

# Genomic analysis of an outbreak of bovine tuberculosis in a man-made multi-host species system: a call for action on wildlife in Brazil

Daiane A. R. Lima<sup>1</sup>, Cristina K. Zimpel<sup>2</sup>, José Salvatore Patané<sup>3</sup>, Taiana T. Silva-Pereira<sup>2</sup>, Rodrigo N. Etges<sup>4</sup>, Rudielle A. Rodrigues<sup>1</sup>, Alberto M.R. Dávila<sup>5</sup>, Cássia Ikuta<sup>6</sup>, José Ferreira Neto<sup>6</sup>, Ana Marcia Guimaraes<sup>2</sup>, and Flávio Araújo<sup>7</sup>

<sup>1</sup>Universidade Federal de Mato Grosso do Sul

<sup>2</sup>Universidade de Sao Paulo Departamento de Microbiologia

<sup>3</sup>Instituto Butantan

<sup>4</sup>Governo do Estado do Rio Grande do Sul

<sup>5</sup>Instituto Oswaldo Cruz

<sup>6</sup>Universidade de Sao Paulo Faculdade de Medicina Veterinaria e Zootecnia

<sup>7</sup>Embrapa Beef Cattle Campo Grande Mato Grosso do Sul Brazil

April 28, 2021

## Abstract

We report on a 15-year-long outbreak of bovine tuberculosis (bTB) in wildlife from a Brazilian safari park. A timeline of diagnostic events and whole-genome sequencing (WGS) of 21 *Mycobacterium bovis* isolates from deer and llamas were analyzed. Accordingly, from 2003 to 2018, at least 16 animals, from 8 species, died due to TB, which is likely an underestimated number. In three occasions since 2013, the deer presented positive tuberculin tests, leading to the park closure and culling of all deer. WGS indicated that multiple *M. bovis* strains were circulating, with at least three founding introductions since the park inauguration in 1977. Recent transmission events between nearby farms and the park were not found based on WGS. Lastly, by discussing socio-economic and environmental factors escaping current regulatory gaps that were determinant of this outbreak, we pledge for the development of a plan to report and control bTB in wildlife in Brazil.

**Title: Genomic analysis of an outbreak of bovine tuberculosis in a man-made multi-host species system: a call for action on wildlife in Brazil**

**Running Title:** Bovine tuberculosis in wildlife from a safari park

**Authors:** Daiane A. R. Lima<sup>1</sup>, Cristina K. Zimpel<sup>2,3</sup>, José Salvatore Patané<sup>4</sup>, Taiana T. Silva-Pereira<sup>2,3</sup>, Rodrigo N. Etges<sup>5</sup>, Rudielle A. Rodrigues<sup>1</sup>, Alberto M.R. Dávila<sup>6</sup>, Cássia Y. Ikuta<sup>3</sup>, José Soares Ferreira Neto<sup>3</sup>, Ana M. S. Guimarães<sup>2,3</sup>, Flávio R. Araújo<sup>7</sup>

<sup>1</sup>Federal University of Mato Grosso do Sul, School of Veterinary Medicine, Graduate Program in Veterinary Sciences, Campo Grande, Mato Grosso do Sul, Brazil

<sup>2</sup>University of São Paulo, Institute of Biomedical Sciences, Department of Microbiology, Laboratory of Applied Research in Mycobacteria, São Paulo, São Paulo, Brazil

<sup>3</sup>University of São Paulo, School of Veterinary Medicine and Animal Sciences, Department of Preventive Veterinary Medicine and Animal Health, São Paulo, São Paulo, Brazil

<sup>4</sup>Butantan Institute, Center for Bioinformatics and Computational Biology, São Paulo, São Paulo, Brazil

<sup>5</sup>Secretary of Agriculture, Livestock and Rural Development, Porto Alegre, Rio Grande do Sul, Brazil

<sup>6</sup>Fiocruz, Oswaldo Cruz Institute, Computational and Systems Biology Laboratory and Graduate Program on Biodiversity and Health, Rio de Janeiro, Rio de Janeiro, Brazil

<sup>7</sup>Embrapa Beef Cattle, Campo Grande, Mato Grosso do Sul, Brazil

Correspondence: Ana Marcia de Sá Guimarães and Flávio Ribeiro Araújo are considered senior authors:

Flávio Ribeiro Araújo, Animal Health Department, Embrapa Beef Cattle, 79106-550, Av. Rádio Maia, 830 - Vila Popular, Campo Grande MS, Brazil; E-mail: [flabio.araujo@embrapa.br](mailto:flabio.araujo@embrapa.br)

Ana Marcia de Sá Guimarães, Department of Microbiology, Institute of Biomedical Sciences, University of São Paulo, 05508-000 Av. Prof Lineu Prestes, 1374, Room 229, São Paulo, SP, Brazil. E-mail: [anamarcia@usp.br](mailto:anamarcia@usp.br)

## Summary

We report on a 15-year-long outbreak of bovine tuberculosis (bTB) in wildlife from a Brazilian safari park. A timeline of diagnostic events and whole-genome sequencing (WGS) of 21 *Mycobacterium bovis* isolates from deer and llamas were analyzed. Accordingly, from 2003 to 2018, at least 16 animals, from 8 species, died due to TB, which is likely an underestimated number. In three occasions since 2013, the deer presented positive tuberculin tests, leading to the park closure and culling of all deer. WGS indicated that multiple *M. bovis* strains were circulating, with at least three founding introductions since the park inauguration in 1977. Recent transmission events between nearby farms and the park were not found based on WGS. Lastly, by discussing socio-economic and environmental factors escaping current regulatory gaps that were determinant of this outbreak, we pledge for the development of a plan to report and control bTB in wildlife in Brazil.

