Aus dem Institut für Mikrobiologie und Tierseuchen
des Fachbereichs Veterinärmedizin
der Freien Universität Berlin

Genomics of infectious bacterial skin diseases in wild non-human primates: Yaws and Leprosy

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

vorgelegt von
Benjamin Mubemba
Tierarzt aus Kitwe - Sambia

Berlin 2020
Journal-Nr.: 4235

mbvberlin | mensch und buch verlag

Aus dem Institut für Mikrobiologie und Tierseuchen

des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Genomics of infectious bacterial skin diseases in wild non-human primates: Yaws and Leprosy

Inaugural-Dissertation

zur Erlangung des Grades eines

Doktors der Veterinärmedizin

an der

Freien Universität Berlin

vorgelegt von

Benjamin Mubemba

Tierarzt

aus Kitwe - Sambia

Berlin 2020

Journal-Nr.: 4235

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin

der Freien Universität Berlin

Dekan:	Univ.-Prof. Dr. Jürgen Zentek
Erster Gutachter:	PD. Dr. Fabian H. Leendertz
Zweiter Gutachter:	Prof. Dr. Alex Greenwood
Dritter Gutachter:	Univ.-Prof. Dr. Marcus Fulde

Deskriptoren (nach CAB-Thesaurus):
pan, skin diseases, *Treponema pallidum*, leprosy, treponematosis, *Mycobacterium leprae*, tuberculiod leprosy, epidemiology, Tanzania, lake manyara, Ivory coast, Guinea-bissau

Tag der Promotion: 05.11.2020

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: 978-3-96729-080-6
Zugl.: Berlin, Freie Univ., Diss., 2020
Dissertation, Freie Universität Berlin
D188

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

Dedicated to

My children *Michelle, Yakira* and *Asher*, for holding up during the times I was away undertaking this research.

Table of contents

List of Figures iv

List of Tables v

List of abbreviations vi

1 Introduction ..... - 1 -

1.1 Study objectives ..... - 3 -

1.1.1 *Treponema pallidum* infection at TNP and other selected field sites ..... - 3 -

1.1.2 Leprosy infection in TNP and CNP ..... - 3 -

2 Background ..... - 4 -

2.1 Emerging and reemerging infectious diseases ..... - 4 -

2.2 Emergence of Infectious diseases in wild non-human primates ..... - 5 -

2.3 Impact of infectious diseases in wild non-human primates ..... - 7 -

2.4 Health monitoring tools in non-human primates ..... - 7 -

2.4.1 Sampling tools ..... - 7 -

2.4.2 Molecular biology tools ..... - 9 -

2.4.3 Bioinformatics analyses ..... - 11 -

2.5 Infectious skin diseases in wild non-human primates ..... - 12 -

2.5.1 Yaws-like disease ..... - 13 -

2.5.2 Leprosy ..... - 19 -

3 Materials and methods ..... - 27 -

3.1 *Treponema pallidum* infections in NHPs in TNP and other field sites ..... - 27 -

3.1.1 Study sites and samples ..... - 27 -

3.1.2 DNA extraction ..... - 29 -

3.1.3 Screening for *Treponema pallidum* ..... - 29 -

3.1.4 Library preparation, genome wide capture and high throughput sequencing ..... - 31 -

3.1.5 Bioinformatics analysis ..... - 34 -

3.2 Leprosy disease chimpanzees in CNP and TNP ..... - 36 -

3.2.1 Site and Samples ..... - 36 -

3.2.2 DNA extraction ..... - 38 -

3.2.3 Screening of *M. leprae* in faecal and necropsy samples ..... - 40 -

3.2.4 Library preparation, genome wide capture and high throughput sequencing ..... - 43 -

3.2.5 Bioinformatics analysis ..... - 43 -

3.2.6 Genotyping of insufficiently covered genome by PCR ..... - 44 -

3.2.7 Metabarcoding for dietary analysis ..... - 44 -

3.2.8 Histopathology ... - 46 -
3.2.9 Detection of PGL-1 antibodies ... - 46 -
4 Results ... - 47 -
4.1 Treponema pallidum infections in NHPs in TNP and other field sites ... - 47 -
4.1.1 Screen PCR and whole genome capture ... - 47 -
4.1.2 Phylogenetic analyses ... - 50 -
4.2 Leprosy disease among chimpanzees in CNP and TNP ... - 55 -
4.2.1 Evaluation of clinical manifestations ... - 55 -
4.2.2 Screening of *M. leprae* in faecal, urine, fruit wedges and necropsy samples ... - 55 -
4.2.3 Genome wide capture and genotyping analyses ... - 56 -
4.2.4 Comparative genomics of *M. leprae* strains from chimpanzees ... - 59 -
4.2.5 Dietary analysis for mammalian DNA ... - 63 -
4.2.6 Rapid lateral flow test for detection of PGL-1 antibodies ... - 63 -
5 Discussion ... - 64 -
5.1 *Treponema pallidum* infections in TNP and other selected field sites ... - 64 -
5.1.1 There is a large genomic diversity of TPE in TNP ... - 64 -
5.1.2 Genomic diversity of TPE infecting NHPs is geographically structured ... - 65 -
5.1.3 Detection of TPE in bone samples from sub-saharan Africa field sites ... - 66 -
5.1.4 The zoonotic potential of TPE ... - 67 -
5.2 Leprosy infection in CNP and TNP ... - 68 -
5.2.1 Sylvatic lepromatous leprosy in wild chimpanzees; the clinical evidence ... - 68 -
5.2.2 Faecal samples are a potential surveillance tool for *M. leprae* in wild NHPs ... - 68 -
5.2.3 Wild chimpanzees are infected with rare *M. leprae* genotypes ... - 70 -
5.2.4 Focussed future leprosy studies ... - 72 -
5.2.5 Coinfection of *M. leprae* and *M. tuberculosis* in a great ape ... - 74 -
6 Conclusions and outlook ... - 75 -
7 Summary ... - 77 -
8 Zusammenfassung ... - 79 -
9 References ... - 81 -
10 Appendices ... - 105 -
10.1 Table S1: Bone samples screened for Treponema pallidum ... - 105 -
10.2 Table S2: Published genomes used in this study ... - 109 -

10.3 Table S3: Patristic distances of simian derived TPE strains analysed in this study using the ML 10X-95% tree ..... - 110 -

10.4 Figure S1: Supplementary trees for TP phylogenetic analysis ..... - 111 -

10.5 Table S4: Faecal, urine and fruit wedge screening of Woodstock for *M. leprae* from onset of symptoms- June 2018 ..... - 115 -

10.6 Table S5: Faecal screening of Woodstock for *M. leprae* for samples collected between 2014 and June 2018 ..... - 116 -

10.7 Table S6: Faecal screening of Zora for *M. leprae* for samples between March 2001 and June 2009 ..... - 116 -

10.8 Table S7: Organ screening Zora necropsy samples for *M. leprae* ..... - 118 -

10.9 Table S8: Screening of all necropsy samples collected in Täi National Park, Côte d'Ivoire, since 2001 for *M. leprae* ..... - 119 -

10.10 Table S9: Faecal screening for *M. leprae* for samples collected in Cantanhez National Park, Guinea Bissau ..... - 121 -

10.11 Table S10: Published and unpublished (from collaborators) *M. leprae* genomes used in this study ..... - 126 -

10.12 Table S11: List of primers used for *M. leprae* genotyping ..... - 132 -

10.13 Pathology report for Zora ..... - 132 -

11 Ethical statement ..... - 135 -

12 Published work ..... - 136 -

13 Acknowledgements ..... - 137 -

14 Selbstständigkeitserklärung ..... - 138 -

List of Figures

Figure		Page
Figure 1	Map showing study sites for the *Treponema pallidum* study	28
Figure 2	Lesions due to Treponema *pallidum* infection in *Cercobus atys* monkeys in Täi National Park, Côte d'Ivoire	28
Figure 3	Map showing study sites for the *Mycobacterium leprae study*	36
Figure 4	Lesions due to *Mycobacterium leprae* in chimpanzees in Cantanhez National Park, Guinea Bissau	37
Figure 5	Lesion progression due to *Mycobacterium leprae* in a chimpanzee in Täi National Park, Côte d'Ivoire	39
Figure 6	Pictures of Zora taken a few months before her death in 2009	40
Figure 7	Maximum clade credibility tree of *Treponema pallidum* strains	51
Figure 8	Maximum clade credibility tree from ingroup analysis of TPE strains	52
Figure 9	Cladograms for phylogenetic read placement of TPE positive bone samples	53
Figure 10	Plots of faecal screening for *M. leprae* results for Woodstock and Zora	57
Figure 11	Maximum Parsimony tree for *Mycobacterium leprae* strains	60
Figure 12	Zoom-in sub-trees for 4N/O and 2F genotypes	61
Figure 13	Results of the rapid leprosy lateral flow tests for both Zora and the positive control	63
Figure S1	Supplementary trees from phylogenetic analysis of *Treponema pallidum* strains	111

List of Tables

Table		Page
Table 1	Primers used for *Treponema pallidum* screening	31
Table 2	Summary for library preparation, capture and sequencing methods used for each sample in the *Treponema pallidum* study	33
Table 3	Primers used in the leprosy study	42
Table 4	Lesion and sample types observed at TNP and PCR screening results for each sample	47
Table 5	Mapping results for non-human primate *Treponema pallidum* subsp. *pertenue* strains	49
Table 6	Summary table of phylogenetic analyses of TP strains	54
Table 7	Genotyping result for animal TNP_566	57
Table 8	Mapping results for chimpanzee derived *M. leprae* genomes	58
Table S1	List of bone samples screened for *Treponema pallidum*	105
Table S2	Published Treponema *pallidum* genomes used in this study	109
Table S3	Patristic distances extracted from maximum likelihood tree for *Treponema pallidum* strains (NHPs only)	110
Table S4	Faecal, urine, fruit wedge sample screening of Woodstock for *Mycobacterium leprae* from onset of symptoms- June 2018	115
Table S5	Faecal screening of Woodstock for *Mycobacterium leprae* for samples between 2014 and June 2018	116
Table S6	Faecal screening of Zora for *Mycobacterium leprae* for samples between March 2001 and June 2009	116
Table S7	Organ screening Zora necropsy samples for *Mycobacterium leprae*	118
Table S8	Screening of all necropsy samples collected in Täi National Park, Côte d'Ivoire, since 2001 for *M. leprae*	119
Table S9	Faecal screening for M. leprae for faecal samples collected from Cantanhez national Park in guinea Bissau	121
Table S10	Published and unpulished (from collaborators) *M. leprae* genomes used in this study	126
Table S11	List of primers used for *M. leprae* genotyping	132

List of abbreviations

AFB	Acid fast acilli
BB	Borderline leprosy
BL	Borderline lepromatous leprosy
BL	Borderline leprosy
BT	Borderline tuberculoid leprosy
CNP	Cantanhez National Park
DLL	Diffuse lepromatous leprosy
DNA	Deoxyribonucleic acid
GIT	Gastrointestinal tract
GNP	Gombe National Park
EID	Emerging infectious disease
IUCN	International Union for Conservation of Nature
LL	Lepromatous leprosy
LMNP	Lake Manyara National Park
LMNP	Lake Manyara National Park
MLST	multi-locus sequence typing
MRCA	Most Recent Common Ancestor
NGS	Next generation sequencing
NHP	Non-human primate
NHPs	Non-human primates
NTDs	Neglected tropical diseases
PCR	Polymerase chain reaction
RKI	Robert Koch Institute
SNP	Single nucleotide polymorphisms

TNP	Taï National Park
TP	*Treponema pallidum*
TPE	*Treponema pallidum* subsp. *Pertenue*
TT	Tuberculoid leprosy
USA	United States of America
WHO	World Health Organisation

1 Introduction

Wildlife diseases have an impact on animal fitness and ultimately conservation [1]. In addition, the majority of emerging human infectious diseases have their origins in animal populations [2]. This is particularly true for non-human primates (NHPs) because of their close phylogenetic relationship to humans. For example, the close phylogenetic relatedness between NHPs and humans coupled with the spatial overlap between primate hosts [3], great apes such as chimpanzees and gorillas share a significant number of pathogens with humans posing public health threats arising from zoonotic infections [4]. Despite their importance, the study of NHP diseases comes with its own particular challenges, such as obtaining high quality samples through invasive sampling due to limited access to these animals as well as ethical considerations [5]. Further, collecting necropsy samples can also be dangerous. This is because these animals sometimes die of deadly zoonotic diseases such as anthrax or Ebola virus [6,7], as they inhabit regions that are hot spots for these infections [8]. Therefore, adapting non-invasive tools that are specific to diseases affecting them that minimizes risks to the investigators and the animals is critical [5]. Besides, diagnostic tools that facilitate the improved health monitoring and the understanding of the prevailing disease ecology are pivotal in complementing other conservation tools in protecting wildlife [5].

In addition, most diseases affecting NHPs have been understudied due to earlier studies on NHPs being focused more on the origin of human immunodeficiency viruse (HIV) in man with diagnostic tools developed then inclined to that [9]. Among the least studied diseases of NHPs are dermatoses, despite several pathogens involved in these pathologies having shown the potential to infect humans [7,10–13]. Most dermatoses that have been described are in captive settings due to easier access or observation of the clinical diseases [14]. In addition, the previous lack of surveillance for them in the wild has also contributed to the little knowledge about their status in the wild. Among the few notable diseases that have been described to have dermatological manifestations in NHPs include; monkeypoxvirus, herpesvirus, simian retroviruses, yaws-like disease and leprosy [14] although leprosy has only been described in captive settings [15–18].

Increasing evidence suggests many NHP species in sub-Saharan Africa are infected with *Treponema pallidum* subsp. *pertenue* (TPE), the bacterium causing yaws in humans [19,20]. Yaws–like disease was observed in Taï National Park (TNP) in January 2014 for the first time in a habituated social group of *Cercocebus atys* monkeys and *Treponema pallidum* subsp. *pertenue* (TPE) was determined to be the cause [19]. Since then, diverse symptoms of yaws-like orofacial and genital lesions in this social group have been noted. However, it is unknown whether the diverse symptoms observed in this ecosystem were all caused by the same subspecies of *Treponema pallidum* (TP) and whether cross-species transmission occurs

between different NHP species in this ecosystem or other sub-Saharan Africa sites where yaws remains endemic. In addition, it is also unknown whether TPE historically infected other NHP species in TNP and NHPs at other sites in sub-Saharan Africa, as both genomic and descriptive evidence of the disease in NHPs from the distant past are lacking. The recent detection of *Treponema pallidum* (TP) DNA in chimpanzee bones collected from TNP gave indication that other species might be affected, though it was not possible, in this study, to identify the particular TP subspecies involved in these infections [21]. Since the recognition of the disease in the 1960s [22], only a single TPE genome isolated from the baboon is available from the distant past [23]. This lack of genomic evidence from the distant past coupled with the poor sampling of TPE across NHP hosts and space has partially precluded an understanding of how the genomic diversity of TPE lineages that do occur in NHPs is distributed across hosts and space. Hence, it is unknown whether TPE lineages that exist have a specific tropism for certain host species and/or are inclined to the habitat of the hosts. Equally, this has subsequently led to the poor understanding of how the disease is transmitted between different NHP hosts, if it is, at all.

In June 2018, another skin pathology involving a chimpanzee (*Pan troglodytes verus*) from a habituated group in TNP was observed. The chimpanzee presented with leprosy-like lesions on the face. This was the first-time leprosy-like symptoms were observed in TNP. At another research site in West Africa, Cantanhez National Park (CNP), Guinea Bissau, the first sightings of chimpanzees with leprosy-like lesions occurred in 2013 in camera trap images. The number of affected individuals at this field site appears to be increasing over the last six years. Sylvatic leprosy has never been described in wild NHPs. All known cases of leprosy in NHPs were diagnosed in research sanctuaries and *Mycobacterium leprae,* one of the pathogens causing leprosy in humans was determined to be the cause. It remains unknown whether *M. leprae* infections in these captive animals were acquired in the wild, or whether they were acquired from humans once they entered captivity. As a result, neither is it currently known whether NHP are infected with these pathogens in the wild, nor has it been established whether the occurrence of the disease in NHPs is incidental to zoonotic spillovers from humans or NHPs are also natural hosts [24]. If spillover occurs, it is also not known whether leprosy persists in wild NHP populations. One challenging aspect of studying leprosy in wild NHPs is the inability to collect biopsies from individuals, precluding standard diagnostic approaches [5,25].

This doctoral dissertation aims to fill some of these knowledge-gaps through an investigation of the aforementioned skin pathologies among wild NHPs. By application of both invasive and non-invasive sampling tools coupled with molecular biology and bioinformatics approaches, useful genomic data was obtained from tissue, bone and faecal samples to elucidate the

prevailing disease ecology of the respective diseases in the above metioned wildlife protected areas.

1.1 Study objectives

The objectives of this doctoral dissertation are presented in two parts;

1.1.1 *Treponema pallidum* infection at TNP and other selected field sites

In the first part, I aimed to document TPE diversity in wild NHPs in TNP and screen for TPE at 12 sites in sub-Saharan Africa. The aim was to elucidate how the genomic diversity of TPE lineages that do occur in NHPs is distributed across hosts and space. The specific objectives were:

a) To test whether the diverse symptoms observed in sooty mangabeys from TNP were caused by a single lineage of TP subspecies.
b) To screen for TP in multiple NHP species from 12 selected sites across sub-Saharan Africa (Including TNP) and test whether cross-species transmission of TP occurs between NHPs within ecosystems.

1.1.2 Leprosy infection in TNP and CNP

The second part of this dissertation represents an investigation into the causative agent of leprosy-like symptoms observed in both TNP and CNP. The specific objectives were:

a) To adapt a suitable non-invasive screening tool as it was not possible to collect invasive samples from these wild great apes.
b) To determine and characterise the causative agent of the skin lesions affecting NHPs observed in TNP and CNP.

2 Background

2.1 Emerging and reemerging infectious diseases

Emerging infectious diseases (EIDs) are diseases that affect a particular population for the first time or have existed in a population but are increasing in incidence or their geographic range [26]. On the other hand, reemerging diseases refer to EIDs that historically infected a particular population but continue to appear in new locations or in drug-resistant forms, or that reappear after eradication [27]. EIDs are caused by infectious agents (whether bacterial, viral, fungi, and parasitic micro-organisms) that can be transmitted between same or different species of hosts [28]. Mechanisms driving the emergence of these diseases include changes in ecological landscapes, behavioural driven activities relating to population demographics, increase in international travel and trade, microbial evolution and adaptation, industrialisation activities and the lack of robust public health systems in some parts of the world to help in effecting control measures [26]. Notably, most of these infectious agents causing pandemics and in general those affecting humans are zoonotic in nature i.e. are of animal origin [2].

For example, simian immunodeficiency viruses (SIVs) crossed host and ecological barriers resulting into the present time SIV diversity with varying disease outcomes across the primate host range. To that effect, at least fourteen unique ancient SIV cross-species transmission events have been identified [29]. One iconic example of such an event is the jumping of SIVs from the NHP hosts into the human population resulting into the present day human immunodeficiency viruses (HIV 1 and 2) [30,31]. In addition, Ayouba *et al.* (2013) observed the continued zoonotic spillovers of SIVs strains from sooty mangabeys into human communities surrounding TNP resulting in a new HIV-2 lineage [32]. Anthropogenic activities such as bush meat hunting, bush meat preparation and consumption, habitat encroachments and other factors that increase contact between human and NHPs have been cited for these continued zoonotic transmission events [8]. Human movement, poverty, inequality and commercial sex work have been cited as mechanisms that have enabled the rapid spread of HIV within the human population after the spillover, especially in sub-saharan Africa [27]. HIV has become the most important EID due to its detrimental population effects. Global prevalence has dramatically increased from 8.7 million people living with HIV in 1990 to 36.8 million people living in 2017 with an estimated 32 million HIV related deaths [33].

Other EIDs that caused epidemics at global stage include: severe acute respiratory syndrome (SARS) caused by a coronavirus; highly pathogenic avian influenza (H5N1) and swine flu (H1N1) that are both caused by influenza A viruses, Ebola virus, Zika virus and rift valley fever virus [2,26,34,35]. In 2002 a corona virus emerged from bats and jumped into the human population causing SARS in China, Hong kong and Vietnam. By 2003 when the epidemic was

declared over, more than 8000 cases of SARS were recorded from 26 countries with a case fatality rate of 10% [36]. The emergence of SARS was associated with the handling and consumption of civet cats and ferret badgers by humans as well as encroachment into bat territories by humans [37]. International air travel was at the center for the rapid spread of epidemic [27] and so were the west Africa 2014 Ebola epidemics and the recent Zika epidermics in the Americas [38–40]. EIDs such as Zika virus and Rift valley fever virus that are vector borne are triggered by climatic shifts that favour the rise in the density of principal vectors [41] though recent evidence suggest that NHPs may also be naturally infected by Zika virus [42]. H5N1 and H1N1 emerged from wild birds and pigs respectively [27,35] and quickly spread via air travel after emergence in the human population [27] as the influenza A viruses also readily spread between people. Close interactions between humans and the respective animals were responsible for these zoonotic spillovers [27,35,43].

The example of simian immunodeficiency viruses (SIVs), also demonstrate the contribution by NHPs to the emergence of EIDs. It has been shown that most zoonotic spillover events from NHPs are geographically spread along the equatorial rainforests of Africa, Asia and Neotropics due to the primate richness of these regions with over 20% of the NHP species being hosts for at least one of the 63 unique zoonoses [3,8]. The genetic similarities between humans and NHPs predisposes both to common infections especially when the two share the same habitat or overlapping geographic range [3]. Devaux *et al.* (2019) argue that the unprecedented rate at which the ecological barriers between humans and NHPs are being broken due to massive destruction of ecosystems by anthropogenic related activities, has led to the increased risk of exposure of humans to NHP pathogens. In addition, the authors identified NHP illegal trade and meat consumption, the use of NHPs in biomedical research, conversion of forest lands in urbanization and ecotourism as drivers for the cross-species transmission of NHP contributed infectious diseases in humans [44]. Examples of pathogens that infect NHPs and present potential EID threat to humans may include: monkeypoxvirus, Ebola virus, *Bacillus anthracis*, herpes B virus, yellow fever virus, and retroviruses [12,30,44–47]

2.2 Emergence of Infectious diseases in wild non-human primates

In wild NHPs, to better understand the disease ecology at play in infectious disease emergence, questions regarding infectious diseases should address the following key fundamental aspects: the drivers for their emergence in NHP communities and the impact on the conservation and fitness of wild primates. In addition, their direct or indirect impacts on human health should be assessable [4,48]. Pedersen *et al*. (2009) identifies three key stages of infectious diseases emergence in primate populations; the first stage is the opportunity for

the disease to emerge which is determined by the biogeographical distance of the host and pathogen; the smaller the distance, the more the infection initiation is likely. This offers the opportunity for the pathogen and the host to interact [3]. This is followed by transmission of the pathogen into a population which should be compatible with evolutionary and ecological barriers as they determine the ability of the pathogen to infect a host [3]. Lastly, the infection should be able to establish itself in a population; this is mostly regulated by host-pathogen contact rates in combination with the demographics of the susceptible population. The higher the contact rate, the more likely the infection will be sustained in a population [3]. Chapman *et al*. (2005) suggests that while host, pathogen and environmental factors are obvious and essential in the successful infection and establishment of the disease in a population, anthropogenic driven factors are crucial in triggering the emergence of infectious diseases in primate hosts. Specifically in NHP populations, research activities, ecotourism, habitat fragmentation and climate change have been cited as some of the leading anthropogenic triggers [48]. For example, research and ecotourism activities were at the centre of anthropozoonotic transmission of human borne pathogens into wild NHP populations, as was the case for the human coronavirus outbreak among chimpanzees in TNP [13]. In a general perspective, it is reported that 55% of the emerging infectious wildlife diseases are attributable to anthropogenic activities [49]. In this regard, the authors concluded that environmental degradation such as habitat fragmentation or total loss of it, was the most important anthropogenic factor driving the emergence of infectious diseases in wildlife populations and NHPs were not an exception [49].

Besides anthropogenic activities, climate change has also been cited as another factor contributing to the emergence or re-emergence of diseases in animal communities [50]. In NHPs, this climatic shift is viewed as one that is more likely to promote favourable conditions for increased disease transmission and increased animal stress making the animals more vulnerable to pathogenic microbes [48]. For instance, based on a data set collected over a period of 30 years, Chapman and co-workers (2005) demonstrated how climate change altered fruiting patterns in Kibale National Park (KNP), Uganda, leading to nutritional stress among NHP populations in some years; a scenario that is well known to increase vulnerability to infection [51]. Furthermore, nutritional stress driven parasite infections have been observed to be more common in NHPs inhabiting fragmented forest systems as opposed to intact ones [52]. This is a classic example of how both climatic shifts and anthropogenic factors complement one another in driving disease emergence. Equally, increased precipitation in some areas has been shown to increase vector density resulting in eruption of vector borne diseases such as malaria [53]. Wu *et al.* (2018) observed that malaria parasite prevalence

among the chimpanzees of TNP had a seasonal variation with a higher disease prevalence being observed in wet seasons [54].

2.3 Impact of infectious diseases in wild non-human primates

The impact of these diseases in wild primate populations can be far-reaching, especially for endangered NHP species whose population sizes have been strongly declining. [5,55,56]. Diseases can indeed lead to either significant mortalities or a reduction in species fecundity [48]. For example, Ebola virus was suggested to be responsible for more than 5,000 mortalities of the highly endangered western gorillas (*Gorilla gorilla*) and untold mortalities among wild chimpanzees in central Africa during the early 2000s Ebola epidemics [7]. In TNP, anthrax has continued to kill chimpanzees and other wildlife species threatening the already dwindling numbers of the endangered west African chimpanzee in the TNP ecosystem [6,47]. Furthermore, respiratory infectious agents were found to be a leading cause of morbidity and mortality in chimpanzees of Gombe National Park (GNP) in Tanzania [57] and TNP [13,58,59]. Overall, mortalities due to infectious diseases combined with threats from bush meat hunting and habitat loss has brought some of the NHPs species to the brink of extinction [48,55] and their survival partly depends on the extent to which these diseases can be prevented or controlled [5]. Infectious agents have also been shown to trigger a cascade of events that leads to depressed species fecundity and ultimately results in jeopardizing population viability. Notably, nutritional stress is among the leading disease induced factors that negatively impacts animal fecundity. One way diseases lead to reduced fecundity is by impairing optimal feeding [60,61]. Nutritional stress was observed to increase inter-birth intervals and constrained survival chances of infants in blue monkeys (*Cercopithecus mitis)* leading to reduced replacements in the troops [62].

2.4 Health monitoring tools in non-human primates

2.4.1 Sampling tools

Invasive sampling approaches such as collection of blood and tissue biopsy are the traditional approaches in health monitoring of captive NHPs. These approaches offer high quality samples for a wide range of molecular and serological screening of pathogens especially those with a systemic phase [5]. However, when it comes to wild-living NHPs, the practicality of these approaches as far as capturing heavy bodied primates in sufficient numbers required for studies is not feasible [5]. Furthermore, the risk of anthropozoonotic transmission of diseases [13] and the danger these capture events pose to both the animals and the capturing teams makes these approaches unpopular as far as conservation guidelines are concerned. In most cases, these are last resort approaches for rare circumstances where human intervention is highly recommended [5].

With an increasing number of infectious diseases threatening NHPs and the risk for cross-species transmission threatening human health, quick and easy to implement health monitoring tools are becoming essential [5]. Leendertz *et al.* (2006) proposed a systematic health monitoring framework making use of non-invasive samples such as urine and faeces complemented with molecular analyses in protecting the health of great apes and other threatened mammals [5]. Over the years, this approach has proved to be effective and conservation-wise friendly as there is minimal disturbance of animals in their natural habitats. For example, faecal samples were used to precisely diagnose some of the notorious anthropozoonotic infections of great apes such as human respiratory syncytial virus (HRSV) [63], human metapneumovirus (HMPV) [64], coronavirus [13] and rhinovirus [65]. Leendertz *et al*. (2004) also demonstrated the use of urine in detection of simian T-lymphotropic virus type 1 antibodies in wild adult chimpanzees of TNP [66]. Additionally, urinary neopterin was recently used to predict survival chances of wild chimpanzees in the face of respiratory outbreaks [67]. These non-invasive approaches avert the risks associated with capturing events and give latitude to screening of large numbers of NHPs so long as the samples are well collected and preserved [5]. In the present study, we used faecal samples to screen and extract useful genomic information to reconstruct *M. leprae* genome. This is the first time a full genome of *M. leprae* is being sequenced from faeces.

Flies and bone samples are emerging sampling tools in the detection of pathogens that cause disease in NHPs. They offer avenues for the detection of pathogens circulating in the areas they are collected from. Recently, we showed that flies collected in TNP carry both the yaws and anthrax pathogens [68]. The anthrax pathogen *Bacillus cereus* biovar *anthracis* (Bcbva) isolated from the flies was viable, an indication that flies might be involved in its dissemination and transmission [68]. Elsewhere [69], DNA of TPE the yaws pathogen: was isolated from flies collected from ecosystems inhabited by yaws diseased NHPs. The above studies show that flies can be used to indirectly monitor NHP diseases such as yaws and anthrax. Non-human primate bone samples can be a source of pathogen information reflective of the pathogen that infected the hosts before death [70]. For example, Gogarten *et al.* (2016) showed that it was possible to extract TP sequences from non-symptomatic old bones, a pathogen causing yaws disease in NHPs [21]. In this present study, building on the study of Gorgaten *et al*. (2016), we were able to extract informative genomic information from bone samples and subsequently used this genomic information to investigate the ecology of yaws disease in NHPs. Overall, the two tools are cost effective and easy to implement as far as sampling is concerned even in a resource limited environment [71].

Carcass monitoring and necropsy examination of dead NHPs by trained personnel is a useful tool and provides preliminary insights into the cause of death [5]. Necropsy examination

provides valuable samples for both molecular and histological diagnostic approaches. The time between death and necropsy should be minimized. Ideally, the shorter the period, the higher the chances of precisely identifying the causative agent [5]. Samples for molecular investigations can be preserved in RNAlater or snap frozen to preserve nucleic acid integrity, while those for histological examination can be fixed in 10% buffered formalin [5]. Carcass monitoring and necropsy examination as complementary tools were employed to investigate mortalities related to sylvatic anthrax caused by Bcbva in the TNP ecosystem. It was revealed that of the 204 carcasses investigated for the period 1996 to 2014, 81 (39.70%) carcasses including NHPs had detectable Bcbva DNA allowing for retrieval of Bcbva whole genomes from positive carcarsses. In addition, histological examination of samples from these carcasses revealed a typical anthrax infection [47]. However, care should be undertaken by personnel performing necropsies to protect themselves against local highly pathogenic organisms such as Ebola and sylvatic anthrax in some African rainforests where these microrganisms are known to have killed NHPs [7,47]. In this doctoral dissertation necropsy samples extracted a decade ago were used to trace leprosy disease in a population of wild chimpanzees back to that time.

2.4.2 Molecular biology tools

Molecular biology analyses are essential tools in the health monitoring of pathogens infecting NHPs and over the years, a number of molecular methods have been developed or adapted to investigate pathogens infecting NHPs. The type of sample material, how it is preserved and the suspected candidate pathogen determines what type molecular approaches would be applicable in the investigation [5]. Furthermore, it is a standard requirement that sample material meant for such analyses are preserved in such a manner that assures the integrity of nuleic acids such as in RNAlater or freshly frozen [5].

End-point polymerase chain reaction (PCR), nested PCR or quantitative PCR (qPCR) systems targeting conserved genomic regions of pathogens are widely applied in investigating NHP infections [5]. In a nested PCR system, product from first round PCR is used as template and is employed when sensitivity of the of PCR assay needs to be scaled up [72]. The PCR systems should be designed in such way that they are specific for the target pathogens [5]. However, where only the screening of a group of pathogens belonging to the same family is desirable, generic PCR systems can be applied. For example, Ehlers and co-workers (1999) developed a PCR system targeting a wide range of Herpesviruses [73]. In addition, generic PCRs were also applied in investigating cytomegaloviruses infecting NHPs [74,75]. In the event were close strains of pathogens need to be differentiated, genotyping PCR assays targeting single nucleotide polymorsim (SNP) locus differentiating the strains (genotypes) can be performed to determine the nucleotide base present at that locus [76]. For instance,

genotying techniques were applied to determine the origin of chikungunya virus circulating among NHPs in Malaysia [77]. In addition, multilocus genotyping approach targeting several loci of target genes was recently used to characterise strain diversity of TPE infecting NHPs in Tanzania [20]. Furthermore, multilocus typing was employed to demonstrate respiratory disease among wild chimpanzees caused by *Pasteurella multocida* in TNP [58]. In most cases, PCR systems are employed as diagnostic or screening tools.

Techniques such as genome wide capture coupled with high throughput sequencing that enable reconstruction of whole genomes or larger genome fragments have revolutionalised the field of molecular microbiology. Such advancements have paved ways for easier downstream analysis of sequence data such as in phylogenetics, evolution, epidemiology, and drug resistance [78]. Deep sequencing of several overlapping long range PCR amplicons is also a common approach in reconstructing whole genomes [79]. Whole genome capture employs synthetic baits or probes to capture target genome or genome fragments. The technique is advantageous in that it enables the robust capturing of minute amounts of pathogen DNA swamped in endogenous or host DNA [78]. For instance, this technique was used to characterise TP and *M.leprae* from ancient samples considered to have decimal amount of pathogen DNA [80–82]. In NHP health monitoring, this tool coupled with the associated bioinformatic tools has been pivotal in pathogen discovery and characterisation of disease ecology by aiding in sequencing of pathogen genomes that remained a challenge previously [13,19,47,83]. In the present study, genome wide capture was the core microbiology technique and was applied to understand the disease ecology of yaws [84] and leprosy in NHPs.

The initial part of the genome wide capture workflow involves the construction of sequencing libraries. Genomic DNA or copy DNA in the case of RNA is fragmented into a desirable length of DNA fragments [85]. This is followed by enzymatic reactions that peform the blunt end repair of the DNA fragments and the adaptor ligation step that tags short synthetic oligonucleotides of a known sequence to both ends on of the DNA fragments [86]. After attachment of adapters, the desired library insert size can be fine tuned to select only fragment inserts of the same size through a magnetic bead cleaning systems [87]. The adaptor sequences enable interaction of the DNA fragments with the sequencing platforms as well as providing binding sites for indexing primers. The indexing reactions have a dual purpose: to both amplify the ligated DNA as well as to attach unique identifier sequences (barcodes/indicies) essential when pooling multiple libraries in a single capture reaction [86]. Library quantification is the final step in the the library construction stage. It determines the available DNA molecules that have been successfully ligated and are available for downstream investigations [78,86,87]. A qPCR described elsewhere [88] was shown to effectively and correctly estimate DNA

concentration at this stage. At this point, constructed libraries suffice for shortgun sequencing meant for metagenomics analysis [89].

Once the minimum concentration of the library required for capture has been achieved , libraries are subjected to hybridization capture at a specific temperature with either RNA or DNA synthetic probes or baits that are specific to the target genomic region [78]. The hybridization is achieved either through in-solution capture where the baits are biotinylated allowing target DNA-bait complexes to be isolated using a magnetic bead cleaning systems or in-solid phase capture where probes are bound to a micro-array system to isolate the target DNA [78,90–92]. Post capture final steps include elution of the target DNA bound to the probes or baits and the subsequent amplification, quantification and pooling in readiness for high throughput sequencing [78,88,92].

Another molecular biology tool that was relevant to the present study include metabarcoding analysis which is a tool that aids in species identification. This is based on PCR assays targeting short DNA fragments (maximum of 800 bp) of particular genes unique to a certain species or organisms. Resulting sequences are compared to reference library using bioinformatic tools to identify the likely species from which the sample was derived from [93]. In NHPs, such analyses have been applied in diet analysis, microbiome analyses and biomonitoring and ecological assessments of habitats [68,71,94–96].

2.4.3 Bioinformatics analyses

Luscombe *et al.* (2001) defines bioinformatics as an interdisciplinary science that involves the application of the “informatics techniques” (derived from disciplines such as applied maths, computer science, and statistics) to understand and organize biological data on a large-scale [97]. Application of bionformatic tools ranges from short and long sequences to small and large data sets. In respect to this dissertation, the bioinformatics workflow reviewed is for processing of high throughput sequence data to infer phylogenetic relationships. However, it should be noted that long enough sequences from PCR amplicons can also be used to infer phylogenetic relationships in NHP infections [59,98]. Bioinformatic analyses make use of computer programs and softwares that have been developed specifically for bionformatics such as those described elsewhere [99]. Briefly, processing of high throughput sequence data to assemble full genomes starts with demultiplexing and identification of reads of the sequenced samples according to the Indices used in the library construction step. This is followed by quality check and adaptor trimming to require that a sequence read should have a minimum number of bases of a good quality to pass [100]. Quality processed reads are then mapped to the reference genome preferably one that was used to design the baits or probes using mappers such as those described elsewhere [101]. Assembled genome(s) are further fine tuned by

prescribing a criterion to call a consensus genome i.e. the number of unique reads that needs to be present at the particular genome position (coverage depth) and the threshold level of identity for the unique reads for a base to be called at that position. For example, most analyses require that three unique reads and a minimum of 50% threshold identity of the reads should be met for the nucleotide base to be called [102].

The ultimate interpretation of sequence data lies in the phylogenetic inference that is applied on a generated genomic data set which in most cases are tree-based inference methods. The generated trees allow for inferring origins of pathogens, or their spread and evolutionary history [103]. Phylogenetic inference methods include distance, parsimony, likelihood and Bayesian methods and a choice for one method from the other largely depends on the question the data set is answering [103]. In the event that the reconstructed genome does not suffice for a tree based phylogenetic pipeline analysis, methods such as phylogenetic read placement algorithms have recently been developed to assign sequence reads to branches of a reference phylogenetic tree based on the highest likelihood of a read to belong to a particular branch [104,105]. In the present study, we applied both read placement approaches to assign TPE partial genomes from bone samples to the TP phylogeny [84] and a phylogenetic pipeline to analyse good quality genomes.

2.5 Infectious skin diseases in wild non-human primates

Literature regarding skin infections in wild NHPs is scantly available with the few that have been reported being in captive settings [14]. This can be attributed to both poor and difficult surveillance of skin disease among wild primates and easier access to those in zoos and sanctuaries. To that effect, veterinarians that encounter NHP skin infections, heavily rely on extrapolations from human, livestock and/or pet medicine disciplines in diagnostic evaluations [14]. Notable pathogens among those that have been reported to have dermatological manifestations in wild primates include: monkeypoxvirus, herpesvirus, simian retroviruses, leprosy, and yaws-like disease.

Briefly, monkeypoxvirus a close relative to the human smallpox virus, is now considered a re-emerging zoonotic pathogen in Africa [106]. In NHPs, the disease is characterised by pustule eruptions on skin, fever, and high case fatality in some instances [12,107]. The virus was widely reported to cause disease among captive NHPs [107] and cases among wild NHPs are also on the rise [12,108]. On the other hand, herpesvirus simplex an anthropozoonotic agent, elicits vesicular lesions on the oral mucosa, genitals and other soft body parts in most hominid species [11,14,109]. Simian retrovirus such as the Simian T-Lymphotropic virus causes skin lymphomas NHP species [10,45,110].

This doctoral thesis was a focus of two bacterial diseases of NHPs with skin manifestations i.e. yaws-like disease and leprosy. Both yaws and leprosy are listed as neglected tropical diseases (NTDs) of humans by the World Health Organization (WHO) and are targeted for elimination [111,112]. However, the success of these elimination exercises partly depends on the extent by which other mammals are infected by the same pathogens; as reservoir neutralization is a must in all eradication efforts [113]. Therefore, understanding the dynamics of these diseases in wild NHPs is critical to the success of the ongoing eradication efforts. The detailed review highlighting their causative agents, clinical disease presentation and epidemiology are in the sections that follow.

2.5.1 Yaws-like disease

2.5.1.1 Overview and aetiology

Yaws-like disease in NHPs is caused by the spirochete bacterium *T. pallidum* [22,114]. The TP spirochete has affected humankind since the late 15th century [115] and has been noted to affect NHPs since the 1960s. To that effect, three pathogenic TP spirochete subspecies that are morphologically similar, but are distinguishable genetically, epidemiologically, and clinically into three distinct disease syndromes are recognised; yaws (subsp. *pertenue;* TPE), venereal syphilis (subsp. *pallidum;* TPA) and bejel (subsp. *endemicum;* TEN) [116–118]. The spirochete bacterium TP causes a large global disease burden in humans [119,120]. For instance, annually nearly 8 million new cases of venereal syphilis are reported globally [121], while in the remaining 13 countries where yaws remains endemic it is estimated that more than 80,000 new cases occur. Exact estimates of the number of bejel cases from the Sahel region and Arabian Peninsula are lacking [122–124]. Though treatable, these treponematoses remain major public health threats across the globe [118–120]. Ongoing efforts aim to reduce the prevalence of these trepanomatoses, particularly for yaws where an ongoing eradication campaign aims to eradicate the disease globally by 2030 [125].

To date, only the subspecies TPE has been described to infect NHPs [19]. The first description of the disease in NHPs dates back to 1960s in a troop of yellow baboons (*Papio cynocephalus cynocephalus*) in Guinea of west Africa [114]. Since then, TPE infection has been reported to affect a number of NHP species in wildlife protected areas across sub-Saharan Africa [19,20]. The disease in NHPs is viewed by experts as one that can potentially undermine the World Health Organisation (WHO) ongoing campaign which was earlier scheduled to eradicate yaws globally by 2020 [112] but has since been rescheduled to 2030 [125]. This is because strains causing the disease in both humans and NHPs are highly similar [19]. However, the bi-directional transmission of these strains between NHPs and humans has not yet been demonstrated.

2.5.1.2 The genus *Treponema*

Most members of the genus *Treponema* (phylum: Spirochaetes, class: Spirochaetia Order: Spirochaetales, Family: *Spirochaetacea*) [126] are non-pathogenic with only the species *T. pallidum, Treponema. carateum* and *Treponema. paraluiscuniculi* being recognized as pathogenic [120,127]. *T. carateum* causes pinta disease in humans, a skin disease known to be endemic to the Americas [120]. Though *T. carateum* is serologically and morphologically similar to *T. pallidum* subspecies, it has not been successfully isolated to enable characterisation for phylogenetic assignment. It is still also unknown whether it either genetically clusters with the sub-species of *T. pallidum* causing skin lesions in humans or is a different species all together [120]. On the other hand, *T. paraluiscuniculi* causes syphilis-like lesions in rabbits and is not infectious to humans or other hominids but is phylogenetically closely related to the members of *T. pallidum*. Genome decay may explain its inability to infect humans since a number of genes coding for virulence factors in *T.pallidum* subspecies are dormant in *T. paraluiscuniculi* [127]. Both *T. carateum* and *T. paraluiscuniculi* have not been reported in NHPs.

2.5.1.3 *Treponema pallidum*

In general, bacteria in the species TP are helical shaped gram-negative spirochetes though staining them can be challenging [128]. Structurally, the bacterium is composed of an inner cytoplasmic layer followed by a film of peptidoglycan layer that is covered by an outer membrane anchored in a corkscrew helical shape around the outer axial filament [129]. It is a motile bacterium whose endo-flagellum lies in the periplasmic space with size ranges of 6-20 µm in length and 18-20 µm in width across the helical shape frame [129]. The whole genome sequence of TP and in particular the simian prototype isolate Fribourg-Blanc has a circular chromosomal genome of about 1,140,481 base pairs (bp) with approximately 1,122 genes coding for various proteins [23].

All the three pathogenic subspecies of TP absolutely depend on mammalian host cells for their viability and growth [129]. The obligatory and parasitic nature of these bacteria has precluded their successful *in vitro* culture and propagation on all known axenic media. To date, routine propagation is only achievable in animal models [130]. However, recently, *in vitro* culture of TPA was experimentally demonstrated on a micro-aerobic rabbit epithelial cell co-incubation system using the well-studied TPA Nichols syphilis strain [131]. In this particular study, the authors reported that the cultured treponemes showed continuous growth with full viability measured by their ability to infect rabbits and remain motile for over 6 months. However, this propagation system has neither been applied to other closely related pathogenic subspecies

of TP such as TPE (yaws) and TEN (bejel), nor has it been commercially adopted for routine laboratory propagation [131].

