award, Medical Biodefense Conference, Munich, Germany.

15 Acknowledgements

This work was undertaken at the Robert Koch Institute from October 2012 to September 2016 and was a cooperation between the departments ZBS2 (Highly pathogenic microorganisms) and P3 (Epidemiology of highly pathogenic microorganisms).

I would like to thank my three brilliant supervisors, who all introduced me to new worlds on their own.

Fabian - thank you for giving me the opportunity to work in this unique environment on this fascinating topic and introducing me to Taï National Park. Your vision, courage and enthusiasm were an inspiration.

Silke, you are such a passionate microbiologist with an answer to every question and an open mind for crazy ideas. It was a pleasure to work with you – and also to buy dresses in tansanian street stalls, to celebrate the soccer world cup in Abidjan or my birthday in the BSL3 lab.

Sébastien supervised the data analysis of this work - the best compliment of many I can give you is that you made me interested in things I never knew I wanted to learn. You are a fantastic teacher.

Constanze - I could not have asked for a better partner in crime. Working with you was always a pleasure, from nights where we had to share a 90 cm bed in the rain forest surrounded by spiders and fire ants to celebratory prosecco on the steps of Josephsplatz. It started with two related PhD subjects, but I found a friend for life.

I was lucky to also work with three great postdocs. Susann, whose creativity in the lab, organizing skills and ability to find a solution for all kinds of problems (even sometimes if it was just sharing a bottle of wine) were invaluable. Kathrin, who can catch every kind of animal and waded with me through the dark and flooded jungle filled with crocodiles and leopards. Susanne, it was a pleasure to explore the mysteries of serology with you – and the world of karaoke.

I would like to thank Roland Grunow for always taking the time to discuss ideas and giving me detailed feedback, which certainly improved this work.

I also thank Lothar Wieler, my supervisor at FU Berlin, for his interest in this work and his continuous support – despite all his many tasks and responsibilities, his students are still a top priority.

Roman Biek, thank you for your input, for always taking time to think through my problems and ideas and for your very valuable feedback.

I am very grateful that Professor Peter-Henning Clausen and Professor Wolfgang Beyer have agreed to review this thesis.

I was lucky to work at RKI in two different groups that I both love very much: P3 and ZBS2.

Everyone in P3 was involved in this project in one way or the other. Working, exchanging ideas and celebrating together with you all turned my years in Berlin into an unforgettable experience. I would especially like to thank Ulla, Kevin and Andreas – the heart of P3 - for their help with DNA extraction and PCR testing of the TNP tissue

samples and generally for having a big heart for the veterinarians in the lab. Ari, thanks for the duiker catching support. Kim, thanks for introducing me to the world of wildlife anesthesia. Hélène, thanks for teaching me how to do a necropsy. Maja, Grit, Jan, Alex, Freda, Siv Aina, Doris, Resi, Livia, Verena, Melanie, thank you for help with logistics, advice on stats and always being open for discussions and after work beers. Markus and Ferda, thanks for making all of this possible by finding samples, materials and generally everything. And last but not least: thanks Modi - I wouldn't have been here without you!

I was also a member of ZBS2 and it was a pleasure to work here. The work in the BSL3 lab wouldn't have been possible without Tanja, Silke, Petra, Sabine and Andy, who went out of their way to help me- even at weekends. Also, some of the older TNP tissue samples were cultured by the ZBS2 team. Anne, I always enjoyed working in your serology lab with you - thank you for all your patience and support with weird wildlife samples. And generally thanks to everyone for your ideas, discussions, cakes, great christmas parties and birthday prosecco.

I would also like to thank our Ivorian partners. Emmanuel and Chantal always welcomed me warmly in Côte d'Ivoire and helped with any kind of problem. It was a pleasure to work with Moussa, Felicité and Gragnon at LANADA and in Berlin. I hope you will make many interesting discoveries in the future.

Jule Hinzmann, Wojciech Dabrowski and Torsten Semmler from RKI and Graham Hamilton from Glasgow Polyomics were very supportive during sequencing and bioinformatic analyses.

Thank you to our partners from the Max Planck Institute for Evolutionary Anthropology, especially the guys that supported me during the field missions: Vera and Sylvain were the camp managers of the Taï Chimpanzee Project during my time in the forest and together with Roman Wittig they enabled us to go to all the interesting places and take all the necessary samples. Thanks also to the Taï Monkey Project staff for helping out and keeping me company.

A big thank you to my team of field assistants in TNP who were always open for all kinds of fly or duiker catching missions. Special thanks to the person that accompanied me during all my time in TNP and that taught my so many things about the forest: Nestor Gouyan Bah. Also none of this would have been possible without the two best drivers in Côte d'Ivoire: Sidiki and Youssouf.

I would like to thank the authorities in Côte d'Ivoire, especially the Ministry of the Environment and Forests and the Ministry of Research, the Ivorian Office of National Parks (Office Ivorien des Parcs et Reserves), the directorship of the Taï National Park and the Centre Suisse de Recherches Scientifiques en Côte d'Ivoire in Abidjan.

I am grateful to the Deutsche Forschungsgemeinschaft for funding this project and my position (grant: DFG KL 2521/1-1).

I thank my parents Ulf and Wiltrud who were certainly not always happy with my travels, but always supported me 100% without questioning- and so did my brother Tom and my sister Lisa. Exploring the world is way more fun when you have a great home to come back to.

Vince, without you this thesis would not exist. May there be many more adventures!

16 Selbständigkeitserklärung

Ich bestätige, dass die vorliegende Dissertation in allen Teilen von mir selbstständig angefertigt wurde. Beiträge von Kollegen und die benutzten Hilfsmittel sind vollständig gekennzeichnet und angegeben worden. Veröffentlichungen von Teilen der vorliegenden Dissertation sind angegeben. Weiter erkläre ich, dass ich nicht schon anderweitig einmal die Promotionsabsicht angemeldet oder ein Promotionsverfahren beantragt habe.

Berlin, den 12.04.2018

Fee Zimmermann

mbvberlin mensch und buch verlag

49,90 Euro | ISBN: 978-3-86387-946-4