**KEYWORDS:** bovine tuberculosis, wildlife, *Mycobacterium bovis*, public health policies, genomics

## 1 INTRODUCTION

Bovine tuberculosis (bTB) is an OIE (World Organisation for Animal Health) notifiable disease with major impact on livestock and wildlife (OIE, 2011). The disease is present in most countries, but with variable prevalence (Cousins, 2018). While certain developed nations have significantly reduced or eradicated the disease, developing countries with absent or inefficient control programs struggle to contain it, leading to a significant toll on livestock producers and/or devastating effects on wildlife (O'Reilly, Daborn, 1995; Ayele et al. 2004; Ferreira Neto, 2019). bTB can also be transmitted to humans through close contact with infected animals or the consumption of unpasteurized milk (Olea-Popelka et al., 2017). It is estimated that 147,000 new cases and 12,500 deaths occur due to zoonotic TB every year (WHO, 2017). Recently, the OIE, WHO (World Health Organization), The Union, and FAO (Food and Agriculture Organization) launched the roadmap for zoonotic TB, urging stakeholders to apply the One Health approach to tackle zoonotic TB and contribute to achieve the sustainable development goal of ending the global TB epidemic by 2030 (WHO, 2017).

Brazil is the second largest beef producer in the world and the number one exporter. In 2001, the Brazilian government launched the National Program for the Control and Eradication of Animal Brucellosis and Tuberculosis (PNCEBT), setting structured strategies of bTB detection and control through tuberculin skin testing (TST) and slaughterhouse and trade surveillance in cattle and buffalo herds. Fifteen years later, the prevalence of bTB in 13 states was assessed through TST, representing 75% of the Brazilian bovine herd (Ferreira Neto et al., 2016). The epidemiologic status was described as heterogeneous within and among states, with the prevalence varying from 0 to 2.5% in cattle and from 0 and 13.9% in herds (Ferreira Neto et al., 2016). In contrast to few other countries, a wildlife reservoir of *Mycobacterium bovis*, the main causative pathogen of bTB, has not been identified in Brazil until now, hence current bTB control efforts are mostly concerned with cattle transit and introduction of infected cattle into uninfected herds.

Molecular epidemiology has been a resourceful tool to understand *M. bovis* populational structure and transmission dynamics in livestock and multi-host systems (Guimaraes & Zimpel, 2020). More recently, whole genome sequencing (WGS) of *M. bovis* has allowed a more precise and in depth understanding of these topics. A recent study of *M. bovis* genomes from multiple countries showed that the historically used clonal complexes (i.e. European 1 and 2, African 1 and 2) (Smith et al., 2011; Müller et al., 2009; Berg et al., 2011; Rodriguez-Campos et al. 2012) do not represent the whole genetic diversity of this pathogen, and the existence of at least four lineages (Lb1-Lb4) and three “unknown groups” of *M. bovis* was suggested (Zimpel et al., 2020). On the other hand, numerous other studies have used WGS of *M. bovis* to evaluate pathogen transmission in cattle herds and wildlife [reviewed by (Guimaraes & Zimpel, 2020)], showing the usefulness of this technique for bTB surveillance systems. WGS allows, for instance, to detect how many *M. bovis* strains are circulating in a particular herd, when these strains were introduced, and the transmission links among animals or farms (Guimaraes & Zimpel, 2020).

In Brazil, despite the rare bTB detection in free-ranging wildlife (Maciel et al., 2018), reports of *M. bovis* infecting captive wild animals are frequent (Zimpel et al., 2017; Rocha et al., 2011, Murakami et al. 2012, Ikuta et al. 2018). The source of infection to these zoo animals, however, remains elusive. In 1977, a privately-owned safari park was inaugurated in the state of Rio Grande do Sul, Southern Brazil for public entertainment and environmental education. This safari comprised an area of 320 acres and harbored many different species of animals, both exotic and native, including deer, capybaras, llamas, camels, hippopotamus, different bird species, horses, non-human primates, among others. This man-made multi-host system posed various sanitary challenges over the years, culminating with its closure as defined by a court order in 2013. Among these challenges, several animals died of TB, with reports dating back to 2003. Therefore, the aims of this study were to compile the TB history of this safari park through official records and to evaluate *M. bovis* strain diversity, time of bTB introduction in the park, and path of transmission using WGS data. To achieve these objectives, we sequenced *M. bovis* genomes of 19 deer culled in 2018 and gathered other two *M. bovis* genomes obtained from llamas in 2012, sequenced in a previous study (Zimpel et al., 2020). The events that took place in the safari park described in this study highlight many regulatory gaps of the current bTB control and animal welfare legislations in the country. While we were able to access the data presented herein only after many years that these events ensued, we take this opportunity to launch a call for action for the control of TB in captive wild animals in Brazil.