2.5.1.4 Clinical disease in non-human primates

Since the early 1960s, NHPs in Guinea and Senegal were reported to be infected with pathogenic *T. pallidum*. These assertions were based on sero-prevalence studies that showed that though no clinical signs were evident in the troops of yellow baboons (*Papio cynocephalus cynocephalus*) at that time, the animals under study had a sero-prevalence rate of 60% in treponemal specific antibodies [22,114,132]. In the late 1960s, mild clinical symptoms in sero-positive guinea baboon troops (*Papio papio*) of Casamance region of Senegal were observed for the first time. The symptoms included localized ulcers to the muzzle and armpits coupled with reactive associated lymph nodes in body areas with ulcers [133]. Since then, a number of clinical symptoms have been described in a number of NHP species across sub-Saharan Africa. In active and overt clinical infections, symptoms observed include: necrotizing dermatitis of orofacial areas, distal extremities and anogenital areas in both sexes [19,134–136].

In the late 1980s in the GNP, Wallis and co-workers (1999) observed genital ulcerations in olive baboons (*Papio anubis*) for the first time [137]. In some cases, the genital ulcerations were so severe that they resulted in case fatalities due to obstructed urinary flow and ensuing urogenital sepsis [137]. Other cases of severe ano-genital lesions in NHP TP infections were later reported in 1994 in Lake Manyara National Park (LMNP) in Tanzania involving olive baboons [138]. Although Harper *et al*. (2012) had earlier argued that the clinical presentation of genital ulceration could have only been unique to olive baboons in East Africa [136]; to the contrary, this clinical manifestation has now been documented to occur in a number of NHP species across sub-Saharan Africa such as in African green monkeys (*Chlorocebus sabaeus*) of Bijilo Forest Park in The Gambia and Niokola –Koba National Park in Senegal [19]. Whether species of the infected NHPs plays a role or not in the ultimate syphilis-like disease clinical presentation, it still remains largely unknown but it is most highly unlikely. In this thesis, we also documented genital ulceration as a clinical manifestation in sooty mangabeys infected with TPE in the TNP.

The occurrence of NHP treponematosis tends to mirror the geographical distribution of human yaws that is confined to the tropical regions [120]. NHP cases have only been reported in countries that have active human yaws cases or have a history of human yaws endemicity [139]. A number of NHP species in the tropical regions of sub-Saharan Africa and south-east Asia have been reported to be infected by the pathogenic TPE spirochetes. Both Knauf *et al.* (2012) and Harper *et al.* (2012) through their study surveys in East Africa , confirmed the

infection with TP of olive baboons (*Papio anubis*) in many sites in Tanzania [135,136]. Recently Chuma and co-workers (2018, 2019), using both serological and molecular tools observed that TPE infections were still geographically widespread in Tanzania affecting olive baboons, yellow baboons*,* vervet monkeys (*Chlorocebus pygerythrus*) and blue monkeys (*Cercopithecus mitis)* [20,134]. Much more recently, the grivet monkeys (*Chlorocebus aethiops*) in the Ethiopian highlands have been confirmed to be infected with TPE [20].

In west Africa, the disease was reported to occur among guinea baboons of Senegal [133,140] and yellow baboons of Guinea [114]. Recently, Knauf and co-workers (2018) confirmed TPE in green monkeys from the Gambia and Senegal, and sooty mangabeys from Cote d´Ivoire [19]. Using serological tools, Klegarth *et al*. (2017), established that treponemal infections were common among free ranging and captive macaques in south east Asia [141].

In contrast to the above studies that employed laboratory tests to confirm the infection in NHPs, Struhsaker and co-workers (2019) reported facial yaws-like and genital syphilis-like lesions in yellow baboons (*Papio anubis*) of Kibale National park in Uganda but the infection was not demonstrated neither by serology nor genetic tests or at the very least, dark field microscopy [142]. Furthermore, Levrero *et al*. (2007) had earlier observed that western lowland gorillas (*G. gorilla*) in the Congo basin of Democratic Republic of Congo had "yaws-like" facial lesions [143]. These observations by Levrero *et al*. (2007) were recently confirmed by molecular tests that indeed gorillas are infected by TPE highlighting the first genetic evidence of TPE infection in a great ape [20]. Other sites where great apes have been seen to have symptoms consistent with TPE infection but no molecular confirmation as yet are; Cameroon (gorilla and chimpanzee), Gabon (gorilla), and Uganda (chimpanzee) [139].

2.5.1.5 Transmission of TPE among NHPs

Transmission of TP among humans is thought to be dictated by behavioral practices and socio-economic status of people. For example, non-venereal trepanomatoses (yaws, bejel, and pinta) are largely considered diseases of poverty as they affect mainly poor communities with poor sanitation in the tropics and dry arid areas (bejel). Hence, socioeconomic status of affected populations seem to shape the epidemiological patterns of these diseases in part [120]. They are believed to spread via contact i.e. either skin to skin contact or via a breach in mucous membranes in the case of bejel [120]. Transmission of TP among the NHPs is also not clearly understood. However, based on the clinical lesions that have been observed over time in various troops and species, some modes of transmissions have been suggested.

TP transmission in NHP populations is to a great extent believed to be via direct contact between infected and naive individuals [134]. This argument is based on the similarities of first clinical manifestations noted in NHPs in the 1960s that mirrored yaws disease in humans that

is believed to be largely spread by contact [133]. In the particular study, the investigators observed presence of “yaws-like” lesions loaded with treponemes localised to the muzzle and armpit region in a number of guinea baboons (*Papio papio)* similar to those noted in human yaws infection [133]. Due to this clinical resemblance, it was argued that transmission modes in both hosts should be similar. However, others have argued that for either intra-species or inter-species transmission to occur via this mode, there must be close interactions that warrant break of skin such as tumble-play, fights and predation among different individuals [143–145] to warrant entry of treponemes into the naive host [146,147].

The sexual transmission route was also suggested in 1989 when syphilis-like lesions localised to the genital areas in sexually active olive baboons (*Papio anubis*) were observed in GNP [137]. Later in 1994, ano-genital ulcerations consistent with yaws infection reported earlier in GNP [137], were reported in olive baboons of LMNP not far from GNP [138]. Further reports highlighting the gross pathology, serological evidence and genetic characterisation of this ano-genital disease at LMNP and other sites in east Africa were documented by both Harper *et al.* (2012) and Knauf *et al.* (2012) [135,136]. Based on this predilection of the syphilis-like lesions for the ano-genital body areas and occurring mostly in sexually active animals, scientists have suggested that sexual transmission could be responsible for the spread of these pathogenic treponemes in these troops. However, more holistic studies interrogating other risk factors such as the roles played by vectors are needed to ascertain this transmission mode [136]. Though ano-genital lesions suggest sexual transmission and resemblance to human venereal syphilis, genetic characterisation of strains isolated from these lesions are TPE strains whose transmission is largely believed to be via contact [19,139].

Vectoral transmission was suggested based on experimental studies that demonstrated the transmission of the yaws bacterium by insect vectors of the genus *Hippelates pallipes* between man and rabbits. This was based on appearance of comparable clinical lesions in the newly inoculated rabbits and the demonstration of viable spirochetes isolated from inoculated rabbits on dark field illumination microscopy [148]. In other observational studies involving primates, it was suggested that mechanical transmission of TP by arthropod vectors would be plausible in settings where NHP have open wounds [143]. The role played by biting flies such as stable flies and sand flies in TP transmission has not been investigated. These biting flies have been shown to be effective in the transmission of a wide range of diseases in both human and animals [149,150].

Though transmission possibilities of the yaws spirochete by arthropods were demonstrated in humans [151], at the point of writing this dissertation, there were no known experiments done to demonstrate vectoral transmission of the TPE bacterium in NHPs. Recently, Knauf and co-workers (2016) isolated TP DNA from flies inhabiting the ecosystems where yaws disease in

NHPs are reported to be prevalent [69]. Gogarten *et al*. (2019) then showed that flies that formed stable associations with NHPs social groups also carry TP DNA, opening the way to their spreading the pathogen [68]. Commenting on the study of Knauf and co-workers (2016), Stamm (2016) raised pertinent questions that future studies should consider to better understand the role flies play in the TPE ecology and these include: What are the sources of fly associated TP DNA?; Are fly associated treponemes viable and in sufficient amounts to warrant inter-species or intra-species transmission?; Is there clinical or serological evidence of TPE infection in humans that live in the borders of these sites?; If so, what effect would fly control have on the incidence/prevalence of yaws in these regions? [152].

2.5.1.6 Zoonosis perspective

Whether inter-species transmission between humans and NHPs exits or not, is largely unknown. However, TPE strains isolated from NHPs and those isolated from humans show highly similar genomes, which is notably compatible with recent circulation/transmission between NHPs and humans [19]. Unfortunately, the paucity of TPE genomes from both human and NHP infections limits further investigations into this zoonotic transmission possibility. If this zoonotic transmission of yaws occurs undetected, it has been viewed by experts as one that can undermine and impede the ongoing yaws campaigns by WHO scheduled to eradicate yaws globally by the year 2030 [19,69,120,139]. Like any other human disease eradication strategy, the success of this campaign is partly dependent on the assumption that no other non-human reservoir requires neutralisation [113].

In an experimental setting, the simian Fribourg-Blanc strain isolated from a guinea baboon (*Papio papio*) induced classical yaws symptoms in humans [153]. Similarly human infecting spirochetes of the species TP were also reported to elicit symptoms in NHPs [154]. Generally, this suggests that host innate immune response and molecular compatibility barriers to interspecies transmission of TPE are low, though other barriers to spillover might certainly exist [155]. In addition, geographical proximity between humans and NHPs seems to be close. For example, all countries where TPE has been reported in NHPs either have a history of yaws endemicity or are presently reporting active infections in humans suggesting a geographical proximity of both human and NHPs infection trends [139,156]. Opportunities for transmission between NHPs and humans exist. In their report, Knauf *et al.* (2012) advanced a theory that bush meat trade is a common practice in most countries where yaws remain endemic and offers a point of contact between humans and NHPs sought for their meat. Furthermore, the rise in the demand of usable land in the wake of increasing agricultural activity and industrialisation leading into encroachment of NHP habitats by humans also increases the likelihood of cross-species transmission [2,8,155]. More studies are therefore

needed to investigate the zoonotic potential of NHP-infecting TP, including NHP-infecting TPE.

2.5.2 Leprosy

2.5.2.1 Overview and aetiology

Leprosy is a chronic infectious disease that predominantly affects the integumentary system, associated peripheral nerves, and the mucosa of the upper respiratory tract [25,157]. If it remains untreated, the disease leads to lameness, blindness, organ failure and ultimately death [25]. Leprosy has affected humankind since time immemorial [80,158,159] and still remains a public health threat of significant proportions in tropical countries where the disease is endemic. For example, in 2016 alone, WHO recorded over 200, 000 of new cases and 80% of these cases were recorded in India, Indonesia and Brazil [160]. Brazil remains the only country above the WHO target of less than one new case per 10,000 population per annum [111,160] .

Also known as Hansen disease, named after the Norwegian physician who first described the pathogen in 1873, leprosy is caused by the bacteria *Mycobacterium. leprae* (*M. leprae*) of the genus *Mycobacterium* [25]. *M. leprae* was the only known cause of leprosy until the year 2008 when *Mycobacterium lepromatosis* (*M. lepromatosis*), was isolated from patients that had died of leprosy [161]. *M. lepromatosis* is the sister species of *M. leprae*, from which it may have diverged ca. 14 million years ago [162].To a greater extent, the two pathogens have been reported in humans [161,163] and to a lesser extent in wildlife [24,111,164,165]. However, only *M. leprae* has been described to infect NHPs [24].

2.5.2.2 The genus *Mycobacterium*

The genus *Mycobacterium* consists of more than 140 species of bacteria (Phylum: Actinobacteria; Order: Actinomycetales Family: *Mycobacteriaceae*) that are acid-fast, rod-shaped, slow growing and non-motile [166,167]. These bacteria inhabit a wide range of environments and hosts such as water bodies, soil and mammals [168]. The mycobacteria of medical and veterinary significance can be classified into three broad categories and include: *Mycobacterium tuberculosis* complex (MTBC) whose members cause tuberculosis in humans (*M. tuberculosis, M. africanum, M. canetti,* and *M. bovis*) and animals (*M. bovis, M. caprae, M. pinnipedi,* and *M. microti*) [169]; leprosy causing mycobacteria in both humans and animals (*M. leprae* and *M. lepromatosis*) [25,158,161]; and the non-tuberculous mycobacteria (*M. marinum, M. abscessus* and *M. avium-intracellulare* complex) that cause a broad range of infections in immunocompromised individuals in humans [170,171]. Of particular interest in this dissertation, are the leprosy causing mycobacteria.

2.5.2.3 *Mycobacterium leprae*

Mycobacterium leprae is rod-shape bacillus that is acid fast and can be demonstrated on both Fite-Faraco and Ziehl Neelson stains. The size ranges measure about 1-8 μm and 0.2–6 μm in length and width respectively [166]. The *M. leprae* bacilli are encapsulated with a rich hydrophobic mycolic acid layer a feature that is believed to assure good DNA preservation in paleopathological studies [82]. The genome of *M. leprae* averages 3.2 million base pairs and is second smallest of the genomes in the genus *Mycobacterium* [172]; owing to the massive genome decay [173]. This reductive genome evolution is estimated to have occurred between 12-20 million years ago and inevitably led to *M. laprae* becoming an obligate intracellular parasite that does not grow outside a living host [173,174]. Consequently, to date, *M. leprae* has not been propagated on any known axenic media [173]. The successful propagation of the bacteria has only been achieved in animal models with the longest doubling time ever known in bacterial propagation of 14 days [175]. As a result of this biological feature, *M. leprae* also has the longest incubation period averaging about 3-5 years in humans and in rare circumstances 30 years [157]. The genome of *M. leprae* is remarkably highly conserved with only about 215 single nucleotide polymorphisms (SNPs) identified among regional reference strains sequenced so far. This translates to 99.99% sequence identity for leprosy genomes available in public domain [174,176].

While this close sequence identity is appreciated, the leprosy phylogeny still resolves into a diversity of four major genotypes with subtypes that are reflective of their geographic origin, medieval human migrations and trade routes [163,176]. Benjak *et al.* (2018) reconciled the genotypes suggested earlier [176] after analysis of 154 genomes from 25 countries representative of the global distribution of leprosy [163]. Genotype 1 has three human derived subgenotypes (A, B and C) that circulate in East Asia, Middle East, South Asia, and single isolates from Central America and Malawi. Genotype 2 also has three human derived subgenotypes (E, F, and H) [163]. Subgenotype 2E consists of isolates predominantly from Ethiopia in East Africa and a few from Yemen in Middle East. On the other hand, subgenotype 2F is representative of strains that circulated in the medieval times in Europe and recent Ethiopian strains [80], while 2H strictly consists of strains from Ethiopia [163]. Genotype 3 has four subgenotypes (I, L, 3K-1, and 3K-0) isolated from humans, red squirrels (*Sciurus vulgaris*) and armadillos (*Dasypus novemcinctus*). Subgenotype 3I consists of strains circulating presently among red squirrels in Europe and those that circulated in the medieval time in Europe [80]. In this cluster also are strains responsible for the zoonotic transmission of leprosy in southern USA between armadillos and humans and also isolates from humans in the Americas [163,177]. Subtype 3L only has a single isolate from Martinique in the Caribbean which was erroneously misclassified as a strain originating from new Caledonia earlier [176].

Subgenotype 3K-1 and 3K-0 consist of exclusively human derived strains circulating in East Asia and Oceania. Lastly, genotype 4 has three subgenotypes (P, O and N) with subgenotypes 4P and 4N having strains from West Africa, Central America and South America. Subgenotype 4O exclusively has only West African strains [163,176]. In addition, genotype 4 had only human strains until recently when strains derived from NHPs captured from West Africa were determined and formed a separate subtype that has come to be known as 4N/O [80,159]. The 4N/O comprises of NHP derived strains and one human isolate from West Africa and two other isolates of the same patient from Brazil [163]

2.5.2.4 *Mycobacterium lepromatosis*

M. lepromatosis is a close relative of the bacterium *M. lepare,* and the two share a deep most recent common ancestor (MRCA) [162]. The two pathogens are distinguishable genetically and on the basis of the variable disease syndromes they cause [178]. The comparative genomic analysis of the 16S rRNA genes of the two pathogens, revealed significant sequence differences that warranted the assignment of *M. lepromatosis* as a separate species [178]. Furthermore, whole genome comparison of the two only yielded 87% sequence identity as opposed to the 99.99% identity that *M. leprae* genomes share among themselves [174,176,179]. However, due to the reductive evolution that occurred to the MRCA 10 million years ago prior to their divergence, both share a number of common biological features that include: both bacteria are acid fast, morphologically similar; uncultivable on axenic media, slow doubling times, obligate intracellular parasites, similar genome size and both show massive genome decay with similar pseudogenes [162]. Despite these similarities, the two cause variable disease spectrums of leprosy in humans [178–180]. In addition, *M. leprae* has wide geographical prevalence whilst *M. lepromatosis* has only been reported in Mexico, Caribbean and the United Kingdom [162,181].

2.5.2.5 Leprosy disease in humans

Few reports detailing leprosy disease spectrums in NHPs are available precluding the good illustration of the disease in NHPs. Therefore, to understand the disease spectrum in NHPs, the comparison to the human disease spectrum that is well illustrated is recommended [15]. In this regard, a review of the human disease spectrum is necessary.

The human leprosy case definition should satisfy at least one of the following; the presence of hypo-pigmented or reddish skin lesion(s); damage to the peripheral nerves as demonstrated by loss of sensation and mobility to hand, feet or face and positive skin smears [25,157]. Clinically, two modes of classifying leprosy were adopted i.e. the Ridley-Jopling system[182] and the Madrid system [183].

2.5.2.5.1 Ridley-Jopling system

In the Ridley-Jopling system, Five categories of disease manifestations are recognised based on the level of host immune response against the pathogen [182]. These categories are arranged in the descending order of immunity of the host. On one polar end of the disease spectrum, is the tuberculoid leprosy (TT) characterised by few and defined asymmetric skin lesions, few bacilli on skin smears, low antibody titres, high cellular response and the demonstration of lymphocytes and epithelioid cells on histology [157,182]. On the other polar end representative of the low immunity of the host is the lepromatous leprosy (LL). This is the extreme form of the disease when the host can no longer mount immunity against the causative agent. At this stage, the disease is characterised by many and diffuse symmetric skin and nerve lesions, many bacilli on microscopy of skin smears, low cellular response, high antibody titres, and demonstration of undifferentiated macrophages on histology [157,182]. Between the two polar ends of the disease spectrum, lies the intermediate forms of the disease that include: borderline tuberculoid leprosy (BT), borderline leprosy (BB), and borderline lepromatous leprosy (BL). The clinico-pathologic characteristics of the intermediate forms tend to be a linear progression of the symptoms of the tuberculoid leprosy towards the lepromatous leprosy in a linear correlation with a declining immunity of the host [157].

2.5.2.5.2 Madrid system

The Madrid system has 3 categories based on the clinical evaluation of the lesions. The 3 categories are tuberculoid, borderline, and lepromatous forms of leprosy. In the tuberculoid leprosy, patients present with few isolated hypo-pigmented or copper to brownish coloured nodular or macular lesions. In some instances, there might be evidence of loss of sensation, nerve thickening in the affected areas and loss of motor function. Microscopy of skin smears from these lesions reveal few acid fast bacilli [25,183]. Borderline leprosy is the intermediate of the tuberculoid and the lepromatous leprosy where patients present with moderate and asymmetrical lesion distribution coupled with definite loss of nerve function to the affected body areas [184]. In this category, the hypo-pigmented lesions are neither isolated as in tuberculoid leprosy nor widely spread as in lepromatous leprosy. In contrast, lepromatous leprosy is characterised by chronic inflammatory tissue involving large skin areas. Skin smears from these lepromatous lesions contain enormous amounts of bacilli arranged in slabs (lepromas) [25]. In addition, there is permanent nerve damage resulting in loss of sensation and motor function to the affected areas [25,183]. In the severe cases that remain untreated, the infection can lead to blindness, deformity and a systemic infection affecting several body systems [184].

2.5.2.6 Recent considerations to leprosy classification

The above two classification systems were defined based on infection with *M. leprae* alone. *Mycobacterium lepromatosis* has a different clinico-pathological manifestation in humans causing diffuse lepromatous leprosy (DLL) [161]. DLL is characterised by wide spread necrotic skin lesions containing large amounts of acid fast bacteria (AFB) [161,185]. In addition, autopsy and diagnostic evidence suggest that wide internal organ involvement is characteristic of *M. lepromatosis* infection [180]. *Mycobacterium lepromatosis* is considered endemic to Mexico and the Caribbean and accounts for more leprosy cases in these regions than it was earlier believed and hence *M. lepromatosis* should always be considered an important differential in all leprosy cases in these regions [186].

2.5.2.7 Leprosy in non-human primates

The clinical course of the disease in NHPs has not been well described due to the limited number of leprosy cases that have been reported. However, based on the few case reports of natural and experimental infection of leprosy, a number of consistent symptoms have been described and the disease in NHPs tend to mirror the disease presentation observed in humans [187]. Furthermore, similar incubation periods of 5-30 years as is the case in humans have been noted too in NHPs [24].

In one case of a chimpanzee taken from Sierra Leone (West Africa) to Japan for research purposes, it manifested lepromatous leprosy after 29 years in the sanctuary. The gross lesions observed included: multiple nodular hypo-pigmented lesions on the lips, around eye orbits, crus and abdomen suggestive of lepromatous leprosy [188]. PCR tests on nasal swab smears collected from this animal confirmed leprosy infection [188]. In addition, other cases involving chimpanzees in a research centres, similar clinical manifestations were described [15,187]. In sooty mangabeys, the disease presentation observed was that of borderline-leprosy to borderline–lepromatous leprosy with neuropathic deformities of limbs [16,18]. In experimental settings; inoculation of sooty mangabeys with viable *M. leprae* strains resulted in diverse clinical symptoms ranging from intermediate to lepromatous leprosy forms with neural manifestations in some individuals. The higher the inoculation doses, the severe the symptoms suggesting a pathogen load dependant manifestation [189].

From cases reported so far, only three NHP species have shown susceptibility to leprosy infection and these include; chimpazees [15,187,188], sooty maangabeys [189] and cynomolgus macaque (*Macaca fascicularis*) [17]. Unfortunately, in all the above cases, authors claimed that the NHPs were infected in countries of origin prior to moving into research sanctuaries abroad, a scenario suggesting natural infection prior to their translocation [16,24,187,188]. However, relating such findings to the disease status in wild NHPs in these

regions is challenging. Recently, Honap *et al.* (2018) compared genomes generated from archival material collected from these case reports [16,17,188] with other published human derived genomes and showed that the NHP derived *M. leprae* genomes were closely related with human *M. leprae* strains circulating in the regions they came from revealing a strong geographical association of the NHPs and their native place of origin [24]. The phylogenetic analysis placed the genomes derived from a chimpanzee and sooty mangabey that originated from West Africa into branch 4 (genotype 4) of the leprosy phylogeny that consists of human strains circulating in West Africa and South America [24]. Similarly, the genome derived from cynomolgus macaque that originated from The Philippines clustered in branch 0 (subgenotype 3K-0: Schuenemann *et al.* 2018) consisting of *M. leprae* strains infecting humans in East Asia [24]. To the best of my knowledge, other than the cases investigated in this study, there are no other known cases of *M. leprae* infection in wild NHPs that have been observed in the wild.

2.5.2.8 Transmission of *M. leprae* among NHPs

The transmission of leprosy in NHPs is not well understood just as it is not well illustrated in humans [173]. This is precluded by the inability to propagate leprosy causative agents on axenic media coupled with slow doubling times in animal models [173,190], thwarting all transmission experiments attempts in all known hosts [111,173]. In NHPs, this is compounded by paucity of genomic data from NHPs and other wild hosts arising from the lack of leprosy surveillance in the wild. Consequently, this has led to the limited understanding of sylvatic leprosy ecology that affects NHPs [111].

In humans, transmission of leprosy is thought to be through the respiratory route during periods of extended contacts between naive and infected individuals especially those of same households [111,184]. In the same vein, transmission by close contact was suggested in a case involving 2 sooty mangabeys that were housed together in a research sanctuary in the USA [16,18]. The authors claimed that the second case of a sooty mangabey [18] acquired the leprosy infection as a result of contact with the index case, [16] as the two animals shared the same housing for a long time and the symptoms were comparable. The symptoms in the second case only appeared 7 years after the index case fitting within the incubation period of leprosy [18]. However, this example does not rule out the possibility of common source infection for both cases and that the second case could just have had a long incubation period as suggested in literature [188].

The other possible route of transmission between different wild NHP species is predation [24]. For example, chimpanzees are known to prey on sooty mangabeys [191,192] a scenario that possibly explains the close phylogenetic relationship of both sooty mangabey and chimpanzee derived genomes in the leprosy phylogeny [24]. To put this example into perspective; in

humans, leprosy is likely to be transmitted by consumption of infected meat as shown by da Silva *et al.* (2018) that established a positive correlation between the consumption of infected armadillos meat and the human leprosy infection in the leprosy endemic region of the Brazilian Amazonia [111].

2.5.2.9 Leprosy in other wildlife species

2.5.2.9.1 Leprosy in armadillos

Humans were considered exclusive hosts of *M. leprae* till a leprosy-like disease was reported in wild nine banded armadillos in the Texas gulf coast of the USA [193,194]. Since then, the natural infection in armadillos has been described to occur overlapping the wide geographical range of the armadillos in many counties of south-eastern USA, Mexico, Argentina and Amazonian basin [111,190]. Overt outward symptoms of the disease are rare in armadillos and the infection is mainly considered a systemic one [195,196]. When gross lesions are present, they are confined to the cooler parts of the animal such as the feet, nose, ears and ventral body and are indicative of an advanced infection [197,198]. Skin lesions are difficult to observe due to the hardy integument covering the body [198]. Histological examinations of tissues from deceased animals reveal a generalised dissemination of acid fast bacilli (AFB) in most body systems a histological finding indicative of a systemic leprosy infection [199]. Due to their susceptibility to infection and ability to produce large quantities of *M. leprae* bacilli despite the slow propagation turnovers, armadillos are a model of choice for *In vivo* propagation [200]. However, it is believed that *M. leprae* was already circulating in the wild armadillos prior to their use as laboratory models [201].

The zoonosis perspective with regards to leprosy infection is not a recent phenomenon. It was already suspected back in the 1980s that exposure to armadillos was associated with some human leprosy cases in the armadillos native regions [202,203]. But due to lack of technological tools such as genotyping back then, the hypothesis could not be proved [190]. However, recently, it has been shown that humans can contract sylvatic leprosy from infected wild armadillos. It was reported that 42% of leprosy patients that participated in a leprosy survey in South-Eastern USA, had *M. leprae* subtype 3I strain associated with the armadillos [177]. Furthermore, several more cases of zoonotic transmission of the sylvatic leprosy attributed to armadillos exposure have been described in Brazil Amazonia [111]. In the Brazilian Amazonia, it was reported that people with repeated or long exposure to the armadillos had high titres of PGL-1 antibodies which is a specific diagnostic serological marker for leprosy infection [111]. The risk factors for this cross-species transmission is thought to be either direct or indirect contact of humans with armadillos through habitat encroachment,

armadillos meat preparation and consumption, and the keeping of armadillos as pets [111,177,190].

2.5.2.9.2 Leprosy in red squirrels

In addition to NHPs and armadillos, British red squirrels are also infected with both *M. leprae* and *M. lepromatosis* bacteria [181,204]. In overtly diseased animals, symptoms observed in both *M. leprae* and *M. lepromatosis* infected red squirrels are indistinguishable and include: wart-like lesions on lips, eyelids, ear pinnae, snout, limb extremities and loss of fur [181]. Histological features of the infection include: neural involvement with presence of AFB in nerve endings, generalised dissemination of AFB in organ depicting a BL and LL for *M. leprae* and DLL for *M. lepromatosis* infection respectively [181].

Recently, red squirrels in England were also implicated in zoonotic leprosy as they were found to be infected with a *M. leprae* subgenotype 3I that infects both humans and armadillos [181]. This subtype has been identified as one responsible for the zoonotic circulation of leprosy in southern USA [177,205]. In addition, this subtype also infected humans in the medieval times in what is the present England and Denmark suggesting the possibility of persistence of *M. leprae* in other reservoir hosts other than humans [80,163,181]. It is thought that the *M. leprae* spillover from humans to the red squirrels happened in the medieval times and was maintained in these rodents since then [181]. On the other hand, the phylogenetic placement of NHP derived *M. leprae* genomes into the respective west African and Asian clades consisting of human infecting strains raises the possibilities of zoonotic circulation of *M. leprae* in these regions as this placement is also reflective of the origin of NHPs [24]. However, in both red squirrels and NHP associated leprosy, further research is warranted to ascertain the level to which this probable zoonosis occurs and the role which these wildlife hosts play in the ecology of the disease [24,181].

3 Materials and methods

3.1 *Treponema pallidum* infections in NHPs in TNP and other field sites

3.1.1 Study sites and samples

The focus area for the study of *Treponema pallidum* infections in NHPs was the TNP (5° 45′ 0″ N, 7° 7′ 0″ W). The study also included 11 other field sites across sub-Saharan Africa where bone samples analyzed in this study were collected from (Figure 1, Appendices: Table S1). The 3,300 km^2 area of TNP lies within the remaining larger parts of the primary evergreen rain forest of west Africa and is located in the western part of Côte d'Ivoire bordering Liberia [206]. For over half of the year, the climate in TNP is characterised by high humidity as a result of high temperatures (~28°C) and high amounts of rainfall (1800mm) [206]. This unique climate supports its rich biodiversity in flora and fauna leading to the park being declared a world heritage site in 1982 [207]. The park is home to a wide range of birds, reptiles, amphibians and mammals that include 11 different NHP species [208,209]. There are three habituated chimpanzee social groups in TNP that are followed for studies (i.e. east, south and north groups) managed under the TNP chimpanzee project [208] and two habituated monkey groups studied under the TNP monkey project [209]. Since 2001, a systematic health monitoring system has been in place in TNP to monitor the health of both habituated chimpanzees and monkey groups and unhabituated groups of NHPs in general [5]. The health monitoring involves: the daily observation of habituated groups by interdisciplinary team of primatologists and a veterinarian; investigations through necropsies; the collection of non-invasive samples (urine and feces) or minimal invasive sampling where necessary, laboratory investigations and the communication of the results to the park management for corrective and preventive measures [5].

In January 2014, Cercocebus atys monkeys from a habituated social group in TNP were observed with orofacial lesions and lesions of their distal extremities; two animals (Animal ID 1863 and 1864: Table 4) were previously sampled and TPE was determined to be the cause of the infection [19]. Over the next two years, other individuals started showing genital ulcerations and necrotising dermatitis on the inner parts of the thighs and ventral abdomen, often with visible yellow crusts. Orofacial and genital lesions were still observed on other animals in the group during the study period (Figure 2). Three more individuals with visible lesions were chemically immobilised using a combination of xylazine (1 mg/kg) and ketamine (10 mg/kg) administered by blowpipe. Biopsy and lesion swab samples were collected from genital and orofacial lesions.

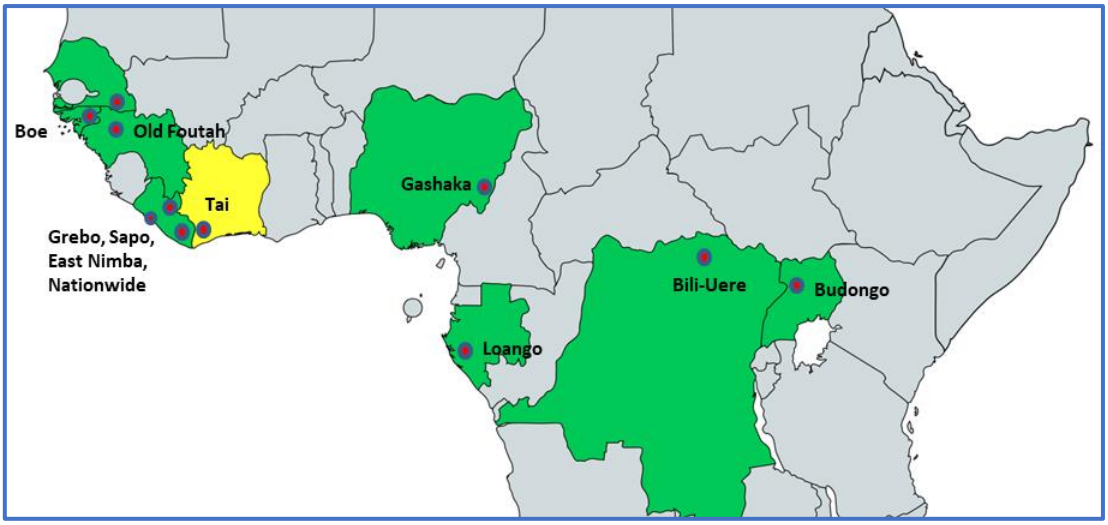


Figure 1: Study sites. Red dots indicate sites where samples analysed for TPE in this study were collected from across sub-Saharan Africa. TNP (in yellow) was the main study site and is the only place where either tissue or swab and bone samples were collected. In the other sites, only bone samples were collected (Map created in; https://mapchart.net/africa.html)

Figure 2: Lesions due to TPE infection in sooty mangabeys. A. Necrotising dermatitis of inner parts of the thighs and ventral abdomen with yellowish crusts**. B.** Necrotic orofacial lesions **C.** genital lesions in females **D.** Genital lesions in males. (*Photos: Leendertz's lab-Robert Koch Institute, Berlin*)

Samples for molecular biology analyses were preserved in RNAlater® (Life technologies, New York). Samples were shipped to the Robert Koch Institute (RKI), Berlin, Germany for laboratory analysis.

To get insights into the NHP trepanomatoses circulating in TNP during the last three decades, we screened non-symptomatic bones collected opportunistically at TNP (*N*=67; Appendices: Table S1). We also screened NHP bones collected from 11 additional field sites in sub-Saharan Africa (*N*=80; Appendices: Table S1; Figure 1) collected through the support of the Pan African Program: The Cultured Chimpanzee (PanAf; www.panafrican.eva.mpg.de). NHP bones were assigned to particular species using molecular methods targeting the 16S gene as described previously [95], coupled with morphological assignment by experts in the field.

3.1.2 DNA extraction

DNA was extracted from both skin biopsies (N=4) and swabs (N=3) using the DNeasy Blood and Tissue kit (QIAGEN, Hilden, Germany) following the manufacturers' instructions. DNA was extracted from bones using a silica-based method. Briefly, non-lesioned bones were drilled using a fine drill bit at slow speed to produce ~150 mg of bone powder. The drilling was done in a designated sealed glove box, both to prevent any contamination of the bones and to prevent exposure of researchers to pathogens that might be present in the bones (e.g., in TNP, cases of sylvatic anthrax (*Bacillus cereus* biovar *anthracis*) are frequent, and this pathogen can be cultured from bones) [47]. The box was UV sterilized and subsequently surfaces were bleached following the drilling of each bone and extraction. Drill bits were changed on every bone to prevent cross contamination. The DNA extraction from the bone powder was done following a protocol modified from [210,211] and described in detail in [212]. Extracted DNA from all samples was quantified with a Qubit fluorometer using the double stranded DNA high sensitivity assay kit (Invitrogen™, California USA) following the manufacturer's instructions. DNA was subsequently stored at -20°C.

3.1.3 Screening for *Treponema pallidum*

To screen for *Treponema pallidum* DNA in swabs and biopsy samples, we performed an end-point polymerase chain reaction (PCR) assay targeting the 67 bp of the *polA* gene fragment, using primers (Table 1) previously developed for screening human clinical specimens [213]. PCR reactions were performed in 25µL reactions; up to 200 ng of DNA was amplified using 1.25 U of high-fidelity Platinum Taq™ polymerase (Invitrogen™), 10x PCR buffer (Invitrogen™), 200µM dNTPs, 4mM MgCl2, 200nM of both forward and reverse primers. The thermal cycling profile was as follows; denaturation at 95°C for 5 min, followed by 40 cycles of 95°C for 15 sec, 60°C for 30 sec, and 72°C for 1 min. An elongation step at 72°C for 10 min was included. Known positive DNA and negative controls were included.

The 67 bp amplified product is too short for direct Sanger's sequencing, so to confirm the results of this initial screening, positive samples were further tested with a semi-nested assay targeting the cytoplasmic filament protein gene (*cfpA*) [136]. These primer pairs (Table 1) yielded a 352 bp outer product in the primary PCR and a 189 bp inner fragment in the nested PCR. In both primary and nested PCR assays, the reactions were performed as follows; about 200 ng of DNA was amplified in a 25μL reaction using 1.25 U of high-fidelity Platinum Taq™ polymerase®, 10x PCR buffer, 200μM dNTPs, 4mM MgCl2, 200nM of both forward and reverse primers. Subsequently, 2 μl of a 1:20 dilution of the first round PCR product was used as input template for the nested PCR. The thermal cycling profile for both the primary and nested PCRs were the same as in the initial screening assay.

PCR products were visualized on a 1.5% agarose gel stained with GelRed® (ThermoFisher Scientific, Massachusetts, USA). Positive bands were purified using the PureLink™ Quick Gel Extraction Kit (Invitrogen™) following the manufacturer`s protocol. Purified products were stored at -20°C until they were sequenced using BigDye Terminator v3.1 Cycle Sequencing Kit and sequences were compared to publicly available sequences in EMBL through BLAST [214]. All samples that tested positive in the confirmatory assay were selected for whole genome in-solution hybridization capture and high throughput sequencing.

Potential DNA degradation in bone samples precluded the use of the above-mentioned confirmation assay. To select the most promising samples, we estimated copy numbers in the bones, using a real-time qPCR also targeting the 67bp fragment of the *polA* gene [213]. Samples were tested in duplicate. Briefly, 5 μl of total DNA was amplified in a 25μL qPCR reaction containing 10x PCR buffer, 200μM dNTPs, 4mM MgCl2, 300nM of both forward and reverse primers, 100nM of the probe and 0.5U of high-fidelity Platinum Taq™ polymerase®. The thermal cycling profile was set as follows; DNA denaturation at 95°C for 10 min followed by 45 cycles at 95°C for 15s and 60°C for 34s. All bone samples that had detectable TP DNA in duplicate reactions were selected for in-solution hybridization capture and high throughput sequencing.

Table 1: Primers used for *Treponema pallidum* screening

Type of PCR	Primer pair 5´-3´	Product size(bp)
polyA PCR	Forward_AGGATCCGGCATATGTCCAA Reverse_GTGAGCGTCTCATCATTCCAAA	67
cfpA nested PCR	Forward_GAGTCCCAATGTGTTTCATCC Reverse_TAGGATGGCAATCTCCTTCG	352
	Forward_GAGCGTCTGGACGTAATGG Reverse_TAGGATGGCAATCTCCTTCG	189
polA qPCR	Forward_AGGATCCGGCATATGTCCAA Reverse_GTGAGCGTCTCATCATTCCAAA Probe_6FAM-ATGCACCAGCTTC+G+A	67

3.1.4 Library preparation, genome wide capture and high throughput sequencing

For biopsy and swab sample extracts, we sheared 1000 ng of DNA per sample to 400 bp fragments using the Covaris S2 (intensity: 4, duty cycle: 10%, cycles per burst: 200, treatment time: 55 seconds and temperature: 4-5°C). Bone samples were not sheared due to the potentially fragmented nature of DNA in these older specimens. Following shearing, two library preparation methods were followed. In one method, the DNA was then converted into double indexed Illumina libraries using the NEBNext® Ultra™ II DNA Library Prep kit (New England Biolabs®, Massachusetts, USA; samples: Tai_105, Boe_092, 22_52, 2117, 5847, 1864). In the other, the Accel-NGS (Swift BIOSCIENCES™, Michigan, USA; samples: 1864, 22_52, 2116) was used to convert the DNA into single indexed Illumina libraries following the respective manufacturer's protocols (Table 2). Samples 1864 and 22_52 were sequenced previously using libraries generated with Accel-NGS kit [19,21], but in this study, we generated new libraries from these samples using NEBNext® Ultra™ II DNA Library Prep kit to improve genome coverage. The fact that two different library preparation methods were used only reflects the adoption of a new method in the laboratory, mostly prompted by higher library conversion rates with the NEBNext® Ultra™ II DNA Library Prep kit; we have no reason to expect that these different methods would systematically impact the genomes generated. All generated libraries were quantified using the KAPA library quantification kit (KAPA Biosystems, Cape Town, South Africa) following the manufacturer's instructions. In all library preparations, chicken DNA was included as a control.

Libraries were enriched for *Treponema pallidum* sequences using in-solution hybridization capture with biotinylated RNA baits and following the manufacturer's protocol (myBaits, Arbor

Biosciences, Michigan, USA). The baits spanned the simian derived Fribourg-Blanc reference genome (RefSeq ID: NC_021179.1) with a 2x tiling. In-solution hybridization capture was done for two rounds of 48 hrs each. After each round of capture, a post capture amplification step was performed using the KAPA HiFi master mix (KAPA Biosystems) with P5 and P7 Illumina primers to generate about 200 ng of enriched DNA per sample. The post-capture amplification thermal profile was as follows: initial hot start at 98˚C for 2 min followed by 12 cycles or what is required to generate 200 ng (12-16 cycles) at 98˚C for 20 sec, 65˚C for 30 sec and 72˚C for 45 sec. Enriched libraries were quantified using the KAPA library quantification kit. Prior to sequencing, libraries were diluted to 4 nM and pooled for sequencing on an Illumina Miseq (Illumina, San Diego, USA) with 300 bp paired end reads (V3 chemistry; Table 2) at the Robert Koch Institute (Berlin, Germany).

Table 2: Summary table for library preparation, capture and sequencing of TPE

Sample ID	NHPs species	Sample type	Library Prep method		Capture method	Sequencing method	NP / Country
			NEBNext	Accel NGS			
Boe_092	*Cercocebus atys*	bone	X		In-solution	Miseq	Boe/GB
2117	*Cercocebus atys*	swab-genital lesion	X		In-solution	Miseq	TNP/CI
5847	*Cercocebus atys*	swab-genital lesion	X		In-solution	Miseq	TNP/CI
2116	*Cercocebus atys*	swab-genital lesion		X	In-solution	Miseq	TNP/CI
1864	*Cercocebus atys*	biopsy-face lesion	X	X	In-solution	Miseq	TNP/CI
Control	chicken DNA	tissue	X		In-solution	Miseq	-
Tai_105	*Piliocolobus badius*	bone	X		In-solution	Miseq	TNP/CI
22_52	*Piliocolobus badius*	bone	X	X	In-solution	Miseq	TNP/CI

Boe= Boe national park, **GB=**Guinea Bissau, **CI=** Côte d'Ivoire

3.1.5 Bioinformatics analysis

Paired-end reads generated here, along with published SRAs from prior TP sequencing efforts (Table 5 and Appendices: Table S2) were trimmed using Trimmomatic v0.38, removing the leading and trailing reads below Q30, clipping any part of the read where the average base quality across 4 bp was less than 30, and removing surviving reads less than 30 bp in length [100]. Surviving read pairs were merged using Clip and Merge version 1.7.8 with default settings [215]. Merged reads and surviving single-end reads were combined and mapped to TPE Fribourg-Blanc (RefSeq ID: NC_021179.1) using BWA-MEM [101] with a minimum seed length of 29. Mapped reads were sorted using Picard's SortSam and subsequently de-duplicated with Picard's MarkDuplicates (https://broadinstitute.github.io/picard/index.html). Alignments with a MAPQ smaller than 30 and a mapping length lower than 30 were also removed using SAMtools [216]. We called two consensus sequences for each genome, using Geneious v11.1.5 [102]: i) the first consensus required a minimum of 10 unique reads (i.e., 10X coverage) and at least 95% identity for a base to be unambiguously called, ii) the second consensus required 5X coverage and at least 95% identity. In the following, we mostly report about the results obtained with the 5X, 95% data set, which includes all the partial genomes generated in this study. To confirm these findings, we repeated the analyses using the 10X, 95% data set that included only the higher-quality genomes available, thereby excluding two partial genomes of lower quality we could assemble from bone material (Tai 105 and Boe 092).

Whole genomes on each set of consensus sequences were aligned using the multiple sequence alignment program MAFFT [217]. We then removed all previously described paralogous and putative recombinant genes (Supplementary material) and selected conserved blocks using the Gblocks tool [218] in SeaView v4 [219]. Phylogenetic inference was performed on the resulting alignments of informative positions after stripping all identical sites and ambiguities in the final data sets: 5X, 95% (1736 positions: 38 sequences including the two TPE genomes from bones) and 10X, 95% (1743 positions: 36 sequences). We ran analyses on versions of these alignments that included or excluded non-TPE sequences.

Final alignments were then uploaded to the online ATGC PhyML-SMS tool (http://www.atgc-montpellier.fr/phyml-sms/) for construction of a maximum likelihood phylogeny using smart model selection [220] with Bayesian Information Criterion and subtree pruning and regrafting (SPR) for the tree improvement and otherwise using default settings. Branch robustness was estimated using Shimodaira-Hasegawa approximate likelihood ratio test (SH-like aLRT) [221]. The maximum likelihood (ML) tree was then rooted using TempEst (version 1.5.1), which estimated the best-fitting root of these phylogenies using the heuristic residual mean squared function, which minimizes the variance of root-to-tip distances [222]. Evolutionary pairwise

distance between TP strains were extracted from the ML phylogeny using the Patristic program [223].

To explore the robustness of our phylogenetic analysis, we also ran Bayesian Markov Chain Monte Carlo (BMCMC) analyses with BEAST (version 1.10.4) using four different models that all assumed the same nucleotide substitution model (identified in the PhyML-SMS analysis) but combined either a strict clock model or an uncorrelated lognormal relaxed clock model with demographic models either assuming a coalescent process and a constant population size or a birth-death process. These models covered a plausible range of clock and tree priors, similar to those explored in another recent publication on syphilis genomics [224]. For all models, we examined the output of multiple runs for convergence and appropriate sampling of the posterior using Tracer (version 1.7.1) [225] before merging runs using Log Combiner (version 1.10.4) [226]. The best representative tree was then identified from the posterior set of trees and annotated with Tree Annotator (version 1.10.4: distributed with BEAST). The resultant maximum likelihood and maximum clade credibility (MCC) tree files were edited using iTOL (https://itol.embl.de/) [227].

For the partial genomes where the aforementioned phylogenetic pipeline was unable to resolve their position in the phylogeny, we performed phylogenetic read placement using the evolutionary placement algorithm tool EPA-ng to determine the position of individual reads on the TP phylogeny [104]. We performed read placement on a red colobus bone sample (22-52; this particular sample had a very long branch in our initial analyses) and two previously sequenced chimpanzee bone samples (11786 and 15028) that were collected in TNP [21], for which the presence of TP was confirmed, but no subspecies assignment had been performed. Briefly, we selected filtered merged and single-end reads that mapped to TPE that were 50 bp or longer. Surviving reads were then aligned to the genomes used to build the MCC tree, using the parsimony-based phylogeny-aware read alignment program (PaPaRa) [228]. The resulting PaPaRa alignment was then split into the query reads and the original alignment using the split function available in the EPA-ng toolkit [228]. To estimate the best fitting evolutionary model for the phylogenetic placement of the query reads, both the reference tree and reference alignment were evaluated using the RaxML-ng toolkit [229]. The best fitting model was then used to place query reads on the reference tree with the EPA-ng tool. The resulting EPA-ng jplace tree files were visualised as heat-trees depicting the percentage of reads placed on each branch using the gappa toolkit [105]. The settings for the tree visualisation were as follows: every query read was treated as a point mass concentrated on the highest-weight placement and multiplicity of each query read set to 1.

3.2 Leprosy disease chimpanzees in CNP and TNP

3.2.1 Site and Samples

3.2.1.1 Cantanhez National Park, Guinea Bissau

Cantanhez National Park (CNP) lies within the 1,057km^2 of Cantanhez forests in the southwestern part of the Republic of Guinea Bissau (Figure 3). It is a home to a number of NHPs including west African chimpanzees (*Pan troglodytes verus*) that live in close proximity with the expanding human farming communities that surround the park [230,231]. As of 2003, it is estimated that there were between 600 to 1000 unhabituated chimpanzees living in the fragmented forests of CNP [232] though current numbers are lacking. The first sightings of chimpanzees (*Pan troglodytes verus*) with leprosy-like lesions of CNP occurred in 2013 using camera traps. By 2017, animals with leprosy-like lesions were observed at the following study sites: Caiquene-Cadique, Lautchande and Cambeqze; all study sites managed by the Centre for Ecology and Conservation, University of Exeter, Penryn, UK in collaboration with the Centre for Research in Anthropology (CRIA NOVA FCSH), Lisbon, Portugal. It is not feasible to anesthetise these wild animals for biopsy collection; in April 2017, fresh faecal samples (N=121) from chimpanzee groups with symptomatic individuals (Figure 4) were collected in RNAlater on a cross-sectional basis following nesting of these groups (Appendices: Table S9). Collected samples were shipped to RKI for molecular analysis. In addition, a selected set of camera trap images were evaluated for clinical manifestations.

Figure 3: Map showing sites in West Africa where chimpanzees were observed to have leprosy-like lesions

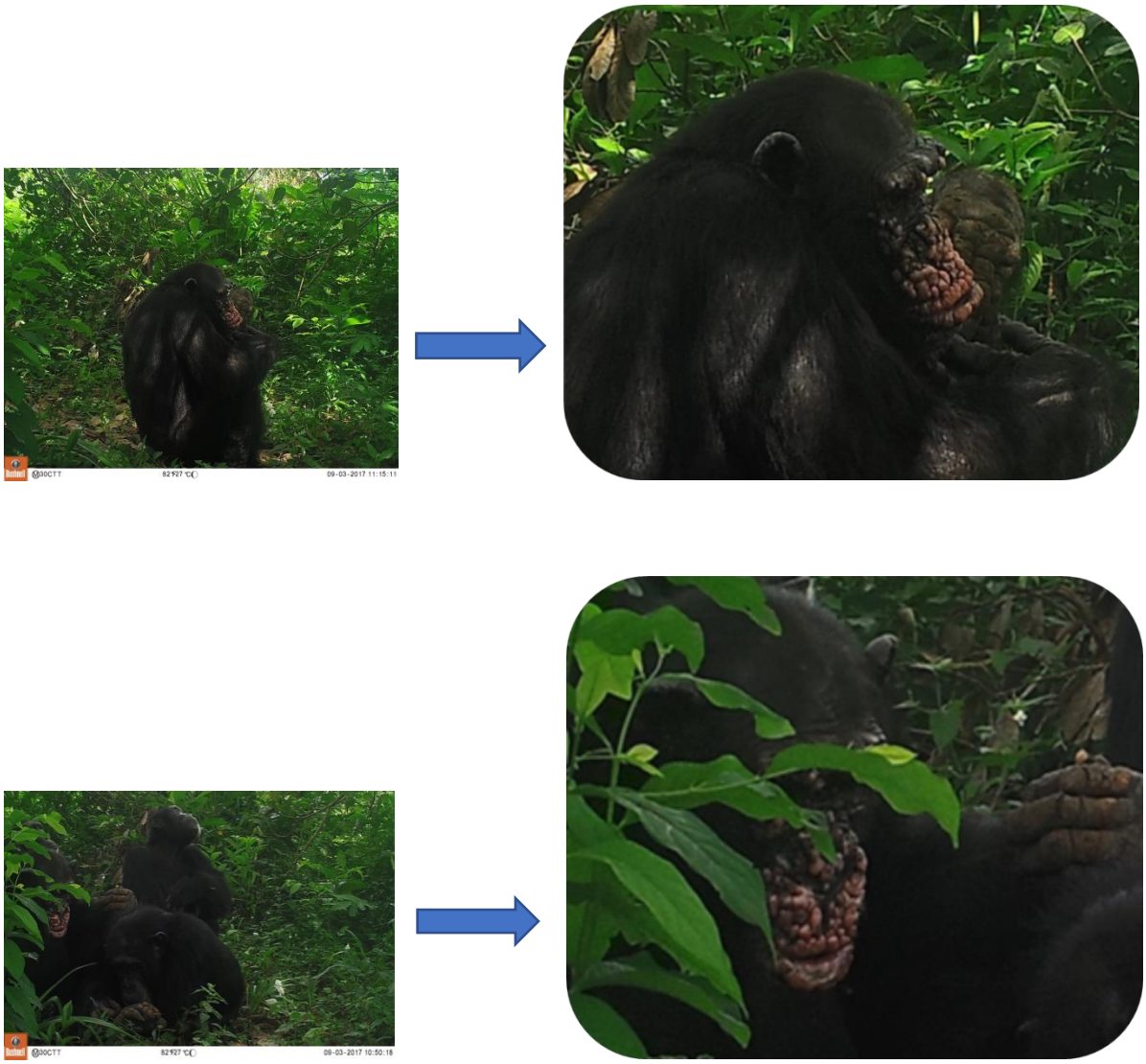


Figure 4: Camera trap images of symptomatic individuals in Cantanhez National Park, Guinea Bissau. Shown here are two individuals with multiple reddish to copper coloured nodular lesions on the maxilla-facial areas, distal extremities of the fore limbs, lips, eyebrows and ear pinnae. Images were taken on 27th March, 2017. (*Camera trap images: Centre for Ecology and Conservation, University of Exeter, Penryn, UK*)

3.2.1.2 Täi National Park, Côte d'Ivoire

Study site details for TNP are described in section 3.1.1. In June 2018, a habituated chimpanzee (referred hereafter as TNP_566/Woodstock) aged 25 years from the south group presented with leprosy-like lesions on the facial areas. Close human interactions are restricted and the affected individual could not be restrained for biopsy and hence, a non-invasive approach to screen faecal, urine and fruit wedge samples targeting the leprosy causative agent was applied. As part of the disease progression monitoring, still photos were taken on a regular basis till May, 2019 (Figure 5). In addition, screening of necropsy samples collected from TNP gave indication that another individual (referred hereafter as TNP_418/Zora) estimated to have been 57 years at the time of her death had leprosy prior her death in 2009. Review of archived photos of Zora revealed leprosy-like lesions (Figure 6).

3.2.1.2.1 Faecal, urine and fruit wedge samples

Faecal (N=16), urine (N=8) and fruit wedge samples (N=2) were collected from Woodstock after the onset of skin lesions from June through July 2018 and January 2019. These samples (Appendices: Table S4). were tested with PCR systems targeting *M. leprae.* Considering that leprosy has an incubation period of up to five years in people and in rare cases apparently up to 30 years in captive NHPs [24,188], we retrospectively screened all available faecal samples of this individual back to 2014 (Appendices: Table S5) to gain insights as to when this animal got infected.

3.2.1.2.2 Necropsy samples

All available necropsy samples (N=40) (Appendices: Table S8) since the start of the health monitoring program in TNP [5] were screened to determine whether leprosy was responsible for any mortalities or had infected any chimpanzees that had died in this community. The spleen was used for initial testing based on research in armadillos showing that leprosy can be detected in the spleen [111].

3.2.2 DNA extraction

DNA extractions were performed at RKI in a laboratory that has never been used for any leprosy investigations. DNA was extracted from faecal samples using the EURx GeneMATRIX stool DNA purification kit (Roboklon, Germany) and following the manufacturer´s protocol. DNA was extracted from necropsy samples using the DNeasy Blood and Tissue kit (QIAGEN, Germany) following the manufacturers' instructions. Extracted DNA was then quantified with a Qubit fluorometer using the double stranded DNA high sensitivity kit (Invitrogen™, Germany) following the manufacturer's instructions and was subsequently stored at -20°C until use.

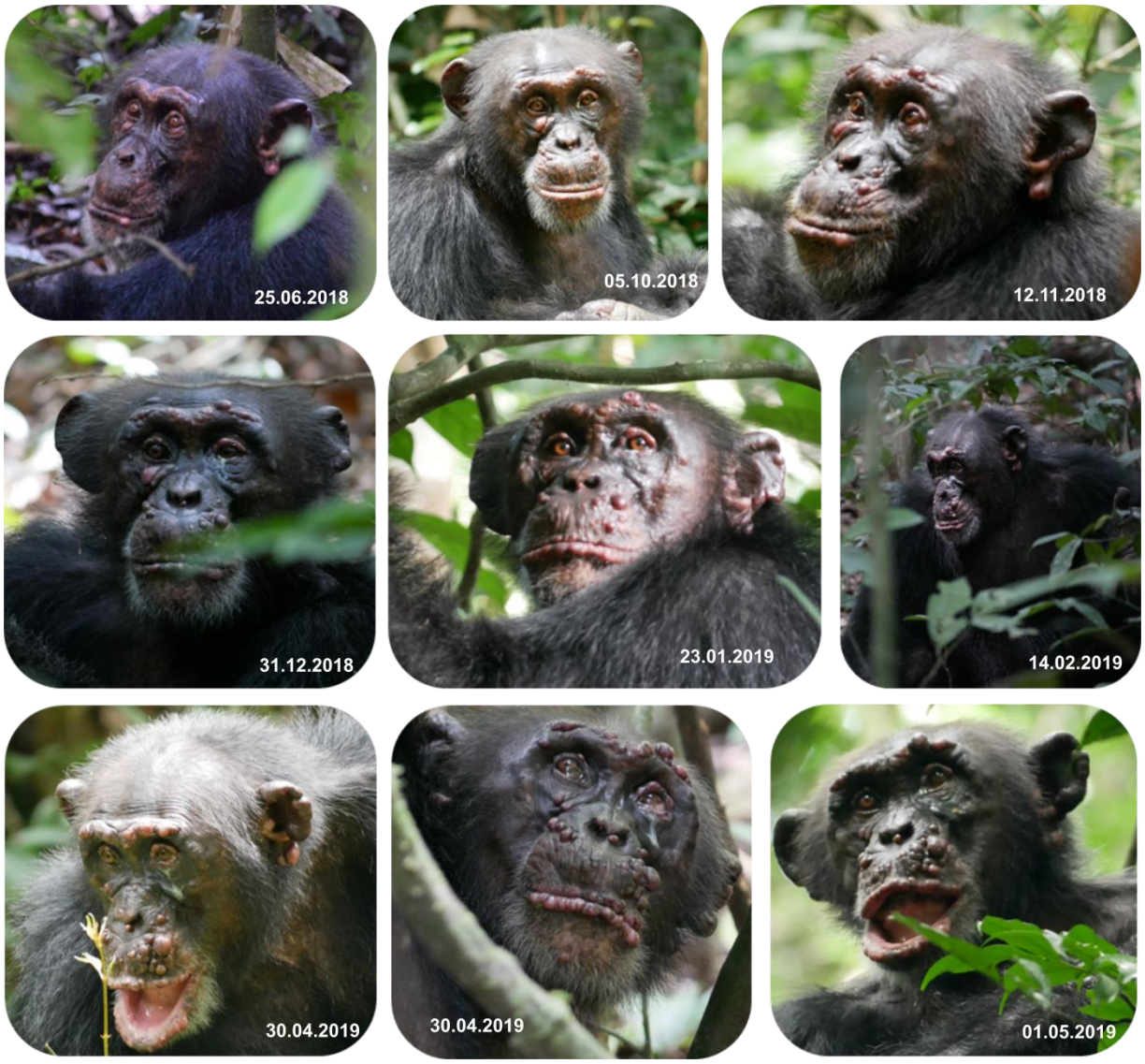


Figure 5: Lesion progression of Woodstock in Täi National Park, Côte d'Ivoire. The hypo-pigmented nodular lesions were increasing with time in number, size and depigmentation. The lesions affected the lips, maxillofacial parts, eyelids with marked ectropion, eyebrows, and ear pinnae. Ocular discharge is observed from the left eye approximately 6 months after onset of lesions. *Photos: Leendertz's lab, RKI, Berlin, Germany*.

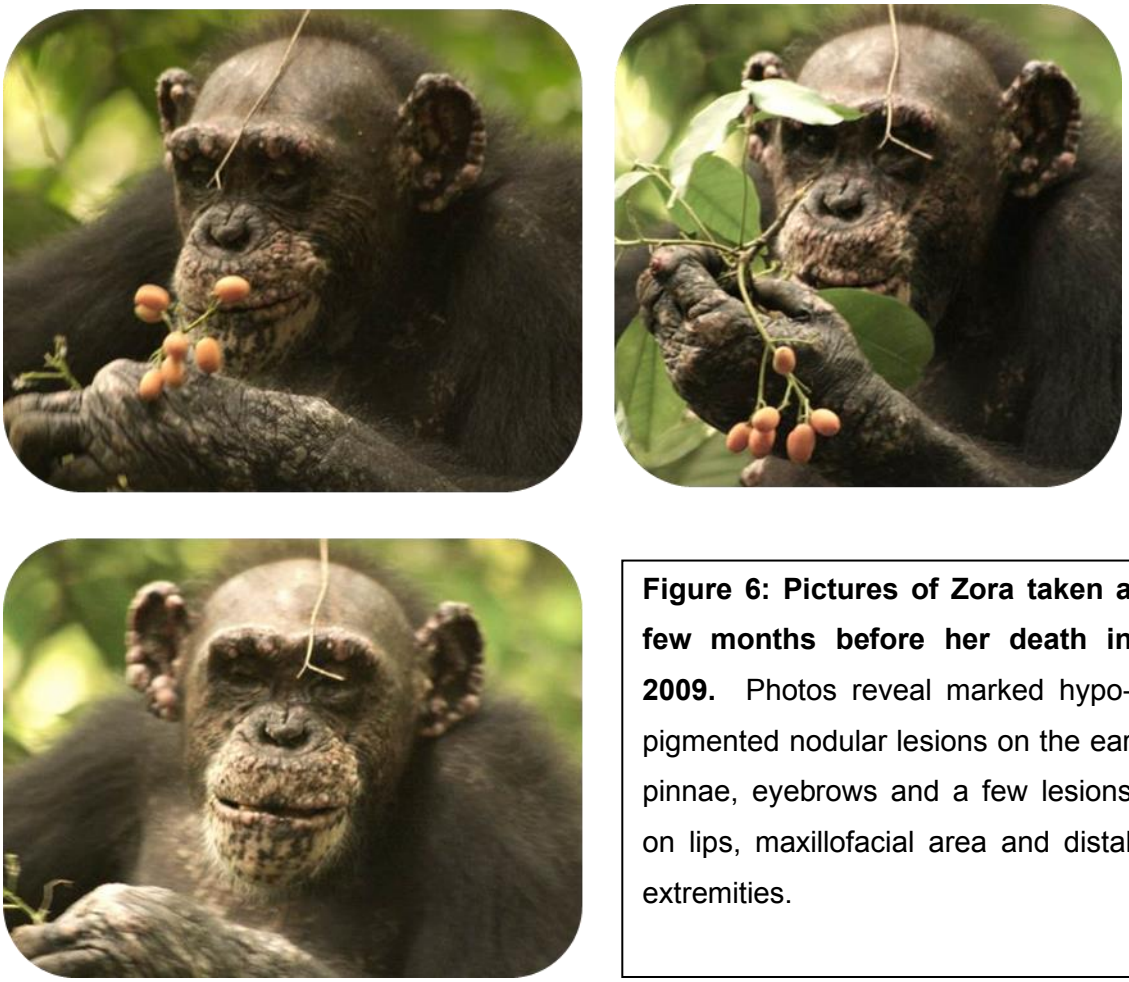


Figure 6: Pictures of Zora taken a few months before her death in 2009. Photos reveal marked hypo-pigmented nodular lesions on the ear pinnae, eyebrows and a few lesions on lips, maxillofacial area and distal extremities.

3.2.3 Screening of *M. leprae* in faecal and necropsy samples

We applied two nested PCR systems using primers described previously for use on tissue samples (Table 3); these systems target two distinct but conserved regions of the *M. leprae* genome [233]. The repetitive element (RLEP) assay targeted the 129bp outer fragment in the primary PCR and a 99bp inner fragment with the nested PCR of the multi copy repetitive element (RLEP) gene. The 18kDA assay targeted the 136bp outer fragment in the primary PCR and a 110bp inner fragment of the 18kDa antigen protein with the nested PCR. To avoid contamination, separate rooms were used for PCR master mix set up, addition of DNA in the primary PCR, and addition of the primary PCR product for the nested PCR.

In both PCR systems, the master mix was comparable for the respective assays, except for the difference in the primer pairs and thermal cycling conditions. The primary PCRs were performed in 20µL reactions: up to 200ng of DNA was amplified using 1.25U of high-fidelity Platinum Taq™ polymerase (Invitrogen®, Germany), 10x PCR buffer (Invitrogen®, Germany), 200µM dNTPs, 4mM MgCl2, 200nM of both forward and reverse primers. The thermal cycling profile for the RLEP primary and nested PCRs was as follows; denaturation at 95°C for 3 min,

followed by 50 cycles of 95°C for 30 sec, 58°C for 30 sec, and 72°C for 1 min. An elongation step at 72°C for 10 min was included. On the other hand, the thermal cycling conditions for the 18kDA assay were set to: 95°C for 3 min, followed by 50 cycles of 95°C for 30 sec, 55°C for 30 sec, and 72°C for 1 min. An elongation step at 72°C for 10 min was included. A 2µL of a 1:20 dilution of the primary PCR product was used as template for the nested assays. Molecular grade water was used as non-template control.

PCR products were visualized on a 1.5% agarose gel stained with GelRed® (ThermoFisher scientific, Germany). Positive bands were purified using the Purelink Gel extraction kit (Invitrogen®, Germany) following the manufacturer`s instruction manual. Both RLEP and 18kDA amplified product are too short for sanger sequencing, so the cleaned amplified products were appended with fusion primers (Primary PCR primers inserted with M13F and M13R primers) (Table 3), through a PCR reaction with conditions as in the primary PCR, but using only 25 cycles. The resulting fused PCR products were then enzymatically cleaned using the ExoSAP-IT™ PCR Product Cleanup assay (ThermoFisher scientific, Germany) and sanger sequenced using M13 primers. Resulting sequences were compared to publicly available sequences in EMBL using BLAST [214]. Both faecal and necropsy positive samples were selected for in-solution hybridisation capture and next generation sequencing

Table 3: Primers used in the leprosy study

PCR system	Primer pair 5´-3´	Product size(bp)
RLEP_Primary PCR	Forward_ TGCATGTCATGGCCTTGAGG Reverse_ CACCGATACCAGCGGCAGAA	129
RLEP _Nested PCR	Forward_ TGAGGTGTCGGCGTGGTC Reverse_ CAGAAATGGTGCAAGGGA	99
Fusion PCR_RLEP	Fusion_M13_Foward_ GTAAAACGACGGCCAGTGAGGTGTCGGCGTGGTC Fusion_M13_Reverse_ CAGGAAACAGCTATGACCAGAAATGGTGCAAGGGA	139
18kDA _Primary PCR	Forward_ TCATAGATGCCTAATCGACTG Reverse_ GGCACATCTGCGGCCAGCA	136
18kDA _Primary PCR	Forward_ ATCGACTGTTGTTTGCGCAAC Reverse_ CCAGCAACCGAAATGTTCGGA	110
Fusion PCR_18kDA	Fusion_M13_Foward_GTAAAACGACGGCCAGATCGACTGTTGTTTGCGCAAC Fusion_M13_Reverse_ CAGGAAACAGCTATGACCCAGCAACCGAAATGTTCGGA	150
Mammal identification PCR	16Smam1_CGGTTGGGGTGACCTCGGA 16Smam2_GCTGTTATCCCTAGGGTAACT 16Smam_Human blocker_CGGTTGGGGCGACCTCGGAGCAGAACCC 16Smam_Pig blocker _CGGTTGGGGTGACCTCGGAGTACAAAAAAC	130
Mammal identification fusion PCR	16Smam1_Illumina_adapter_TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCGGTTGGGGTGACCTCGG 16Smam2_Illumina_adapter_GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCTGTTATCCCTAGGGTAACT 16Smam_Human blocker_CGGTTGGGGCGACCTCGGAGCAGAACCC 16Smam_Pig blocker _CGGTTGGGGTGACCTCGGAGTACAAAAAAC	160

3.2.4 Library preparation, genome wide capture and high throughput sequencing

Selected *M. leprae* positive faecal and necropsy samples were converted into doubled indexed libraries as in section 3.1.5 (using the NEBNext® Ultra™ II DNA Library Prep kit (New England Biolabs®, USA). To reconstruct whole genomes, libraries were target enriched for *M. leprae* using in-solution hybridisation capture and following MYbaits' protocol (Arbor Biosciences, USA) as previously done [24]. Around 1.5µg of each DNA library was captured in singleplex or pooled (Results: Table 7). The enrichment hybridisation capture was done for two rounds of 24 hrs each. Post capture amplification step was performed after each capture round using the KAPA HiFi master mix (KAPA Biosystems, South Africa) and P5 and P7 Illumina primers to generate about 200ng of enriched DNA per sample and/or pool. The post-capture amplification thermal profile was as follows: initial hot start at 98˚C for 2 min followed by a variable number of cycles required to generate at least 200 ng of DNA at 98˚C for 20 sec, 65˚C for 30 sec and 72˚C for 45 sec. Enriched libraries were purified using AMPure beads followed by quantification with the KAPA library quantification kit (KAPA Biosystems, South Africa). Prior to sequencing, libraries were then diluted to 4nM and pooled for sequencing on an Illumina Nextseq 500 (2x150bp v2 chemistry)

3.2.5 Bioinformatics analysis

Raw reads were processed as described elsewhere (3). Briefly, all raw reads were analysed individually using the following method: In parallel, raw reads from the same individuals were also combined in the same FASTQ file prior to analysis. Reads were adapter- and quality-trimmed with Trimmomatic v0.35 [100]. The quality settings were "SLIDINGWINDOW: 5:15 MINLEN: 40". Pre-processed reads were mapped onto the *M. leprae* TN reference genome (GenBank AL450380.1) with Bowtie2 v2.3.4.2 [234]. Reads were filtered out with mapping quality below 8 and duplicate reads were omitted from downstream analyses to avoid false SNP calls. Repetitive regions and ribosomal RNA genes in the reference sequence were omitted. SNP calling was done using VarScan v2.3.9 [235]. To avoid false-positive SNP calls, the following cut offs were applied: minimum overall coverage of five non-duplicated reads, minimum of three non-duplicated reads supporting the SNP, mapping quality score >8, base quality score >15, and an SNP frequency above 80%. All variant call format (VCF) files were analyzed using snpEff v4.3 [236].

Called SNPs of the newly sequenced genomes from chimpanzees (Results: Table 7) were compared to the 234 publicly available Illumina reads and 44 (Appendices: Table S10) unpublished genomes from six countries in West Africa obtained from collaborators to enable

gain insights into the dynamics of the disease in the region. Reads were processed and mapped to the *M. leprae* TN reference as described above (section 3.2.5). The phylogeny was performed using a concatenated SNP alignment. Maximum Parsimony (MP) trees were constructed in MEGA7 [237] from the two new genomes from this study (Results: Table 7) and 234 previously published genomes (Appendices: Table S10) [80,163] and 44 genomes obtained from our collaborators (Appendices: Table S10) using 500 bootstrap replicates and *M. lepromatosis* [161] as outgroup. Sites with missing data were partially deleted (80% coverage cut-off), resulting in 4040 variable sites used for the tree calculation.

3.2.6 Genotyping of insufficiently covered genome by PCR

Two genotyping PCR assays were designed to genotype RLEP faecal samples from TNP_566 that had a poor coverage. In the first PCR, specific SNPs combinations (Appendices: Table S11) were determined from GB_64 sequencing (Results: Table 7) to design two sets of primers that were able to identify if the strain belonged to the branch 4 or other branches. Once identified in the branch 4, additional primers were designed to differentiate between the brach 4 subtypes i.e. 4N, 4O and 4P or the 4N/O subtype (including GB64, Ch1, SM1, Ng13-33, 2188-2007 and 2188-2014) and determine whether the TNP_566 strain belonged to the primate branch.

For the second PCR assay, another set of SNPs combination (Appendices: Table S11) were determined from TNP_418 sequencing (Results: Table 7) and primers were designed to genotype TNP_566 on assumption that the two animals belonged to the same social group and were likely to be infected by the same genotype of *M. leprae*. All PCR conditions were the same as in the leprosy screening PCRs except for the primer sets (Appendices: Table S11) and associated annealing temperatures.

3.2.7 Metabarcoding for dietary analysis

Chimpanzees often feed on other animals, including other primates [192]; DNA of both prey and pathogens infecting these animals could potentially pass through the digestive track of the chimpanzee and be detected in their faeces. To determine whether chimpanzee samples from CNP (N=121) contained mammalian prey DNA that might play a role in the leprosy ecology, a metabarcoding analysis targeting mammalian DNA in these stool samples was implemented. The diet and hunting behaviour of chimpanzees in TNP is already well illustrated and was not done in this study [192,238].

In the first PCR, amplicons were generated using a primer pair targeting the 130 bp fragment of the mitochondrial DNA of the 16S gene (Table 3), in combination with human and pig blocking primers [95,96]. The PCR master mix contained 1.25U of Platinum Taq™ enzyme, x10 PCR buffer, 200µM of dNUTPs, 4mM of $MgCl_2$, 200nM of both forward and reverse

primers, 1000nM µl of each blocking primer (10µM), 0.3U of AmpErase® Uracil N-Glycosylase (UNG) enzyme (Life Technologies, Germany) and was made to a final volume of 25µl with molecular grade water. The thermal cycling conditions were set to: UNG activity at 45°C for 7 min and its deactivation at 95°C for 5 min; DNA denaturation at 95°C for 5 min followed by 42 cycles of 95°C for 30 sec, 64°C for 30 sec, 72°C for 1 min and an elongation step at 72°C for 10 min. Following visualisation on 1.5% agarose gel stained with GelRed®, positive bands were purified using the Purelink Gel extraction kit following the manufacturer`s protocol.

In the second PCR, the cleaned amplicons were appended with fusion primers (16S mammal primers fused with Illumina specific adapter sequence) (Table 3) to convert the amplicons into Illumina libraries. A total of 12.5µL of 1:50 dilution of the cleaned amplicon product was amplified in a PCR reaction mix containing 1.25 U of Platinum Taq™ enzyme (Invitrogen®), x10 PCR buffer, 200µM of dNUTPs, 4mM of $MgCl_2$, 200nM of both forward and reverse primers, 10µM of each blocking primer. The reaction volume was brought to 25µL. The thermal cycling profile was as follows; 95°C for 5 min followed by 15 cycles of 95°C for 30 sec, 64°C for 30 sec, 72°C for 1 min and an elongation step at 72°C for 10 min. This was followed by cleaning of the amplicon product using the MagSi-NGS PREP Plus bead cleaning system (AMSBIO, UK) and following the manufacturer's protocol. Thereafter, 5µL of each cleaned amplicon product appended with Illumina specific adapters was dual indexed in a 50µL PCR reaction containing 5µL of each Nextera Index Primer (Illumina) and 25µL of 2X KAPA HiFi HotStart ReadyMix (KAPA Biosystems, SA). The thermal cycling conditions were set to; 95°C for 3 min followed by 8 cycles of 95°C for 30 sec, 55°C for 30 sec, 72°C for 30 sec and an elongation step at 72°C for 5 min. Finally, the dual indexed amplicon libraries were cleaned with a bead system as in the prior step and quantified using the TECAN reader spectophometry assay (TECAN).

The quantified libraries were then normalised to equimolar amounts and pooled for sequencing on the Illumina NextSeq (2x150-cycles, v2 chemistry).

Sequenced paired-end reads were joined using the *illuminapairedend* function in the software package OBITools (v 1.1.22), with a minimum alignment score of 40 [239]. Read pairs that were not merged were removed from subsequent analysis. Primer sequences were removed from surviving read pairs using the ngsfilter function in OBITools [239]. Reads were then quality-trimmed using Trimmomatic (v0.36), using a minimum average quality score of 30 over a sliding window of four bases and a minimum length of 80bp [100]. Surviving reads were de-replicated using the OBITools function obiuniq [239], and to avoid contamination we removed all sequences for which less than ten copies were detected in a sample.

A reference database was built by downloading the NCBI nucleotide collection (nt_v5: downloaded on August 23rd, 2019) and selecting mammalian sequences based on their taxid using the blastdbcmd command [240]. Mammalian sequences were converted to an EcoPCR database using the OBITools obiconvert function, using the NCBI taxonomy. An *in silico* PCR was performed using the OBITools function EcoPCR, allowing for up to three mismatches between the primers and reference sequences and a maximum amplicon length of 800bp [239]. Taxonomic assignment to this database was performed using the OBITools ecotag function, that uses the Needleman–Wunsch algorithm to map each sequence against the ecoPCR database (minimum sequence identity of 95%), generating a taxonomic assignment based on the last common ancestor of sequences with a similar mapping score [239,241]. We considered taxonomic assignments at the species and genus levels.

3.2.8 Histopathology

To further confirm the infection, organ samples of Zora that were submitted to the Germany Primate Center in Göttingen, Germany for pathological investigations earlier in 2010 were reanalyzed for leprosy. The pathology report revealing the findings of pathological investigations is attached as an **appendix 10.13/Pathology report for Zora**.

3.2.9 Detection of PGL-1 antibodies

To further confirm the disease at TNP, heart blood from Zora collected at necropsy in 2009 was evaluated for, the presence of antibodies against the phenolic glycolipid-I (PGL-I) of *M. leprae* using the rapid lateral flow test [242]. The PGL-I is a diagnostic marker that is known to be highly specific for *M. leprae* [243]. Briefly, whole blood was diluted in a ratio of 1:10 with phosphate buffered saline and 60µl of the resulting dilution was loaded on the lateral flow test strip. *M. leprae* positive lyophilized human serum was used as a positive control. The results of both the control and test lanes for both tests were evaluated after 3 minutes from loading the samples.

4 Results

4.1 Treponema pallidum infections in NHPs in TNP and other field sites

4.1.1 Screen PCR and whole genome capture

All symptomatic animals sampled with biopsies or swabs tested positive for *Treponema pallidum* in at least one of the sample types collected (Table 4). Sequences generated for the respective assays were all identical (a representative sequence was uploaded to Zenodo: doi.org/10.5281/zenodo.3540499). In addition, based on the *polA* gene qPCR assay, we detected TP DNA in NHP bones in the following field sites and species/subspecies: Boe in Guinea Bissau (*Cercocebus atys*), TNP in Côte d'Ivoire (*Cercopithecus diana, Pan troglodytes verus* and *Piliocolobus badius*), East Nimba and through the Nationwide survey in Liberia (*Pan troglodytes verus*), Bili-Uere in the Democratic Republic of Congo (*Pan troglodytes schweinfurthii),* and Budongo in Uganda (*Cercopithecus mitis, Colobus guereza*; Appendices Table S1).

Table 4: Lesion and sample types observed at TNP and PCR screening results for each sample

Animal ID	Species	Lesions observed	Sample type	*PolA PCR*	*cfpA PCR*
5847	*Cercocebus atys*	Anogenital lesions	biopsy	Negative	Negative
		Anogenital lesions	swab	Positive	Positive
		Facial lesions	biopsy	Negative	Negative
		Body ventral lesions	biopsy	Negative	Negative
2116	*Cercocebus atys*	Genital lesions	swab	Positive	Positive
2117	*Cercocebus atys*	Genital lesions	swab	Positive	Positive
1864	*Cercocebus atys*	Orofacial lesions	biopsy	Positive	Positive

From the *Cercocebus atys* study group in TNP, we were able to sequence three new TPE genomes from biopsy and swab samples with 5X coverage of 47.5% - 97.5% of the genome (Table 5). We previously sequenced a partial genome from one *Cercocebus atys* with facial lesions in this study group (1864), [19]; here, we improved the 5X genome coverage from 82.39% to 95.5%. Further, we recovered partial TPE genomes from a *Piliocolobus badius* (Tai 105) bone from TNP, Côte d'Ivoire, as well as one *Cercocebus atys* (Boe 092) bone from Boe in Guinea Bissau with 76% and 69.7% at 1X genome coverage or 26.9% and 21.7% at 5X genome coverage respectively. The control library (Chicken DNA) had only a single read surviving the read quality control processing, translating to 0.005% 1X genome coverage. All genomes were distinct, including those obtained from individuals belonging to a single social

group of *Cercocebus atys* monkeys. None of the genomes generated in this study had the A2058G and A2059G mutations in the 23S ribosomal RNA gene, which have been demonstrated to confer antimicrobial resistance to macrolides antibiotics [244].

Table 5: Mapping results for non-human primate *Treponema pallidum* subsp. *pertenue* strains generated in this study.

Sample ID	NHPs species	Sample type	Deduplicated reads mapped	Pos. covered 1X	Pos. covered 5X	% genome 1X	% genome 5X
Boe 092	*Cercocebus atys*	bone	36407	793392	247484	69.57	21.7
2117	*Cercocebus atys*	swab-genital lesion	44143	1009366	541728	88.50	47.5
5847	*Cercocebus atys*	swab-genital lesion	457063	1117566	1111969	97.99	97.5
2116	*Cercocebus atys*	swab-genital lesion	131550	1119023	1066350	98.12	93.5
1864	*Cercocebus atys*	biopsy-face lesion	202416	1116210	1089159	97.87	95.5
Control	chicken	tissue	1	52	0	0.005	0
Tai 105	*Piliocolobus badius*	bone	42719	867318	448209	76.05	39.3
22_52	*Piliocolobus badius*	bone	43373	711608	306789	62.40	26.9

4.1.2 Phylogenetic analyses

Phylogenetic analysis of reconstructed genomes from this study and all other TPE and TEN genomes and a selection of TPA genomes in GenBank **(**Appendices Table S2) yielded tree topologies largely consistent in both the PhyML-SMS and BEAST based approaches. Overall, the ML and MCC tree topologies resolved into distinct reciprocally monophyletic groups representative of the TP subspecies (TPA, TPE, and TEN: Figure 7 and Appendices: supplementary trees). The TPE clade included both human and all NHP infecting strains, while TPA and TEN clades consisted only of human isolates (Figure 7 and Appendices: supplementary trees). Read placement of bone derived TP reads from one *Piliocolobus badius* specimen (22_52) and two *Pan troglodytes verus* specimens (11786 and 15028) from TNP, resulted in the majority of reads falling within the TPE clade, and more particularly within the TNP clade (Figure 9).

While these analyses clearly showed that all NHP derived strains were TPE, the relationships within the TPE clade were poorly resolved, potentially due to the extremely long branches separating these clades. Such long branches have the potential to complicate phylogenetic analyses [245], so to help resolve the uncertainty of the relationship between the different TPE strains, the analysis was then focused on the TPE clade and ingroup phylogenetic analysis was implemented.

The ingroup analysis of TPE diversity revealed that strains isolated from TNP clustered together in ML and BEAST analyses under a relaxed clock model (Figure 8 and Appendices: supplementary trees). However, statistical support was relatively weak, and in BEAST analyses under a strict clock model this monophyly was no longer observed (Table 6and Appendices: supplementary trees). Furthermore, no signal for monophyly for TPE strains based on clinical manifestations was observed in the current study; in that TNP strains causing different clinical manifestations in *Cercocebus atys* monkeys and simian strains in general, did not form statistically supported, reciprocally monophyletic groups based on clinical manifestation (Figures 7, 8 and Table 6). Rather, simian strains formed reciprocally monophyletic groups based on geography (i.e. strains from the same location clustered together). These geographic clusters were supported in both ML and BEAST analyses under a relaxed clock model (Figure 7,8 and Appendices: supplementary trees). These strains had an average patristic distance of 0.0300%, compared to 0.0017% and 0.0000% for the strains infecting *Chlorocebus sabaeus* monkeys in Senegal and the Gambia respectively (Appendices: Table S3).

Considering larger spatial scales, we observed that the *Cercocebus atys* bone derived TPE (Boe_092) collected in Guinea Bissau and the Fribourg-Blanc strain isolated from a baboon in Guinea clustered within the clade of TPE strains infecting *Chlorocebus sabaeus* monkeys from the neighbouring Senegal and Gambia (Figure 8). All analyses and models retrieved this monophyly, in most cases with strong statistical support (Table 6).

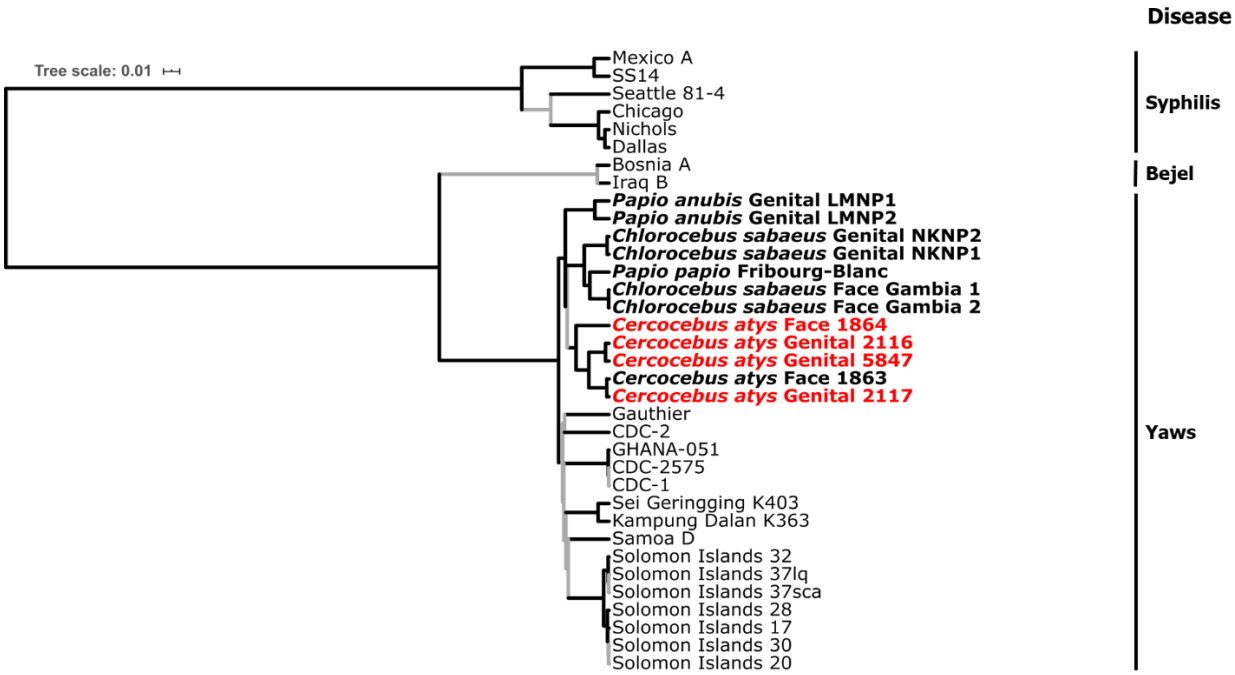


Figure 7: Maximum clade credibility tree of *Treponema pallidum* strains (relaxed clock model assuming a birth-death process). All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red with a minimum coverage of 10X to call a base, and a threshold of 95% identity for a base to be called. Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree are indicated in gray. The scale shows nucleotide substitutions per variable site. Raw sequence read files for all new genomes appearing in this tree were archived in NCBI under BioProject PRJNA588802

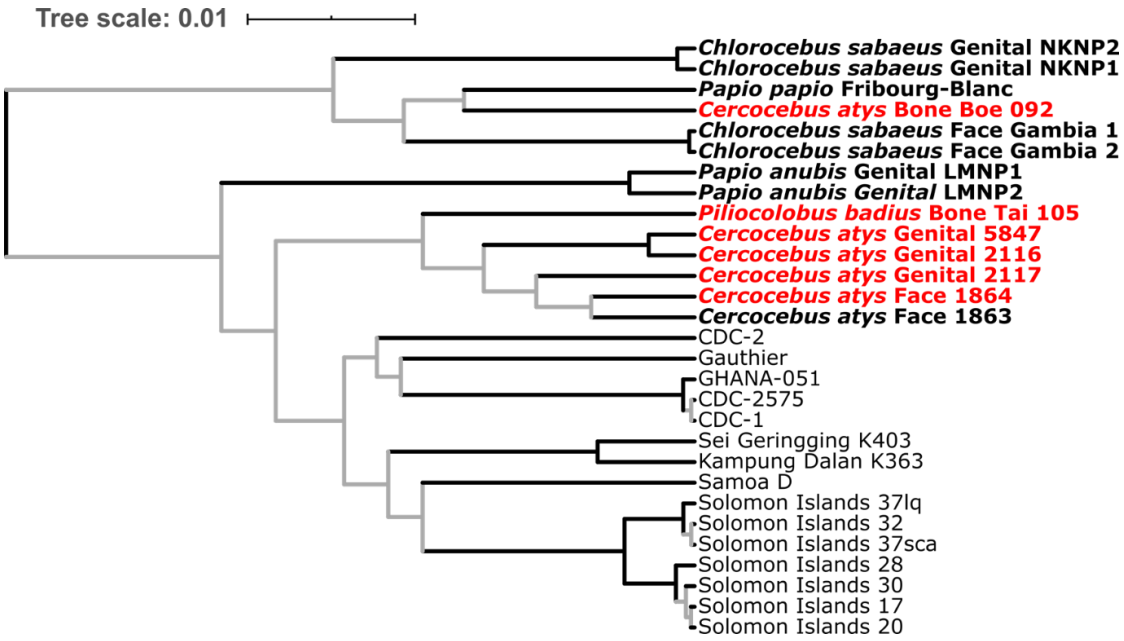


Figure 8: Maximum clade credibility tree from ingroup analysis of TPE strains (relaxed clock model assuming a birth-death process). All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red with a minimum coverage of 5X to call a base, and a threshold of 95% identity for a base to be called. This analysis included two partial low coverage genomes derived from bone samples (Boe 092 and Tai-105). Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree are indicated in gray. The scale shows nucleotide substitutions per variable site.