## 2 MATERIALS AND METHODS

### 2.1 Study site, official records and deer sampling

This study was carried out in a privately-owned safari park located in the state of Rio Grande do Sul, Southern Brazil (Figure 1). To understand and describe the history of bTB in this park, we accessed records of the official state veterinary service that was used as a basis to close the park by court order in 2013 due to sanitary problems. In this official record, necropsy and laboratory reports from different public universities and private diagnostic laboratories dating back from 2003 were compiled.

We had access to tissue samples of deer that were euthanized after 16 of them (out of 51 tested) presented positive results in comparative cervical tests using *M. bovis* and *M. avium* PPD (purified protein derivative). Briefly, TST was carried out in 51 deer by the official veterinary service of the state following the United States Department of Agriculture (USDA) protocol on three different occasions from 2013 to 2015. A total of 16 deer tested positive, indicating the presence of bTB in the population. After judicial procedures, in 2018, all 281 deer present in the park were mandatorily slaughtered. Retropharyngeal, submandibular, and mesenteric lymph nodes were conveniently collected from 21 deer (*Cervus unicolor*, *Cervus elaphus* and *Dama dama*) with lesions suggestive of tuberculosis (LST). From these, 19 isolates of *M. bovis* were obtained as described previously (Lima et al., 2021). Unfortunately, the name of the host species of each sample were not recorded at the time of sampling.

### 2.2 DNA extraction of *M. bovis* from deer

DNA was extracted from the 19 *M. bovis* isolates from deer following an adapted protocol (van Embden et

al., 1993) at the Animal Immunology Laboratory, a biosafety level 3 facility located in the Department of Animal Health, Embrapa Beef Cattle, Campo Grande, Mato Grosso do Sul, Brazil.

### 2.3 Whole genome sequencing of *M. bovis* from deer

DNA samples of *M. bovis* isolates from deer were sent to Oswaldo Cruz Institute (Fiocruz), Manguinhos, Rio de Janeiro, Brazil, for WGS. The concentration and quality of DNA were measured using Qubit fluorometer (Invitrogen, California, USA) and Nanodrop 2000c spectrophotometer (Thermo Fisher Scientific, Massachusetts, USA). Paired-end genomic libraries were constructed using Nextera DNA Flex Library Prep kit (Illumina, California, USA), according to the manufacturer's instructions. Genomic libraries were then sequenced in HiSeq 2500 platform (Illumina) using HiSeq Rapid SBS Kit V2 (200 cycles). Illumina sequencing reads were made available in the Sequence Read Archive (SRA), NCBI (Appendix Table 1).

### 2.4 *Mycobacterium bovis* genomes from llamas

*Mycobacterium bovis* was isolated from two llamas of the same safari park in 2012 as part of an unrelated study (Ikuta, 2015) and their genomes sequenced in 2019 (Zimpel et al., 2020) (Illumina, paired-end, 100 bp – llama #1 and 250 bp - llama #2) . Both genomes were included in this study (Appendix Table 1) and compared with *M. bovis* genomes obtained from the deer.

### 2.5 *Mycobacterium bovis* genomes from cattle herds in Rio Grande do Sul

For comparison purposes, 71 sequencing reads (Illumina, paired-end, 100 bp) obtained from 13 cattle herds in Rio Grande do Sul, the same state of the safari park, were included in the analysis. These *M. bovis* isolates were obtained and sequenced in an unrelated study (Rodrigues et al., submitted). Figure 1 indicates the location of the safari park in relation to these cattle herds.

### 2.6 Quality control and reads mapping

Sequencing files were trimmed and the adapters were removed using Trimmomatic 0.39 (sliding window 5:20) (Bolger, Lohse & Usadel, 2014). The quality of the trimmed reads was evaluated using FastQC (Simon, 2010). Genomes had to meet the following quality criteria to be included in the study: GC content of ~65% without multiple or anomalous peaks, coverage of at least 15X, median read length Reads were mapped using Burrows-Wheeler Aligner 0.7.17 (BWA-MEM) (Li, 2010) against *Mycobacterium tuberculosis* H37Rv for the identification of RDs (regions of difference) and to check that genomes had a sequencing coverage of at least 95%. Accordingly, the resulting bam files were evaluated for the absence of RD4 and presence of RD1 (*M. bovis* specific and to rule out BCGs, respectively) using an algorithm previously described (Zimpel, 2020). Quality-approved reads were then mapped against *M. bovis* AF2122/97 using BWA-MEM, and Picard v2.18.23 (<https://github.com/broadinstitute/picard>) was used to remove duplicates from resulting files.

### 2.7 Variant calling

SNPs and INDELs (insertions and deletions) were called as previously described (Zimpel, 2020) using samtools mpileup (Li, 2011) and VarScan mpileup2cns (Koboldt et al., 2012) with parameters of minimum read depth of 7, mapping quality and minimum base quality of 20, and strand bias filter on, followed by annotation using SnpEff (Cingolani et al., 2012). INDELs as well as SNPs from repetitive regions (PE/PPE, transposases, integrases, maturase, phage and repetitive family 13E12 genes) were removed from the analysis (Zimpel, 2020). Genomes were excluded if the number of heterogeneous SNPs exceeded 15% of the total number of detected SNPs in a genome.