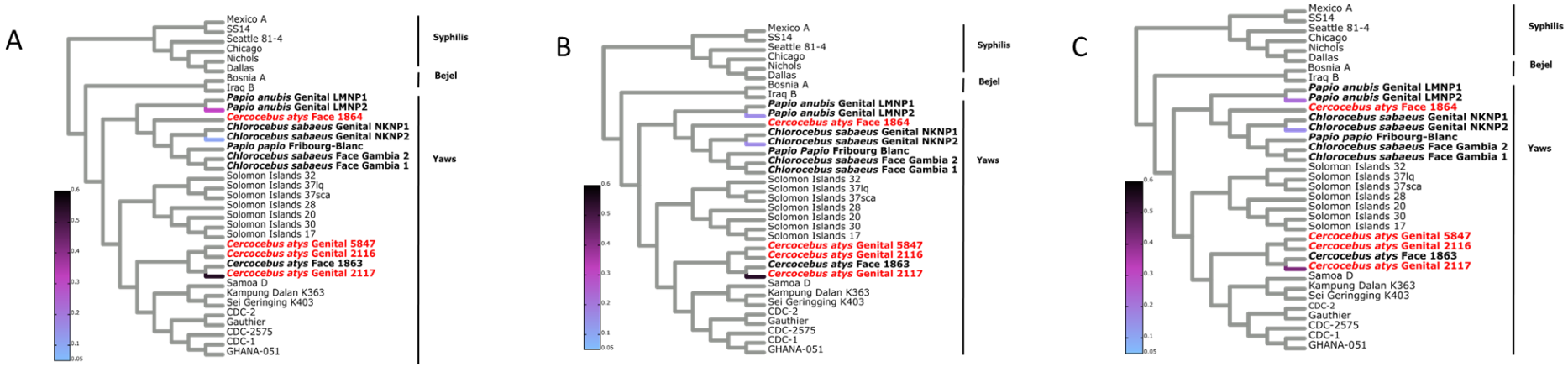


Figure 9: Phylogenetic placement of bone samples. Heat-trees visualization of phylogenetic placement of TPE mapped reads from bone samples on to the TP MCC (10X overage and 95% threshold) reference tree using the evolutionary placement algorithm (epa-ng). The approximate percentage of reads placed on to a particular branch of the *Treponema pallidum* cladogram is shown as a linearly scaled color density. Genomes generated in this study are shown in red. **A** Sample 11786 (Total number of reads = 517), **B** Sample 15028 (Total number of reads = 3,581), **C** Sample 22_52 (Total number of reads = 19,389). Raw sequence read files for all all bone samples placed on this tree were archived in NCBI under BioProject PRJNA588802

Table 6: A summary table of phylogenetic analyses implemented in this study showing the results of the monophyletic groups of simian TPE strains from West Africa that form within the TPE clade based on the different tree probabilistic methods, molecular clocks and priors. The results discussed in this study are based on the phylogenetic trees from analyses highlighted in gray.

Analysis					New strains belong to TPE	Senegal/Gambia/Guinea Bissau clade	Tai National Park clade	Monophyletic groups based on clinical manifestation
Dataset	**Analysis including TPA and TEN strains**	**Probabilistic method**	**Clock**	**Tree prior**				
X5_95	Yes	ML			✓	✓	**X**	**X**
		BMCMC	Strict	Coalescent	✓	✓	**X**	**X**
		BMCMC	Strict	Speciation	✓	✓	**X**	**X**
		BMCMC	Relaxed	Coalescent	✓	✓	✓ **(weak)**	**X**
		BMCMC	Relaxed	Speciation	✓	✓	✓ **(weak)**	**X**
X10_95	Yes	ML			✓	✓	**X**	**X**
		BMCMC	Strict	Coalescent	✓	✓	✓ **(weak)**	**X**
		BMCMC	Strict	Speciation	✓	✓	✓ **(weak)**	**X**
		BMCMC	Relaxed	Coalescent	✓	✓	✓ **(weak)**	**X**
		BMCMC	Relaxed	Speciation	✓	✓	✓	**X**
X5_95	No	ML			NA	✓	✓ **(weak)**	**X**
		BMCMC	Strict	Coalescent	NA	✓	**X**	**X**
		BMCMC	Strict	Speciation	NA	✓	**X**	**X**
		BMCMC	Relaxed	Coalescent	NA	✓ **(weak)**	✓ **(weak)**	**X**
		BMCMC	Relaxed	Speciation	NA	✓ **(weak)**	✓ **(weak)**	**X**
X10_95	No	ML			NA	✓	✓	**X**
		BMCMC	Strict	Coalescent	NA	✓	✓ **(weak)**	**X**
		BMCMC	Strict	Speciation	NA	✓	✓ **(weak)**	**X**
		BMCMC	Relaxed	Coalescent	NA	✓	✓	**X**
		BMCMC	Relaxed	Speciation	NA	✓	✓	**X**

4.2 Leprosy disease among chimpanzees in CNP and TNP

4.2.1 Evaluation of clinical manifestations

From the evaluation of camera trap images of chimpanzees taken on 9th March, 2017 at CNP sites (Figure 4), a few individuals showed evidence of leprosy disease. Typical gross lesions observed included: multiple reddish to copper coloured nodular (hypo-pigmented) lesions on the maxilla-facial areas, distal extremities of the fore limbs, lips, eyebrows and ear pinnae with no evidence of the nose outline in some individuals.

At TNP, Woodstock initially presented with a few copper-coloured nodular (hypo-pigmented) lesions on the lips, eye orbits, eyelids, ear pinnae. However, from photo images taken between June, 2018 and May, 2019, the lesions were observed to be progressive in size, depigmentation and in number. By the 31st December, 2018, a mucopurulent discharge from the left eye, severe madarosis and ectropion were apparent, suggesting ocular complications in leprosy infection [246] (Figure 5). By the end of period covered in the scope of this evaluation on 1st May, 2019, the veterinarian on site observed neither gait abnormalities nor neurological deficiencies.

Based on literature review [15,18,175,187–189] and evaluation of lesions using camera trap images, still photos and videos, a tentative diagnosis of lepromatous leprosy was reached clinically in both chimpanzee groups.

4.2.2 Screening of *M. leprae* in faecal, urine, fruit wedges and necropsy samples

Of the 121 faecal samples analyzed from CNP, one faecal sample was positive in both the RLEP and 18kDA assays, while additional two were only positive with the RLEP assay (Appendices: Table S9). At TNP, *M. leprae* DNA was first detected in the faecal samples of Woodstock when he started manifesting facial skin lesions in June, 2018. Of the 11 samples collected from him between June 2018 and January 2019, 10 were positive in both assays (Appendices: Table S4). Retrospective screening of faecal samples from this individual back to 2014 yielded only negative results (Appendices: Table S5). Furthermore, only a single urine sample was positive on RLEP assay and the two fruit wedges were negative.

From the screening of necropsy samples collected since 2001 at TNP (Appendices: Table S8), *M. leprae* DNA was detected in the spleen sample of an individual identified as Zora (TNP_418). Zora died in 2009 of tuberculosis complications and *M. tuberculosis* was isolated [247]. On detection of *M. leprae* DNA in this spleen sample, all other available necropsy organs of Zora were also tested for *M. leprae* to determine the extent of systemic infection. In addition to the spleen sample, *M. leprae* DNA was detected in the following organs in both assays:

lung, liver and the mesenteric lymph node while the kidney and heart were only positive in the RLEP assay (Appendices: Tables S7). To gain insights on the earliest time the infection could be detected in Zora, a retrospective screening of all available faecal samples (N=38) of this individual from 2009 the year the individual died going back to 2001 (Appendices: Table S6) was performed. From this longitudinal screening, the first detection of *M. leprae* DNA was observed in 2002 in the RLEP assay, and faeces from this animal were sporadically positive in this assay until June 2006 when both screening assays yielded consistently positive results (12 of 14 samples) up to the time the individual died in 2009 (Appendices: Table S6; Figure 9). To obtain deeper insights on these results and also to be able to characterise the pathogen, the lung, liver, and spleen samples from Zora were also included in the hybridisation capture for *M. leprae.*

Reviewing archived photos of Zora, it was observed that photos taken a few months prior her death, had pathognomonic gross symptoms of leprosy. These symptoms included: hypo-pigmented nodular lesions on the ear pinnae, eyebrows and a few lesions on lips, maxillofacial area and distal extremities (Figure 6). Furthermore, to ascertain whether other individuals in this social group were infected at the time of Zora's death, a cross-sectional screening of contact animals in this social group was performed by testing all available faecal samples collected between 2009 and 2010. All other animals were negative during this period including Woodstock.

4.2.3 Genome wide capture and genotyping analyses

From faecal samples collected from the chimpanzee community at CNP, 99.9% of the *M. leprae* genome was recovered from sample GB_64 with mean coverage of 28.3-fold while while samples GB_26 and GB_68 yielded only about 1% of the genome with mean coverage of 0.08- and 0.09-fold respectively. Combined libraries from TNP_566 (Woodstock) yielded 23.5 % of the genome with a mean coverage of 0.9-fold while a total of 99.7 % of the *M. leprae* genome was recovered from TNP_418 (Zora).

TNP_566 was negative for the SNPs tragetting branch 4 suggetin that the animal was infected with a different strain other than branch 4 strains that are common to West Africa (Table 6). However, SNPs that were deciphered from the genome-wide comparison of TNP_418 and manually checked and visualized in TNP_566 using IGV, showed that of the four, two SNPs were covered by available merged high throughput sequencing data of TNP_566. the other two SNPs that were checked by genotyping PCR revealed that the two animals had identical SNPs suggesting that both were infected by the same strain on *M. leprae* (Table 6).

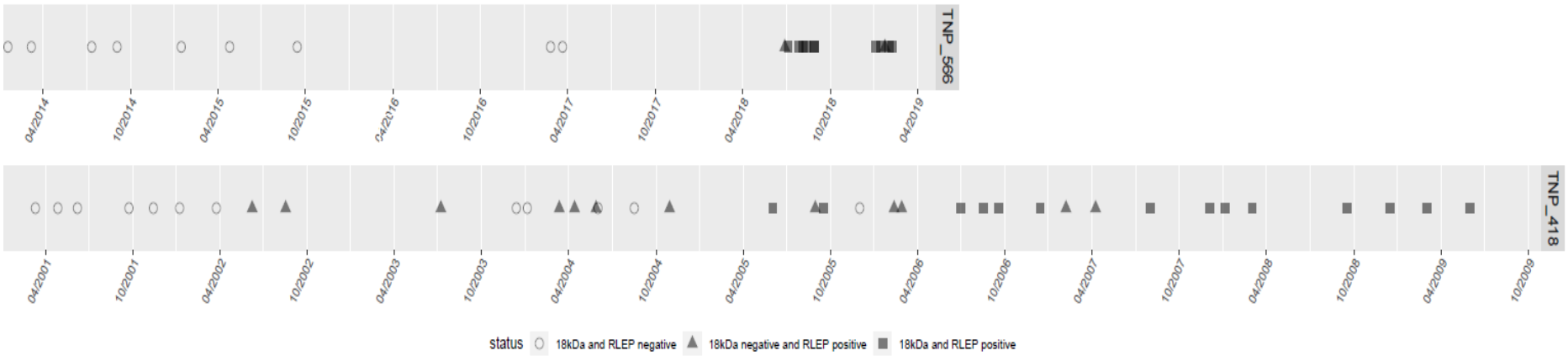


Figure 10: Plots of faecal screening for *M. leprae* results for Woodstock and Zora. Samples are plotted according to the day they were collected

Table 7: Genotyping results for the for the Animal TNP_566

Targeted gene	Genome position	Amplicon size (bp)	Expected mutation	Observed mutation	Interpretation
ML0283	365373	201	T>C	T	T: another branch; C: branch 4
ML2356	2815502	202	G>C	G	G: another branch: C: branch 4
ML0411	508986	228	C>G	G	If branch 4; G: genotype 4N/O; A: genotype 4N, 40 or 4P
ML0048	60123	169	C>T	T	T: genotype 2F; C: other genotype
ML0565	683097	194	C>T	T	T: genotype 2F; C: other genotype

Table 8: Mapping results for chimpanzee derived *M. leprae* genomes

Animal/faecal ID	Library_ID	Raw reads	Processed reads	Mapped reads	deduplicated reads	Mean fold coverage	Genome coverage %
Woodstock (TNP_566)	CIVW05-07		12568086			0.1	
	CIVW05-08		66065153			0.3	
	CIVW11-08		15171685			0.1	
	CIVW28-07		14434880			0.1	
	CIVW28-08		46191246			0.2	
	CIVW29-08		16687734			0.1	
	CIVW_all		171118784			0.6	
	TNP566		4902921			1.12	
	CIVW_TNP566_merged	**4980562**	**4897997**	**821098**	**25971**	**0.9**	**23.5**
Zora (TNP_418)	TNP418_liver		1003882			0.5	
	TNP418_spleen1		12954817			12.6	
	TNP418_spleen2		13994119			13.4	
	TNP418_lung		779316			0.3	
	TNP418_merged	**29332537**	**28732134**	**25267818**	**294922**	**25.8**	**99.7**
GB_64	GB64_1		47805995			6.3	
	GB64_2		101399379			31.6	
	GB64_1+2_merged	**102844426**	**101333236**	**3025429**	**595745**	**28.3**	**99.9**
GB_26	GB26	26929407	26829543	312104	3373	0.08	1.06
GB_68	GB68	30061283	29922965	307517	4082	0.09	1.31

4.2.4 Comparative genomics of *M. leprae* strains from chimpanzees

Phylogenomic analyses inferred using the maximum parsimony approach resulted in a tree topology that resolved into genotype defined clades consistent with *M. leprae* phylogenies previously described [80,159,163]. The genome sequenced from the chimpanzee faecal sample GB_64 (CNP, Guinea Bissau) clustered into branch 4 (genotype 4) and specifically clustered within the recently described subgenotype 4N/O consisting of two other NHP derived genomes from NHPs obtained from West Africa (Sierra Leone and Nigeria) but manifested leprosy symptoms in captivity abroad [16,24,188]. Subgenotype 4N/O also consists of a human strain from Niger (Ng13-33) and two human strains from Brazil (2188-2007 and 2188-2014) isolated from the same patient of a relapse case of leprosy (Figure 10 and 11) [24,163]. Among humans, this subgnotype seems to be a rare one as it has only been isolated from a single human out of the 83 human cases sequenced from West Africa and two separate isolations of the same human relapse case out of the 34 cases from Brazil that group with genotype 4 strains. Furthermore, *M. leprae* strains within 4N/O did not form a monophyletic group, hence relationships among themselves were unresolved and it was difficult to determine whether GB_64 was closely related to previously sequenced NHP or human derived *M. leprae* genomes in the 4N/O clade.

Surprisingly and interestingly, comparative genomics revealed that the *M. leprae* genome isolated from the spleen necropsy sample of TNP_418 (Zora), clustered within subgenotype 2F clade consisting of strains that circulated from the medieval time in Europe and modern human Ethiopian strains isolated in 2015 [80]. The genome TNP_418 was closely related to Jorgen_749, an *M. leprae* genome isolated from ancient sample in Denmark of a strain estimated to have circulated between 1223- 1279 [80] (Figure 10, 11). This subgenotype shared a last MRCA with subgenotype 2E and all subgenotypes within genotype 1 (A, B, D). Until now, 2F subgenotype was never reported in West Africa and was only recently isolated from two patients in Ethiopia in 2015 in East Africa. The only single human derived *M. leprae* genome sequenced so far from Côte d'Ivoire clusteres within branch 4 like all other West African human *M. leprae* isolates. From the foregoing, it is clear that the 2F subgenotype is presently uncommon in the human population and equally its prevalence in the wild is unknown as no surveillance for it exists. In addition, the genotyping PCR results tying both TNP_566 and TNP_418 together suggests that only the 2F subgenotype circulates at TNP.

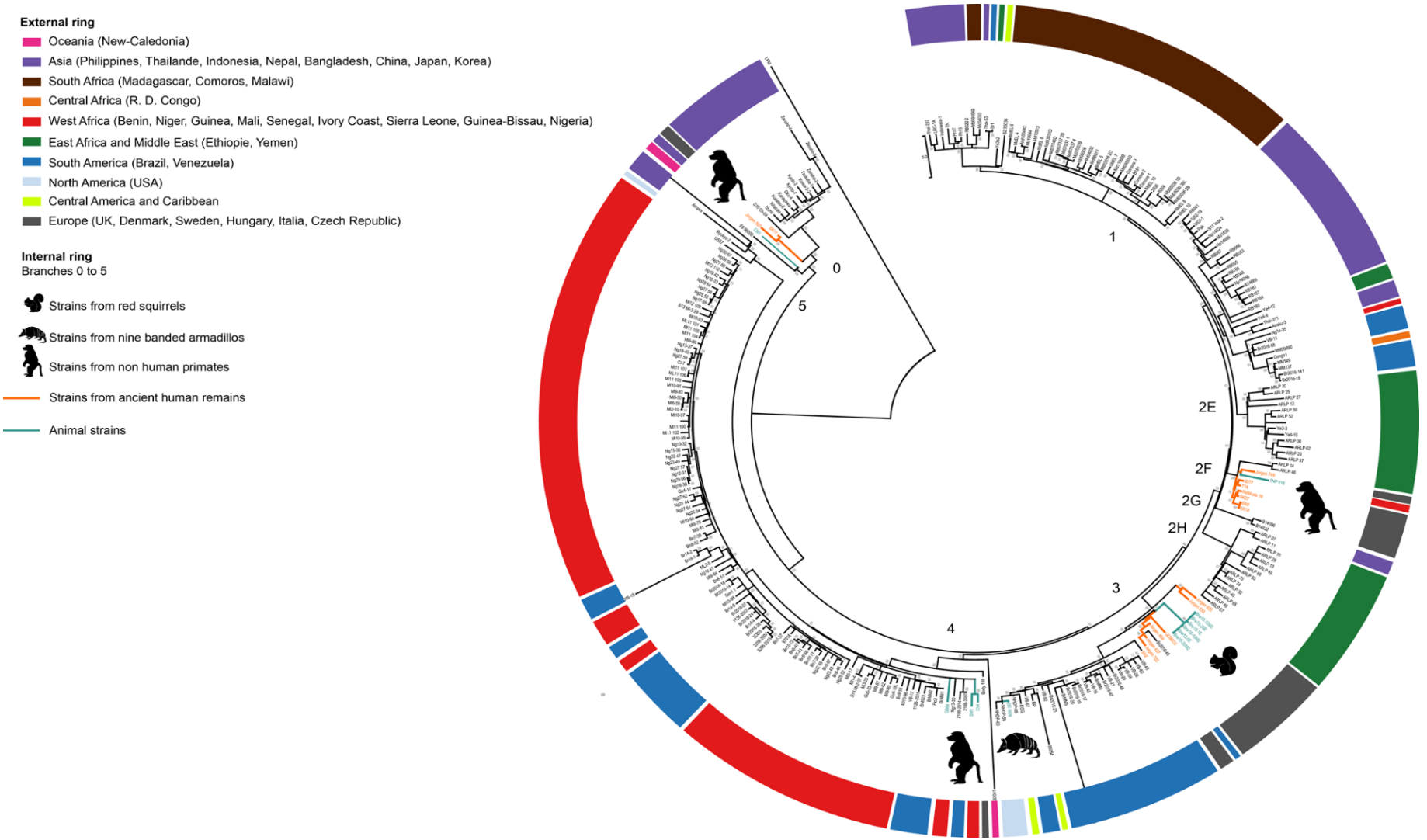
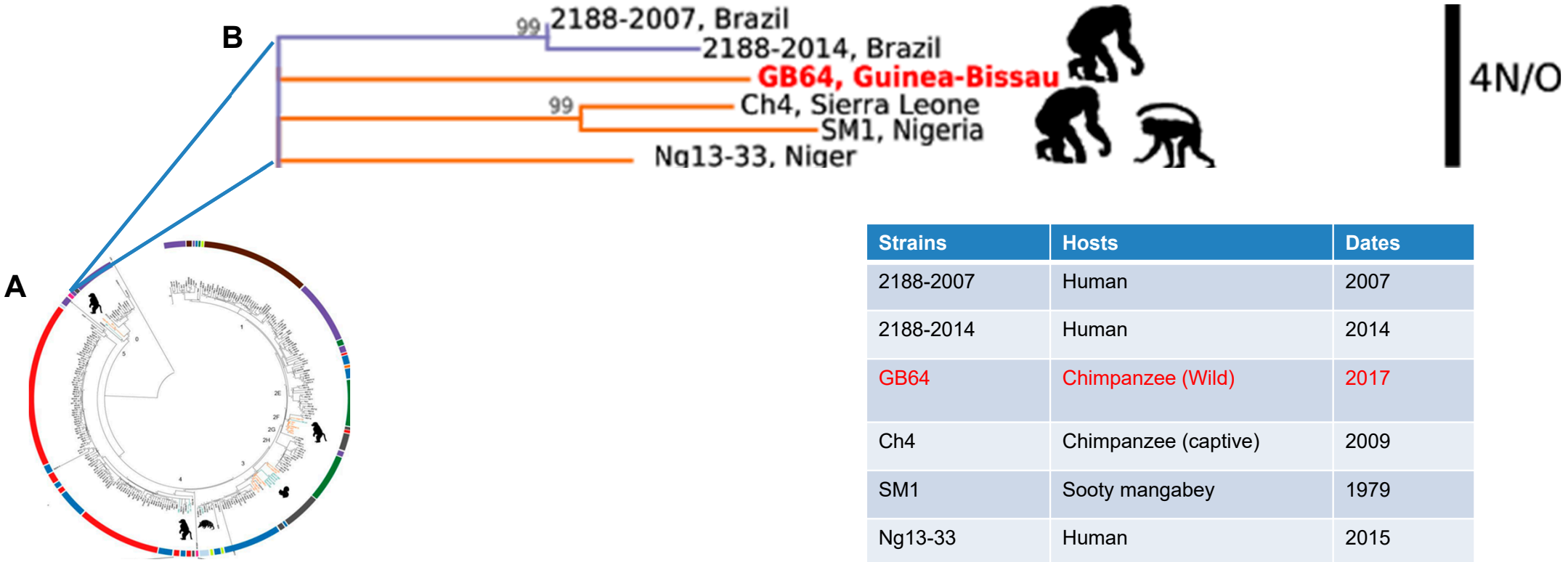


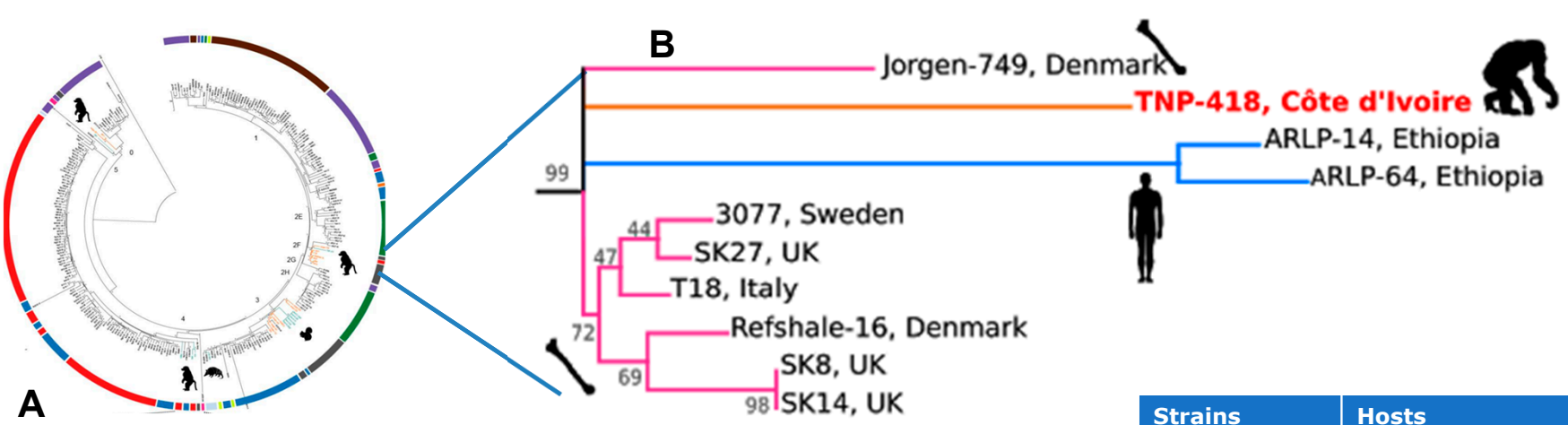
Figure 11: Phylogenetic relationship of newly sequenced *M. leprae* strains from chimpanzees and humans from West Africa: Maximum parsimony tree, including 280 *M. leprae* genomes was constructed using 500 bootstraps replicates and *M. lepromatosis* as outgroup; the scale is represented by single nucleotide polymorphism. Branch numbers from 0 to 5 based on the previously define branching system (Schueneman et al., 2018) are included in the inner ring. The geographical origin of each strain is represented by colors, as indicated in the top left caption. Strains from ancient human remains and animal are highlighted in orange and blue respectively.



Strains	Hosts	Dates
2188-2007	Human	2007
2188-2014	Human	2014
GB64	Chimpanzee (Wild)	2017
Ch4	Chimpanzee (captive)	2009
SM1	Sooty mangabey	1979
Ng13-33	Human	2015

- **Figure 12:** Phylogenetic relationship of newly sequenced *M. leprae* strains from chimpanzees and humans from West Africa. **A:** Maximum parsimony tree, including 280 *M. leprae* genomes. The tree was constructed using 500 bootstraps replicates and *M. lepromatosis* as outgroup; the scale is represented by single nucleotide polymorphism. Sites with missing data were partially deleted (80% coverage cut off), resulting in 4379 variable sites used for the tree calculation. **B:** Zoom into the Branch 4, genotype 4N/O (Stefanie *et al.,* 2017) containing the newly sequenced genome GB_64; strains isolated from west Africa are represented by brown branches and those from Brazil are represented by purple branches; The table shows dates when the strain isolated from the host.

- (1) Schueneman *et al*, 2015; (2) Schueneman *et al*, 2018, (3) Honap *et al,* 2018 (4) Stefanie *et al*, 2018 (5) Benjak *et al*, 2018 (6) Mendum *et al,* 2014



Strains	Hosts	Dates
Jorgen-749	Human	1223-1279
TNP-418	Chimpanzee (wild)	2009
ARLP-14	Human	2015
ARLP-64	Human	2015
3077	Human	1032-1155
SK27	Human	1000-1200
T18	Human	600-700
Refshale-16	Human	1046-1163
SK8	Human	10101160
SK14	Human	1000-1200

- **Figure 12:** Phylogenetic relationship of newly sequenced *M. leprae* strains from chimpanzees and humans from West Africa. **A:** Maximum parsimony tree, including 280 *M. leprae* genomes. The tree was constructed using 500 bootstraps replicates and *M. lepromatosis* as outgroup; the scale is represented by single nucleotide polymorphism. Sites with missing data were partially deleted (80% coverage cut off), resulting in 4379 variable sites used for the tree calculation **B:** Zoom into the Branch 2F containing the newly sequenced genome TNP_418; The tables shows dates when the strain was isolated from the host. The strains isolated from UK are represented by pink branches and those from Ethiopia are represented by blue branches while the strain isolated from west Africa are represented by brown branches.

- (1) Schueneman *et al*, 2015; (2) Schueneman *et al*, 2018, (3) Honap *et al,* 2018 (4) Stefanie *et al*, 2018 (5) Benjak *et al*, 2018 (6) Mendum *et al,* 2014

4.2.5 Dietary analysis for mammalian DNA

Of the 121 chimpanzee faecal samples from CNP in Guinea Bissau analysed for mammalian DNA, three samples were positive. However, none of the mammalian DNA positive faecal samples were positive for *M. leprae*. Further, only a single faecal sample was positive for each of the following wildlife species; diana monkeys (*Cercopithecus diana*), spotted forest genets (*Genetta pardina*), and red river hogs (*Potamochoerus porcus*).

4.2.6 Rapid lateral flow test for detection of PGL-1 antibodies

Test bands were observed on both the test and control lanes for both the positive control and blood from Zora confirming the presence of the immunoglobulin M antibodies to PGL-I of *M. leprae.* This offers further evidence of leprosy infection in Zora and evidence that lateral leprosy tests designed for screening humans can be applied for surveillance of leprosy in NHPs especially in captive settings were collection of bllod is feasible. (Figure 12).

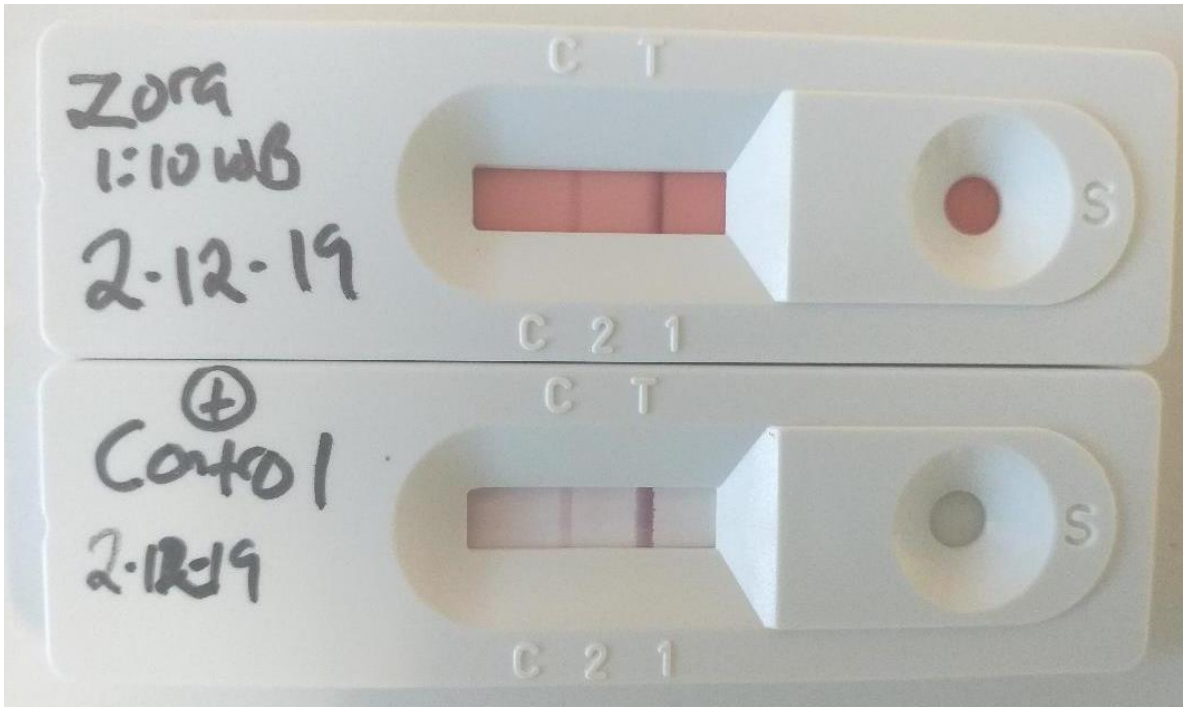


Figure 13: Results of the rapid leprosy lateral flow tests for both Zora and the positive control. For the test to be valid, both the C (control lane) and the T (test lane) should have visible bands.

5 Discussion

5.1 *Treponema pallidum* infections in TNP and other selected field sites

5.1.1 There is a large genomic diversity of TPE in TNP

This present study confirms the presence of TPE in orofacial and genital lesions of *Cercocebus atys* monkeys, joining a growing body of evidence that TPE causes a diversity of symptoms in NHPs. Orofacial and genital lesions caused by TPE infections have been reported in other NHP species [114,133–136,248], however, this study presents genomic evidence for the occurrence of both symptoms in a single NHP social group inhabiting the same ecosystem. There was no conclusive evidence to suggest that TPE strains causing the two different pathologies (orofacial and genital lesions) in the *Cercocebus atys* monkeys social group formed separate monophyletic groups based on lesions the animals manifested. More broadly, phylogenetic analyses from these reconstructed genomes revealed a tree topology where all *Cercocebus atys* isolates consistently belonged to the TPE clade. This observation was also true for all other NHP genomes isolated from other ecosystems. In addition, results presented here expand the spatial distribution of where genital lesions are observed as a symptom of TPE infection in NHPs. This is opposed to the earlier belief that genital symptoms were only confined to East Africa and more common in olive baboons [136]. However, the description of the TPE infection characterised by genital-like lesions among *Cercocebus atys* monkeys of TNP noted in this study, in addition to the earlier description among green monkeys in Niokolo-Koba National Park in Senegal, gives a contrary view [19].

Notably, in all TPE strains sampled from the TNP ecosystem, there is no indication that strains were epidemiologically linked, as no identical genomes were sampled. This was also the case for two TPE strains sampled from the Lake Manyara National Park ecosystem (LMNP, Tanzania). Indeed, the average evolutionary distance between strains was comparable in TNP and LMNP (Appendices: Table S3). In contrast, in NHP species in some other ecosystems, nearly identical genomes were observed in smaller sample sizes, potentially indicative of an epidemiological link: particularly the two *Chlorocebus sabaeus* monkey populations in Bijilo Forest Park (The Gambia) and Niokolo National Park (Senegal) [19]. The factors driving this large genetic diversity of TPE in some ecosystems and not others are unknown, underscoring the poor understanding of TPE ecology and evolution in wild NHP populations. This level of genetic diversity in TNP is a fascinating one, especially that all genomes were isolated from the same social group of *Cercocebus atys* monkeys inhabiting the same ecosystem. Observations made within a single social group of sooty mangabeys do not support the epizootic spread of a single clone, but rather points towards frequent independent

introductions of the bacterium, which can cause orofacial and genital lesions. This could be one possible explanation for the lack of epidemiological link among genomes sampled from the TNP ecosystem. However, sources of these independent introductions into the social group and the implications of such with regards to the genetic diversity of TPE strains in this ecosystem, remain an important area of future research.

5.1.2 Genomic diversity of TPE infecting NHPs is geographically structured

Results from this present study support the findings of a recent study of TPE strain diversity among NHPs from Tanzania that used multi-locus sequence typing (MLST) [20] and found that simian derived genomes did not form monophyletic clades based on host species, but rather clustered based on geography. This study from Tanzania highlighted considerable genotype diversity in NHPs sharing an ecosystem and that the incongruence between the host phylogeny and that of their TPE strains is suggestive of rare intra-species transmission of TPE [20]. In addition, the authors also argued that the geographic clusters observed in Tanzania were a good indicator of temporal and spatial stability of TPE strains infecting NHPs [20].

In the present study, which used TPE genomes rather than MLST and generated data from different ecosystems in sub-Saharan Africa, three geographical clades were observed; the east African clade consisting of isolates from Tanzania and the two West African clades that formed well-supported monophyletic groups. TPE isolates from TNP clustered separately from isolates from NHPs in far West Africa, particularly Guinea, Guinea Bissau, Senegal and The Gambia, which are all close neighbouring countries (Figure 7 and 8). The far West Africa clade consisted of TPE strains isolated from *Chlorocebus sabaeus* monkeys, *Papio papio*, and a bone derived sequence from a *Cercocebus atys* monkey, all from ecosystems in close proximity to one another. The formation of these clades irrespective of the symptoms observed or the host species from which the strains were derived from, suggests that geography and the ecosystems NHPs inhabit may be important factors influencing the diversification of NHP TPE strains [249]. The phylogenetic read placement of TPE reads derived from chimpanzee (*Pan troglodytes verus*: 11786 and 15028) and red colobus (*Piliocolobus badius*: 22-52) bone samples from TNP (Figure 9) onto the TNP *Cercocebus atys* dominated clade provides further support to the suggestion that geography shapes the phylogenetic relationships among simian derived TPE strains. Such a clustering pattern indicate cross-species transmission of yaws between NHP species sharing the same habitat or infection of NHP sharing a habitat from some shared unknown source [3,4]. Pederson *et al.* (2009) observed that where cross-species transmission of a pathogen within an ecosystem occurs, it is governed by geographic proximity of the hosts in the ecosystem, phylogenetic relatedness of the hosts, and the biology of both the host and the pathogen [3]. Specifically in TNP, both geographic proximity of the hosts and the phylogenetic relatedness of the hosts would favour the transmission of TPE as all the NHP

species identified to be infected by TPE in this ecosystem, are known to have overlapping home ranges [192,208,209]. This scenario increases the likelihood of pathogen transmission from one host to the other and the ease by which the pathogen establishes itself in NHP social groups [3].

One transmission mode that has been suggested is vectorial transmission. Under experimental conditions, viable TP spirochetes were transmitted by flies between different host species causing clinical disease [148,151]. Recently, Gogarten *et al*. (2019) showed that flies carrying the yaws pathogen formed high-density persistent associations with NHP social groups in the TNP. These stable high-density fly-NHP associations are likely to provide opportunities for TPE transmission by flies [68]. Knauf *et al*. (2016) also isolated TP DNA from necrophagous flies in ecosystems where TP infections in NHPs are common [69]. However, further experimental work is needed to confirm whether flies can actually transmit TPE in the wild [152].

Interspecies interactions that could facilitate transmission via direct contact between NHP species inhabiting TNP are well-documented; these include a strong predator-prey relationship between chimpanzees and red colobus monkeys, while the overlapping home ranges and large amounts of time spent in mixed-species associations provides opportunities for cross-species transmission through various kinds of direct contact (e.g., grooming, fighting, play, mating) [46,192]. A sexual transmission mode has been suggested due to the predilection of genital lesions in individuals of reproductive age [137], though younger animals are also infected in the TNP *Cercocebus atys* group. Further research is also needed to understand the direct transmission of TPE between NHP species.

5.1.3 Detection of TPE in bone samples from sub-saharan Africa field sites

This study joins a growing body of evidence that human and wildlife bones may contain a treasure trove of data on treponemes [21,250]. TPE DNA was detected in NHP bones from TNP and four other additional sites across sub-Saharan Africa (Appendices: Table S1) and partial genomes could be reconstructed from some TPE positive bones, sufficient to inform our understanding of TPE ecology and evolution. The finding of TPE in bones confirms that NHP infections existed in TNP for at least three decades (Appendices: Table S1), complementing clinical evidence that only started accumulating from 2014 onward [19]. In addition, despite the lack of clinical evidence, the successful phylogenetic placement of bone derived TPE reads derived from chimpanzee (*Pan troglodytes verus*) and red colobus (*Piliocolobus badius*) monkey bones onto the TNP clade, suggests that *Cercocebus atys* monkeys are not the only NHP species affected in the park and lends support to local transmission of TPE within ecosystems. Finally, the detection of TPE DNA in bones from sub-

Saharan Africa sites where no clinical cases have been reported in NHPs, suggests that TPE circulation in NHPs is underreported. For example, recently, three more NHP species were described to be infected with TPE and include: Ethiopian grivet monkeys (*Chlorocebus aethiops*), vervet monkeys (*Chlorocebus pygerythrus*) from Tanzania and western lowland gorillas (*Gorilla gorilla gorilla*) from the Congo basin in the Republic of the Congo [20]. In this regard, increased surveillance of TPE in wildlife is needed and will expand our knowledge of the ecological niche of this pathogen and could also be useful for informing eradication efforts. Despite the knowledge that TPE has been infecting NHPs since at least the late 1960s [114,133], only a single isolate dating back to this period is available [23]. Future studies extracting genomic data from old NHP bones or human medical archival materials from yaws endemic regions may provide important insights into TPE ecology and evolution. For example, a recent study, demonstrated the retrieval of TPE genomes from ancient specimens providing the first evidence ever of yaws infections in medieval Northern Europe, far from the tropical regions in which present-day yaws is typically endemic [251].

5.1.4 The zoonotic potential of TPE

The wide spatio-temporal relationships between human and simian derived TPE strains has precluded the testing of whether cross-species transmission occurs between the two host species. Indeed, genomic data from human infections is largely lacking from these countries where NHP infections are common i.e. in countries where NHP TPE infections have been characterised and sequenced, no genomic data from human infections is currently available. For example, to date no human strains from Côte d'Ivoire and Tanzania have been sequenced, despite ongoing human infections [19]. Given the spatial signal among simian isolates detected in this study and elsewhere [20], future studies should investigate human yaws infections in these regions and determine whether such a spatial signal also exists for human-infecting yaws-causing bacteria on a larger scale. Such data will help to test and determine whether inter-species transmission of TPE between humans and NHPs occurs, particularly as the sampling of NHP TPE genetic diversity has greatly improved over the last several years. In fact, of the thirty TPE genomes included in the current phylogenomic analyses of this presnt study, fourteen originate from NHP communities in sub-Saharan Africa. If zoonotic spillover of TPE from NHP to humans occurs at undetectable levels [19], one positive finding is that all NHP derived TPE strains characterized to date [20], including those generated in this study, have no mutations in the 23S ribosomal RNA gene that are known to confer antimicrobial resistance to macrolides antibiotics [244]. This suggest that if zoonotic transmission occurs, at least currently available antibiotic treatments would be expected to be effective against these strains in the ongoing yaws eradication campaign.

5.2 Leprosy infection in CNP and TNP

5.2.1 Sylvatic lepromatous leprosy in wild chimpanzees; the clinical evidence

A chimpanzee, a sooty mangabey (*Cercocebus atys*) and a macaque obtained from the wild were described to have leprosy disease after long stays in captivity [15–17]. *Mycobacterium leprae* genomes from these cases were recently determined and are reflective of the geographic origin of the NHPs. This suggested that leprosy exists in wild NHPs in these regions (West Africa and The Philippines) though no clinical cases were observed in the wild until now [24]. This present study brings to light matching clinical and molecular evidence that two populations of wild Western African chimpanzees in West Africa are infected by *M. leprae;* the aetiological agent of leprosy. This becomes the first study to report sylvatic leprosy in NHPs and particulary in chimpanzees.

Though an exact number of affected animals could not be determined, camera trap efforts at CNP indicated that a several chimpanzees are affected. In affected individuals, the nodular hypo-pigmented lesions were wide spread with some individuals having complete loss of the face suggestive of lepromatous form of leprosy [15]. For the case of Woodstock at TNP that was monitored closely over a period of time, within a period of six months, the lesions had increased in number, size, depigmentation and distribution covering almost the whole maxillofacial area, eyebrows, and ear pinnae; a gross pathologic pattern indicative of lepromatous leprosy [15]. A similar gross pathological pattern could be deciphered for the earlier case of Zora after reviewing archived photos taken few months before her death. This rapid progression of symptoms of Woodstock could be suggestive of a fast-declining immunity. In humans, progression from tuberculoid leprosy (few lesions) to lepromatous leprosy (wide spread) was found to be in a linear correlation with a declining immunity of affected individuals [25,184]. In addition, the detection of *M. leprae* DNA in most internal organs (spleen, liver, lung, mesenteric lymph node, heart and kidney) of Zora coupled with the demonstration of massive infiltration of acid fast bacilli in the necropsy spleen sample on histological examination (Appendice 10.13/pathology report for Zora) and the detection of antibodies against PGL-I [157,242] in the blood of Zora, points to a systemic phase of leprosy that is pathognomonic of lepromatous leprosy [181].

5.2.2 Faecal samples are a potential surveillance tool for *M. leprae* in wild NHPs

Given the non-practicality of testing biopsies from affected individuals as routinely done in humans [163], results from the devised non-invasive approach to confirm these leprosy-like disease cases indicated that faecal samples can be used for PCR screening of *M. leprae*.

While the detection rate for the faecal sample screening at CNP among the unhabituated chimpanzees was very low, the longitudinal investigations with faecal samples from both Zora and Woodstock from a habituated TNP social group suggested otherwise. *Mycobacterium leprae* DNA could be detected in faecal samples of Woodstock from the time this individual started manifesting symptoms in June, 2018 (Appendices: Table S5). The detection rate of *M. leprae* DNA was consistent for the whole period under study. In addition, *M. leprae* DNA could be detected in faecal samples of Zora going back to 2002 indicating that seven years prior to her death, she was already shading pathogen DNA in the faeces (Appendices: Table S6). Initially, the detection was sporadic and only on RLEP assay, but from 2006 till 2009 the year the individual died; the detection rate was consistent on both RLEP and 18kDA PCR diagnostic assays. The discrepancy in the assays' detection ability can be explained by the difference in the sensitivity of the assays as RLEP targets a multi-copy gene, hence able to detect even minute amount of leprosy bacilli DNA [233]. From both longitudinal screens of Zora and Woodstock, it is evident that the detection on both PCR systems gets better as the disease progresses. Overall, these results suggest that non-invasive faecal samples that are easily collected, can be used for surveillance of leprosy in wild NHPs especially in habituated communities that allow for the immediate collection, preservation and linking of samples to individual animals. Furthermore, the successful sequencing of both a full and a partial genome of *M. leprae* from chimpanzee faecal samples provides additional support that indeed faecal samples are useful tools with a potential to be applied on a larger scale in leprosy surveillance for wild NHP communities. This is the first time a full genome of *M. leprae* is being sequenced from faeces.

Two theories regarding the likely sources of *M. leprae* DNA detected in the faeces can be speculated; in the first theory, the direct discharge of *M. leprae* bacilli into the gastrointestinal tract (GIT) lumen is hypothesised. In humans where the clinical course of the disease is well illustrated, it is known that body immunity is at the lowest in lepromatous leprosy allowing for systemic spread of lesions [184] coupled with a large invasion of the lympho-reticular system (spleen, bone marrow, lymph nodes and lymphoid tissue) and liver by the *M. leprae* bacilli forming lepromas in these systems [25]. If leprosy lesions in gut-associated lymphoid tissue which are an integral part of the GIT do occur, they are likely to aid in the discharge of the leprosy bacilli in the GIT lumen. This was shown to occur with other Mycobacteria such as *M. avium paratuberculosis* in lambs [252]. The detection of *M. leprae* DNA in almost all major organs of Zora including the mesenteric lymph nodes that drain the gut-associated lymphoid tissue lends support to this theory (Appendices: Table S7). In a lepromatous human patient, ulcerative leprosy lesions along the GIT were documented at autopsy [253]. Though no internal gross lesions for NHP leprosy have been documented so far, it can be argued that

gastrointestinal leprosy associated lesions likely exists in overtly diseased lepromatous chimpanzees; this is on the basis that the outward leprosy pathology is comparable between humans and chimpanzees [15] and hence, internal pathology is also likely to be the same. GIT lesions would then enable the easy release of the *M. leprae* bacilli into gastrointestinal contents enabling detection in faeces. However, comprehensive studies are certainly needed to ascertain GIT gross and histopathological lesions that exist as a result of leprosy infection in NHPs.

The second theory is premised on the basis that in overtly diseased lepromatous human patients, upper respiratory mucous membranes and oral lymphoid tissue excrete large amounts of leprosy bacilli [25,184]. This scenario presents a high possibility of the upper respiratory excretions being swallowed into the GIT due to close proximity of the upper respiratory and GIT systems. This would enable the subsequent detection of the *M. leprae* DNA that passes down the GIT system. Detection of pathogen DNA in faeces has been effectively demonstrated for respiratory pathogens in chimpanzee populations such as the coronavirus, human metapneumovirus (HMPV) and human respiratory syncytial virus (HRSV) [13,98]. If this occurs with leprosy, then grooming behaviour in chimpanzees [208], particularly grooming of body lesions would likely enhance the oral route transmission of the pathogen. Recently, grooming of ulcerating lesions was found to play a role in the spread of monkeypoxvirus in chimpanzees' of TNP via the oral route and could be detected in faeces [108]. This example, gives supporting evidence that DNA of *M. leprae* swallowed into the GIT can be detected in faeces of NHPs.

5.2.3 Wild chimpanzees are infected with rare *M. leprae* genotypes

The detection of two rare genotypes (2F and 4N/O) in chimpanzee populations of West Africa raises questions regarding how these wild NHPs got infected (Figure 11 and 12). Traditionally, leprosy has been believed to be spread by close contact of infected and naive individuals in humans through the respiratory root [25] though this notion has been recently questioned by a number of human leprosy cases without previous contact to infected individuals, highlighting the importance of the animal reservoirs or unknown ecological niches [111,254]. Both affected populations comprise of West African chimpanzees that are indigenous to West Africa [208]. The chimpanzees at TNP where genotype 2F was detected, has been habituated since 1979 [208] and there is no record of importation from East Africa where the only two cases of genotype 2F in Africa were recently sequenced [163] or contact to Europe where 2F strains circulated in the medieval times [80,82]. This finding was both surprising and interesting as it was the first time a genotype 2 strain was being reported in West Africa; a region dominated by genotype 4 strains. Until now, genotype 2F only comprised of *M. leprae* strains from the medieval time in the present Denmark, Sweden and UK in Europe and modern strains from

Ethiopia isolated and sequenced in 2015 [80,163]. This suggests that an unknown niche of this pathogen is within the proximity of this chimpanzee population and the pathogen could be circulating within the social group. The results of the present study, coupled with the fact that the only human isolate from Côte d'Ivoire is a genotype 4 (Appendices: Table S10) like all other human isolates from West Africa, potentially erases the possibility of anthropozoonotic spill overs of this genotype from humans into this chimpanzee community though more sampling of human cases in Côte d'Ivoire is needed to interrogate this opinion with certainity. Moreover, human contact is highly restricted with this chimpanzee community [208]. These observations in large part are suggestive of chimpanzees being reservoir hosts of this rare *M. leprae* genotype in this region.

The clustering of GB_64 (Guinea Bissau) sequenced from a NHP faecal sample within the 4N/O genotype clade was not surprising (Figure 11 and 12). This finding was expected as all *M. leprae* isolates sequenced previously from West Africa including those derived from NHPs captured from West Africa are all genotype 4 [163]. However, the present study now offers both genetic and clinicopathologicl evidence, that remained elusive for a long time, that indeed, *M. leprae* exists among wild NHPs in West Africa as earlier suspected [15,24]. This finding, coupled with studies that postulate that the current genotypes of *M. leprae* are defined by geography [159,163], confirms the assertion of Honap and co-workers (2018) who argued that the *Cercocebus atys* monkey that was shipped from Nigeria to the USA [16] and a chimpanzee that was shipped from the Sierra Leone to Japan [188] nearly four decades ago were already infected with leprosy prior to translocation into foreign sanctuaries [24]. Moreover, *M. leprae* genotype 4 has never been documented to occur either in the USA or Japan erasing the possibility of human to animal transmission [163]. In addition, the clustering together of GB_64 genome from Guinea Bissau with other genomes from NHP traced back to Nigeria and Sierra Leone in the same clade, also suggests that this genotype has a wide spatial span among wild NHPs in West Africa though recent clinical and matching molecular evidence for sylvatic leprosy in NHPs are still lacking in these countries.

Though the relationships among the isolates in the 4N/O clade were unresolved and it could not be determined if GB_64 was closely related to the human infecting or the NHP infecting strains in this clade, the phylogeny presented in this study shows that genotype 4N/O is certainly rare among humans. This subgenotype has only been reported in two human cases i.e. one case from Niger and the other case from Brazil [163]. The rareness of this subgenotype in humans could suggest the maintenance of this subgenotype in the wild NHPs too. Additionally, the account of the two cases of a chimpanzee [188] and a *Cercocebus atys* monkey [16] that are believed to have been infected in West Africa prior to their translocation into foreign sanctuaries four decades ago [24], is compatible with the existence of this

subgenotype in the wild since then. It is highly unlikely that the disease in NHPs was as result of a spillover event from humans as earlier suggested [24], as there is only a single human case of this sugenotype in West Africa that has been documented since that time [163]. Furthermore, the detection of 2F in West Africa where the genotype has never been reported in humans, questions the human to NHPs transmission theory too. The results pesented here points to the hypothesis that chimpanzees might be reservoir hosts of these rare *M. leprae* strains that are circulating in the wild.

In contrast, the data presented here does not completely rule out the slightest slim possibility of ancient transmission of these genotypes from humans to NHPs. This could have been followed by their persistence in the wild NHPs over generations and their failure to thrive in a human population. This theory of animal reservoir maintenance of *M.leprae* was advanced to explain the resurgence of leprosy of genotype 3I among red squirrels that circulated in medieval times in Europe [164,181] and among armadillos in the southern USA [177,190,201]. As the disease is not widely observed in NHPs (at least for habituated troops in sub-Saharan Africa), this ancient cross-species transmission could have been a single event coupled with the inability of *M. leprae* to be highly communicable [177]. Alternatively, the infection of the NHPs from an environment source can be adduced. For example, recently, *M. leprae* DNA was detected in soil samples collected from leprosy endemic regions particularly from habitats that are in close proximity to symptomatic humans, armadillos, and red squirrels [255]. However, given the detection of *M. leprae* DNA in faecal samples of symptomatic NHPs reported in the present study, the detection in soil samples could also be as a result of environmental contamination by the excreta from these animals and not necessarily an environment niche for the pathogen. In addition, viable and virulent *M. leprae* was found to survive for a long time in Amoeba species that are common in the environment [256]. This observation in part suggests that *M. leprae* can survive in the environment and can be a source of infection for other hosts. Given these observations made here, future studies should ascertain whether *M. leprae* DNA detected in NHP faeces is from viable bacterial organisms and whether Amoebas in this ecosystem absorb and maintain them in the environment. Such studies would help in revealing the ecological niche of this pathogen in NHP habitats if at all, it exists.

5.2.4 Focussed future leprosy studies

Dietary analysis was performed to explore whether the chimpanzees were being exposed through what they eat. It has been shown in humans that consumption of infected armadillos meat increases the likelihood of leprosy infection [111]. Additionally, chimpanzees usually prey on other mammals within their home range [192] and are frequently infected by pathogens infecting their prey [5,257]. Diet analysis of chimpanzees at CNP that were screened for

leprosy revealed that they prey upon diana monkeys (*Cercopithecus diana*), spotted forest genets, and red river hogs (*potamochoerus porcus*). Though the above mammalian species are not known to be infected by *M. leprae* and none of the *M. leprae* positive faecal samples were positive for these mammalian species, the search for other mammalian *M. leprae* host reservoirs in these habitats should include these wildlife species.

The grouping together of NHP infecting isolate from Guinea Bissau (GB_64) and a human infecting isolate (Ng13-33) from Niger isolated in 2015 in the same clade coupled with their close spatio-temporal spans might also suggest a probable circulation of this rare genotype between humans and NHPs in West Africa. This is the only single human isolate out of the 83 genotype 4 genomes sequenced so far from West Africa that clusters in the 4N/O clade (Appendices Table S10). Unfortunately, both genomic and epidemiological data from human leprosy infections in Guinea Bissau, where GB_64 was isolated from are lacking [258]. This scenario precluded further understanding of this intriguing relationship. Given these observations, efforts should be made to establish the status of the disease among humans in Guinea Bissau to determine whether humans in this country are infected with the same genotype. Furthermore, more intense sampling and sequencing of human leprosy cases in Côte d'Ivoire should equally be done to ascertain whether the *M. leprae* 2F genotype detected in chimpanzees in TNP is common, especially in the human population within the vicinity of TNP. Besides, screening of old human and NHP bones for *M. leprae* as previously demonstrated [80,82] might help to elucidate whether the rare genotypes infecting NHPs were prevalent in some distant past in these regions or other leprosy endemic countries in Africa. These studies would also help generate more genomic data required to understand whether cross-species transmission of these rare genotypes occurred in the past.

Cross-species events of *M. leprae* have already been documented. For example, cross-species transmission events of genotype 3I between armadillos and humans have been reported to occur in southern USA [177]. Genotype 3I also circulated in the medieval time in Europe [80] and is currently circulating among red squirrels in the United Kingdom posing a threat of re-introduction of leprosy among humans in Europe [181]. In this respect, efforts for surveillance in humans especially in places were critical genomic and epidemiological data is lacking such as in Guinea Bissau and Côte d'Ivoire coupled with intensive surveillance among wildlife species should be encouraged; as it is now evident that leprosy is not only a human disease as other animals are equally naturally infected. Such efforts might be useful in informing strategies and policies in the ongoing campaign by WHO aiming to reduce leprosy prevalence to one new case per 10,000 population per annum at a global scale [111,160].

5.2.5 Coinfection of *M. leprae* and *M. tuberculosis* in a great ape

The detection of *M. leprae* DNA in most of Zora's internal organs brings to light the co-infection of *M. leprae* and *M. tuberculosis* in a great ape. These findings were surprising based on the fact that at the time of her death, Zora was diagnosed with a *M. tuberculosis* infection with viable *M. tuberculosis* being isolated and sequenced from mesenteric lymph nodes [259]. It can be hypothesised that the declining immunity of Zora at the time due to leprosy infection that had gone systemic, could have given opportunity for the activation of *M. tuberculosis* which is usually latent in healthy individuals [25,157]. In humans, *M. tuberculosis* is a leading infectious agent killer in immune compromised individuals [260]. From a historical perspective, Donoghue *et al.* (2005) demonstrated the co-infection of both pathogens in human archaeological specimens with palaeopathological signs of leprosy dating back to the thirteenth century [70]. The authors concluded that the depressed immunity that is associated with leprosy progression led to activation of latent *M. tuberculosis* causing mortalities among leprosy patients hence the historical decline of leprosy in Europe around the same time [70]. The two pathogens have overlapping geographic endemicity, hence the co-infection of the two in humans does occur but prevalence proportions are unknown [261]. Equally, the co-infection status in great apes is largely unknown. However, due to the low immunity associated with leprosy infection, tuberculosis infection in NHPs should also be considered likely and intervention strategies should be informed with this in mind.

6 Conclusions and outlook

- There is a large diversity of TPE strains infecting NHPs in TNP. The observations made within a single social group of *Cercocebus atys* monkeys does not support the epizootic spread of a single clone, but rather points towards frequent independent introductions of the bacterium, which can cause syphilis- like and yaws-like lesions.
- On a larger scale, the geographic clustering of TPE genomes observed in this study might be compatible with cross-species transmission of TPE within ecosystems or environmental exposure leading to acquisition of closely related strains. Future studies of TPE should consider exploring possible transmission routes between NHPs in these ecosystems.
- This study offers evidence that human and wildlife bones may contain a treasure trove of data on treponemes. The finding of TPE in bones collected from TNP and the successful phylogenetic read placement of TPE mapped reads from them, confirmed the existence of NHP infections in other species at least three decades ago complementing clinical evidence which only started accumulating from 2014 onwards.
- The detection of TPE DNA in bones from sub-Saharan Africa sites where no clinical cases have been reported in NHPs, suggests that TPE circulation in NHPs is under reported and warrants increased surveillance of TPE in wildlife. Such efforts will expand our knowledge on the ecological niche of this pathogen and could be useful for informing eradication efforts in the ongoing eradication campigns by WHO.
- This study generated four genomes reconstructed from swabs/biopsies and two partial genomes from bone samples adding to the limited number of TPE genomes existing today. The paucity of TPE genomes with a wide spatio-temporal span from both NHPs and humans has limited further investigations into whether cross-species transmission between humans and NHPs does occur. Therefore, any addition of new TPE genomes is a step forward towards the demonstration of this very important aspect. The cross-species transmission aspect has remained elusive for a very long time despite its importance in view of its potential to undermine the ongoing yaws eradication efforts by WHO.
- This study provided the first ever matching clinical and molecular evidence of sylvatic leprosy infection in two wild chimpanzee populations in West Africa.
- *Mycobacterium leprae* DNA could be detected in faecal samples. In addition, both a single full and partial *M. leprae* genome were sequenced from positive faecal samples for the first time implying that faecal samples can be used as surveillance tool for monitoring leprosy in the wild NHPs and has great potential for scalability.

- Wild chimpanzees are infected with rare *M. leprae* strains pointing to them being potential reservoir hosts of these genotypes though it remains unclear whether this pathogen can directly be transmitted between humans and animals or indirectly via unknown environmental niches.
- Future studies sampling for these rare genotypes in human populations as well as intensive surveillance of the disease in the wild in West Africa is warranted. This is to ascertain the extent to which these rare *M. leprae* strains affects wild primates and other wild mammalian species as well as establish whether humans in these regions are infected with the same genotypes. This will subsequently aid to elucidate the ecology of the disease in these regions and establish whether cross-species transmission of these strains occur between humans and NHPs or other wildlife.
- Prior to this study, there were only two NHP derived *M. leprae* genomes available in the public domain (GenBanK)and this study generated two additional NHP derived *M. leprae* genomes. The more NHP derived genomes are generated, the more likely it will enable future studies to test the role NHPs play in the ecology of the disease.

7 Summary

Genomics of infectious bacterial skin diseases in wild non-human primates: Yaws and Leprosy

Yaws-like disease in non-human primates (NHPs) caused by *Treponema pallidum* subsp. *pertenue* (TPE) has been largely reported in sub-Saharan Africa with diverse dermatological manifestations. However, it remained unclear how the genomic diversity of TPE lineages that do occur in NHPs is distributed across hosts and space. In Taï National Park (TNP), Côte d'Ivoire, symptomatic *Cercocebus atys* monkeys were observed to have yaws- and syphilis-like lesions. The present study investigated the TPE diversity in wild NHPs in TNP and other sites where the disease occurs in sub-Saharan Africa.

Phylogenomic analyses from this study revealed that dermatological pathology observed among sooty mangabeys in TNP were caused by a large diversity of TPE lineages. All TPE genomes determined from these sooty mangabeys were different and exhibited divergence levels not observed in other field sites where the disease seems to be epizootic. The results presented in this dissertation, do not support the epizootic spread of a single TPE clone at TNP, but rather points towards frequent independent introductions of the bacterium, which can cause orofacial and genital-like lesions within a single social group of *Cercocebus atys* monkeys. Besides, simian TPE isolates determined in this study and those available in the public domain, did not form monophyletic clades based on host species or the type of symptoms caused by an isolate, but rather clustered based on geography. This phylogenetic pattern is compatible with cross-species transmission of TPE within ecosystems where this disease occurs, though it remains unclear how often and by what means this transmission occurs. This study now lays basis for future studies on the potential occurrence of TPE transmission among different NHP species and future studies should aim to reveal transmission pathways of TPE in NHP habitats.

This work also demonstrates that informative TPE genomic data can be obtained from old NHPs bones collected from endemic areas; helping to trace the disease back into the past as well as identify new sites that might be affected by the disease where no clinical cases are currently observed. Thus, using bone derived TPE short sequences coupled with phylogenetic read placement algorithm, this study showed that TPE infected NHPs in TNP for at least three decades ago, complementing clinical evidence which only started accumulating from 2014. Additionally, detection of TPE in field sites were no clinical cases are observed indicates that TPE infections in NHPs are underreported, warranting the intensive surveillance of the disease in wild NHPs.

In the second part, this study aimed to investigate the causative agent of leprosy-like symptoms observed in NHPs in both Cantanhez National Park and Täi National Park. The investigations involved adapting a suitable non-invasive screening tool to facilitate the determination and characterisation of the causative agent of the observed skin lesions.

Using faecal samples, this study determined that *M. leprae* was the cause of the leprosy- like lesions confirming the first ever observed sylvatic leprosy cases among NHPs in West Africa. This finding shows that faecal samples of NHP that are easily collected, can be used for surveillance of leprosy in wild NHPs and have the potential for scalability in terms of screening huge populations over large areas. Further, this study also presents sequence data highlighting that wild chimpanzees are infected with two genotypes of *Mycobacterium leprae* that are rare in human populations i.e. 4N/O and 2F and could potentially serve as reservoir hosts of these rare genotypes. For the first time, genotype 2F of *M. leprae* that circulated mostly in the medieval time in Europe was detected in West Africa. The detection of 2F in a region dominated by genotype 4 *M. leprae* strains questions the current understanding that the infection in NHPs was a result of a human strains spill overs. Finally, it is now evident that *M. leprae* is not only a human pathogen since red squirrels, armadillos and NHPs are also naturally infected. Therefore, investigations into the role these animal reservoir hosts play in the ecology of leprosy disease might be useful in informing strategies on the ongoing leprosy eradication campaign by WHO.

8 Zusammenfassung

Genomik infektiöser bakterieller Hauterkrankungen bei wilden nichtmenschlichen Primaten: Gieren und Lepra

Es gibt eine vielzahl von Berichten über eine gierartige Krankheit bei nicht-menschlichen Primaten (NHPs), die durch Treponema pallidum subsp. pertenue (TPE) verursacht wird und in ihren verschiedenen dermatologischen Ausprägungen aus weiten Teilen der Subsahararegion Afrikas. Es blieb jedoch bislang unklar wie die genomische Vielfalt der NHP-spezifischen TPE-Linien über Wirte und Raum verteilt ist. Im Taï-Nationalpark (TNP), Côte d'Ivoire, wurden symptomatische Rußmangaben (Cercocebus atys atys atys) mit gier- und syphilisähnlichen Läsionen beobachtet. Im Zusammenhang dazu untersucht die vorliegende Studie die Diversität von TPE bei wilden NHPs in TNP und anderen Standorten, an denen die Krankheit in Subsahara-Afrika auch zuvor aufgetreteten is.

Phylogenomische Analysen aus dieser Studie zeigten, dass das dermatologische Krankheitsbild von Rußmangaben in TNP, durch eine große Vielfalt von TPE-Linien verursacht wird. Alle TPE-Genome, die aus dieser Mangabenart isoliert werden konnten, waren unterschiedlich und zeigten deutlich andere Divergenzniveaus als bisher für eine einzige Spezies an einem Feldforschungsort, an dem die Krankheit vermutlich epizootisch ist, beobachtet werden konnte. Die in dieser Arbeit vorgestellten Ergebnisse unterstützen nicht die epizootische Ausbreitung eines einzelnen TPE-Klons, sondern weisen stattdessen auf häufige unabhängige Einführungen des Bakteriums hin, die syphilis- und gierähnliche Läsionen innerhalb einer einzigen sozialen Gruppe von Rußmangaben verursachen können. Desweiteren bilden die in dieser Studie bearbeiteten Isolate von simianem TPE keine monophyletischen Kladen mit den bislang veröffentlichten Vergleichsgenomen auf der Grundlage ihrer Wirtsarten oder Art der Symptome, sondern in Bezug auf ihre Geographie. Dieses phylogenetische Muster ist einhergehend mit der artenübergreifenden Übertragung von TPE in denjenigen Ökosystemen, in denen die Krankheit auftritt. Es bleibt allerdings unklar, wie oft und auf welche Weise diese Übertragung stattfindet. Diese Studie bildet nun die Grundlage für zukünftige Studien über das mögliche Auftreten der TPE-Übertragung bei NHPs.

Diese Arbeit zeigt auch, dass wertvolle TPE Genomdaten aus alten NHP-Knochen aus endemischen Gebieten gewonnen werden können. Dies kann helfen die Krankheit in die Vergangenheit zurückverfolgen und neue Standorte zu identifizieren, die in der Gegenwart oder Zukunft von der Krankheit betroffen sein könnten, auch wenn zum jetzigen Zeitpunkt keine klinischen Fälle bekannt sind. So zeigte diese Studie unter Verwendung von knochenabgeleiteten TPE-Kurzsequenzen in Verbindung mit einem phylogenetischen

Leseplatzierungsalgorithmus , dass TPE in TNP für mindestens drei Jahrzehnte NHPs infiziert hat und ergänzt damit klinische Beweise, die sich erst ab 2014 zu sammeln begannen. Darüber hinaus zeigt der Nachweis von TPE in Freilandanlagen, in denen offiziell keine klinischen Fälle beobachtet werden, dass TPE-Infektionen in NHPs weitgehend unbemerkt bleiben, was eine intensive Überwachung der Krankheit bei wilden NHPs rechtfertigt.

Im zweiten Teil zielte diese Studie darauf ab, den Erreger lepraähnlicher Symptome zu untersuchen, die bei NHPs im Cantanhez Nationalpark und im Täi Nationalpark beobachtet wurden. Die Untersuchungen umfassten die Anpassung eines geeigneten nicht-invasiven Screeningtools, um die Bestimmung und Charakterisierung des Erregers der beobachteten Hautläsionen zu erleichtern.

Anhand von Stuhlproben konnte in dieser Studie festgestellt werden, dass *M. leprae* die Ursache für die lepraähnlichen Läsionen war und somit die ersten jemals beobachteten sylvatischen Leprafälle bei NHPs in Westafrika bestätigten. Dieses Ergebnis zeigt, dass Fäkalproben von NHP, die ohne große Schwierigkeiten gesammelt werden können, zur Überwachung der Lepra bei wilden NHPs verwendet werden können und das Potenzial für Skalierbarkeit beim Screening großer Populationen über große Gebiete haben. Darüber hinaus stellt diese Studie auch Sequenzdaten vor, die zeigen, dass wilde Schimpansen mit zwei Genotypen von Mycobacterium leprae infiziert sind, die in menschlichen Populationen selten sind. Dabei handelt es sich um 4O-P und 2F, und somit besteht die Möglichkeit, dass Schimpansen als Reservoirwirte dieser seltenen Genotypen dienen könnten. Zum ersten Mal wurde in Westafrika der Genotyp 2F von *M. leprae* entdeckt, der hauptsächlich im Mittelalter in Europa zirkulierte. Der Nachweis von 2F in einer Region, die vom Genotype 4 *M. leprae* dominiert wird, stellt das aktuelle Verständnis in Frage, dass die Infektion in NHPs das Ergebnis einer Übertragung vom Menschen waren. Abschließend ist inzwischen evident, dass *M. leprae* nicht nur ein menschlicher Erreger ist, da auch rote Eichhörnchen, Gürteltiere und NHPs auf natürlichem Wege infiziert werden können. Somit könnten Untersuchungen über die Rolle dieser Tiere als Reservoirwirte bei der Ökologie der Leprakrankheit nützlich sein, um Strategien für die Leprabekämpfung zu entwickeln.

9 References

1. Belant JL, Deese AR. Importance of wildlife disease surveillance. Human-Wildlife Interact. 2010;4: 165–169.

2. Taylor LH, Latham SM, woolhouse MEJ. Risk factors for human disease emergence. Woolhouse MEJ, Dye C, editors. Philos Trans R Soc London Ser B Biol Sci. 2001;356: 983–989. doi:10.1098/rstb.2001.0888

3. Pedersen AB, Davies TJ. Cross-Species Pathogen Transmission and Disease Emergence in Primates. Ecohealth. 2009;6: 496–508. doi:10.1007/s10393-010-0284-3

4. Davies TJ, Pedersen AB. Phylogeny and geography predict pathogen community similarity in wild primates and humans. Proceedings Biol Sci. 2008;275: 1695–701. doi:10.1098/rspb.2008.0284

5. Leendertz FH, Pauli G, Maetz-Rensing K, Boardman W, Nunn C, Ellerbrok H, et al. Pathogens as drivers of population declines: The importance of systematic monitoring in great apes and other threatened mammals. Biol Conserv. 2006;131: 325–337. doi:10.1016/J.BIOCON.2006.05.002

6. Leendertz FH, Ellerbrok H, Boesch C, Couacy-Hymann E, Mätz-Rensing K, Hakenbeck R, et al. Anthrax kills wild chimpanzees in a tropical rainforest. Nature. 2004;430: 451–452. doi:10.1038/nature02722

7. Bermejo M, Rodríguez-Teijeiro JD, Illera G, Barroso A, Vilà C, Walsh PD. Multiple Ebola Virus Transmission Events and Rapid Decline of Central African Wildlife. Science (80-). 2006;303: 387–390. doi:10.1126/science.1092528

8. Han BA, Kramer AM, Drake JM. Global patterns of zoonotic disease in mammals. Trends Parasitol. 2016;32: 565–577.

9. Evans DT, Silvestri G. Nonhuman primate models in AIDS research. Current Opinion in HIV and AIDS. 2013. pp. 255–261. doi:10.1097/COH.0b013e328361cee8

10. Sasseville VG, Veazey R, Knight HL, Caunt D, Lackner AA, Rottman JB, et al. Characterization of the Cutaneous Exanthem in Macaques Infected with a Nef Gene Variant of SIVmac239. J Invest Dermatol. 1998;110: 894–901. doi:10.1046/J.1523-1747.1998.00205.X

11. Heldstab A, Rüedi D, Sonnabend W, Deinhardt F. Spontaneous generalized Herpesvirus hominis infection of a lowland gorilla (Gorilla gorilla gorilla). J Med

Primatol. 1981;10: 129–35.

12. Radonić A, Metzger S, Dabrowski PW, Couacy-Hymann E, Schuenadel L, Kurth A, et al. Fatal Monkeypox in Wild-Living Sooty Mangabey, Côte d'Ivoire, 2012. Emerg Infect Dis. 2014;20: 1009–1011. doi:10.3201/eid2006.13-1329

13. Patrono L V., Samuni L, Corman VM, Nourifar L, Röthemeier C, Wittig RM, et al. Human coronavirus OC43 outbreak in wild chimpanzees, Côte d´Ivoire, 2016. Emerg Microbes Infect. 2018;7: 1–4. doi:10.1038/s41426-018-0121-2

14. Bernstein JA, Didier PJ. Nonhuman primate dermatology: a literature review. Vet Dermatol. 2009;20: 145–56. doi:10.1111/j.1365-3164.2009.00742.x

15. Donham KJ, Leininger JR. Spontaneous Leprosy-Like Disease in a Chimpanzee. J Infect Dis. 1977;136: 132–136. doi:10.1093/infdis/136.1.132

16. Meyers WM, Walsh GP, Brown HL, Binford CH, Imes GD, Hadfield TL, et al. Leprosy in a mangabey monkey--naturally acquired infection. Int J Lepr Other Mycobact Dis. 1985;53: 1–14.

17. Valverde CR, Canfield D, Tarara R, Esteves MI, Gormus BJ. Spontaneous leprosy in a wild-caught cynomolgus macaque. Int J Lepr Other Mycobact Dis. 1998;66: 140–8.

18. Gormus BJ, Wolf RH, Baskin GB, Ohkawa S, Gerone PJ, Walsh GP, et al. A second sooty mangabey monkey with naturally acquired leprosy: first reported possible monkey-to-monkey transmission. Int J Lepr Other Mycobact Dis. 1988;56: 61–5.

19. Knauf S, Gogarten JF, Schuenemann VJ, Nys HM De, Düx A, Strouhal M, et al. Nonhuman primates across sub-Saharan Africa are infected with the yaws bacterium Treponema pallidum subsp. pertenue. Emerg Microbes Infect. 2018;7: 1–4. doi:10.1038/s41426-018-0156-4

20. Chuma IS, Roos C, Atickem A, Bohm T, Anthony Collins D, Grillová L, et al. Strain diversity of Treponema pallidum subsp. pertenue suggests rare interspecies transmission in African nonhuman primates. Sci Rep. 2019;9: 14243. doi:10.1038/s41598-019-50779-9

21. Gogarten JF, Düx A, Schuenemann VJ, Nowak K, Boesch C, Wittig RM, et al. Tools for opening new chapters in the book of Treponema pallidum evolutionary history. Clin Microbiol Infect. 2016;22: 916–921. doi:10.1016/j.cmi.2016.07.027

22. Fribourg-Blanc A, Mollaret HH, Niel G. [Serologic and microscopic confirmation of treponemosis in Guinea baboons]. Bull Soc Pathol Exot Filiales. 1966;59: 54–9.

23. Zobaníková M, Strouhal M, Mikalová L, Čejková D, Ambrožová L, Pospíšilová P, et al. Whole Genome Sequence of the Treponema Fribourg-Blanc: Unspecified Simian Isolate Is Highly Similar to the Yaws Subspecies. PLoS Negl Trop Dis. 2013;7: e2172. doi:10.1371/journal.pntd.0002172

24. Honap TP, Pfister L-A, Housman G, Mills S, Tarara RP, Suzuki K, et al. Mycobacterium leprae genomes from naturally infected nonhuman primates. Small PLC, editor. PLoS Negl Trop Dis. 2018;12: e0006190. doi:10.1371/journal.pntd.0006190

25. Walker SL, Withington SG, Lockwood DNJ. 41 - Leprosy. In: Farrar J, Hotez PJ, Junghanss T, Kang G, Lalloo D, White NJBT-MTID (Twenty-third E, editors. London: W.B. Saunders; 2014. pp. 506-518.e1. doi:https://doi.org/10.1016/B978-0-7020-5101-2.00042-X

26. Morse SS. Factors in the Emergence of Infectious Diseases. Emerg Infect Dis. 1995;1: 7–15. doi:10.3201/eid0101.950102

27. Morens DM, Fauci AS. Emerging infectious diseases: threats to human health and global stability. PLoS Pathog. 2013;9: e1003467. doi:10.1371/journal.ppat.1003467

28. Nunn CL, Altizer SM. Infectious diseases in primates : behavior, ecology and evolution. Oxford University Press; 2006.

29. Bell SM, Bedford T. Modern-day SIV viral diversity generated by extensive recombination and cross-species transmission. Silvestri G, editor. PLOS Pathog. 2017;13: e1006466. doi:10.1371/journal.ppat.1006466

30. Gao F, Bailes E, Robertson DL, Chen Y, Rodenburg CM, Michael SF, et al. Origin of HIV-1 in the chimpanzee Pan troglodytes troglodytes. Nature. 1999;397: 436–441. doi:10.1038/17130

31. Chen Z, Luckay A, Sodora DL, Telfer P, Reed P, Gettie A, et al. Human immunodeficiency virus type 2 (HIV-2) seroprevalence and characterization of a distinct HIV-2 genetic subtype from the natural range of simian immunodeficiency virus-infected sooty mangabeys. J Virol. 1997;71: 3953–60.

32. Ayouba A, Akoua-Koffi C, Calvignac-Spencer S, Esteban A, Locatelli S, Li H, et al. Evidence for continuing cross-species transmission of SIVsmm to humans: characterization of a new HIV-2 lineage in rural Côte d'Ivoire. AIDS. 2013;27: 2488–91. doi:10.1097/01.aids.0000432443.22684.50

33. Frank TD, Carter A, Jahagirdar D, Biehl MH, Douwes-Schultz D, Larson SL, et al. Global, regional, and national incidence, prevalence, and mortality of HIV, 1980–2017, and forecasts to 2030, for 195 countries and territories: a systematic analysis for the Global Burden of Diseases, Injuries, and Risk Factors Study 2017. Lancet HIV. 2019. doi:10.1016/s2352-3018(19)30196-1

34. Gorshkov K, Shiryaev SA, Fertel S, Lin Y-W, Huang C-T, Pinto A, et al. Zika Virus: Origins, Pathological Action, and Treatment Strategies. Front Microbiol. 2018;9: 3252. doi:10.3389/fmicb.2018.03252

35. Morens DM, Folkers GK, Fauci AS. Emerging infections: a perpetual challenge. The Lancet Infectious Diseases. 2008. pp. 710–719. doi:10.1016/S1473-3099(08)70256-1

36. Christian MD, Poutanen SM, Loutfy MR, Muller MP, Low DE. Severe Acute Respiratory Syndrome. Clin Infect Dis. 2004;38: 1420–1427. doi:10.1086/420743

37. Reflections on SARS. Lancet Infectious Diseases. 2004. p. 651. doi:10.1016/S1473-3099(04)01183-1

38. Fasina FO, Shittu A, Lazarus D, Tomori O, Simonsen L, Viboud C, et al. Transmission dynamics and control of Ebola virus disease outbreak in Nigeria, July to September 2014. Euro Surveill. 2014;19: 20920. doi:10.2807/1560-7917.es2014.19.40.20920

39. Gardner LM, Bóta A, Gangavarapu K, Kraemer MUG, Grubaugh ND. Inferring the risk factors behind the geographical spread and transmission of Zika in the Americas. Forshey BM, editor. PLoS Negl Trop Dis. 2018;12: e0006194. doi:10.1371/journal.pntd.0006194

40. Bogoch II, Creatore MI, Cetron MS, Brownstein JS, Pesik N, Miniota J, et al. Assessment of the potential for international dissemination of Ebola virus via commercial air travel during the 2014 west African outbreak. Lancet. 2015;385: 29–35. doi:10.1016/S0140-6736(14)61828-6

41. Gubler DJ, Vasilakis N, Musso D. History and Emergence of Zika Virus. J Infect Dis. 2017;216: S860–S867. doi:10.1093/infdis/jix451

42. Terzian ACB, Zini N, Sacchetto L, Rocha RF, Parra MCP, Del Sarto JL, et al. Evidence of natural Zika virus infection in neotropical non-human primates in Brazil. Sci Rep. 2018;8. doi:10.1038/s41598-018-34423-6

43. Wolfe ND, Escalante AA, Karesh WB, Kilbourn A, Spielman A, Lal AA. Wild primate populations in emerging infectious disease research: The missing link? Emerging

Infectious Diseases. Centers for Disease Control and Prevention (CDC); 1998. pp. 149–158. doi:10.3201/eid0402.980202

44. Devaux CA, Mediannikov O, Medkour H, Raoult D. Infectious Disease Risk Across the Growing Human-Non Human Primate Interface: A Review of the Evidence. Front Public Heal. 2019;7: 1–22. doi:10.3389/fpubh.2019.00305

45. Sintasath DM, Wolfe ND, LeBreton M, Jia H, Garcia AD, Diffo JLD, et al. Simian T-lymphotropic virus diversity among nonhuman primates, Cameroon. Emerg Infect Dis. 2009;15: 175–184. doi:10.3201/eid1502.080584

46. Gogarten JF, Akoua-Koffi C, Calvignac-Spencer S, Leendertz SAJ, Weiss S, Couacy-Hymann E, et al. The ecology of primate retroviruses – An assessment of 12 years of retroviral studies in the Taï national park area, Côte d′Ivoire. Virology. 2014;460–461: 147–153. doi:10.1016/J.VIROL.2014.05.012

47. Hoffmann C, Zimmermann F, Biek R, Kuehl H, Nowak K, Mundry R, et al. Persistent anthrax as a major driver of wildlife mortality in a tropical rainforest. Nature. 2017;548: 82–86. doi:10.1038/nature23309

48. Chapman CA, Gillespie TR, Goldberg TL. Primates and the Ecology of their Infectious Diseases: How will Anthropogenic Change Affect Host-Parasite Interactions? Evol Anthropol Issues, News, Rev. 2005;14: 134–144. doi:10.1002/evan.20068

49. Dobson A, Foufopoulos J. Emerging infectious pathogens of wildlife. Philos Trans R Soc Lond B Biol Sci. 2001;356: 1001–12. doi:10.1098/rstb.2001.0900

50. Forman S, Hungerford N, Yamakawa M, Yanase T, Tsai H-J, Joo Y-S, et al. Climate change impacts and risks for animal health in Asia. Rev Sci Tech. 2008;27: 581–97.

51. Chapman CA, Chapman LJ, Struhsaker TT, Zanne AE, Clark CJ, Poulsen JR. A long-term evaluation of fruiting phenology: importance of climate change. J Trop Ecol. 2005;21: 31–45. doi:10.1017/S0266467404001993

52. Chapman C, Hodder S, Rothman J. Host-parasite dynamics: connecting primate field data to theory. 2009. pp. 463–483.

53. Patz JA, Frumkin H, Holloway T, Vimont DJ, Haines A. Climate change: challenges and opportunities for global health. JAMA. 2014;312: 1565–80. doi:10.1001/jama.2014.13186

54. Wu DF, Löhrich T, Sachse A, Mundry R, Wittig RM, Calvignac-Spencer S, et al. Seasonal and inter-annual variation of malaria parasite detection in wild chimpanzees.