### 2.8 Phylogenetic reconstruction and principal component analysis of *M. bovis* from the safari park

A core-SNP matrix of *M. bovis* genomes from the safari was generated using a previously described algorithm (Zimpel, 2020) and subjected to ascertainment bias correction (ASC) using IQ-Tree -fconst (Nguyen, Schmidt, von Haeseler & Minh, 2015). The ASC-corrected matrix was used in the ModelFinder program (Kalyaanamoorthy et al., 2017) to select the best substitution model according to Bayesian Information

Criterion (BIC). The best model, TVM+F+R10, was fixed to reconstruct the phylogeny using maximum likelihood (ML) algorithm with 1,000 UFBoot pseudoreplicates (Hoang et al., 2017). *Mycobacterium bovis*AF2122/97 was included in the analysis and five representative genomes of *Mycobacterium caprae* were used as outgroup (ERR1462591, ERR1462625, ERR1462617, ERR1462581).

To evaluate clustering among *M. bovis* isolates, a PCA (principal component analysis) was constructed based on the SNP matrix. The function princomp was used in R software 3.5.0 to generate 2- and 3-dimensions PCA graphs.

## 2.9 Clonal complexes and spoligotyping

The clonal complexes African 1 and 2, and European 1 (Smith et al., 2011; Müller et al., 2009; Berg et al., 2011; Rodriguez-Campos et al., 2012) were identified using the bam files generated from mapping the reads against *Mycobacterium tuberculosis* H37Rv as previously described (Zimpel et al., 2020). The SNP in the *guaA* gene, predictive of clonal complex European 2, was investigated in the same bam files by checking the position 3,813,236 (Rodriguez-Campos et al., 2012). Spoligotyping patterns were predicted with SpoTyping software (Xia, Teo & Ong, 2016) using the fastq files. The SB numbers were retrieved from the online database (*Mycobacterium bovis*Spoligotype Database - mbovis.org).

## 2.10 Dating estimates

Dating analyses were performed in BEAST v1.10.4 (Suchard et al., 2018). The phylogenetic tree was estimated altogether with time inference of nodes. Isolation dates were used for tip-dating (Rieux & Balloux, 2016), with a relaxed clock model (uncorrelated lognormal rates across branches). The prior for the clock rate (parameter *ucln.mean*) was set as [1e-9; 2e-7] substitutions/site/branch/year (s/s/b/y), a conservative range encompassing values from previous studies (Lillebaek et al., 2016; Pepperell et al., 2013; Kay et al., 2015; Eldholm et al., 2015; Menardo et al., 2019). Markov Chain Monte Carlo (MCMC) chains were run to avoid convergence to local optima, each run for 10<sup>7</sup> generations, or until convergence of parameters and effective sample sizes (ESS)  $\geq 200$  (checked in Tracer v1.7.1 (Rambaut et al., 2018)). Burnin trees were discarded, and the remaining ones were annotated using TreeAnnotator (Rambaut & Drummond, 2010), generating a maximum clade credibility tree with median heights. Intervals for ages are reported for each node of the phylogenetic tree, as [minimum between two runs and maximum between two runs], to be conservative. The same five genomes of *M. caprae* described above were used as outgroup.

## 2.11 Pairwise SNP distance and minimum spanning tree

The core-SNP matrix using *M. bovis* genomes from the safari park and from cattle herds was used to construct a matrix of SNP distances between genomes and a minimum spanning tree using PHYLOViZ 2.0 (Francisco et al., 2012) with default parameters.

# 3 RESULTS

## 3.1 bTB history in the park

From the investigation of official records, we delineated a timeline of all TB cases that culminated in animal deaths from 2003 to 2018 in the safari park (Figure 2). In brief, from 2003 to 2015, at least 16 animals, from 8 different species, died or were humanely euthanized due to TB. Except for two llamas in 2012, TB was only diagnosed based on necropsy findings and Ziehl-Neelsen staining of tissue samples (Figure 2). It is possible to observe an increase in the number of cases starting around 2005, which coincides with management changes in the park (personal communication of a previous employee). These changes resulted in overpopulation (particularly of deer) and poor feeding practices. In 2013, due to widespread TB, other sanitary problems and selling of deer to properties in different Brazilian states, the official veterinary service, through a court order, demanded the closure of the park to visitations and halting of animal trade. From 2013 onwards, many tuberculin skin tests were performed in deer, cattle and buffalo. However, as the PNCEBT solely covers the latter two species, only these were mandatorily culled. The deer tested positive by TST in 2013, 2014 and 2015, and recommendations to contain the outbreak were made to the park. In 2018, after the

official veterinary service noted the park did not follow through with the recommendations, they ordered depopulation of the park deer based on interstate trade of potentially infected animals. Tissue from necropsy of the two llamas and the culled deer were the only animal samples ever subjected to *M. bovis* culture. A more complete history of all TB cases of the park is described in Supplementary material.

### 3.2 Phylogenetic reconstruction of *M. bovis* from the safari park

Out of the 19 *M. bovis* genomes obtained from deer, two were excluded because they presented only 35% and 43% coverage against *M. tuberculosis* H37Rv. All remaining genomes were confirmed to be *M. bovis* based on the absence of RD4 (i.e. region is deleted) and presence of RD1 (i.e. region is not deleted). Interestingly, one genome (TB002S2L) presented 35.77% of heterogeneous SNPs, suggesting the existence of a mixed strain infection in a deer. The remaining 16 *M. bovis* isolates from deer and both isolates obtained from both llamas were further used to construct a ML phylogenetic tree and a PCA using a core-SNP matrix (Figure 3).