Malar J. 2018;17: 38. doi:10.1186/s12936-018-2187-7

55. IUCN. Primates in peril : the world's 25 most endangered primates, 2016-2018 | IUCN. 2017 [cited 5 Oct 2019]. Available: https://www.iucn.org/content/primates-peril-worlds-25-most-endangered-primates-2016-2018

56. IUCN. The IUCN Red List of Threatned species. In: version 2019-2 [Internet]. 2019 [cited 11 Nov 2019]. Available: https://www.iucnredlist.org

57. Williams JM, Lonsdorf E V., Wilson ML, Schumacher-Stankey J, Goodall J, Pusey AE. Causes of death in the Kasekela chimpanzees of Gombe National Park, Tanzania. Am J Primatol. 2008;70: 766–777. doi:10.1002/ajp.20573

58. Köndgen S, Leider M, Lankester F, Bethe A, Lübke-Becker A, Leendertz FH, et al. Pasteurella multocida Involved in Respiratory Disease of Wild Chimpanzees. Moreno E, editor. PLoS One. 2011;6: e24236. doi:10.1371/journal.pone.0024236

59. Köndgen S, Kühl H, N'Goran PK, Walsh PD, Schenk S, Ernst N, et al. Pandemic Human Viruses Cause Decline of Endangered Great Apes. Curr Biol. 2008;18: 260–264. doi:10.1016/J.CUB.2008.01.012

60. Coop RL, Holmes PH. Nutrition and parasite interaction. Int J Parasitol. 1996;26: 951–962. doi:10.1016/S0020-7519(96)80070-1

61. Hing S, Narayan EJ, Thompson RCA, Godfrey SS. The relationship between physiological stress and wildlife disease: consequences for health and conservation. Wildl Res. 2016;43: 51. doi:10.1071/WR15183

62. Foerster S, Cords M, Monfort SL. Seasonal energetic stress in a tropical forest primate: proximate causes and evolutionary implications. PLoS One. 2012;7: e50108. doi:10.1371/journal.pone.0050108

63. Grützmacher KS, Köndgen S, Keil V, Todd A, Feistner A, Herbinger I, et al. Codetection of Respiratory Syncytial Virus in Habituated Wild Western Lowland Gorillas and Humans During a Respiratory Disease Outbreak. Ecohealth. 2016;13: 499–510. doi:10.1007/s10393-016-1144-6

64. Kaur T, Singh J, Tong S, Humphrey C, Clevenger D, Tan W, et al. Descriptive epidemiology of fatal respiratory outbreaks and detection of a human-related metapneumovirus in wild chimpanzees (Pan troglodytes) at Mahale Mountains National Park, Western Tanzania. Am J Primatol. 2008;70: 755–765. doi:10.1002/ajp.20565

65. Scully EJ, Basnet S, Wrangham RW, Muller MN, Otali E, Hyeroba D, et al. Lethal Respiratory Disease Associated with Human Rhinovirus C in Wild Chimpanzees, Uganda, 2013. Emerg Infect Dis. 2018;24: 267–274. doi:10.3201/eid2402.170778

66. Leendertz FH, Boesch C, Ellerbrok H, Rietschel W, Couacy-Hymann E, Pauli G. Non-invasive testing reveals a high prevalence of simian T-lymphotropic virus type 1 antibodies in wild adult chimpanzees of the Tai National Park, Cote d'Ivoire. J Gen Virol. 2004;85: 3305–3312. doi:10.1099/vir.0.80052-0

67. Wu DF, Behringer V, Wittig RM, Leendertz FH, Deschner T. Urinary neopterin levels increase and predict survival during a respiratory outbreak in wild chimpanzees (Taï National Park, Côte d'Ivoire). Sci Rep. 2018;8: 13346. doi:10.1038/s41598-018-31563-7

68. Gogarten JF, Düx A, Mubemba B, Pléh K, Hoffmann C, Mielke A, et al. Tropical rainforest flies carrying pathogens form stable associations with social nonhuman primates. Mol Ecol. 2019; mec.15145. doi:10.1111/mec.15145

69. Knauf S, Raphael J, Mitjà O, Lejora IAV, Chuma IS, Batamuzi EK, et al. Isolation of Treponema DNA from Necrophagous Flies in a Natural Ecosystem. EBioMedicine. 2016;11: 85–90. doi:10.1016/j.ebiom.2016.07.033

70. Donoghue HD, Marcsik A, Matheson C, Vernon K, Nuorala E, Molto JE, et al. Co-infection of Mycobacterium tuberculosis and Mycobacterium leprae in human archaeological samples: a possible explanation for the historical decline of leprosy. Proceedings Biol Sci. 2005;272: 389–94. doi:10.1098/rspb.2004.2966

71. Gogarten JF, Hoffmann C, Arandjelovic M, Sachse A, Merkel K, Dieguez P, et al. Fly-derived DNA and camera traps are complementary tools for assessing mammalian biodiversity. Environ DNA. 2019; edn3.46. doi:10.1002/edn3.46

72. Hyong SK, Kim DM, Neupane GP, Lee YM, Yang NW, Sook JJ, et al. Comparison of conventional, nested, and real-time PCR assays for rapid and accurate detection of Vibrio vulnificus. J Clin Microbiol. 2008;46: 2992–2998. doi:10.1128/JCM.00027-08

73. Ehlers B, Borchers K, Grund C, Frölich K, Ludwig H, Buhk HJ. Detection of new DNA polymerase genes of known and potentially novel herpesviruses by PCR with degenerate and deoxyinosine-substituted primers. Virus Genes. 1999;18: 211–220. doi:10.1023/A:1008064118057

74. Leendertz FH, Deckers M, Schempp W, Lankester F, Boesch C, Mugisha L, et al. Novel cytomegaloviruses in free-ranging and captive great apes: phylogenetic

evidence for bidirectional horizontal transmission. J Gen Virol. 2009;90: 2386–2394. doi:10.1099/vir.0.011866-0

75. Anoh AE, Murthy S, Akoua-Koffi C, Couacy-Hymann E, Leendertz FH, Calvignac-Spencer S, et al. Cytomegaloviruses in a Community of Wild Nonhuman Primates in Taï National Park, Côte D'Ivoire. Viruses. 2017;10. doi:10.3390/v10010011

76. Syvanen AC. Toward genome-wide snp genotyping. Nat Genet. 2005;37: S5–S10. doi:10.1038/ng1558

77. Suhana O, Nazni WA, Apandi Y, Farah H, Lee HL, Sofian-Azirun M. Insight into the origin of chikungunya virus in Malaysian non-human primates via sequence analysis. Heliyon. 2019;5. doi:10.1016/j.heliyon.2019.e02682

78. Gaudin M, Desnues C. Hybrid capture-based next generation sequencing and its application to human infectious diseases. Frontiers in Microbiology. Frontiers Media S.A.; 2018. doi:10.3389/fmicb.2018.02924

79. Maurier F, Beury D, Fléchon L, Varré JS, Touzet H, Goffard A, et al. A complete protocol for whole-genome sequencing of virus from clinical samples: Application to coronavirus OC43. Virology. 2019;531: 141–148. doi:10.1016/j.virol.2019.03.006

80. Schuenemann VJ, Avanzi C, Krause-Kyora B, Seitz A, Herbig A, Inskip S, et al. Ancient genomes reveal a high diversity of Mycobacterium leprae in medieval Europe. Monack DM, editor. PLOS Pathog. 2018;14: e1006997. doi:10.1371/journal.ppat.1006997

81. Schuenemann VJ, Kumar Lankapalli A, Barquera R, Nelson EA, Iraíz Hernández D, Acuña Alonzo V, et al. Historic Treponema pallidum genomes from Colonial Mexico retrieved from archaeological remains. Norris SJ, editor. PLoS Negl Trop Dis. 2018;12: e0006447. doi:10.1371/journal.pntd.0006447

82. Schuenemann VJ, Singh P, Mendum TA, Krause-Kyora B, Jäger G, Bos KI, et al. Genome-wide comparison of medieval and modern Mycobacterium leprae. Science. 2013;341: 179–83. doi:10.1126/science.1238286

83. Patrono LV, Pleh K, Samuni L, Ulrich M, Roethemeier C, Sachse A, et al. Monkeypox emergence in wild chimpanzees reveals distinct clinical outcomes and viral diversity. 2020. doi:10.5281/ZENODO.3373886

84. Mubemba B, Gogarten JF, Schuenemann VJ, Düx A, Lang A, Nowak K, et al. Geographically structured genomic diversity of non-human primate-infecting

Treponema pallidum subsp. pertenue. bioRxiv. 2019; 848382. doi:10.1101/848382

85. Knierim E, Lucke B, Schwarz JM, Schuelke M, Seelow D. Systematic comparison of three methods for fragmentation of long-range PCR products for next generation sequencing. PLoS One. 2011;6. doi:10.1371/journal.pone.0028240

86. Head SR, Kiyomi Komori H, LaMere SA, Whisenant T, Van Nieuwerburgh F, Salomon DR, et al. Library construction for next-generation sequencing: Overviews and challenges. Biotechniques. 2014;56: 61–77. doi:10.2144/000114133

87. Quail MA, Kozarewa I, Smith F, Scally A, Stephens PJ, Durbin R, et al. A large genome center's improvements to the Illumina sequencing system. Nat Methods. 2008;5: 1005–1010. doi:10.1038/nmeth.1270

88. Meyer M, Briggs AW, Maricic T, Höber B, Höffner B, Krause J, et al. From micrograms to picograms: Quantitative PCR reduces the material demands of high-throughput sequencing. Nucleic Acids Res. 2008;36. doi:10.1093/nar/gkm1095

89. Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. Microb Inform Exp. 2012;2. doi:10.1186/2042-5783-2-3

90. Gnirke A, Melnikov A, Maguire J, Rogov P, LeProust EM, Brockman W, et al. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat Biotechnol. 2009;27: 182–189. doi:10.1038/nbt.1523

91. Albert TJ, Molla MN, Muzny DM, Nazareth L, Wheeler D, Song X, et al. Direct selection of human genomic loci by microarray hybridization. Nat Methods. 2007;4: 903–905. doi:10.1038/nmeth1111

92. Kozarewa I, Armisen J, Gardner AF, Slatko BE, Hendrickson CL. Overview of Target Enrichment Strategies. Curr Protoc Mol Biol. 2015;112: 7.21.1-7.21.23. doi:10.1002/0471142727.mb0721s112

93. Savolainen V, Cowan RS, Vogler AP, Roderick GK, Lane R. Towards writing the encyclopaedia of life: An introduction to DNA barcoding. Philosophical Transactions of the Royal Society B: Biological Sciences. Royal Society; 2005. pp. 1805–1811. doi:10.1098/rstb.2005.1730

94. Gogarten JF, Davies TJ, Benjamino J, Gogarten JF, Graf J, Mielke A, et al. Factors influencing bacterial microbiome composition in a wild non-human primate community in Taï National Park, Côte d'Ivoire. ISME J. 2018;12: 1. doi:10.1038/s41396-018-0166-1

95. Calvignac-Spencer S, Merkel K, Kutzner N, Kühl H, Boesch C, Kappeler PM, et al. Carrion fly-derived DNA as a tool for comprehensive and cost-effective assessment of mammalian biodiversity. Mol Ecol. 2013;22: 915–924. doi:10.1111/mec.12183

96. Schubert G, Stockhausen M, Hoffmann C, Merkel K, Vigilant L, Leendertz FH, et al. Targeted detection of mammalian species using carrion fly-derived DNA. Mol Ecol Resour. 2015;15: 285–294. doi:10.1111/1755-0998.12306

97. Luscombe NM, Greenbaum D, Gerstein M. What is bioinformatics? A proposed definition and overview of the field. Methods Inf Med. 2001;40: 346–358. doi:10.1055/s-0038-1634431

98. Köndgen S, Schenk S, Pauli G, Boesch C, Leendertz FH. Noninvasive Monitoring of Respiratory Viruses in Wild Chimpanzees. Ecohealth. 2010;7: 332–341. doi:10.1007/s10393-010-0340-z

99. Kumar S, Dudley J. Bioinformatics software for biologists in the genomics era. Bioinformatics. 2007;23: 1713–1717. doi:10.1093/bioinformatics/btm239

100. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30: 2114–2120.

101. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. ArXiv. 2013;1303.

102. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics. 2012;28: 1647–1649. doi:10.1093/bioinformatics/bts199

103. Yang Z, Rannala B. Molecular phylogenetics: Principles and practice. Nature Reviews Genetics. 2012. pp. 303–314. doi:10.1038/nrg3186

104. Barbera P, Kozlov AM, Czech L, Morel B, Darriba D, Flouri T, et al. EPA-ng: Massively Parallel Evolutionary Placement of Genetic Sequences. Posada D, editor. Syst Biol. 2019;68: 365–369. doi:10.1093/sysbio/syy054

105. Czech L, Barbera P, Stamatakis A. Genesis and Gappa: Library and Toolkit for Working with Phylogenetic (Placement) Data. bioRxiv. 2019; 647958. doi:10.1101/647958

106. Jordan R, Goff A, Frimm A, Corrado ML, Hensley LE, Byrd CM, et al. ST-246 antiviral efficacy in a nonhuman primate monkeypox model: determination of the minimal

effective dose and human dose justification. Antimicrob Agents Chemother. 2009;53: 1817–22. doi:10.1128/AAC.01596-08

107. Arita I, Henderson DA. Smallpox and monkeypox in non-human primates. Bull World Health Organ. 1968;39: 277–83.

108. Patrono L V., Pléh K, Samuni L, Ulrich M, Röthemeier C, Sachse A, et al. Monkeypox virus emergence in wild chimpanzees reveals distinct clinical outcomes and viral diversity. Nat Microbiol. 2020. doi:10.1038/s41564-020-0706-0

109. Lefaux B, Duprez R, Tanguy M, Longeart L, Gessain A, Boulanger E. Nonhuman primates might be highly susceptible to cross-species infectivity by human alpha-herpesviruses. Vet Pathol. 2004;41: 302–304. doi:10.1354/vp.41-3-302-a

110. Yakovleva LA, Lennert K, Chikobava MG, Indzhiia L V, Klotz IN, Lapin BA. Morphological characteristics of malignant T-cell lymphomas in baboons. Virchows Arch A Pathol Anat Histopathol. 1993;422: 109–20. doi:10.1007/bf01607162

111. da Silva MB, Portela JM, Li W, Jackson M, Gonzalez-Juarrero M, Hidalgo AS, et al. Evidence of zoonotic leprosy in Pará, Brazilian Amazon, and risks associated with human contact or consumption of armadillos. PLoS Negl Trop Dis. 2018;12. doi:10.1371/journal.pntd.0006532

112. Asiedu K, Fitzpatrick C, Jannin J. Eradication of yaws: historical efforts and achieving WHO's 2020 target. PLoS Negl Trop Dis. 2014;8: e3016.

113. Teshome H. Review on Principles of Zoonoses Prevention, Control and Eradication. Am J Biomed Sci Res. 2019;3: 188–197. doi:10.34297/AJBSR.2019.03.000660

114. Fribourg-Blanc A, Mollaret HH. Natural treponematosis of the African primate. Primates Med. 1969;3: 113–121.

115. Armelagos GJ, Zuckerman MK, Harper KN. The Science Behind Pre-Columbian Evidence of Syphilis in Europe: Research by Documentary. Evol Anthropol Issues, News, Rev. 2012;21: 50–57. doi:10.1002/evan.20340

116. Čejková D, Zobaníková M, Chen L, Pospíšilová P, Strouhal M, Qin X, et al. Whole genome sequences of three Treponema pallidum ssp. pertenue strains: Yaws and syphilis treponemes differ in less than 0.2% of the genome sequence. PLoS Negl Trop Dis. 2012;6. doi:10.1371/journal.pntd.0001471

117. Centurion-Lara A, Molini BJ, Godornes C, Sun E, Hevner K, Van Voorhis WC, et al. Molecular differentiation of Treponema pallidum subspecies. J Clin Microbiol.

2006;44: 3377–80. doi:10.1128/JCM.00784-06

118. Marks M, Solomon AW, Mabey DC. Endemic treponemal diseases. Trans R Soc Trop Med Hyg. 2014;108: 601–7. doi:10.1093/trstmh/tru128

119. Lithgow K V, Hof R, Wetherell C, Phillips D, Houston S, Cameron CE. A defined syphilis vaccine candidate inhibits pallidum. Nat Commun. 2017;8: 1–10. doi:10.1038/ncomms14273

120. Giacani L, Lukehart SA. The endemic treponematoses. Clin Microbiol Rev. 2014;27: 89–115. doi:10.1128/CMR.00070-13

121. Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, et al. Global Estimates of the Prevalence and Incidence of Four Curable Sexually Transmitted Infections in 2012 Based on Systematic Review and Global Reporting. Meng Z, editor. PLoS One. 2015;10: e0143304. doi:10.1371/journal.pone.0143304

122. WHO. ACCELERATING WORK TO OVERCOME THE GLOBAL IMPACT OF NEGLECTED TROPICAL. 2012; 1–22.

123. WHO. First WHO report on neglected tropical diseases: working to overcome the global impact of neglected tropical diseases. 2010. doi:WHO/HTM/NTD/2010.1

124. WHO. Global health sector strategy on Sexually Transmitted Infections, 2016-2021. 2016. Report No.: WHO/RHR/16.09.

125. Dyson L, Mooring EQ, Holmes A, Tildesley MJ, Marks M. Insights from quantitative and mathematical modelling on the proposed 2030 goals for Yaws. Gates Open Res. 2019;3: 1576. doi:10.12688/gatesopenres.13078.1

126. Longo LD. Classic pages in obstetrics and gynecology. Vorlaufiger Bericht uber das Vorkommen von Spirochaeten in syphilitischen Krankheitsprodukten und bei Papillomen. Fritz Richard Schaudinn and Erich Hoffmann. Arbeiten aus dem kaiserlichen Gesundheitsamtes (Berli. Am J Obstet Gynecol. 1978;132: 571–572.

127. Šmajs D, Zobaníková M, Strouhal M, Čejková D, Dugan-Rocha S, Pospíšilová P, et al. Complete genome sequence of Treponema paraluiscuniculi, strain Cuniculi A: The loss of infectivity to humans is associated with genome decay. PLoS One. 2011;6. doi:10.1371/journal.pone.0020415

128. Xuedong Zhou YL. Subgingival Microbes. Atlas Oral Microbiol. 2015; 67–93. doi:10.1016/B978-0-12-802234-4.00004-5

129. LaFond R, Lukehart S. Biology of Syphilis. Clin Microbiol Rev. 2006;35: 33–82.

doi:10.1128/CMR.19.1.29

130. Antal GM, Lukehart SA, Meheus AZ. The endemic treponematoses. Microbes Infect. 2002;4: 83–94. doi:10.1016/S1286-4579(01)01513-1

131. Edmondson DG, Hu B, Norris SJ. Long-Term In Vitro Culture of the Syphilis Spirochete Treponema pallidum subsp. pallidum. mbio.asm.org. 2018;9: 1–18.

132. FRIBOURG-BLANC A, NIEL G, MOLLARET HH. [NOTE ON SOME IMMUNOLOGICAL ASPECTS OF THE AFRICAN CYNOCEPHALUS. 1. ANTIGENIC RELATIONSHIP OF ITS GAMMA GLOBULIN WITH HUMAN GAMMA GLOBULIN. 2. GUINEAN ENDEMIC FOCUS OF TREPONEMATOSIS]. Bull Soc Pathol Exot Filiales. 1963;56: 474–85.

133. Baylet R, Thivolet J, Sepetjian M, Nouhouay Y, Baylet M. [Natural open treponematosis in the Papio papio baboon in Casamance]. Bull Soc Pathol Exot Filiales. 1971;64: 842–846.

134. Chuma IS, Batamuzi EK, Collins DA, Fyumagwa RD, Hallmaier-Wacker LK, Kazwala RR, et al. Widespread Treponema pallidum Infection in Nonhuman Primates, Tanzania. Emerg Infect Dis. 2018;24: 1002–1009. doi:10.3201/eid2406.180037

135. Knauf S, Batamuzi EK, Mlengeya T, Kilewo M, Lejora IAV, Nordhoff M, et al. Treponema Infection Associated With Genital Ulceration in Wild Baboons. Vet Pathol. 2012;49: 292–303. doi:10.1177/0300985811402839

136. Harper KN, Fyumagwa RD, Hoare R, Wambura PN, Coppenhaver DH, Sapolsky RM, et al. Treponema pallidum Infection in the Wild Baboons of East Africa: Distribution and Genetic Characterization of the Strains Responsible. PLoS One. 2012;7. doi:10.1371/journal.pone.0050882

137. Wallis J& RL. Primate Conservation: The Prevention of Disease Transmission. Int J Primatol. 1999;20: 803–826. doi:https://doi.org/10.1023/A:1020879700286

138. Mlengeya T. Distribution pattern of a sexually trasmitted disease (std) of olive baboon in lake manyara national park, tanzania. 2004.

139. Knauf S, Liu H, Harper KN, Sascha Knauf, Hsi Liu and KNH. Treponemal Infection in Nonhuman Primates as Possible Reservoir for Human Yaws. Emerg Infect Dis. 2012;19: 2058–2059. doi:10.3201/eid1912.130863

140. Knauf S, Barnett U, Maciej P, Klapproth M, Ndao I, Frischmann S, et al. High prevalence of antibodies against the bacterium treponema pallidum in Senegalese

Guinea Baboons (Papio papio). PLoS One. 2015;10: 1–13. doi:10.1371/journal.pone.0143100

141. Klegarth AR, Ezeonwu CA, Rompis A, Lee BPYH, Aggimarangsee N, Chalise M, et al. Survey of treponemal infections in free-ranging and captive Macaques, 1999–2012. Emerg Infect Dis. 2017;23: 816–819. doi:10.3201/eid2305.161838

142. Struhsaker TT, Angedakin S, Landsmann A. Facial and genital lesions in baboons (Papio anubis) of Kibale National Park, Uganda. Primates. 2019; 2–5. doi:10.1007/s10329-019-00715-2

143. Levrero F, Gatti S, Gautier-Hion A, Menard N. Yaws disease in a wild gorilla population and its impact on the reproductive status of males. Am J Phys Anthropol. 2007;132: 568–575. doi:10.1002/ajpa.20560

144. Hausfater G. Predatory Behavior of Yellow Baboons. Behaviour. 1976;56: 44–68.

145. Rose MD. Interspecific play between free ranging guerezas (Colobus guereza) and vervet monkeys (Cercopithecus aethiops). Primates. 1977;18: 957–964. doi:10.1007/BF02382946

146. Hackett CJ. On the epidemiology of yaws in African miners (1942). Trans R Soc Trop Med Hyg. 1984;78: 536–538.

147. HACKETT CJ. Extent and nature of the yaws problem in Africa. Bull World Health Organ. 1953;8: 110–129.

148. Kumm HW, Turner TB. The Transmission of Yaws from Man to Rabbits by an Insect Vector, Hippelates Pallipes Loew1. Am J Trop Med Hyg. 1936;s1-16: 245–271. doi:https://doi.org/10.4269/ajtmh.1936.s1-16.245

149. Baldacchino F, Muenworn V, Desquesnes M, Desoli F, Charoenviriyaphap T, Duvallet G. Transmission of pathogens by Stomoxys flies (Diptera, Muscidae): a review. Parasite. 2013;20: 26. doi:10.1051/parasite/2013026

150. Goddard J. Sand Fly-Transmitted Diseases. Infectious Diseases and Arthropods. Totowa, NJ: Humana Press; 2000. pp. 139–149. doi:10.1007/978-1-59259-721-5_6

151. Satchell GH, Harrison RA. Experimental observations on the possibility of transmission of yaws by wound-feeding Diptera, in Western Samoa. Trans R Soc Trop Med Hyg. 1953;47: 148–153.

152. Stamm L V. Flies and Yaws: Molecular Studies Provide New Insight. EBioMedicine. 2016;11: 9–10. doi:10.1016/j.ebiom.2016.08.028

153. Smith JL, David NJ, Indgin S, Israel CW, Levine BM, Justice J, et al. Neuro-ophthalmological study of late yaws and pinta. II. The Caracas project. Sex Transm Infect. 1971;47: 226–251. doi:10.1136/sti.47.4.226

154. Lawton Smith J, Israel CW. Recovery of spirochagetes in the monkey by passive transfer from human late sero-negative syphilis. Br J Vener Dis. 1968;44: 109–15. doi:10.1136/sti.44.2.109

155. Plowright RK, Parrish CR, McCallum H, Hudson PJ, Ko AI, Graham AL, et al. Pathways to zoonotic spillover. Nat Rev Microbiol. 2017;15: 502.

156. WHO. Yaws. 2019. Available: https://www.who.int/news-room/fact-sheets/detail/yaws

157. FINE PEM. LEPROSY: THE EPIDEMIOLOGY OF A SLOW BACTERIUM. Epidemiol Rev. 1982;4: 161–188. doi:10.1093/oxfordjournals.epirev.a036245

158. D. Donoghue H, Michael Taylor G, A. Mendum T, R. Stewart G, Rigouts L, Y-C. Lee O, et al. The Distribution and Origins of Ancient Leprosy. Hansen's Disease - The Forgotten and Neglected Disease. IntechOpen; 2019. doi:10.5772/intechopen.75260

159. Monot M, Honoré N, Garnier T, Araoz R, Coppée J-Y, Lacroix C, et al. On the origin of leprosy. Science. 2005;308: 1040–2. doi:10.1126/science/1109759

160. WHO. Global leprosy update, 2016: accelerating reduction of disease burden. In: Weekly epidemiological record. 2017 p. No. 35, 2017, 92, 501–520.

161. Han XY, Seo YH, Sizer KC, Schoberle T, May GS, Spencer JS, et al. A new Mycobacterium species causing diffuse lepromatous leprosy. Am J Clin Pathol. 2008;130: 856–864. doi:10.1309/AJCPP72FJZZRRVMM

162. Singh P, Benjak A, Schuenemann VJ, Herbig A, Avanzi C, Busso P, et al. Insight into the evolution and origin of leprosy bacilli from the genome sequence of Mycobacterium lepromatosis. Proc Natl Acad Sci U S A. 2015;112: 4459–64. doi:10.1073/pnas.1421504112

163. Benjak A, Avanzi C, Singh P, Loiseau C, Girma S, Busso P, et al. Phylogenomics and antimicrobial resistance of the leprosy bacillus Mycobacterium leprae. Nat Commun. 2018;9. doi:10.1038/s41467-017-02576-z

164. Schilling A-K, Avanzi C, Ulrich RG, Busso P, Pisanu B, Ferrari N, et al. British Red Squirrels Remain the Only Known Wild Rodent Host for Leprosy Bacilli. Front Vet Sci. 2019;6: 6–11. doi:10.3389/fvets.2019.00008

165. Avanzi C, Benjak A, Stevenson K, Simpson VR, Busso P, Mcluckie J, et al. Infected

With Leprosy Bacilli. Science (80-). 2016;354: 744–748. doi:10.1126/science.aah3783

166. Eisenstadt J, Hall GS. Microbiology and classification of mycobacteria. Clin Dermatol. 1995;13: 197–206. doi:10.1016/0738-081X(95)00021-7

167. Tortoli E, Fedrizzi T, Meehan CJ, Trovato A, Grottola A, Giacobazzi E, et al. The new phylogeny of the genus Mycobacterium: The old and the news. Infect Genet Evol. 2017;56: 19–25. doi:10.1016/J.MEEGID.2017.10.013

168. Hartmans S, de Bont JAM, Stackebrandt E. The Genus Mycobacterium--Nonmedical BT - The Prokaryotes: Volume 3: Archaea. Bacteria: Firmicutes, Actinomycetes. In: Dworkin M, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. New York, NY: Springer New York; 2006. pp. 889–918. doi:10.1007/0-387-30743-5_33

169. BañUls AL, Sanou A, Van Anh NT, Godreuil S. Mycobacterium tuberculosis: Ecology and evolution of a human bacterium. J Med Microbiol. 2015;64: 1261–1269. doi:10.1099/jmm.0.000171

170. Schluger NW. Tuberculosis and nontuberculous mycobacterial infections in older adults. Clin Chest Med. 2007;28: 773–vi. doi:10.1016/j.ccm.2007.07.004

171. Wolinsky E. Mycobacterial Diseases Other Than Tuberculosis. Clin Infect Dis. 1992;15: 1–12.

172. Benjak A, Avanzi C, Benito Y, Breysse F, Chartier C, Boschiroli M-L, et al. Highly Reduced Genome of the New Species Mycobacterium uberis , the Causative Agent of Nodular Thelitis and Tuberculoid Scrotitis in Livestock and a Close Relative of the Leprosy Bacilli . mSphere. 2018;3. doi:10.1128/msphere.00405-18

173. Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, et al. Massive gene decay in the leprosy bacillus. Nature. 2001;409: 1007–11. doi:10.1038/35059006

174. Singh P, Cole ST. Mycobacterium leprae: genes, pseudogenes and genetic diversity. Future Microbiol. 2011;6: 57–71. doi:10.2217/fmb.10.153

175. Bhat RM, Prakash C. Leprosy: An overview of pathophysiology. Interdisciplinary Perspectives on Infectious Diseases. 2012. doi:10.1155/2012/181089

176. Monot M, Honoré N, Garnier T, Zidane N, Sherafi D, Paniz-Mondolfi A, et al. Comparative genomic and phylogeographic analysis of Mycobacterium leprae. Nat Genet. 2009;41: 1282–1289. doi:10.1038/ng.477

177. Sharma R, Singh P, Loughry WJ, Lockhart JM, Inman WB, Duthie MS, et al. Zoonotic Leprosy in the Southeastern United States. Emerg Infect Dis. 2015;21: 2127–34. doi:10.3201/eid2112.150501

178. Han XY, Sizer KC, Thompson EJ, Kabanja J, Li J, Hu P, et al. Comparative sequence analysis of Mycobacterium leprae and the new leprosy-causing Mycobacterium lepromatosis. J Bacteriol. 2009;191: 6067–74. doi:10.1128/JB.00762-09

179. Han XY, Mistry NA, Thompson EJ, Tang H-L, Khanna K, Zhang L. Draft Genome Sequence of New Leprosy Agent Mycobacterium lepromatosis. Genome Announc. 2015;3. doi:10.1128/genomeA.00513-15

180. Han XY, Seo Y-H, Sizer KC, Schoberle T, May GS, Spencer JS, et al. A New *Mycobacterium* Species Causing Diffuse Lepromatous Leprosy. Am J Clin Pathol. 2008;130: 856–864. doi:10.1309/AJCPP72FJZZRRVMM

181. Avanzi C, Del-Pozo J, Benjak A, Stevenson K, Simpson VR, Busso P, et al. Red squirrels in the British Isles are infected with leprosy bacilli. Science. 2016;354: 744–747. doi:10.1126/science.aah3783

182. Ridley DS, Jopling WH. Classification of leprosy according to immunity. A five-group system. Int J Lepr Other Mycobact Dis. 1966;34: 255–73.

183. Classification Committee on Technical resolutions: classification. Int J Lepr. 1953.

184. Lastória JC, Abreu MAMM de. Leprosy: review of the epidemiological, clinical, and etiopathogenic aspects - part 1. An Bras Dermatol. 2014;89: 205–18. doi:10.1590/abd1806-4841.20142450

185. Vera-Cabrera L, Escalante-Fuentes WG, Gomez-Flores M, Ocampo-Candiani J, Busso P, Singh P, et al. Case of diffuse lepromatous leprosy associated with "Mycobacterium lepromatosis". J Clin Microbiol. 2011;49: 4366–8. doi:10.1128/JCM.05634-11

186. Han XY, Sizer KC, Velarde-Félix JS, Frias-Castro LO, Vargas-Ocampo F. The leprosy agents Mycobacterium lepromatosis and Mycobacterium leprae in Mexico. Int J Dermatol. 2012;51: 952–9. doi:10.1111/j.1365-4632.2011.05414.x

187. Hubbard GB, Lee DR, Eichberg JW, Gormus BJ, Xu K, Meyers WM. Spontaneous Leprosy in a Chimpanzee (Pan troglodytes). Vet Pathol. 1991;28: 546–548. doi:10.1177/030098589102800617

188. Suzuki K, Udono T, Fujisawa M, Tanigawa K, Idani G, Ishii N. Infection during infancy

and long incubation period of leprosy suggested in a case of a chimpanzee used for medical research. J Clin Microbiol. 2010;48: 3432–4. doi:10.1128/JCM.00017-10

189. Gormus BJ, Xu K, Baskin GB, Martin LN, Bohm RP, Blanchard JL, et al. Experimental leprosy in monkeys. I. Sooty mangabey monkeys: transmission, susceptibility, clinical and pathological findings. Lepr Rev. 1995;66: 96–104.

190. Truman R. Leprosy in wild armadillos. Lepr Rev. 2005;76: 198–208.

191. Trevathan W. The chimpanzees of Gombe. Patterns of behavior. By Jane Goodall. American Journal of Physical Anthropology. John Wiley & Sons, Ltd; 1987. pp. 409–410. doi:10.1002/ajpa.1330730313

192. Boesch C, Boesch H. Hunting behavior of wild chimpanzees in the Taï National Park. Am J Phys Anthropol. 1989;78: 547–573. doi:10.1002/ajpa.1330780410

193. Walsh GP, Storrs EE, Burchfield HP, Cotrell EH, Vidrine MF, Binford CH. Leprosy-like disease occurring naturally in armadillos. J Reticuloendothel Soc. 1975;18: 347–51.

194. Smith JH, Folse DS, Long EG, Christie JD, Crouse DT, Tewes ME, et al. Leprosy in wild armadillos (Dasypus novemcinctus) of the Texas Gulf Coast: epidemiology and mycobacteriology. J Reticuloendothel Soc. 1983;34: 75–88.

195. Binford CH, Storrs EE, Walsh GP. Disseminated infection in the nine-banded armadillo (Dasypus novemcinctus) resulting from inoculation with M. leprae. Observations made on 15 animals studied at autopsy. Int J Lepr Other Mycobact Dis. 1976;44: 80–3.

196. Binford CH, Meyers WM, Walsh GP, Storrs EE, Brown HL. Naturally acquired leprosy-like disease in the nine-banded armadillo (Dasypus novemcinctus): histopathologic and microbiologic studies of tissues. J Reticuloendothel Soc. 1977;22: 377–88.

197. Truman Baton Rouge, LA) R (LSU, Sanchez R. Armadillos: models for leprosy. Lab animal (USA). 1993.

198. Sharma R, Lahiri R, Scollard DM, Pena M, Williams DL, Adams LB, et al. The armadillo: a model for the neuropathy of leprosy and potentially other neurodegenerative diseases. Dis Model Mech. 2013;6: 19–24. doi:10.1242/dmm.010215

199. Job CK, Sanchez RM, Hastings RC. Manifestations of Experimental Leprosy in the Armadillo. Am J Trop Med Hyg. 1985;34: 151–161. doi:10.4269/ajtmh.1985.34.151

200. Kirchheimer WF, Storrs EE. Attempts to establish the armadillo (Dasypus

novemcinctus Linn.) as a model for the study of leprosy. I. Report of lepromatoid leprosy in an experimentally infected armadillo. Int J Lepr Other Mycobact Dis. 1971;39: 693–702.

201. Truman RW, Shannon EJ, Hagstad H V., Hugh-Jones ME, Wolff A, Hastings RC. Evaluation of the Origin of Mycobacterium Leprae Infections in the Wild Armadillo, Dasypus Novemcinctus. Am J Trop Med Hyg. 1986;35: 588–593. doi:10.4269/ajtmh.1986.35.588

202. Thomas DA, Mines JS, Thomas DC, Mack TM, Rea TH. Armadillo Exposure Among Mexican-Born Patients with Lepromatous Leprosy. J Infect Dis. 1987;156: 990–992. doi:10.1093/infdis/156.6.990

203. Lumpkin LR, Cox GF, Wolf JE. Leprosy in armadillo handlers. J Am Acad Dermatol. 1984;10: 1073. doi:10.1016/S0190-9622(84)80352-7

204. Simpson V, Hargreaves J, Butler H, Blackett T, Stevenson K, McLuckie J. Leprosy in red squirrels on the Isle of Wight and Brownsea Island. Vet Rec. 2015;177: 206–207. doi:10.1136/vr.h4491

205. Truman RW, Singh P, Sharma R, Busso P, Rougemont J, Paniz-Mondolfi A, et al. Probable zoonotic leprosy in the southern United States. N Engl J Med. 2011;364: 1626–33. doi:10.1056/NEJMoa1010536

206. Doran D. Influence of Seasonality on Activity Patterns, Feeding Behavior, Ranging, and Grouping Patterns in Taï Chimpanzees. Int J Primatol. 1997;18: 183–206. doi:10.1023/A:1026368518431

207. UNESCO. Taï National Park - UNESCO World Heritage Centre. [cited 26 Oct 2019]. Available: https://whc.unesco.org/en/list/195/

208. Boesch C, Boesch-Achermann H. The chimpanzees of the Taï Forest: Behavioural ecology and evolution. Oxford: Oxford Univ. Pr.; 2000.

209. Mcgraw W, Zuberbühler K, Noë R. Monkeys of the Taï Forest. An African Monkey Community. 2007.

210. Rohland N, Hofreiter M. Ancient DNA extraction from bones and teeth. Nat Protoc. 2007;2: 1756–1762.

211. Gamba C, Hanghøj K, Gaunitz C, Alfarhan AH, Alquraishi SA, Al-Rasheid KAS, et al. Comparing the performance of three ancient DNA extraction methods for high-throughput sequencing. Mol Ecol Resour. 2015.

212. Orlando L, Ginolhac A, Raghavan M, Vilstrup J, Rasmussen M, Magnussen K, et al. True single-molecule DNA sequencing of a pleistocene horse bone. Genome Res. 2011;21: 1705–1719. doi:10.1101/gr.122747.111

213. Leslie DE, Azzato F, Karapanagiotidis T, Leydon J, Fyfe J. Development of a real-time PCR assay to detect Treponema pallidum in clinical specimens and assessment of the assay's performance by comparison with serological testing. J Clin Microbiol. 2007;45: 93–96. doi:10.1128/JCM.01578-06

214. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215: 403–410. doi:10.1016/S0022-2836(05)80360-2

215. Peltzer A, Jäger G, Herbig A, Seitz A, Kniep C, Krause J, et al. EAGER: efficient ancient genome reconstruction. Genome Biol. 2016;17: 60.

216. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25: 2078–2079.

217. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30: 772–80. doi:10.1093/molbev/mst010

218. Talavera G, Castresana J. Improvement of Phylogenies after Removing Divergent and Ambiguously Aligned Blocks from Protein Sequence Alignments. Kjer K, Page R, Sullivan J, editors. Syst Biol. 2007;56: 564–577. doi:10.1080/10635150701472164

219. Gouy M, Guindon S, Gascuel O. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Mol Biol Evol. 2010;27: 221–224. doi:10.1093/molbev/msp259

220. Lefort V, Longueville J-E, Gascuel O. SMS: Smart Model Selection in PhyML. Mol Biol Evol. 2017;34: 2422–2424. doi:10.1093/molbev/msx149

221. Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol. 2010;59: 307–321. doi:10.1093/sysbio/syq010

222. Rambaut A, Lam TT, Carvalho LM, Pybus OG. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). Virus Evol. 2016;2: vew007. doi:10.1093/VE/VEW007

223. Fourment M, Gibbs M. PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. BMC Evol Biol. 2006;6: 1.

doi:10.1186/1471-2148-6-1

224. Arora N, Schuenemann VJ, Jäger G, Peltzer A, Seitz A, Herbig A, et al. Origin of modern syphilis and emergence of a pandemic Treponema pallidum cluster. Nat Microbiol. 2016;2: 16245. doi:10.1038/nmicrobiol.2016.245

225. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol. 2018;67: 901–904.

226. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7: 214.

227. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 2019. doi:10.1093/nar/gkz239

228. Berger SA, Stamatakis A. Aligning short reads to reference alignments and trees. Bioinformatics. 2011;27: 2068–2075. doi:10.1093/bioinformatics/btr320

229. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. Wren J, editor. Bioinformatics. 2019. doi:10.1093/bioinformatics/btz305

230. Hockings KJ, Sousa C. Human-Chimpanzee Sympatry and Interactions in Cantanhez National Park, Guinea-Bissau: Current Research and Future Directions. Primate Conserv. 2013;26: 57–65. doi:10.1896/052.026.0104

231. Vieira WF, Kerry C, Hockings KJ. A comparison of methods to determine chimpanzee home-range size in a forest–farm mosaic at Madina in Cantanhez National Park, Guinea-Bissau. Primates. 2019;60: 355–365. doi:10.1007/s10329-019-00724-1

232. Spartaco G, Dell'Omo G. Primates of Guinea-Bissau, West Africa: distribution and conservation status. Primate Conserv. 2003;19: 73–77.

233. DONOGHUE HD, HOLTON J, SPIGELMAN M. PCR primers that can detect low levels of Mycobacterium leprae DNA. J Med Microbiol. 2001;50: 177–182. doi:10.1099/0022-1317-50-2-177

234. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9: 357–9. doi:10.1038/nmeth.1923

235. Koboldt DC, Chen K, Wylie T, Larson DE, McLellan MD, Mardis ER, et al. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics. 2009;25: 2283–2285. doi:10.1093/bioinformatics/btp373

236. Cingolani P, Platts A, Wang LLL, Coon M, Nguyen T, Wang LLL, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly (Austin). 2012;6: 80–92. doi:10.4161/fly.19695

237. Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33: 1870–1874. doi:10.1093/molbev/msw054

238. Bi Z, Wittig R. Long-term diet of the chimpanzees (Pan troglodytes verus) in Taï National Park: interannual variations in consumption. 2019. pp. 242–260. doi:10.1017/9781108674218.016

239. Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E. obitools : a unix - inspired software package for DNA metabarcoding. Mol Ecol Resour. 2016;16: 176–182. doi:10.1111/1755-0998.12428

240. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, et al. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10: 421. doi:10.1186/1471-2105-10-421

241. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970;48: 443–453. doi:10.1016/0022-2836(70)90057-4

242. Bührer-Sékula S, Smits HL, Gussenhoven GC, van Leeuwen J, Amador S, Fujiwara T, et al. Simple and fast lateral flow test for classification of leprosy patients and identification of contacts with high risk of developing leprosy. J Clin Microbiol. 2003;41: 1991–5. doi:10.1128/jcm.41.5.1991-1995.2003

243. Stefani MM, Martelli CM, Morais-Neto OL, Martelli P, Costa MB, de Andrade AL. Assessment of anti-PGL-I as a prognostic marker of leprosy reaction. Int J Lepr Other Mycobact Dis. 1998;66: 356–64.

244. Stamm L V, Bergen HL. A point mutation associated with bacterial macrolide resistance is present in both 23S rRNA genes of an erythromycin-resistant Treponema pallidum clinical isolate. Antimicrob Agents Chemother. 2000;44: 806–807. doi:10.1128/aac.44.3.806-807.2000

245. Bergsten J. A review of long-branch attraction. Cladistics. 2005;21: 163–193. doi:10.1111/j.1096-0031.2005.00059.x

246. Grzybowski A, Nita M, Virmond M. Ocular leprosy. Clinics in Dermatology. Elsevier Inc.; 2015. pp. 79–89. doi:10.1016/j.clindermatol.2014.07.003

247. Coscolla M, Lewin A, Metzger S, Maetz-Rennsing K, Calvignac-Spencer S, Nitsche A, et al. Novel Mycobacterium tuberculosis complex isolate from a wild chimpanzee. Emerg Infect Dis. 2013;19: 969–76. doi:10.3201/eid1906.121012

248. Mubemba B, Chanove E, Mätz-Rensing K, Gogarten JF, Düx A, Merkel K, et al. Yaws Disease Caused by Treponema pallidum subspecies pertenue in Wild Chimpanzee, Guinea, 2019. Emerging Infectious Diseases. Centers for Disease Control and Prevention (CDC); 2020. pp. 1283–1286. doi:10.3201/EID2606.191713

249. Parks DH, Porter M, Churcher S, Wang S, Blouin C, Whalley J, et al. GenGIS: A geospatial information system for genomic data. Genome Res. 2009;19: 1896–1904. doi:10.1101/GR.095612.109

250. Fabricius T, Winther C, Ewertsen C, Kemp M, Nielsen SD. Osteitis in the dens of axis caused by Treponema pallidum. BMC Infect Dis. 2013;13: 347. doi:10.1186/1471-2334-13-347

251. Majander K, Pfrengle S, Neukamm J, Kocher A, Plessis L du, Pla-Díaz M, et al. Ancient bacterial genomes reveal a formerly unknown diversity of Treponema pallidum strains in early modern Europe. bioRxiv. 2020; 2020.06.09.142547. doi:10.1101/2020.06.09.142547

252. Lugton I. Mucosa-associated lymphoid tissues as sites for uptake, carriage and excretion of tubercle bacilli and other pathogenic mycobacteria. Immunol Cell Biol. 1999;77: 364–372. doi:10.1046/j.1440-1711.1999.00836.x

253. Taylor CE. Leprosy of the large intestine and gall-bladder: Case report with review of the literature. Trans R Soc Trop Med Hyg. 1945;39: 125–132. doi:10.1016/0035-9203(45)90003-4

254. Bonnar PE, Cunningham NP, Boggild AK, Walsh NM, Sharma R, Davis IRC. Leprosy in Nonimmigrant Canadian Man without Travel outside North America, 2014. Emerg Infect Dis. 2018;24: 165–166. doi:10.3201/eid2401.170547

255. Tió-Coma M, Wijnands T, Pierneef L, Schilling AK, Alam K, Roy JC, et al. Detection of Mycobacterium leprae DNA in soil: multiple needles in the haystack. Sci Rep. 2019;9: 3165. doi:10.1038/s41598-019-39746-6

256. Wheat WH, Casali AL, Thomas V, Spencer JS, Lahiri R, Williams DL, et al. Long-term

Survival and Virulence of Mycobacterium leprae in Amoebal Cysts. Vinetz JM, editor. PLoS Negl Trop Dis. 2014;8: e3405. doi:10.1371/journal.pntd.0003405

257. Bailes E, Gao F, Bibollet-Ruche F, Courgnaud V, Peeters M, Marx PA, et al. Hybrid Origin of SIV in Chimpanzees. Science (80-). 2003;300: 1713 LP – 1713. doi:10.1126/science.1080657

258. WHO. Leprosy; sitautiona and trends. In: Global Health Observatory Data [Internet]. World Health Organization; 2019 [cited 13 Nov 2019]. Available: https://www.who.int/gho/neglected_diseases/leprosy/en/

259. Coscolla M, Lewin A, Metzger S, Maetz-Rennsing K, Calvignac-Spencer S, Nitsche A, et al. Novel *Mycobacterium tuberculosis* Complex Isolate from a Wild Chimpanzee. Emerg Infect Dis. 2013;19: 969–976. doi:10.3201/eid1906.121012

260. Manosuthi W, Chottanapand S, Thongyen S, Chaovavanich A, Sungkanuparph S. Survival Rate and Risk Factors of Mortality Among HIV/Tuberculosis-Coinfected Patients With and Without Antiretroviral Therapy. JAIDS J Acquir Immune Defic Syndr. 2006;43: 42–46. doi:10.1097/01.qai.0000230521.86964.86

261. Mangum L, Kilpatrick D, Stryjewska B, Sampath R. Tuberculosis and Leprosy Coinfection: A Perspective on Diagnosis and Treatment. Open Forum Infect Dis. 2018;5. doi:10.1093/OFID/OFY133

10 Appendices

10.1 Table S1: Bone samples screened for Treponema pallidum

Sample ID	YOD	Species	country	Field site	Test 1/copy number	Test 2/copy number	Av. Copy number
94-9	1994	*Cercocebus atys*	CI	Tai	0.0127	No Ct	0.0127
2023	2000	*Cercocebus atys*	CI	Tai	No Ct	No Ct	No Ct
24-2	2004	*Cercocebus atys*	CI	Tai	No Ct	No Ct	No Ct
26-3	2006	*Cercocebus atys*	CI	Tai	No Ct	No Ct	No Ct
Djo_276		*Cercocebus atys*	CI	Djoroutou	No Ct	No Ct	No Ct
Tai _E_025	2013	*Cercocebus atys*	CI	Tai	No Ct	No Ct	No Ct
Tai_R_443		*cercocebus atys*	CI	Tai	No Ct	No Ct	No Ct
Djo_435		*Cercopithecus*	CI	Djoroutou	No Ct	No Ct	No Ct
Tai_R_024		*Cercopithecus*	CI	Tai	No Ct	No Ct	No Ct
2012	2000	*Cercopithecus campbelli*	CI	Tai	No Ct	No Ct	No Ct
2018	2000	*Cercopithecus diana*	CI	Tai	No Ct	0.8341	0.8341
22-10	2002	*Cercopithecus diana*	CI	Tai	7.03	11.13	9.08
27-5	2007	*Cercopithecus diana*	CI	Tai	No Ct	No Ct	No Ct
94-1		*Cercopithecus diana*	CI	Tai	No Ct	No Ct	No Ct
Tai_E_217	2013	Colobinae	CI	Tai	No Ct	No Ct	No Ct
Tai_R_447		Colobinae	CI	Tai	No Ct	No Ct	No Ct
18	1994	*Colobus polykomos*	CI	Tai	No Ct	No Ct	No Ct
23-8	2003	*Colobus polykomos*	CI	Tai	No Ct	No Ct	No Ct
13434	1999	*Pan troglodytes verus*	CI	Tai	No Ct	0.0254	0.0254
unknown chimp	2014	*Pan troglodytes verus*	CI	Tai	0.11	0.0163	0.06315
11777	1991	*Pan troglodytes verus*	CI	Tai	0.6797	0.4222	0.55095
11786	1992	*Pan troglodytes verus*	CI	Tai	No Ct	0.8607	0.8607*
15027	2009	*Pan troglodytes verus*	CI	Tai	1.01	No Ct	1.01*
11778	1994	*Pan troglodytes verus*	CI	Tai	0.8422	1.44	1.44
11792	1994	*Pan troglodytes verus*	CI	Tai	1.53	No Ct	1.53
Nino (not sure)		*Pan troglodytes verus*	CI	Tai	1.53	No Ct	1.53
11783	1992/93	*Pan troglodytes verus*	CI	Tai	1.42	2.23	1.825
11780	1994	*Pan troglodytes verus*	CI	Tai	2.635	No Ct	2.635
11781	1994	*Pan troglodytes verus*	CI	Tai	1.34	4.3	2.82
15028	1998	*Pan troglodytes verus*	CI	Tai	2.97	3.91	3.44
13430	2000	*Pan troglodytes verus*	CI	Tai	3.63	3.51	3.57
11789	1994	*Pan troglodytes verus*	CI	Tai	No Ct	8.53	8.53
11775	1992/93	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11780	1994	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11785	1993	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11791	1992	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11793	1989	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11800	1992/93	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
11903	1994	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct

12175	1996	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
12176	1996	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
13133	1999	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
13429	2000	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
13438	?	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
13439	1999	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
14994	2004	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
14995	2004	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
14996	2002	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
15001	2005	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
15002	2006?	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
15005	2001	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
15012	2002	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
Gogol	2008	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
Lilou	2009	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
Odrie	2009	*Pan troglodytes verus*	CI	Tai	No Ct	No Ct	No Ct
2101	2001	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	0.24
Tai_R_105	2013	*Piliocolobus badius*	CI	Tai	0.007	0.968	0.4875
24-4	2004	*Piliocolobus badius*	CI	Tai	1.26	0.721	0.9905
94-10	1994	*Piliocolobus badius*	CI	Tai	1.57	No Ct	1.57
22-52	2002	*Piliocolobus badius*	CI	Tai	1.57	6.37	3.97
22-2	2002	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
23-6	2003	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
10-11	2010	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
22-50	2002	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
94-30	1994	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
97-1	1997	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
Tai_E_140	2013	*Piliocolobus badius*	CI	Tai	No Ct	No Ct	No Ct
Gangu_skull 2		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	1.39	No Ct	1.39
Gangu_skull 10	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	1.46	1.46
Gangu_skull 7	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	0.4257	2.95	1.68785
Gangu_skull 15	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	1.71	1.71
Gangu_skull 18	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	3.78	3.78
Gangu_skull 1	2005	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 11	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 12	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 13	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 16	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 17	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 19	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct

Gangu_skull 20		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 21		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 22		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 23		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 24		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 3		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 4		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 5	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 6	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 8	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull 9	2013	*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Gangu_skull25		*Pan troglodytes schweinfurthii*	DRC	Bili-Uere	No Ct	No Ct	No Ct
Loa ape 15		*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 19	old skull	*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa_ape 7		*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa_ape_13		*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa_ape_14		*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa_ape_16		*Gorilla gorilla*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 1		*Pan troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 17		*Pan troglodytes troglodytes*	Gabon	Loango	0.0598	No Ct	0.0598
loa ape 23		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	0.284	0.284
Loa 3		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 10		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 12		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 18		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 2		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 20		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
loa ape 22		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
loa ape 24		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
loa ape 25		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 26		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 27		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 29		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 4		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct

Loa ape 5		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 6	2014	*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 8		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape 9		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa ape_21		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Loa_ape_11		*Pan troglodytes troglodytes*	Gabon	Loango	No Ct	No Ct	No Ct
Gco_B_019	2011	*unknown Primate*	Guinea	Old Foutah	No Ct	No Ct	No Ct
Wpt 626	5/18/2012	*unknown Primate*	Guinea	Old Foutah	No Ct	No Ct	No Ct
Boe_092_b	2013	*Cercocebus atys*	Guinea-Bissau	Boe	21.29	23.35	22.32
Boe_018	2013	*Pan troglodytes*	Guinea-Bissau	Boe	No Ct	No Ct	No Ct
Boe_019_a	2013	*Pan troglodytes*	Guinea-Bissau	Boe	No Ct	No Ct	No Ct
Boe_074_b	2013	*Papio papio*	Guinea-Bissau	Boe	No Ct	No Ct	No Ct
Boe_180_b	2014	*Papio papio*	Guinea-Bissau	Boe	No Ct	No Ct	No Ct
Boe_519	2014	*Papio papio*	Guinea-Bissau	Boe	No Ct	No Ct	No Ct
GBO_032	2013	*Cercocebus atys*	Liberia	Grebo	No Ct	No Ct	No Ct
GBO_306_B	2013	*Cercopithecus*	Liberia	Grebo	No Ct	No Ct	No Ct
Lib_A_1065	2011	*Cercopithecus petaurista*	Liberia	Nationwide	3.02	2.18	2.6
GBO_558	2013	*Colobus guereza*	Liberia	Grebo	No Ct	No Ct	No Ct
Nim 600	1990	*Pan troglodytes*	Liberia	East Nimba	3.85	0.297	2.0735
Lib_C_49b?	2010	*Pan troglodytes verus*	Liberia	Sapo	0.3313	No Ct	0.3313
Lib_C_239	2011	*Pan troglodytes verus*	Liberia	Sapo	1.06	No Ct	1.06
Lib_B_557	2012	*Pan troglodytes verus*	Liberia	Nationwide	5.24;	0.6774	2.9587
GBO_385	2013	*Pan troglodytes verus*	Liberia	Grebo	No Ct	No Ct	No Ct
Lib_C_319	2011	*Pan troglodytes verus*	Liberia	Sapo	No Ct	No Ct	No Ct
Gas_451	2013	*Colobus guereza*	Nigeria	Gashaka	No Ct	0.0505	0.0505
Gas_404	2013	*Pan troglodytes ellioti*	Nigeria	Gashaka	No Ct	No Ct	No Ct
Gas_412	2013	*Papio anubis*	Nigeria	Gashaka	No Ct	No Ct	No Ct
KAY_398	2013	*Pan troglodytes verus*	Senegal		No Ct	No Ct	No Ct
Bud_D_003_a	6/23/2012	*Cercopithecus ascanius*	Uganda	Budongo	No Ct	No Ct	No Ct
Bud_D_011_a	7/4/2012	*Cercopithecus mitis*	Uganda	Budongo	1.07	7.455	4.2625
Bud_A_138_a	8/13/2012	*Cercopithecus mitis*	Uganda	Budongo	No Ct	No Ct	No Ct
Bud_D_002_a	6/21/2012	*Colobus guereza*	Uganda	Budongo	No Ct	0.7597	0.7597
Bud_A_452_a	8/23/2012	*Colobus guereza*	Uganda	Budongo	No Ct	No Ct	No Ct
Bud_D_022	7/23/2012	*Pan troglodytes schweinfurthii*	Uganda	Budongo	No Ct	No Ct	No Ct
Bud_D_048	9/11/2012	*Pan troglodytes schweinfurthii*	Uganda	Budongo	No Ct	No Ct	No Ct

- Notes:
  - *Included in the hybridization capture based on [21]
  - CI – Côte d'Ivoire, DRC- Democratic Republic of Congo

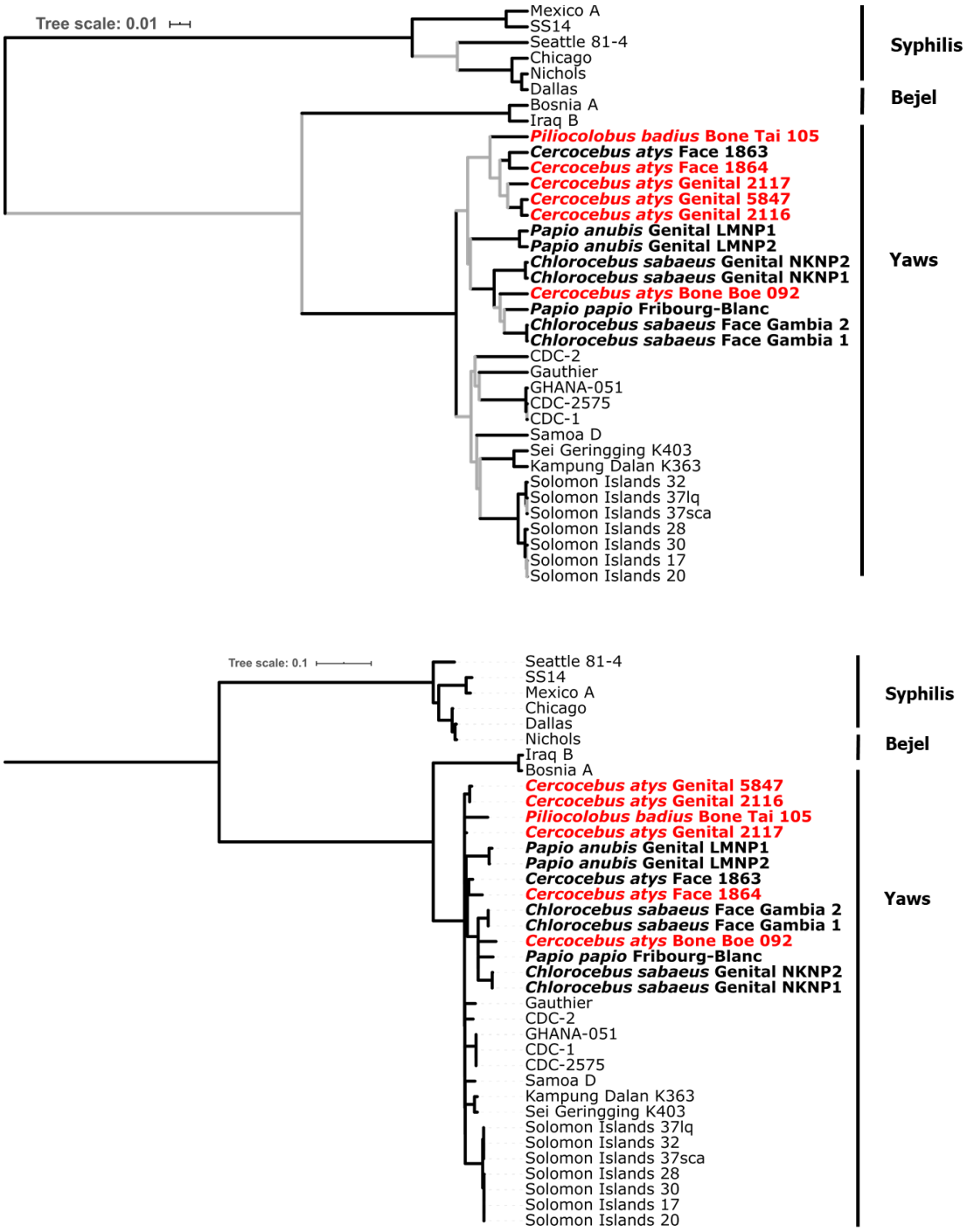
10.2 Table S2: Published genomes used in this study

Isolate ID	RefSeq Accession ID	Host	TP spectrum
Bosnia A	CP007548.1	*Homo Sapiens*	Bejel
Iraq_B	CP032303.1	*Homo Sapiens*	Bejel
Nichols	NC_021490.2	*Homo Sapiens*	Syphilis
SS14	NC_021508.1	*Homo Sapiens*	Syphilis
Chicago	NC_017268.1	*Homo Sapiens*	Syphilis
Mexico A	NC_018722.1	*Homo Sapiens*	Syphilis
Dallas	NC_016844.1	*Homo Sapiens*	Syphilis
Seattle 81-4	CP003679.1	*Homo Sapiens*	Syphilis
Fribourg-Blanc	NC_021179.1	*Papio cynocephalus*	Yaws
Samoa D	NC_016842.1	*Homo Sapiens*	Yaws
Gauthier	NC_016843.1	*Homo Sapiens*	Yaws
CDC-1	CP024750.1	*Homo Sapiens*	Yaws
CDC-2	NC_016848.1	*Homo Sapiens*	Yaws
CDC_2575	CP020366	*Homo Sapiens*	Yaws
Ghana-051	CP020365	*Homo Sapiens*	Yaws
Kampung_Dalan_K363	CP024088.1	*Homo Sapiens*	Yaws
Sei_Geringging_K403	CP024089.1	*Homo Sapiens*	Yaws
Solomon Islands 03	ERR1470343	*Homo Sapiens*	Yaws
Solomon Islands 17	ERR1470344	*Homo Sapiens*	Yaws
Solomon Islands 20	ERR1470335	*Homo Sapiens*	Yaws
Solomon Islands 28	ERR1470338	*Homo Sapiens*	Yaws
Solomon Islands 30	ERR1470334	*Homo Sapiens*	Yaws
Solomon Islands 32	ERR1470342	*Homo Sapiens*	Yaws
Solomon Islands 37 liq	ERR1470330	*Homo Sapiens*	Yaws
Solomon Islands 37 sca	ERR1470331	*Homo Sapiens*	Yaws
Gambia-1	SRR4308597	*Chlorocebus sabeus*	Yaws
Gambia-2	SRR4308605	*Chlorocebus sabaeus*	Yaws
Senegal NKNP-1	SRR4308606	*Chlorocebus sabaeus*	Yaws
Senegal NKNP-2	SRR4308607	*Chlorocebus sabaeus*	Yaws
LMNP-1	CP021113.1	*Papio anubis*	Yaws
LMNP-2_BS5	SRR4308598	*Papio anubis*	Yaws
LMNP-2_BS6	SRR4308599	*Papio anubis*	Yaws
LMNP-2_BS7	SRR4308601	*Papio anubis*	Yaws
LMNP-2_BS8	SRR4308602	*Papio anubis*	Yaws
1863-Hato	SRR4308604	*Cercocebus atys*	Yaws

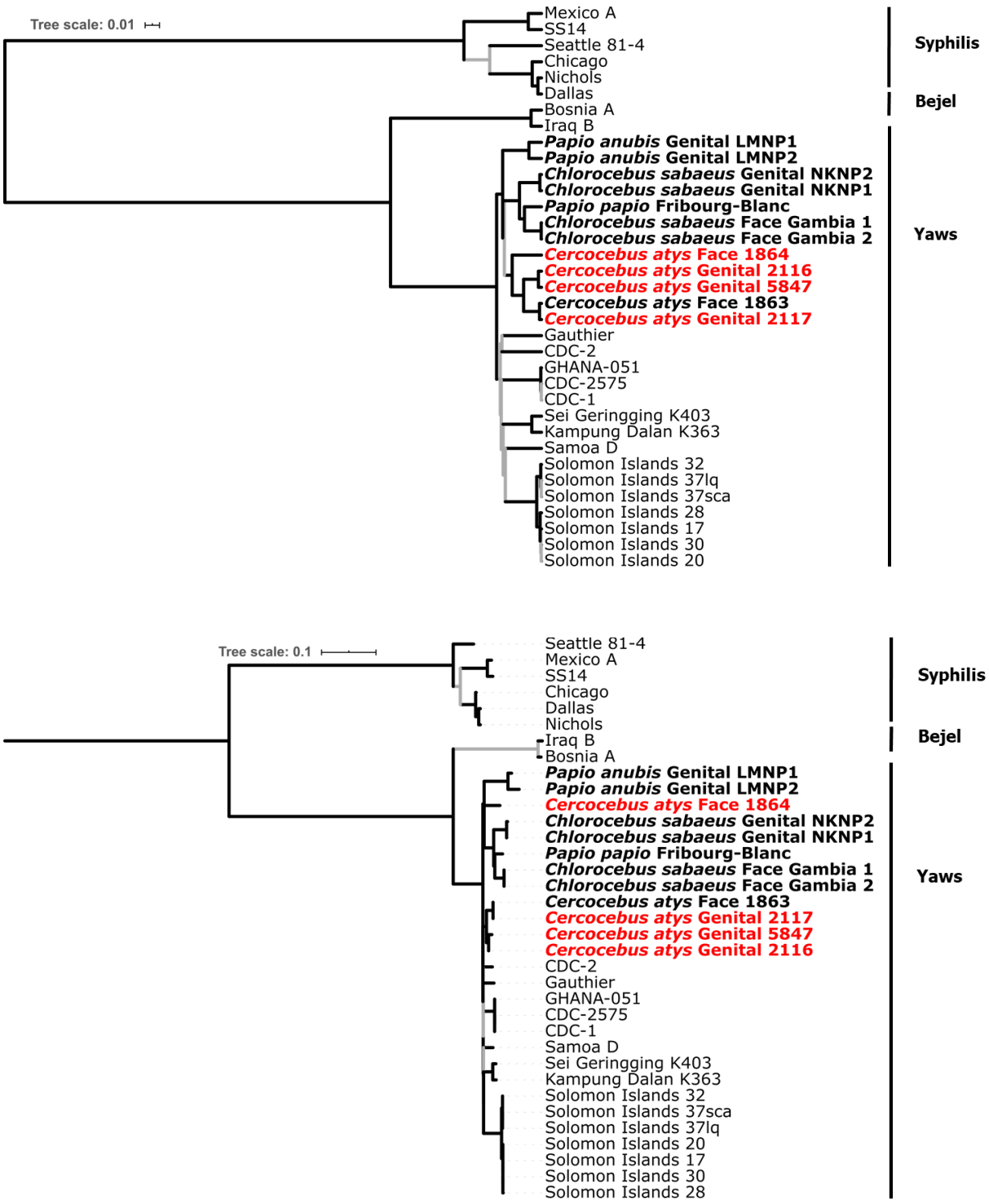
10.3 Table S3: Patristic distances of simian derived TPE strains analysed in this study using the ML 10X-95% tree

NHPs Strain ID	Papio anubis Genital LMNP1	Papio.anubis_Genital_LMNP2	Chlorocebus.sabaeus_Genital_NKNP1	Chlorocebus.sabaeus_Genital_NKNP2	Papio papio Fribourg-Blanc	Chlorocebus.sabaeus_Face_Gambia_1	Chlorocebus.sabaeus_Face_Gambia_2	Cercocebus.atys_Face_1864	Cercocebus.atys_Genital_2116	Cercocebus.atys_Genital_5847	Cercocebus.atys_Genital_2117	Cercocebus.atys_Face_1863
Papio anubis Genital LMNP1	0	0.0271182 3	0.09101906	0.09268256	0.084205 18	0.08687132	0.08687129	0.080959 1	0.0646365	0.0692560 9	0.0664403	0.066440 54
Papio.anubis_Genital_LMNP2	0.027118 23	0	0.10495723	0.10662073	0.098143 35	0.10080949	0.10080946	0.094897 2	0.0785747	0.0831942 6	0.0803785	0.080378 71
Chlorocebus.sabaeus_Genital_NKNP1	0.091019 06	0.1049572 3	0	0.00166352	0.040070 34	0.04273648	0.04273645	0.071804 4	0.0554819	0.0601014 5	0.0572857	0.057285 9
Chlorocebus.sabaeus_Genital_NKNP2	0.092682 56	0.1066207 3	0.00166352	0	0.041733 84	0.04439998	0.04439995	0.073467 9	0.0571454	0.0617649 5	0.0589492	0.058949 4
Papio papio Fribourg-Blanc	0.084205 18	0.0981433 5	0.04007034	0.04173384	0	0.03130374	0.03130371	0.064990 5	0.048668	0.0532875 7	0.0504718	0.050472 02
Chlorocebus.sabaeus_Face_Gambia_1	0.086871 32	0.1008094 9	0.04273648	0.04439998	0.031303 74	0	5.00E-08	0.067656 7	0.0513341	0.0559537 1	0.0531379	0.053138 16
Chlorocebus.sabaeus_Face_Gambia_2	0.086871 29	0.1008094 6	0.04273645	0.04439995	0.031303 71	5.00E-08	0	0.067656 6	0.0513341	0.0559536 8	0.0531379	0.053138 13
Cercocebus.atys_Face_1864	0.080959 05	0.0948972 2	0.07180441	0.07346791	0.064990 53	0.06765667	0.06765664	0	0.0399617	0.0445812 4	0.0417654	0.041765 69
Cercocebus.atys_Genital_2116	0.064636 51	0.0785746 8	0.05548187	0.05714537	0.048667 99	0.05133413	0.0513341	0.039961 7	0	0.0046196	0.0172566	0.017256 83
Cercocebus.atys_Genital_5847	0.069256 09	0.0831942 6	0.06010145	0.06176495	0.053287 57	0.05595371	0.05595368	0.044581 2	0.0046196	0	0.0218762	0.021876 41
Cercocebus.atys_Genital_2117	0.066440 29	0.0803784 6	0.05728565	0.05894915	0.050471 77	0.05313791	0.05313788	0.041765 4	0.0172566	0.0218761 6	0	7.30E-07
Cercocebus.atys_Face_1863	0.066440 54	0.0803787 1	0.0572859	0.0589494	0.050472 02	0.05313816	0.05313813	0.041765 7	0.0172568	0.0218764 1	7.30E-07	0

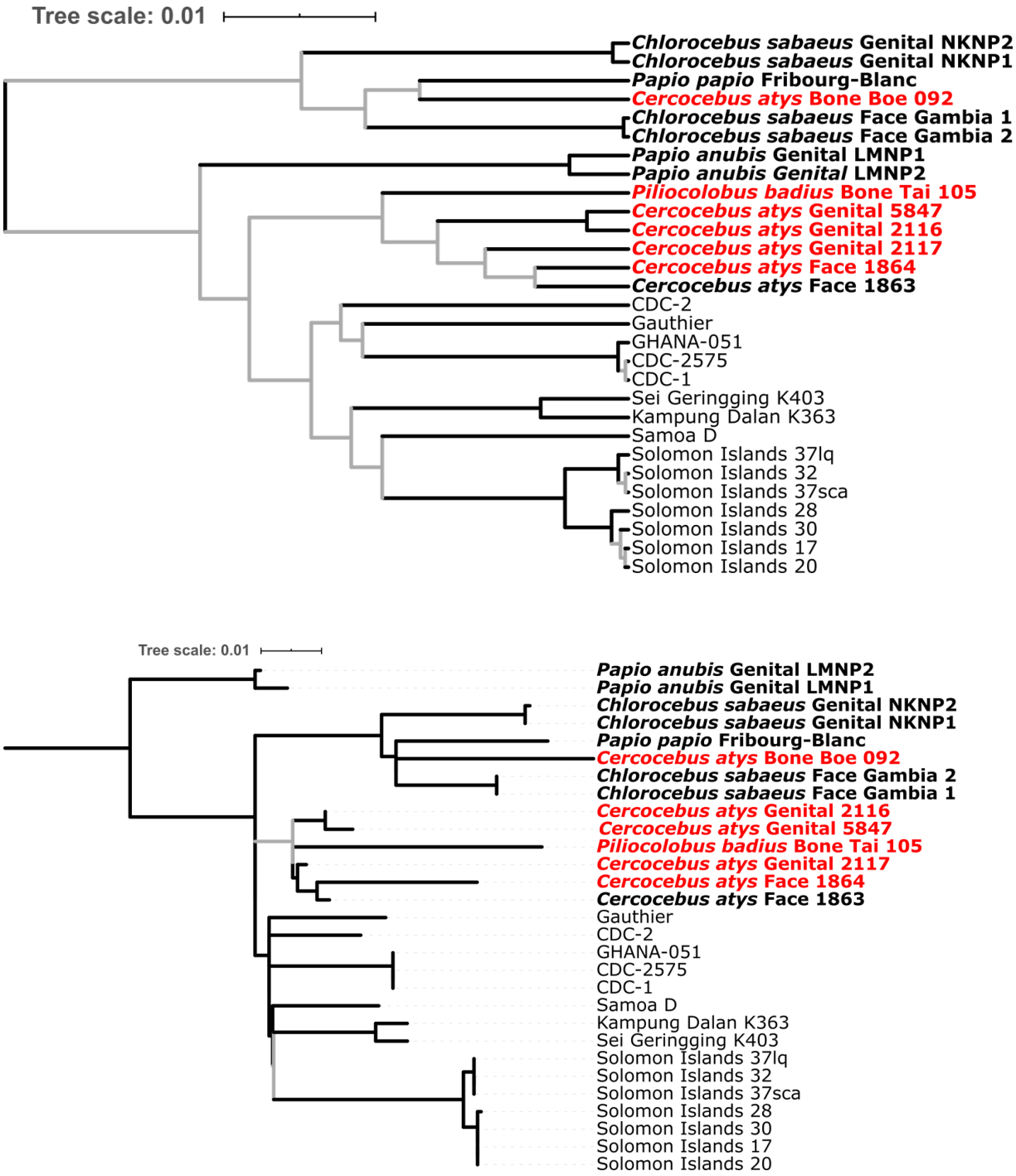
10.4 Figure S1: Supplementary trees for TP phylogenetic analysis



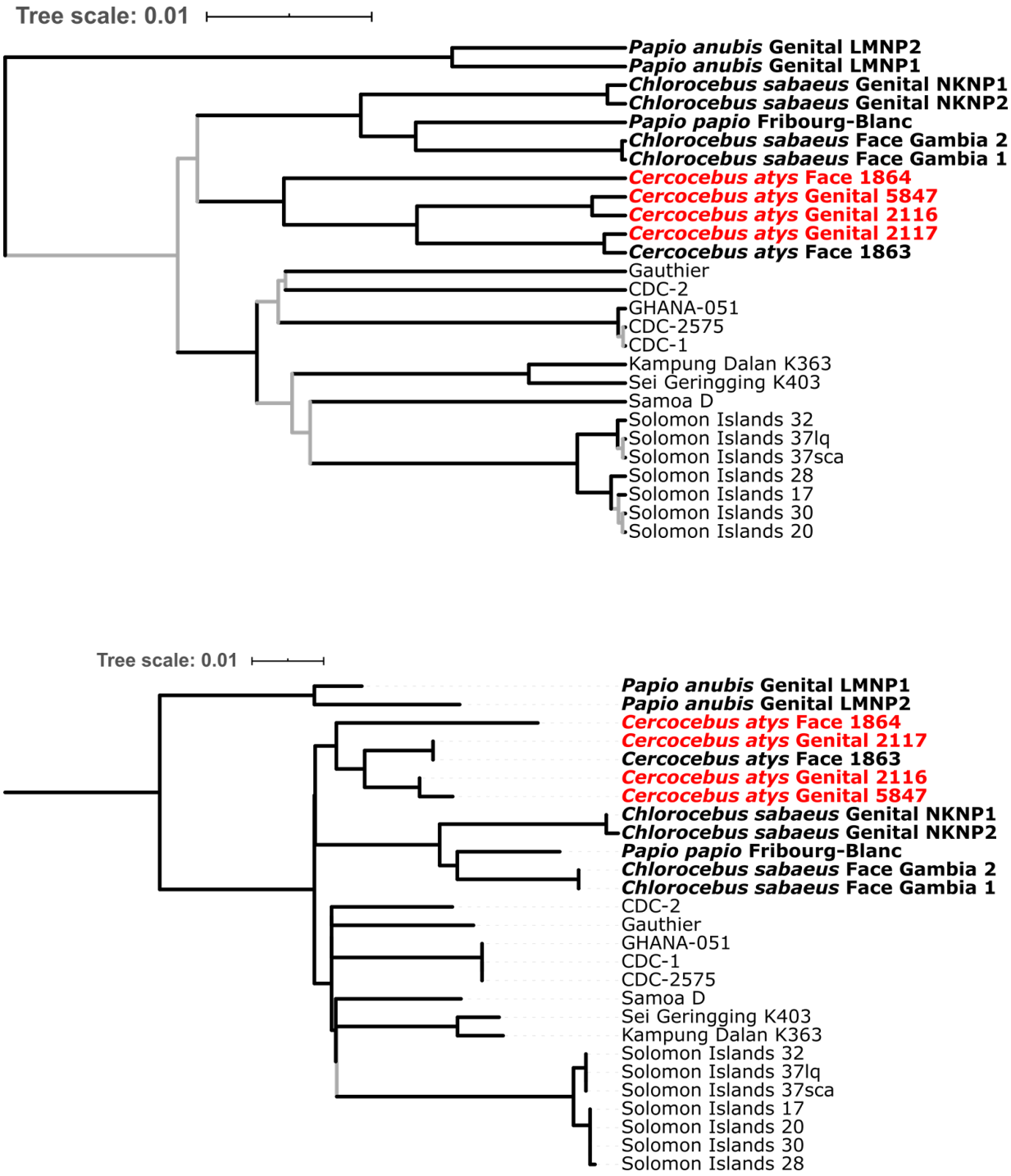
Maximum clade credibility (top: relaxed clock model assuming a birth-death process) and maximum likelihood (bottom) trees at 5X coverage and 95% threshold. All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red. Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree are indicated in gray. The scale shows nucleotide substitutions per variable site.



Maximum clade credibility (top: relaxed clock model assuming a birth-death process) and maximum likelihood (bottom) trees at 10X coverage and 95% threshold. All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red. Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree indicated in gray. The scale shows nucleotide substitutions per variable site.



Maximum clade credibility (top: relaxed clock model assuming a birth-death process) and maximum likelihood (bottom) trees from ingroup analysis of TPE strains at 5X coverage and 95% threshold. All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red. Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree indicated in gray. The scale shows nucleotide substitutions per variable site.



Maximum clade credibility (top: relaxed clock model assuming a birth-death process) and maximum likelihood (bottom) trees from ingroup analysis of TPE strains at 10X coverage and 95% threshold. All simian infecting strains are shown in bold with tip labels showing the species sampled, location of the lesion biopsied or swabbed and sample ID. Genomes generated in this study are shown in red. Branches supported by SH-like aLRT values < 0.90 in the maximum likelihood tree and posterior probabilities < 0.95 in the Bayesian Markov chain Monte Carlo tree indicated in gray. The scale shows nucleotide substitutions per variable site.

10.5 Table S4: Faecal, urine and fruit wedge screening of Woodstock for *M. leprae* from onset of symptoms- June 2018

Animal_ID Field_ID	day	month	year	material	PCR assay	Result (pos/neg)	Sequencing
TNP_566	5	7	2018	feces	18kDA	positive	positive
TNP_566	5	7	2018	feces	RLEP	positive	positive
TNP_566	28	7	2018	feces	18kDA	positive	positive
TNP_566	28	7	2018	feces	RLEP	positive	positive
TNP_566	5	8	2018	feces	18kDA	positive	positive
TNP_566	5	8	2018	feces	RLEP	positive	positive
TNP_566	11	8	2018	feces	18kDA	positive	positive
TNP_566	11	8	2018	feces	RLEP	positive	positive
TNP_566	28	8	2018	feces	18kDA	positive	positive
TNP_566	28	8	2018	feces	RLEP	positive	positive
TNP_566	29	8	2018	feces	18kDA	positive	positive
TNP_566	29	8	2018	feces	RLEP	positive	positive
TNP_566	3	1	2019	feces	18kDA	positive	positive
TNP_566	3	1	2019	feces	RLEP	positive	positive
TNP_566	14	1	2019	feces	18kDA	positive	positive
TNP_566	14	1	2019	feces	RLEP	positive	positive
TNP_566	23	1	2019	feces	18kDA	negative	negative
TNP_566	23	1	2019	feces	RLEP	positive	positive
TNP_566	31	1	2019	feces	18kDA	positive	positive
TNP_566	31	1	2019	feces	RLEP	positive	positive
TNP_566	7	2	2019	feces	18kDA	positive	positive
TNP_566	7	2	2019	feces	RLEP	positive	positive
TNP_566	3	1	2019	urine	18kDA	negative	negative
TNP_566	3	1	2019	urine	RLEP	negative	negative
TNP_566	14	1	2019	urine	18kDA	negative	negative
TNP_566	14	1	2019	urine	RLEP	negative	negative
TNP_566	23	1	2019	urine	18kDA	negative	negative
TNP_566	23	1	2019	urine	RLEP	positive	positive
TNP_566	31	1	2019	urine	18kDA	negative	negative
TNP_566	31	1	2019	urine	RLEP	negative	negative
TNP_566	1	12	2018	fruit wedge	18kDA	negative	negative
TNP_566	1	12	2018	fruit wedge	RLEP	negative	negative

10.6 Table S5: Faecal screening of Woodstock for *M. leprae* for samples collected between 2014 and June 2018

Animal_ID Field_ID	day	month	year	material	PCR assay	Result (pos/neg)	Sequencing
TNP_566	18	1	2014	feces	18kDA	negative	negative
TNP_566	18	1	2014	feces	RLEP	negative	negative
TNP_566	8	3	2014	feces	18kDA	negative	negative
TNP_566	8	3	2014	feces	RLEP	negative	negative
TNP_566	12	7	2014	feces	18kDA	negative	negative
TNP_566	12	7	2014	feces	RLEP	negative	negative
TNP_566	3	9	2014	feces	18kDA	negative	negative
TNP_566	3	9	2014	feces	RLEP	negative	negative
TNP_566	15	1	2015	feces	18kDA	negative	negative
TNP_566	15	1	2015	feces	RLEP	negative	negative
TNP_566	26	4	2015	feces	18kDA	negative	negative
TNP_566	26	4	2015	feces	RLEP	negative	negative
TNP_566	14	9	2015	feces	18kDA	negative	negative
TNP_566	14	9	2015	feces	RLEP	negative	negative
TNP_566	24	2	2017	feces	18kDA	negative	negative
TNP_566	24	2	2017	feces	RLEP	negative	negative
TNP_566	21	3	2017	feces	18kDA	negative	negative
TNP_566	21	3	2017	feces	RLEP	negative	negative
TNP_566	28	6	2018	feces	18kDA	negative	negative
TNP_566	28	6	2018	feces	RLEP	positive	positive

10.7 Table S6: Faecal screening of Zora for *M. leprae* for samples between March 2001 and June 2009

Animal_ID Field_ID	day	month	year	material	PCR assay	Result (pos/neg)	Sequencing
TNP_418	11	3	2001	feces	18kDA	negative	negative
TNP_418	11	3	2001	feces	RLEP	negative	negative
TNP_418	27	4	2001	feces	18kDA	negative	negative
TNP_418	27	4	2001	feces	RLEP	negative	negative
TNP_418	7	6	2001	feces	18kDA	negative	negative
TNP_418	7	6	2001	feces	RLEP	negative	negative
TNP_418	23	9	2001	feces	18kDA	negative	negative
TNP_418	23	9	2001	feces	RLEP	negative	negative
TNP_418	13	11	2001	feces	18kDA	negative	negative
TNP_418	13	11	2001	feces	RLEP	negative	negative
TNP_418	7	1	2002	feces	18kDA	negative	negative
TNP_418	7	1	2002	feces	RLEP	negative	negative

TNP_418	25	3	2002	feces	18kDA	negative	negative
TNP_418	25	3	2002	feces	RLEP	negative	negative
TNP_418	8	6	2002	feces	18kDA	negative	negative
TNP_418	8	6	2002	feces	RLEP	positive	positive
TNP_418	17	8	2002	feces	18kDA	negative	negative
TNP_418	17	8	2002	feces	RLEP	positive	positive
TNP_418	8	7	2003	feces	18kDA	negative	negative
TNP_418	8	7	2003	feces	RLEP	positive	positive
TNP_418	13	12	2003	feces	18kDA	negative	negative
TNP_418	13	12	2003	feces	RLEP	negative	negative
TNP_418	5	1	2004	feces	18kDA	negative	negative
TNP_418	5	1	2004	feces	RLEP	negative	negative
TNP_418	12	3	2004	feces	18kDA	negative	negative
TNP_418	12	3	2004	feces	RLEP	positive	positive
TNP_418	13	4	2004	feces	18kDA	negative	negative
TNP_418	13	4	2004	feces	RLEP	positive	positive
TNP_418	28	5	2004	feces	18kDA	negative	negative
TNP_418	28	5	2004	feces	RLEP	positive	positive
TNP_418	1	6	2004	feces	18kDA	negative	negative
TNP_418	1	6	2004	feces	RLEP	negative	negative
TNP_418	16	8	2004	feces	18kDA	negative	negative
TNP_418	16	8	2004	feces	RLEP	negative	negative
TNP_418	29	10	2004	feces	18kDA	negative	negative
TNP_418	29	10	2004	feces	RLEP	positive	positive
TNP_418	2	6	2005	feces	18kDA	positive	positive
TNP_418	2	6	2005	feces	RLEP	positive	positive
TNP_418	30	8	2005	feces	18kDA	negative	negative
TNP_418	30	8	2005	feces	RLEP	positive	positive
TNP_418	17	9	2005	feces	18kDA	positive	positive
TNP_418	17	9	2005	feces	RLEP	positive	positive
TNP_418	1	12	2005	feces	18kDA	negative	negative
TNP_418	1	12	2005	feces	RLEP	negative	negative
TNP_418	11	2	2006	feces	18kDA	negative	negative
TNP_418	11	2	2006	feces	RLEP	positive	positive
TNP_418	27	2	2006	feces	18kDA	negative	negative
TNP_418	27	2	2006	feces	RLEP	positive	positive
TNP_418	30	6	2006	feces	18kDA	positive	positive
TNP_418	30	6	2006	feces	RLEP	positive	positive
TNP_418	16	8	2006	feces	18kDA	positive	positive
TNP_418	16	8	2006	feces	RLEP	positive	positive
TNP_418	19	9	2006	feces	18kDA	positive	positive
TNP_418	19	9	2006	feces	RLEP	positive	positive
TNP_418	14	12	2006	feces	18kDA	positive	positive
TNP_418	14	12	2006	feces	RLEP	positive	positive
TNP_418	6	2	2007	feces	18kDA	negative	negative

TNP_418	6	2	2007	feces	RLEP	positive	positive
TNP_418	9	4	2007	feces	18kDA	negative	negative
TNP_418	9	4	2007	feces	RLEP	positive	positive
TNP_418	1	8	2007	feces	18kDA	positive	positive
TNP_418	1	8	2007	feces	RLEP	positive	positive
TNP_418	4	12	2007	feces	18kDA	positive	positive
TNP_418	4	12	2007	feces	RLEP	positive	positive
TNP_418	5	1	2008	feces	18kDA	positive	positive
TNP_418	5	1	2008	feces	RLEP	positive	positive
TNP_418	2	3	2008	feces	18kDA	positive	positive
TNP_418	2	3	2008	feces	RLEP	positive	positive
TNP_418	16	9	2008	feces	18kDA	positive	positive
TNP_418	16	9	2008	feces	RLEP	positive	positive
TNP_418	15	12	2008	feces	18kDA	positive	positive
TNP_418	15	12	2008	feces	RLEP	positive	positive
TNP_418	3	3	2009	feces	18kDA	positive	positive
TNP_418	3	3	2009	feces	RLEP	positive	positive
TNP_418	1	6	2009	feces	18kDA	positive	positive
TNP_418	1	6	2009	feces	RLEP	positive	positive

10.8 Table S7: Organ screening Zora necropsy samples for *M. leprae*

Animal_ID Field_ID	day	month	year	material	PCR assay	Result (pos/neg)	Sequencing
TNP_418.1	8	6	2009	Spleen	18kDA	positive	positive
TNP_418.1	8	6	2009	Spleen	RLEP	positive	positive
TNP_418.2	8	6	2009	lung	18kDA	positive	positive
TNP_418.2	8	6	2009	lung	RLEP	positive	positive
TNP_418.3	8	6	2009	liver	18kDA	positive	positive
TNP_418.3	8	6	2009	liver	RLEP	positive	positive
TNP_418.4	8	6	2009	Mesenteric LN	18kDA	positive	positive
TNP_418.4	8	6	2009	Mesenteric LN	RLEP	positive	positive
TNP_418.5	8	6	2009	LN abscess	18kDA	positive	positive
TNP_418.5	8	6	2009	LN abscess	RLEP	positive	positive
TNP_418.6	8	6	2009	Kidney	18kDA	negative	negative
TNP_418.6	8	6	2009	Kidney	RLEP	positive	positive
TNP_418.7	8	6	2009	heart	18kDA	negative	negative
TNP_418.7	8	6	2009	heart	RLEP	positive	positive
TNP_418.8	8	6	2009	muscle	18kDA	negative	negative
TNP_418.8	8	6	2009	muscle	RLEP	negative	positive
TNP_418.9	8	6	2009	colon	18kDA	negative	negative
TNP_418.9	8	6	2009	colon	RLEP	negative	negative

10.9 Table S8: Screening of all necropsy samples collected in Täi National Park, Côte d'Ivoire, since 2001 for *M. leprae*

Animal_ID	day	month	year	material	PCR assay	Gel_result	sequence_result
1	11	12	2001	spleen	18kDA	negative	
1	11	12	2001	spleen	RLEP	negative	
2	14	2	2002	spleen	18kDA	negative	
2	14	2	2002	spleen	RLEP	negative	
4	21	10	2001	spleen	18kDA	negative	
4	21	10	2001	spleen	RLEP	negative	
5	10	5	1999	spleen	18kDA	negative	
5	10	5	1999	spleen	RLEP	negative	
6	8	6	1999	spleen	18kDA	negative	
6	8	6	1999	spleen	RLEP	negative	
7	14	5	1999	spleen	18kDA	negative	
7	14	5	1999	spleen	RLEP	negative	
8	15	12	1998	spleen	18kDA	negative	
8	15	12	1998	spleen	RLEP	negative	
10	13	2	2002	spleen	18kDA	negative	
10	13	2	2002	spleen	RLEP	negative	
11	26	9	2000	spleen	18kDA	negative	
11	26	9	2000	spleen	RLEP	negative	
14	13	6	2002	spleen	18kDA	negative	
14	13	6	2002	spleen	RLEP	negative	
20	19	3	2004	spleen	18kDA	negative	
20	19	3	2004	spleen	RLEP	negative	
21	10	3	2004	spleen	18kDA	negative	
21	10	3	2004	spleen	RLEP	negative	
22	10	3	2004	spleen	18kDA	negative	
22	10	3	2004	spleen	RLEP	negative	
76	7	2	2006	spleen	18kDA	negative	
76	7	2	2006	spleen	RLEP	negative	
77	10	2	2006	spleen	18kDA	negative	
77	10	2	2006	spleen	RLEP	negative	
78	9	2	2006	spleen	18kDA	negative	
78	9	2	2006	spleen	RLEP	negative	
259	7	4	2008	spleen	18kDA	negative	
259	7	4	2008	spleen	RLEP	negative	
361	19	4	2009	spleen	18kDA	negative	
361	19	4	2009	spleen	RLEP	negative	
363	22	4	2009	spleen	18kDA	negative	
363	22	4	2009	spleen	RLEP	negative	

364	19	4	2009	spleen	18kDA	negative	
364	19	4	2009	spleen	RLEP	negative	
418	6	8	2009	spleen	18kDA	positive	positive
418	6	8	2009	spleen	RLEP	positive	positive
429	8	12	2009	spleen	18kDA	negative	
429	8	12	2009	spleen	RLEP	negative	
557	7	12	2009	spleen	18kDA	negative	
557	7	12	2009	spleen	RLEP	negative	
558	7	12	2009	spleen	18kDA	negative	
558	7	12	2009	spleen	RLEP	negative	
559	11	12	2009	spleen	18kDA	negative	
559	11	12	2009	spleen	RLEP	negative	
560	2	12	2009	spleen	18kDA	negative	
560	2	12	2009	spleen	RLEP	negative	
696	12	4	2013	spleen	18kDA	negative	
696	12	4	2013	spleen	RLEP	negative	
797	22	8	2011	spleen	18kDA	negative	
797	22	8	2011	spleen	RLEP	negative	
798	22	8	2011	spleen	18kDA	negative	
798	22	8	2011	spleen	RLEP	negative	
845	24	8	2011	spleen	18kDA	negative	
845	24	8	2011	spleen	RLEP	negative	
846	15	1	2012	spleen	18kDA	negative	
846	15	1	2012	spleen	RLEP	negative	
1697	14	4	2013	spleen	18kDA	negative	
1697	14	4	2013	spleen	RLEP	negative	
1858	8	1	2014	spleen	18kDA	negative	
1858	8	1	2014	spleen	RLEP	negative	
1885	16	6	2014	spleen	18kDA	negative	
1885	16	6	2014	spleen	RLEP	negative	
1900	2	5	2014	spleen	18kDA	negative	
1900	2	5	2014	spleen	RLEP	negative	
2105	21	7	2015	spleen	18kDA	negative	
2105	21	7	2015	spleen	RLEP	negative	
2112	5	3	2016	spleen	18kDA	negative	
2112	5	3	2016	spleen	RLEP	negative	
2114	10	11	2015	spleen	18kDA	negative	
2114	10	11	2015	spleen	RLEP	negative	
2134	7	3	2016	spleen	18kDA	negative	
2134	7	3	2016	spleen	RLEP	negative	
2189	13	3	2017	spleen	18kDA	negative	
2189	13	3	2017	spleen	RLEP	negative	