The phylogenetic reconstruction suggests the existence of four distinct clusters circulating in the safari park. The most basal of these clusters is formed by a single *M. bovis* isolate from a deer (TB013S13L, purple cluster, Figure 3), while the three other clusters are formed by six *M. bovis* isolates from deer (blue cluster, Figure 3), 4 *M. bovis* isolates from llamas and deer (black cluster, Figure 3), and seven *M. bovis* from deer (green cluster, Figure 3). The PCA analysis indicates the separation of at least three clusters, while two of them appeared very closely related (green and black), as also depicted in the phylogenetic tree (Figure 3, Appendix Figure 1). All *M. bovis* isolates of the safari park were identified as carrying the European 2 marker, being thus identified as lineage Lb3 according to recent classification (14). Interestingly, the spoligotype pattern was the same, SB1401, for all *M. bovis* genomes from deer, while both llama isolates had spoligotype pattern SB0295 (Figure 3).

### 3.3 *M. bovis* from the safari park and cattle farms

We then evaluated the 18 *M. bovis* genomes of deer and llamas in the context of 71 additional genomes obtained from cattle herds located from 37 to 308 km away of the safari park. The phylogenetic tree and associated CCs, lineages and spoligotype patterns are shown in figure 4. While 17 genomes of *M. bovis* from the safari park were part of the same clade (includes clusters 1, 2 and 3 of the safari park), the most distantly related genome TB013S13L branched into a separated clade with other *M. bovis* isolates from cattle. In contrast, three *M. bovis* genomes from cattle herds appeared within the same main clade of the deer from the safari park. Both lineages Lb3 and Lb4, represented by CC markers of Eu2 and Eu1, respectively, were detected in the dataset, with Lb4 constituted by *M. bovis* isolates from cattle farms only. In addition, the phylogenetic clustering did not follow spoligotype patterns, as these appeared dispersed throughout the phylogenetic tree. This includes both llamas and three cattle isolates with SB0295 clustering with *M. bovis* isolates obtained from deer with SB1401. The phylogenetic relatedness of *M. bovis* genomes from cattle farms and the safari park indicates the safari park was affected by strains already circulating in the region during some point in time.

### 3.4 Dating estimates

The park was inaugurated in 1977 and personal communications from a previous employee indicates that TB cases in the park date back from the 1980's. Thus, to better elucidate *M. bovis* introduction, we estimated the ages of the MRCAs (most recent common ancestor) of the *M. bovis* strains circulating in the safari park and results are shown in Table 1. Accordingly, the dating estimates of the MRCAs of each of the three *M. bovis* clusters (1 to 3, Figures 3 and 4) from the safari park ranged from 10 to 40 years before present (BP) (Table 1). In addition, the MRCA of clusters 1 and 2 dated from 23 to 47 years ago, which is also within the time frame of the park existence. In contrast, the MRCA of clusters 1, 2 and 3 dated from 50 to 103 years ago, which is older than the park inauguration. Thus, for clusters 1 and 2, and for cluster 3, it is possible that they are products of two different introduction events at the beginning of the safari park, evolving into different strains over time through persistent animal to animal transmission. Finally, another introduction event pertains to isolate TB013S13L, with a MRCA dating back from 25 to 77 years ago. The MRCA of *M. bovis* and *M. caprae* was dated from 630 to 1,644 years BP, similar to a previous study (Zimpel et al., 2020).

### 3.5 Minimum spanning tree

Next, we used the core-SNP matrix of all *M. bovis* isolates to infer a distance matrix (Appendix Figures 2 and 3, Appendix Table 2) and a minimum spanning tree of the dataset (Figure 5). Traditionally, SNP distances between genomes have been used to infer transmission of *M. tuberculosis* between individuals. Thresholds varying from 5 and 12 SNPs have been proposed to determine if two genomes are from the same *M. tuberculosis* strain (Meehan et al., 2019). By using this same SNP threshold, at least four transmission links in animals from the safari park, involving five deer, were detected (Figure 5). The SNP distance between these genomes were: 4 ( $n=2$ ), 8 ( $n=1$ ) and 10 ( $n=1$ ) (Table 2). No transmission links between llamas and deer were detected using this approach. This finding is further corroborated by the fact that *M. bovis* from deer and from llamas carry distinct spoligotype patterns (SB1401 and SB0295), which are highly divergent in terms of the presence of spacer sequences (Appendix Figure 4). Multiple other transmission links ( $n=28$ ) were detected involving cattle from vicinity herds, but no recent transmission could be inferred between the safari park and cattle herds (Figure 5 and Table 1). The two closest *M. bovis* genomes of cattle to *M. bovis* genomes of the safari park differed by 92 and 64 SNPs, which are 5 to 7 times higher than the SNP cut-off to infer recent transmission (Figure 5). Along with the dating estimates, these results indicate that multiple *M. bovis* strains were circulating in the safari park at the time of sample collection, resulting from at least three different past introductions.