10.10 Table S9: Faecal screening for *M. leprae* for samples collected in Cantanhez National Park, Guinea Bissau

Faecal ID	site	material	Analysis	Result (pos/neg)	Sequencing
1	Caiquene-Cadique	feces	18kDA	negative	
1	Caiquene-Cadique	feces	RLEP	negative	
2	Caiquene-Cadique	feces	18kDA	negative	
2	Caiquene-Cadique	feces	RLEP	negative	
3	Caiquene-Cadique	feces	18kDA	negative	
3	Caiquene-Cadique	feces	RLEP	negative	
4	Caiquene-Cadique	feces	18kDA	negative	
4	Caiquene-Cadique	feces	RLEP	negative	
5	Caiquene-Cadique	feces	18kDA	negative	
5	Caiquene-Cadique	feces	RLEP	negative	
6	Caiquene-Cadique	feces	18kDA	negative	
6	Caiquene-Cadique	feces	RLEP	negative	
7	Caiquene-Cadique	feces	18kDA	negative	
7	Caiquene-Cadique	feces	RLEP	negative	
8	Caiquene-Cadique	feces	18kDA	negative	
8	Caiquene-Cadique	feces	RLEP	negative	
9	Caiquene-Cadique	feces	18kDA	negative	
9	Caiquene-Cadique	feces	RLEP	negative	
10	Caiquene-Cadique	feces	18kDA	negative	
10	Caiquene-Cadique	feces	RLEP	negative	
11	Caiquene-Cadique	feces	18kDA	negative	
11	Caiquene-Cadique	feces	RLEP	negative	
12	Caiquene-Cadique	feces	18kDA	negative	
12	Caiquene-Cadique	feces	RLEP	negative	
13	Caiquene-Cadique	feces	18kDA	negative	
13	Caiquene-Cadique	feces	RLEP	negative	
14	Caiquene-Cadique	feces	18kDA	negative	
14	Caiquene-Cadique	feces	RLEP	negative	
15	Caiquene-Cadique	feces	18kDA	negative	
15	Caiquene-Cadique	feces	RLEP	negative	
16	Caiquene-Cadique	feces	18kDA	negative	
16	Caiquene-Cadique	feces	RLEP	negative	
17	Caiquene-Cadique	feces	18kDA	negative	
17	Caiquene-Cadique	feces	RLEP	negative	
18	Caiquene-Cadique	feces	18kDA	negative	
18	Caiquene-Cadique	feces	RLEP	negative	
19	Caiquene-Cadique	feces	18kDA	negative	
19	Caiquene-Cadique	feces	RLEP	negative	
20	Caiquene-Cadique	feces	18kDA	negative	
20	Caiquene-Cadique	feces	RLEP	negative	

21	Caiquene-Cadique	feces	18kDA	negative	
21	Caiquene-Cadique	feces	RLEP	negative	
22	Caiquene-Cadique	feces	18kDA	negative	
22	Caiquene-Cadique	feces	RLEP	negative	
23	Caiquene-Cadique	feces	18kDA	negative	
23	Caiquene-Cadique	feces	RLEP	negative	
24	Caiquene-Cadique	feces	18kDA	negative	
24	Caiquene-Cadique	feces	RLEP	negative	
25	Caiquene-Cadique	feces	18kDA	negative	
25	Caiquene-Cadique	feces	RLEP	negative	
26	Caiquene-Cadique	feces	18kDA	negative	
26	Caiquene-Cadique	feces	RLEP	positive	positive
27	Caiquene-Cadique	feces	18kDA	negative	
27	Caiquene-Cadique	feces	RLEP	negative	
28	Caiquene-Cadique	feces	18kDA	negative	
28	Caiquene-Cadique	feces	RLEP	negative	
29	Caiquene-Cadique	feces	18kDA	negative	
29	Caiquene-Cadique	feces	RLEP	negative	
30	Caiquene-Cadique	feces	18kDA	negative	
30	Caiquene-Cadique	feces	RLEP	negative	
31	Caiquene-Cadique	feces	18kDA	negative	
31	Caiquene-Cadique	feces	RLEP	negative	
32	Caiquene-Cadique	feces	18kDA	negative	
32	Caiquene-Cadique	feces	RLEP	negative	
33	Caiquene-Cadique	feces	18kDA	negative	
33	Caiquene-Cadique	feces	RLEP	negative	
34	Caiquene-Cadique	feces	18kDA	negative	
34	Caiquene-Cadique	feces	RLEP	negative	
35	Caiquene-Cadique	feces	18kDA	negative	
35	Caiquene-Cadique	feces	RLEP	negative	
36	Caiquene-Cadique	feces	18kDA	negative	
36	Caiquene-Cadique	feces	RLEP	negative	
37	Caiquene-Cadique	feces	18kDA	negative	
37	Caiquene-Cadique	feces	RLEP	negative	
38	Caiquene-Cadique	feces	18kDA	negative	
38	Caiquene-Cadique	feces	RLEP	negative	
39	Caiquene-Cadique	feces	18kDA	negative	
39	Caiquene-Cadique	feces	RLEP	negative	
40	Caiquene-Cadique	feces	18kDA	negative	
40	Caiquene-Cadique	feces	RLEP	negative	
41	Caiquene-Cadique	feces	18kDA	negative	
41	Caiquene-Cadique	feces	RLEP	negative	
42	Caiquene-Cadique	feces	18kDA	negative	
42	Caiquene-Cadique	feces	RLEP	negative	
43	Caiquene-Cadique	feces	18kDA	negative	

43	Caiquene-Cadique	feces	RLEP	negative	
44	Caiquene-Cadique	feces	18kDA	negative	
44	Caiquene-Cadique	feces	RLEP	negative	
45	Caiquene-Cadique	feces	18kDA	negative	
45	Caiquene-Cadique	feces	RLEP	negative	
46	Caiquene-Cadique	feces	18kDA	negative	
46	Caiquene-Cadique	feces	RLEP	negative	
47	Caiquene-Cadique	feces	18kDA	negative	
47	Caiquene-Cadique	feces	RLEP	negative	
48	Caiquene-Cadique	feces	18kDA	negative	
48	Caiquene-Cadique	feces	RLEP	negative	
49	Caiquene-Cadique	feces	18kDA	negative	
49	Caiquene-Cadique	feces	RLEP	negative	
50	Caiquene-Cadique	feces	18kDA	negative	
50	Caiquene-Cadique	feces	RLEP	negative	
51	Caiquene-Cadique	feces	18kDA	negative	
51	Caiquene-Cadique	feces	RLEP	negative	
52	Caiquene-Cadique	feces	18kDA	negative	
52	Caiquene-Cadique	feces	RLEP	negative	
53	Caiquene-Cadique	feces	18kDA	negative	
53	Caiquene-Cadique	feces	RLEP	negative	
54	Caiquene-Cadique	feces	18kDA	negative	
54	Caiquene-Cadique	feces	RLEP	negative	
55	Caiquene-Cadique	feces	18kDA	negative	
55	Caiquene-Cadique	feces	RLEP	negative	
56	Caiquene-Cadique	feces	18kDA	negative	
56	Caiquene-Cadique	feces	RLEP	negative	
57	Caiquene-Cadique	feces	18kDA	negative	
57	Caiquene-Cadique	feces	RLEP	negative	
58	Caiquene-Cadique	feces	18kDA	negative	
58	Caiquene-Cadique	feces	RLEP	negative	
59	Caiquene-Cadique	feces	18kDA	negative	
59	Caiquene-Cadique	feces	RLEP	negative	
60	Caiquene-Cadique	feces	18kDA	negative	
60	Caiquene-Cadique	feces	RLEP	negative	
61	Caiquene-Cadique	feces	18kDA	negative	
61	Caiquene-Cadique	feces	RLEP	negative	
62	Caiquene-Cadique	feces	18kDA	negative	
62	Caiquene-Cadique	feces	RLEP	negative	
63	Caiquene-Cadique	feces	18kDA	negative	
63	Caiquene-Cadique	feces	RLEP	negative	
64	Caiquene-Cadique	feces	18kDA	positive	positive
64	Caiquene-Cadique	feces	RLEP	positive	positive
65	Caiquene-Cadique	feces	18kDA	negative	
65	Caiquene-Cadique	feces	RLEP	negative	

66	Caiquene-Cadique	feces	18kDA	negative	
66	Caiquene-Cadique	feces	RLEP	negative	
67	Caiquene-Cadique	feces	18kDA	negative	
67	Caiquene-Cadique	feces	RLEP	negative	
68	Caiquene-Cadique	feces	18kDA	negative	
68	Caiquene-Cadique	feces	RLEP	positive	positive
69	Caiquene-Cadique	feces	18kDA	negative	
69	Caiquene-Cadique	feces	RLEP	negative	
70	Caiquene-Cadique	feces	18kDA	negative	
70	Caiquene-Cadique	feces	RLEP	negative	
71	Caiquene-Cadique	feces	18kDA	negative	
71	Caiquene-Cadique	feces	RLEP	negative	
72	Caiquene-Cadique	feces	18kDA	negative	
72	Caiquene-Cadique	feces	RLEP	negative	
73	Caiquene-Cadique	feces	18kDA	negative	
73	Caiquene-Cadique	feces	RLEP	negative	
74	Caiquene-Cadique	feces	18kDA	negative	
74	Caiquene-Cadique	feces	RLEP	negative	
75	Caiquene-Cadique	feces	18kDA	negative	
75	Caiquene-Cadique	feces	RLEP	negative	
76	Caiquene-Cadique	feces	18kDA	negative	
76	Caiquene-Cadique	feces	RLEP	negative	
77	Caiquene-Cadique	feces	18kDA	negative	
77	Caiquene-Cadique	feces	RLEP	negative	
78	Caiquene-Cadique	feces	18kDA	negative	
78	Caiquene-Cadique	feces	RLEP	negative	
79	Caiquene-Cadique	feces	18kDA	negative	
79	Caiquene-Cadique	feces	RLEP	negative	
80	Caiquene-Cadique	feces	18kDA	negative	
80	Caiquene-Cadique	feces	RLEP	negative	
81	Caiquene-Cadique	feces	18kDA	negative	
81	Caiquene-Cadique	feces	RLEP	negative	
82	Caiquene-Cadique	feces	18kDA	negative	
82	Caiquene-Cadique	feces	RLEP	negative	
83	Caiquene-Cadique	feces	18kDA	negative	
83	Caiquene-Cadique	feces	RLEP	negative	
84	Caiquene-Cadique	feces	18kDA	negative	
84	Caiquene-Cadique	feces	RLEP	negative	
85	Caiquene-Cadique	feces	18kDA	negative	
85	Caiquene-Cadique	feces	RLEP	negative	
PN 1 L	Lautchande	feces	18kDA	negative	
PN 1 L	Lautchande	feces	RLEP	negative	
PN 2 L	Lautchande	feces	18kDA	negative	
PN 2 L	Lautchande	feces	RLEP	negative	
PN 3 L	Lautchande	feces	18kDA	negative	

PN 3 L	Lautchande	feces	RLEP	negative	
PN 4 L	Lautchande	feces	18kDA	negative	
PN 4 L	Lautchande	feces	RLEP	negative	
PN 5 L	Lautchande	feces	18kDA	negative	
PN 5 L	Lautchande	feces	RLEP	negative	
PN 6 L	Lautchande	feces	18kDA	negative	
PN 6 L	Lautchande	feces	RLEP	negative	
PN 7 L	Lautchande	feces	18kDA	negative	
PN 7 L	Lautchande	feces	RLEP	negative	
PN 8 L	Lautchande	feces	18kDA	negative	
PN 8 L	Lautchande	feces	RLEP	negative	
PN 9 L	Lautchande	feces	18kDA	negative	
PN 9 L	Lautchande	feces	RLEP	negative	
PN 10 L	Lautchande	feces	18kDA	negative	
PN 10 L	Lautchande	feces	RLEP	negative	
PN 11 L	Lautchande	feces	18kDA	negative	
PN 11 L	Lautchande	feces	RLEP	negative	
PN 12 L	Lautchande	feces	18kDA	negative	
PN 12 L	Lautchande	feces	RLEP	negative	
PN 13 L	Lautchande	feces	18kDA	negative	
PN 13 L	Lautchande	feces	RLEP	negative	
PN 14 L	Lautchande	feces	18kDA	negative	
PN 14 L	Lautchande	feces	RLEP	negative	
PN 15 L	Lautchande	feces	18kDA	negative	
PN 15 L	Lautchande	feces	RLEP	negative	
PN Cz 1	Cambeqze	feces	18kDA	negative	
PN Cz 1	Cambeqze	feces	RLEP	negative	
PN Cz 2	Cambeqze	feces	18kDA	negative	
PN Cz 2	Cambeqze	feces	RLEP	negative	
PN Cz 3	Cambeqze	feces	18kDA	negative	
PN Cz 3	Cambeqze	feces	RLEP	negative	
PN Cz 4	Cambeqze	feces	18kDA	negative	
PN Cz 4	Cambeqze	feces	RLEP	negative	
PN Cz 5	Cambeqze	feces	18kDA	negative	
PN Cz 5	Cambeqze	feces	RLEP	negative	
PN Cz 6	Cambeqze	feces	18kDA	negative	
PN Cz 6	Cambeqze	feces	RLEP	negative	
PN Cz 7	Cambeqze	feces	18kDA	negative	
PN Cz 7	Cambeqze	feces	RLEP	negative	
PN Cz 8	Cambeqze	feces	18kDA	negative	
PN Cz 8	Cambeqze	feces	RLEP	negative	
PN Cz 9	Cambeqze	feces	18kDA	negative	
PN Cz 9	Cambeqze	feces	RLEP	negative	
PN Cz 10	Cambeqze	feces	18kDA	negative	
PN Cz 10	Cambeqze	feces	RLEP	negative	

PN Cz 11	Cambeqze	feces	18kDA	negative	
PN Cz 11	Cambeqze	feces	RLEP	negative	
PN Cz 12	Cambeqze	feces	18kDA	negative	
PN Cz 12	Cambeqze	feces	RLEP	negative	
PN Cz 13	Cambeqze	feces	18kDA	negative	
PN Cz 13	Cambeqze	feces	RLEP	negative	
PN Cz 14	Cambeqze	feces	18kDA	negative	
PN Cz 14	Cambeqze	feces	RLEP	negative	
PN Cz 15	Cambeqze	feces	18kDA	negative	
PN Cz 15	Cambeqze	feces	RLEP	negative	
PN Cz 16	Cambeqze	feces	18kDA	negative	
PN Cz 16	Cambeqze	feces	RLEP	negative	
PN Cz 17	Cambeqze	feces	18kDA	negative	
PN Cz 17	Cambeqze	feces	RLEP	negative	
PN Cz 18	Cambeqze	feces	18kDA	negative	
PN Cz 18	Cambeqze	feces	RLEP	negative	
X	not known	feces	18kDA	negative	
X	not known	feces	RLEP	negative	
Y	not known	feces	18kDA	negative	
Y	not known	feces	RLEP	negative	
Z	not known	feces	18kDA	negative	
Z	not known	feces	RLEP	negative	

10.11 Table S10: Published and unpublished (from collaborators) *M. leprae* genomes used in this study

SN	Isolate ID	Country	Dating	DNA extraction method	Genotype
1	S2-95034	Antilles	1995	Schuenemann et al, Science, 2013	1B
2	Bn7-39	Benin	2015	Benjak et al, Nat Com, 2018	4N
3	Bn7-41	Benin	2015	Benjak et al, Nat Com, 2018	4N
4	Bn8-46	Benin	2015	Benjak et al, Nat Com, 2018	4N
5	Bn8-47	Benin	2015	Benjak et al, Nat Com, 2018	4N
6	Bn8-51	Benin	2015	Benjak et al, Nat Com, 2018	4N
7	Bn8-52	Benin	2015	Benjak et al, Nat Com, 2018	4N
8	Bn10-71	Benin	2017	Collaborators	4N
9	Bn10-72	Benin	2017	Collaborators	4N
10	Bn7-37	Benin	2015	Collaborators	4N
11	Bn7-38	Benin	2015	Collaborators	4N
12	Bn9-59	Benin	2016	Collaborators	4N
13	Bn9-66	Benin	2016	Collaborators	4N
14	Bn9-67	Benin	2016	Collaborators	4N
15	1126-2007	Brazil	2016	Stefani et al. PNTD, 2017	4N

16	1126-2011	Brazil	2016	Stefani et al. PNTD, 2017	4P
17	2188-2007	Brazil	2016	Stefani et al. PNTD, 2017	4N/O
18	2188-2014	Brazil	2016	Stefani et al. PNTD, 2017	4N/O
19	2DDS	Brazil	2015	Benjak et al, Nat Com, 2018	4N
20	3208-2007	Brazil	2016	Stefani et al. PNTD, 2017	4N
21	3208-2015	Brazil	2016	Stefani et al. PNTD, 2017	4N
22	BP	Brazil	2013	Benjak et al, Nat Com, 2018	3I
23	Br1	Brazil	2013	Benjak et al, Nat Com, 2018	4N
24	Br14-1	Brazil	2014	Benjak et al, Nat Com, 2018	4N
25	Br14-2	Brazil	2014	Benjak et al, Nat Com, 2018	4N
26	Br14-3	Brazil	2014	Benjak et al, Nat Com, 2018	3I
27	Br14-4	Brazil	2014	Benjak et al, Nat Com, 2018	4N
28	Br14-5	Brazil	2014	Benjak et al, Nat Com, 2018	4N
29	Br2016-14	Brazil	2016	Benjak et al, Nat Com, 2018	4N
30	Br2016-15	Brazil	2016	Benjak et al, Nat Com, 2018	4N
31	Br2016-16	Brazil	2016	Benjak et al, Nat Com, 2018	4N
32	Br2016-17	Brazil	2016	Benjak et al, Nat Com, 2018	3I
33	Br2016-18	Brazil	2016	Benjak et al, Nat Com, 2018	1D
34	Br2016-19	Brazil	2016	Benjak et al, Nat Com, 2018	3I
35	Br2016-20	Brazil	2016	Benjak et al, Nat Com, 2018	3I
36	Br2016-21	Brazil	2016	Benjak et al, Nat Com, 2018	3I
37	Br2016-24	Brazil	2016	Benjak et al, Nat Com, 2018	4N
38	Br2016-26	Brazil	2016	Benjak et al, Nat Com, 2018	4N
39	Br2016-27	Brazil	2016	Benjak et al, Nat Com, 2018	4N
40	Br2016-45	Brazil	2016	Benjak et al, Nat Com, 2018	3I
41	Br2016-46	Brazil	2016	Benjak et al, Nat Com, 2018	3I
42	Br2016-47	Brazil	2016	Benjak et al, Nat Com, 2018	3I
43	Br4923	Brazil	2004	Monot et al, Nat Genet 2009	4P
44	BrMM1	Brazil	2013	Benjak et al, Nat Com, 2018	4P
45	BrMM2	Brazil	2013	Benjak et al, Nat Com, 2018	4P
46	BrMM4	Brazil	2013	Benjak et al, Nat Com, 2018	3I
47	BrMM5	Brazil	2013	Benjak et al, Nat Com, 2018	3I/2
48	Fio3	Brazil	2012	Benjak et al, Nat Com, 2018	4P
49	S10-Ch-04	China	2006	Schuenemann et al, Science, 2013	3K
50	Body-188	Czech Republic	800-1200	Schueneman et al, Plos, 2018	3M
51	Jorgen-404	Danemark	1219-1276	Schueneman et al, Plos, 2018	3I
52	Jorgen-427	Danemark	1256-1258	Schueneman et al, Plos, 2018	3I
53	Jorgen-507	Danemark	1058-1253	Schueneman et al, Plos, 2018	3K
54	Jorgen-533	Danemark	1044-1214	Schueneman et al, Plos, 2018	3I
55	Jorgen-625	Danemark	1293-1386	Schuenemann et al, Science, 2013	3I
56	Jorgen-722	Danemark	1256-1377	Schueneman et al, Plos, 2018	3I
57	Jorgen-749	Danemark	1223-1279	Schueneman et al, Plos, 2018	2F

58	Refshale-16	Danemark	1046-1163	Schuenemann et al, Science, 2013	2F
59	Brw15-10M2	England	2015	Avanzi et al, Science, 2016	3I
60	Brw15-12M2	England	2015	Avanzi et al, Science, 2016	3I
61	Brw15-1E	England	2015	Avanzi et al, Science, 2016	3I
62	Brw15-20M2	England	2015	Avanzi et al, Science, 2016	3I
63	Brw15-25E	England	2015	Avanzi et al, Science, 2016	3I
64	Brw15-5E	England	2015	Avanzi et al, Science, 2016	3I
65	GC96CU	England	415-545	Schueneman et al, Plos, 2018	3I
66	SK14	England	1000-1200	Mendum et al, BMC Genomics, 2014	2F
67	ARLP-07	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
68	ARLP-08	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
69	ARLP-10	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
70	ARLP-11	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
71	ARLP-12	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
72	ARLP-13	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
73	ARLP-14	Ethiopia	2015	Benjak et al, Nat Com, 2018	2F
74	ARLP-20	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
75	ARLP-23	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
76	ARLP-25	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
77	ARLP-27	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
78	ARLP-29	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
79	ARLP-30	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
80	ARLP-32	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
81	ARLP-37	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
82	ARLP-40	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
83	ARLP-46	Ethiopia	2015	Benjak et al, Nat Com, 2018	2F
84	ARLP-48	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
85	ARLP-49	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
86	ARLP-52	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
87	ARLP-57	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
88	ARLP-62	Ethiopia	2015	Benjak et al, Nat Com, 2018	2E
89	ARLP-63	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
90	ARLP-65	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
91	ARLP-68	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
92	ARLP-73	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
93	ARLP-74	Ethiopia	2015	Benjak et al, Nat Com, 2018	2H
94	97016	French west Indies	2000	Benjak et al, Nat Com, 2018	4N
95	1262-16	Germany (Pakistan)	2016	Benjak et al, Nat Com, 2018	1D
96	Gu4-17	Guinea	2014	Benjak et al, Nat Com, 2018	4N
97	Gu4-19L	Guinea	2014	Benjak et al, Nat Com, 2018	4O
98	Gu5-23	Guinea	2014	Benjak et al, Nat Com, 2018	4O

99	Ml2-10	Guinea	2016	Avanzi et al, CID, 2016	4N
100	Ml6-50	Guinea	2016	Avanzi et al, CID, 2016	4N
101	Ml6-55	Guinea	2016	Avanzi et al, CID, 2016	4N
102	SK11	Hungary	700-800	Schueneman et al, Plos, 2018	3K
103	S11-Inde-2	India	2000	Schuenemann et al, Science, 2013	1D
104	TN	India	1990	Cole et al, Nature, 2001	1A
105	Indonesia-1	Indonesia	2000	Benjak et al, Nat Com, 2018	1A
106	T18	Italy	600-700	Honap et al, PNTD, 2018	2F
107	CI-7	Ivory Coast	2000	Collaborators	4N
108	Airaku-3	Japan	2000	Benjak et al, Nat Com, 2018	1D
109	Amami	Japan	2000	Benjak et al, Nat Com, 2018	3K
110	Izumi	Japan	2000	Benjak et al, Nat Com, 2018	3K
111	Kanazawa	Japan	2000	Benjak et al, Nat Com, 2018	3K
112	Kitasato	Japan	2000	Benjak et al, Nat Com, 2018	3K
113	Kusatsu-6	Japan	2000	Benjak et al, Nat Com, 2018	3K
114	Kyoto-1	Japan	2000	Benjak et al, Nat Com, 2018	3K
115	Kyoto-2	Japan	2000	Singh et al, PNAS, 2015	3K
116	LRC-1A	Japan	2000	Benjak et al, Nat Com, 2018	1A
117	Oku-4	Japan	2000	Benjak et al, Nat Com, 2018	3K
118	Ryukyu-2	Japan	2000	Benjak et al, Nat Com, 2018	3K
119	Tsukuba-1	Japan	2000	Benjak et al, Nat Com, 2018	3K
120	Zensho-2	Japan	2000	Benjak et al, Nat Com, 2018	3K
121	Zensho-4	Japan	2000	Benjak et al, Nat Com, 2018	3K
122	Zensho-5	Japan	2000	Benjak et al, Nat Com, 2018	3K
123	Zensho-9	Japan	2000	Benjak et al, Nat Com, 2018	3K
124	Korea-3-2	Korea	2000	Benjak et al, Nat Com, 2018	3K
125	2936	Malawi	2000	Benjak et al, Nat Com, 2018	1D
126	Ml10-91	Mali	2016	Benjak et al, Nat Com, 2018	4N
127	Ml10-93	Mali	2016	Benjak et al, Nat Com, 2018	4N
128	Ml10-94	Mali	2016	Benjak et al, Nat Com, 2018	4N
129	Ml10-95	Mali	2016	Benjak et al, Nat Com, 2018	4N
130	Ml10-96	Mali	2016	Benjak et al, Nat Com, 2018	4P
131	Ml10-97	Mali	2016	Benjak et al, Nat Com, 2018	4N
132	Ml10-98	Mali	2016	Benjak et al, Nat Com, 2018	4N
133	Ml10-99	Mali	2016	Benjak et al, Nat Com, 2018	4N
134	ML2-5	Mali	2012	Benjak et al, Nat Com, 2018	4N
135	Ml9-79	Mali	2014	Benjak et al, Nat Com, 2018	4N
136	Ml9-80	Mali	2014	Benjak et al, Nat Com, 2018	4-0
137	Ml9-81	Mali	2014	Benjak et al, Nat Com, 2018	4N
138	Ml9-82	Mali	2014	Benjak et al, Nat Com, 2018	4O
139	Ml9-83	Mali	2014	Benjak et al, Nat Com, 2018	4N
140	Ml9-84	Mali	2014	Benjak et al, Nat Com, 2018	4N

141	Ml9-86	Mali	2014	Benjak et al, Nat Com, 2018	4N
142	Ml9-87	Mali	2014	Benjak et al, Nat Com, 2018	4O
143	S13-Ml-3-28	Mali	2013	Schuenemann et al, Science, 2013	4N
144	S14-Ml-2-07	Mali	2013	Schuenemann et al, Science, 2013	4O
145	Ml1-4	Mali	2012	Collaborators	4O
146	Ml11-100	Mali	2016	Collaborators	4N
147	ML11-101	Mali	2016	Collaborators	4N
148	Ml11-102	Mali	2016	Collaborators	4N
149	Ml11-103	Mali	2016	Collaborators	4N
150	Ml11-104	Mali	2016	Collaborators	4N
151	Ml11-105	Mali	2016	Collaborators	4N
152	ML11-106	Mali	2016	Collaborators	4N
153	Ml11-107	Mali	2016	Collaborators	4N
154	Ml12-109	Mali	2016	Collaborators	4N
155	Ml12-110	Mali	2016	Collaborators	4N
156	Ml3-17	Mali	2013	Collaborators	4O
157	Ml3-26	Mali	2013	Collaborators	4O
158	US57	Marshall Islands	2000	Benjak et al, Nat Com, 2018	3K
159	85054	Martinique	1990	Benjak et al, Nat Com, 2018	3I
160	S15-92041	Martinique	1992	Schuenemann et al, Science, 2013	3L
161	EGG	Mexico	2016	Benjak et al, Nat Com, 2018	3I
162	S9-96008	New Caledonia	1996	Schuenemann et al, Science, 2013	3K
163	Ng12-33	Niger	2015	Benjak et al, Nat Com, 2018	4N
164	Ng13-32	Niger	2015	Benjak et al, Nat Com, 2018	4N
165	Ng13-33	Niger	2015	Benjak et al, Nat Com, 2018	4-0
166	Ng14-35	Niger	2015	Benjak et al, Nat Com, 2018	1D
167	Ng15-36	Niger	2015	Benjak et al, Nat Com, 2018	4N
168	Ng15-37	Niger	2015	Benjak et al, Nat Com, 2018	4N
169	Ng16-38	Niger	2015	Benjak et al, Nat Com, 2018	4N
170	Ng17-39	Niger	2015	Benjak et al, Nat Com, 2018	4N
171	Ng12-31	Niger	2015	Collaborators	4N
172	Ng18-40	Niger	2016	Collaborators	4N
173	Ng19-41	Niger	2016	Collaborators	4N
174	Ng19-42	Niger	2016	Collaborators	4N
175	Ng21-44	Niger	2017	Collaborators	4N
176	Ng22-45	Niger	2017	Collaborators	4N
177	Ng22-47	Niger	2017	Collaborators	4N
178	Ng23-48	Niger	2017	Collaborators	4N
179	Ng23-49	Niger	2017	Collaborators	4N
180	Ng25-52	Niger	2017	Collaborators	4N
181	Ng25-53	Niger	2017	Collaborators	4N

182	Ng26-54	Niger	2017	Collaborators	4N
183	Ng26-56	Niger	2017	Collaborators	4N
184	Ng27-57	Niger	2017	Collaborators	4N
185	Ng27-58	Niger	2017	Collaborators	4N
186	Ng27-59	Niger	2017	Collaborators	4N
187	Ng27-60	Niger	2017	Collaborators	4N
188	Ng27-61	Niger	2017	Collaborators	4N
189	Ng27-62	Niger	2017	Collaborators	4N
190	Ng29-64	Niger	2017	Collaborators	4N
191	Ng29-66	Niger	2017	Collaborators	4N
192	Ng30-67	Niger	2017	Collaborators	4N
193	Pak	Pakistan	2010	Benjak et al, Nat Com, 2018	1D
194	Ch4	Philippines	2009	Honap et al, PNTD, 2018	4N/O
195	Sen1-1	Senegal	2018	Collaborators	4N
196	CM1	Sierra Leone	1994	Honap et al, PNTD, 2018	3K
197	3077	Sweden	1032-1155	Benjak et al, Nat Com, 2018	2F
198	Thai-237	Thailand	2000	Benjak et al, Nat Com, 2018	1A
199	Thai-311	Thailand	2000	Benjak et al, Nat Com, 2018	1D
200	Thai-53	Thailand	2005	Monot et al, Nat Genet 2009	1A
201	SK2	U.K	1268-1283	Schuenemann et al, Science, 2013	3I
202	SK27	U.K	1000-1200	Mendum et al, BMC Genomics, 2014	2F
203	SK8	U.K	1010-1160	Schuenemann et al, Science, 2013	2F
204	I30-W09	USA	2005	Truman et al, NEJM, 2011	3I
205	NHDP-55	USA	2004	Truman et al, NEJM, 2011	3I
206	NHDP-63	USA	2004	Monot et al, Nat Genet 2009	3I
207	NHDP-98	USA	2004	Truman et al, NEJM, 2011	3I
208	SM1	West Africa	1979	Honap et al, PNTD, 2018	4N/O
209	Ye2-3	Yemen	2014	Benjak et al, Nat Com, 2018	2E
210	Ye3s2	Yemen	2013	Benjak et al, Nat Com, 2018	1B
211	Ye4-10	Yemen	2015	Benjak et al, Nat Com, 2018	2E
212	Ye4-11	Yemen	2015	Benjak et al, Nat Com, 2018	2E
213	Ye4-12	Yemen	2015	Benjak et al, Nat Com, 2018	1D
214	Ye4-8	Yemen	2015	Benjak et al, Nat Com, 2018	1D

10.12 Table S11: List of primers used for *M. leprae* genotyping

Targeted gene	Genome position (ref TN)	Primers (5'-3')	Amplicon size	Mutation
ML0283	365373	F_CGAGGTATTTACCGCAATGG R_ TTAATTGCCAGGCTGTCCTC	201bp	T>C
ML2356	2815502	F_AAACGCTCACAGCATCACTG R_ AGAAATTGCGATTTGCGTCT	202bp	G>C
ML0411	508986	F_TTGGAAGCCGACTACGATCT R_ CTGAGCCACCGAGTCATACA	228bp	C>G
ML0049		F_CGTTTTTCGGAAGCAAACAT R_ CAGCGAAATTCCAGTCCTCT	169	C>T
ML0565		F_TCGCAAGAGTGTTCGGTATG R_ GTCCCGGTTCAGTCAGTGAT	194	C>T

10.13 Pathology report for Zora

Untersuchungsprotokoll
Dr. K. Mätz-Rensing
Deutsches Primatenzentrum
Abteilung Infektionspathologie

Tel: 0551-3851-386
Fax: 0551-3851-442
email: kmaetz@gwdg.de

Tier-Nr.: Zora
Datum: 27.12.2010
Sekant: K. Mätz-Rensing
Spezies: Pan
Geschlecht: weiblich

G-Nummer: 8352

Einsender: Dr. Leendertz, RKI
Geburtsdatum: -
Gewicht: -
Proben: Spektrum F 10%ig; -
Vorbericht:.
Siehe Anschreiben

Histologischer Befund:
Fortgeschrittene Autolyse, hgr. Gefrierartefakte, soweit beurteilbar:
ZNS: nicht untersucht;
Leber: hgr. chronisch-fibrosierende gemischtzellige Pericholangitis (3,4);
Niere: Glomerula: überwiegend o.b.B.; Tubuli: überwiegend o.b.B.; Interstitium: überwiegend o.b.B. (3)
Milz: hgr. granulomatöse Splenitis unter Ausbildung multipler Granulome. Am Aufbau der Granulome beteiligt sind mehrkernige Riesenzellen vom Langhanstyp, Histiozyten und neutrophile Granulozyten. Mithilfe der Ziehl-Neelsen Färbung sind säurefeste Stäbchen in den Histiozyten sowie in den Riesenzellen nachweisbar; ggr. follikuläre Hyperplasie (3,4,5);

Lunge: hgr. alveoläres Emphysem, ggr. Alveolarhistiozytose (5);
Herz: o.b.B. (4);
Magen: überwiegend o.b.B. (5);
Dünndarm: überwiegend o.b.B. (7;
Dickdarm: überwiegend o.b.B. (8);
Ovar: o.b.B. (6);
Uterus: o.b.B. (6):
Muskulatur: o.b.B. (6);
Haut: unregelmäßige Epidermishyperplasie, orthokeratotische Hyperkeratose, fokal herdförmige hgr. oberflächliche gemischtzellige Dermatitis (10);
Sonstiges: Ausbildung multipler unterschiedlich großer von einer dickwandigen bindegewebigen Kapsel begrenzter Granulome im Mesenterium. Am Aufbau der Granulome beteiligt sind mehrkernige Riesenzellen vom Langhanstyp, Histiozyten und neutrophile Granulozyten, zentral finden sich großflächige Nekroseherde. Mithilfe der Ziehl-Neelsen Färbung sind säurefeste Stäbchen in den Histiozyten sowie in den Riesenzellen nachweisbar. (1,2,9)
Mikrobiologie: nicht eingeleitet

Parasitologie: nicht eingeleitet

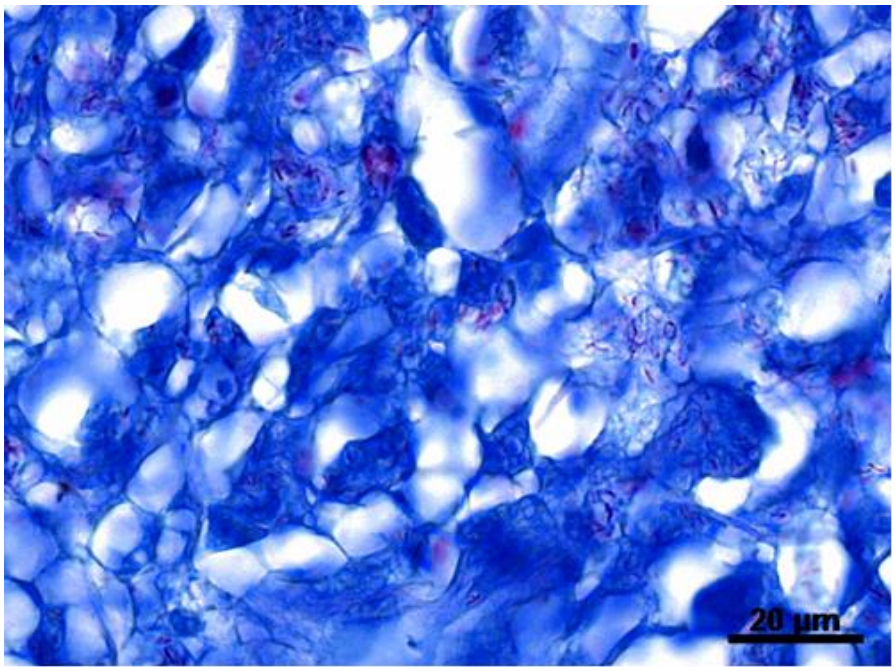
Spezialfärbungen: ,1,2,9: Giemsa.negativ, PAS-Reaktion: negativ, Ziehl-Neelsen: positiv

Zweite Benachrichtigung

Zusammenfassende Bewertung:
Bei der histologischen Untersuchung zeigte sich das Vorliegen multipler granulomatöser Entzündungsprozesse in Milz und Mesenterium. Die Granulome wiesen einen für Tuberkulose typischen Aufbau mit Beteiligung von Langhanszellen und Histiozyten auf, in denen massenhaft säurefeste Stäbchen darstellbar waren. Basierend auf diesen histologischen Befunden kann die Diagnose einer Mykobakteriose gestellt werden. Zur endgültigen Diagnostik muss der Erreger typisiert werden.
Die immunhistochemische Untersuchung unter Verwendung eines Mycobakterium bovis spezifischen Antikörpers war aufgrund massiver autolysebedingter Artefakte leider nicht auswertbar.
Nachtrag Haut: nach erneuter Präparation des restlichen Einsendematerial ergeben sich folgende Befunde:
In einzelnen nicht näher zuordnungsbaren Hautarealen fanden sich eine fokal granulomatöse bis bandartige Entzündungszellinfiltration überwiegend in Dermis und Hypodermis. Die Basalschicht der Epidermis bzw. die Epidermis war nicht alteriert oder grenzt sich klar gegen den Entzündungsprozess ab. In veränderten Arealen waren nur rudimentäre Elemente der Haarbälge, Schweißdrüsen und Talgdrüsen darstellbar.
Die entzündlichen Infiltrate setzen sich aus schaumigen bis vakuolären histiozytären Zellen und deutlich weniger lymphozytären Zellelementen zusammen. In den histiozytären oder epitheloidartigen Zellen sind einzelne oder verklumpte Bakterienaggregate, die in der Fite-Faraco Färbung säurefest reagieren, darstellbar. Das histologische Bild zusammen mit dem Nachweis von säurefesten Stäbchen ist typisch für Lepra.

Kopie an: Einsender

.....
(Dr. Kerstin Mätz-Rensing)


20 µm

11 Ethical statement

All procedures performed on the *Cercocebus atys* monkeys in Tai National Park were approved by the Ministry of Environment and Forests as well as the Ministry of Research, the Office Ivoirien des Parcs et Réserves, and the director of Tai National Park. A veterinarian carried out all sampling following good veterinarian practice. Animal welfare was considered in all procedures carried out and anesthetized animals were monitored for vital functions and remained under close supervision from the time of induction until full recovery and until the animals were able to reunite with their social group. Additionally, all fecal samples imported into Germany for the purspose to be used in this research had valid import permits from the country of origin.

12 Published work

Publications

- Gogarten, J. F., Düx, A., **Mubemba, B**., Pléh, K., Hoffmann, C. *et al.* (2019). Tropical rainforest flies carrying pathogens form stable associations with social nonhuman primates. Molecular Ecology, mec.15145. https://doi.org/10.1111/mec.1514
- **Benjamin Mubemba** and Emeline Chanove *et al*. (2020). Yaws Disease Caused by *Treponema pallidum* subsp. *pertenue* in Wild chimpanzee (Pan troglodytes verus), Guinea, 2019. **Emerging Infectious Diseases.** doi.org/10.3201/eid2606.191713
- **Benjamin Mubemba**, Jan F. Gogarten, Verena J. Schuenemann, Ariane Düx, Alexander Lang, *et al*. (2019). Geographically structured genomic diversity of non-human primate-infecting *Treponema pallidum* subsp. *pertenue.* **Microbial Genomics Journal.** doi.org/10.1099/mgen.0.000463

Posters and conference presentations

- **Mubemba Benjamin**., Gogarten Jan., Patrono Livia., Leendertz Fabian (2018). Dermatoses of non-human primates-Chair/organiser-Round able. International Primatological Society Congress. 19-25th August, 2018, Nairobi Kenya
- **Benjamin Mubemba**., Jan Gogarten., Ariane Düx., Markus Ulrich., Kamilla Pléh., Fabian Leendertz., Sebastien Calvignac-Spence (2019). Genetic diversity of Treponema pallidum subsp. pertenue infecting wild non-human primates in Taï National Park (TNP), Coted'Ivoire. International Symposium on Zoonoses Research, Berlin, Germany.
- Jan Frederik Gogarten, Ariane Düx, **Benjamin Mubemba,** Kamilla Pléh, Constanze Hoffmann, Alexander Mielke, Jonathan Müller Tiburtius, Andreas Sachse, Roman Wittig, Sébastien Calvignac-Spencer, FabianLeendertz (2019). Tropica lrainforest flies carrying pathogens form stable associations with social lnon-human primates. International Symposium on Zoonoses Research, Berlin, Germany.

13 Acknowledgements

I would like to express my heartfelt appreciation to the following institutions and people for their great assistance during this doctoral research work;

- The Robert Koch Institute for giving me an opportunity to undertake this research and for financing my stay in Germany.
- My supervisor PD. Dr. Fabian Leendertz for his guidance, criticism, ideas, on the diseertation write up and for accepting to review the dissertation
- Dr. Sébastien Calvignac-Spencer, for his very invaluable technical guidance in the laboratory analyses, in-silico analyses and critical review of draft manuscripts.
- Dr. Jan F. Gogarten for his guidance in the in-silico analyses stage and his critical review of draft manuscripts.
- Prof Verena J. Schuenemann for hosting me at the Institute of Evolutionary Medicine, University of Zurich, Zurich, Switzerland and offering me part of the laboratory skill training needed to undertake this study.
- Dr. Charlotte Avanzi for supplying the positive controls for the leprosy study and her assistance in the in-silico analyses of the leprosy comparative genomics.
- The Technicians at the Epidemiology of Highly Pathogenic Organisms, Robert Koch-Institut, Berlin, Germany; for the training they offered me at the beginning of my research work and for their passionate assistance whenever they were called upon.
- Maja Kovacev-Wegener at the Epidemiology of Highly Pathogenic Organisms, Robert Koch-Institut, Berlin, Germany; for her administrative support through and through.
- To my fellow postgraduate students at the Epidemiology of Highly Pathogenic Organisms, Robert Koch-Institut, Berlin, Germany for their worthy companionship and moral support and in some instances, others offered technical help.
- My loving wife Abigail, our children *Michelle, Yakira* and *Asher*, my parents and my sister Muyane for their unwavering support, encouragement, love and for being family.
- The Copperbelt University, Zambia; for granting me study leave to undertake my doctoral studies.
- Above all, to God almighty for his provisions, love and divine enablement

14 Selbstständigkeitserklärung

Ich erkläre, dass diese Dissertation, die hiermit an der Freien Universität für den Doktortitel *Doctor Medicinae Veterinariae* eingereicht wird, nicht zuvor von mir oder jemand anderem für den Abschluss an dieser oder einer anderen Universität eingereicht wurde und dass es sich um meine eigene Arbeit in Ausführung handelt und dass alle hierin enthaltenen Materialien ordnungsgemäß zitiert wurden.

Berlin, den 05.11.2020

Benjamin Mubemba

mbvberlin mensch und buch verlag

49,90 Euro | ISBN: 978-3-96729-080-6