## 4 Discussion

In this study, we describe an outbreak of tuberculosis caused by *M. bovis* in a Brazilian safari park that lasted for at least 15 years. This outbreak was a result of multiple pathogen introductions and culminated with the death of many animal specimens. There are two main reasons that can explain why this outbreak was so severe. First, starting in 2005 the park passed through management changes and financial hurdles that led to animal overcrowding and malnutrition. Many animals of different species shared a small enclosure and feeding practices were not adequate to serve them all. Overcrowding and malnutrition are known risk factors of bTB (Pollock & Neill, 2002). Second, the park owners often acquired and introduced animals of unknown bTB status (i.e. without diagnostic testing) into the premises, particularly cattle and buffalo. Collectively, these factors may have facilitated pathogen introduction and spread in the property as well as the development of clinical disease by the wild animals.

Our results suggest two possible timeline scenarios for the dynamics of this bTB outbreak. One scenario involves at least three different *M. bovis* introductions occurring in the time period between the park inauguration in 1977 up until around 2003. The founding *M. bovis* subsequently evolved into many different strains over time, fueled by persistent animal to animal transmission. These strains were likely circulating at low levels or asymptotically until the above-described conditions of overcrowding and malnutrition became conducive of a deadly outbreak. As *M. bovis* is endemic in the region (5.2% of infected herd prevalence (Queiroz et al., 2016)), a second scenario involves the introduction of many different strains over the years, possibly through cattle, contributing to the high *M. bovis* diversity observed in the property. Although cattle and wildlife were separated by a fence, escape and mixing of animals in both enclosures were reportedly common events. Alternatively, pathogen introduction may have occurred with the acquisition of specimens of infected wild animals, since TB cases occur in wild animals in captivity in Brazil (Zimpel et al., 2017; Rocha et al., 2011; Murakami et al., 2018; Ikuta et al., 2018; Murakami et al., 2012).

As a wild animal reservoir of bTB has not been identified in Brazil so far, in its inception the PNCEBT was designed to cover cattle and buffalo only. Hence notifications of TB cases in wildlife are not mandatory, diagnostics are not standardized, and there are no official guidelines for outbreak resolution or prevention in wildlife. These factors likely influenced the course of the outbreak described herein. Thus, we believe the inclusion of wildlife in the PNCEBT can bring great benefits in terms of animal welfare, while increasing biosafety in zoos and parks. In addition, regulation of the diagnosis of tuberculosis in wild animals would have an impact on reducing outbreaks that may occur through sale or trade of infected animals. The many outbreaks seen in Brazilian captive populations over the years (Zimpel et al., 2017; Rocha et al., 2011; Murakami et al., 2018; Ikuta et al., 2018; Murakami et al., 2012), along with the one reported herein, should

form the rationale necessary to develop this nationwide plan to control and report TB in wildlife, as also suggested previously (Valvassoura & Ferreira Neto, 2014). Notwithstanding, for these policies to become effective, it is necessary to integrate federal and state surveillance services and environmental regulatory agencies.

Our study adds to the importance of evaluating potential paths of transmission to understand bTB outbreaks. Using phylogenetic analysis and *M. tuberculosis* -based cut-offs for the number of SNPs, many different strains of *M. bovis* were found circulating in the park, with four recent transmission events involving five deer. The low SNP distance between these five genomes suggests that the same *M. bovis* strain was transmitted from one animal to another, and the amount of accumulated genetic changes is just a reflection of microevolution. In addition, one of the animals was co-infected with multiple strains, as it presented many heterogeneous SNPs. This mixed infection occurs when the individual is exposed to a single infection event carrying multiple strains, or to several infection events with different strains throughout its life, resulting in a superinfection. Such condition is normally seen in conditions of high disease endemicity, as described in this park.

Using the SNP-cutoff approach, recent transmission events between the surveyed cattle farms and the safari park were not observed in this study. It is possible that the sampling of cattle farms was not comprehensive enough to capture the true *M. bovis* diversity of Rio Grande do Sul state. The origin of the cattle maintained at the safari park was unknown to the authors of this study; thus, it is also possible that the region from which they originated was not covered in the cattle sampling. Alternatively, the *M. bovis* strains infecting deer and llamas have been introduced a long time ago in the park (as suggested by the dating analysis) and evolved in geographic isolation. Nevertheless, the phylogenetic relatedness among strains indicates that *M. bovis* was introduced in the park from somewhere in the region at some point in time.

Although commonly used as a measure of *M. tuberculosis* or *M. bovis* transmission between individuals, the SNP cutoff approach was never standardized for *M. bovis*, is sensitive to the pipeline used to detect SNPs (Guimaraes & Zimpel, 2020), and also a simplistic approach to follow the complex evolution of tuberculous mycobacteria (Menardo et al., 2019). The BEAST analysis performed herein rejected a strict clock evolution model of *M. bovis*, corroborating previous studies on *M. bovis* (Zimpel et al., 2020; Crispell et al., 2019; Salvador, 2019) and *M. tuberculosis* (Menardo et al., 2019; Rutaihua et al., 2019). We are also unaware if evolutionary rates of *M. bovis* are influenced by the disease form (e.g. latent *versus* clinical disease) or host species. Out of the 32 possible transmission links, we observed discordant spoligotype patterns in five, from which three were due to loss of spacers, but two either gained or lost spacers (Appendix Figure 5). Therefore, results reported regarding recent transmission events should be interpreted with caution; and further studies should be conducted to address how we can best use genomics to infer transmission of *M. bovis* strains and how these can be compared to traditional genotyping techniques such as spoligotyping and MIRU-VNTR.

This study has limitations. Unfortunately, *M. bovis* isolates from all other animals that died during the outbreak were not obtained, precluding our ability to evaluate the true genetic diversity of the pathogen in the park. We were also unable to control sample collection at the slaughterhouse where deer were culled, thus preventing a better description of the lesions found, number of affected animals, and host species. And finally, we did not have access to records of animal introductions that occurred in the park, impeding a more comprehensive analysis of pathogen introduction.

## 5 FINAL CONSIDERATIONS

By combining epidemiological history and WGS of *M. bovis* strains, we show that the increase in fatal cases observed in the safari park coincided with changes in populational density and nutritional status, and that many pathogen introductions occurred over time. These observations demonstrate that wildlife populations found to be chronically infected and frequently asymptomatic may suddenly suffer from acute, fatal disease following populational disturbances, such as the ones observed herein. The negligence of successive pathogen introductions likely potentiated the outbreak, contributing to disease spread in different animal species. The *M. bovis* strains circulating in the park were classified as Lb3 (CC Eu2), a common lineage found in Brazil, and share phylogenetic relationships with extant *M. bovis* strains circulating in farmed cattle. Thus, at some

point in time, transmission occurred between cattle and the wildlife maintained in the safari park. The use of WGS to delineate the history of this outbreak highlights the importance of genomic surveillance to follow disease introduction into animal populations and transmission between individuals.

With this study we make a pledge for action. It is imperative that reporting of TB cases in wildlife, captive or free-ranging, becomes mandatory in Brazil as to understand the true size of the problem. This should follow increased investments in pathogen identification, as wildlife are affected by both *M. bovis* and *M. tuberculosis* (Murakami et al., 2012). Finally, a national plan for control of TB in wildlife should be designed to incorporate specific guidelines to contain and solve outbreaks and prevent possible zoonotic transmission to veterinarians and caretakers, considering species-specific conditions and difficulties in diagnosing the disease in wild animals, particularly when dealing with endangered species.

## ACKNOWLEDGMENTS

CKZ fellowship is funded through São Paulo Research Foundation (FAPESP grant # 2019/10896-8, 2017/04617-3). TTSP fellowship is funded through CNPq (grant # 443235/2014-7). This study was funded in part by Morris Animal Foundation (grant #D17ZO-307). This study was funded in part by CAPES (Finance code 001). This study was funded in part by CNPq under grant numbers 407826/2018-1 and 306675/2018-8, and FUNDECT, under grant number 59/300.104/2017 (SIAFEM: 027282).

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ORCID

## REFERENCES

World Organisation for Animal Health (OIE) (2011). Bovine tuberculosis. General Disease Information Sheets, 2011, 1-6. Cousins, S. (2018). End TB Strategy. The challenges of preventing bovine tuberculosis. Bulletin of the World Health Organization, 96(2), 82-83. O'Reilly, L.M., Daborn, C.J. (1995). The epidemiology of *Mycobacterium bovis* infections in animals and man: a review. Tubercle and Lung Disease, 76 Suppl 1, 1-46. Ayele, W.Y., Neill, S.D., Zinsstag, J., Weiss, M.G., Pavlik, I. (2004). Bovine tuberculosis: An old disease but a new threat to Africa. International Journal of Tuberculosis and Lung Disease, 8(8), 924-937. Ferreira Neto, J.S. (2019). Brucellosis and tuberculosis in cattle in South America. Brazilian Journal of Veterinary Research, 55(2), 1-23. Olea-Popelka, F., Muwonge, A., Perera, A., Dean, A.S., Mumford, E., Erlacher-Vindel, E., et al. (2017). Zoonotic tuberculosis in human beings caused by *Mycobacterium bovis* - a call for action. Lancet Infectious Diseases, 17(1), e21-25. World Health Organization (WHO), Food and Agriculture Organization of the United Nations (FAO), World Organisation for Animal Health (OIE) (2017). Roadmap for zoonotic tuberculosis. Geneva: World Health Organization, 22 p. Ferreira Neto, J.S., Da Silveira, G.B., Rosa, B.M., Gonçalves, V.S.P., Gris-Filho, J.H.H., Amaku, M., et al. (2016). Analysis of 15 years of the national program for the control and eradication of animal brucellosis and tuberculosis, Brazil. Semina: Ciências Agrárias, 37(5):3385-3402. Guimaraes, A.M.S., Zimpel, C.K (2020). *Mycobacterium bovis*: from genotyping to genome sequencing. Microorganisms, 8(5), 667. Smith, N.H., Berg, S., Dale, J., Allen, A., Rodriguez, S., Romero, B., et al. (2011). European 1: A globally important clonal complex of *Mycobacterium bovis*. Infection, Genetics and Evolution, 11(6), 1340-1351. Müller, B., Hilty, M., Berg, S., Garcia-Pelayo, M.C., Dale, J., Boschiroli, M.L., et al. (2009). African 1, an epidemiologically important clonal complex of *Mycobacterium bovis* dominant in Mali, Nigeria, Cameroon, and Chad. Journal of Bacteriology, 191(6), 1951-1960. Berg, S., Garcia-Pelayo, M.C., Müller, B., Hailu, E., Asimwe, B., Kremer, K., et al. (2011). African 2, a clonal complex of *Mycobacterium bovis* epidemiologically important in East Africa. Journal of Bacteriology, 193(3), 670-678. Rodriguez-Campos, S., Schürch, A.C., Dale, J., Lohan, A.J., Cunha, M.V., Botelho, A., et al. (2012). European 2 - A clonal complex of *Mycobacterium bovis* dominant in the Iberian Peninsula. Infect Genet Evol., 12(4):866-872. Zimpel, C.K., Patané, J.S.L., Guedes, A.C.P., de Souza, R.F., Silva-Pereira, T.T., Camargo, N.C.S., et al. (2020). Global distribution and evolution of *Mycobacterium bovis* lineages. Frontiers in Microbiology, 11, 843. Maciel, A.L.G., Loiko, M.R., Bueno, T.S., Moreira, J.G.,

Coppola, M., Dalla Costa, E.R., et al. (2018). Tuberculosis in Southern Brazilian wild boars (*Sus scrofa*): First epidemiological findings. *Transboundary and Emerging Diseases*, 65(2), 518–526. Zimpel, C.K., Brum, J.S., de Souza Filho, A.F., Biondo, A.W., Perotta, J.H., Dib, C.C., et al. (2017). *Mycobacterium bovis* in a European bison (*Bison bonasus*) raises concerns about tuberculosis in Brazilian captive wildlife populations: a case report. *BMC Research Notes*, 10(1), 91. Rocha, V.C.M., Corrêa, S.H.R., Oliveira, E.M.D., Rodriguez, C.A.R., Fedullo, J.D., Matrone, M., et al. (2011). Tuberculosis determined by *Mycobacterium bovis* in captive waterbucks (*Kobus ellipsiprymnus*) in São Paulo, Brazil. *Brazilian Journal of Microbiology*, 42(2):726–728. Murakami, P.S., Monego, F., Ho, J.L., Gibson, A., Vilani, R.G.D. de C., Soresini, G.C.G., et al. (2012). An outbreak of tuberculosis by *Mycobacterium bovis* in coatis (*Nasua nasua*). *Journal of Zoo and Wildlife Medicine*, 43(2), 338–341. Ikuta, C.Y., Reisfeld, L., Silvatti, B., Salvagni, F.A., De Paula, C.D., Pessier, A.P., et al. (2018). Tuberculosis caused by *Mycobacterium bovis* infection in a captive-bred American bullfrog (*Lithobates catesbeiana*). *BMC Veterinary Research*. 14(1):1–4. Lima, D.A.R., Rodrigues, R.A., Etges, R.N., Araújo, F.R (2021). Bovine tuberculosis in a safari park in Brazil. *Pesquisa Veterinária Brasileira*, 41, e06719. van Embden, J.D., Cave, M.D., Crawford, J.T., Dale, J.W., Eisenach, K.D., Gicquel, B., et al. (1993). Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *Journal of Clinical Microbiology*, 31(2), 406–409. Ikuta, C.Y. Estudo de micobactérias em animais silvestres mantidos em cativeiro [thesis]. São Paulo, Faculdade de Medicina Veterinária e Zootecnia, 2015 [cited 2021-04-09]. doi:10.11606/T.10.2016.tde-16092015-110534. Bolger, A.M., Lohse, M., Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30(15):2114–2120. Simon, A. FastQC: A Quality Control tool for High Throughput Sequence Data. 2010. Li, H., Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics*, 26(5), 589–595. Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetic parameter estimation from sequencing data. *Bioinformatics*, 27(21):2987–2993. Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., et al. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Research*, 22(3), 568–576. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. *Fly (Austin)*, 6(2),80–92. Nguyen, L-T., Schmidt, H.A., von Haeseler, A., Minh, B.Q. (2015). IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, 32(1) , 268–274. Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., Jermini, L.S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature Methods*, 14(6), 587–589. Hoang, D.T., Chernomor, O., von Haeseler, A., Minh, B.Q., Vinh, L.S. (2017). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, 35(2), 518–522. Xia, E., Teo, Y-Y., Ong, RT-H. (2016). SpoTyping: fast and accurate *in silico* *Mycobacterium* spoligotyping from sequence reads. *Genome Medicine*, 8(1), 19. Suchard, M.A., Lemey, P., Baele, G., Ayres, D.L., Drummond, A.J., Rambaut, A. (2018). Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10. *Virus Evolution*, 4(1), vey016. Rieux, A. & Balloux, F. (2016). Inferences from tip-calibrated phylogenies: a review and a practical guide. *Molecular Ecology*. 25(9), 1911–1924. Lillebaek, T., Norman, A., Rasmussen, E.M., Marvig, R.L., Folkvardsen, D.B., Andersen, Å.B., et al. (2016). Substantial molecular evolution and mutation rates in prolonged latent *Mycobacterium tuberculosis* infection in humans. *International Journal of Medical Microbiology*, 306(7), 580–585. Pepperell, C.S., Casto, A.M., Kitchen, A., Granka, J.M., Cornejo, O.E., Holmes, E.C., et al. (2013). The Role of Selection in Shaping Diversity of Natural *Mycobacterium tuberculosis* Populations. Sassetti CM, editor. *PLoS Pathogens*, 9(8), e1003543. Kay, G.L., Sergeant, M.J., Zhou, Z., Chan, JZ-M., Millard, A., Quick, J., et al. (2015). Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nature Communications*, 6(1), 6717. Eldholm, V., Monteserin, J., Rieux, A., Lopez, B., Sobkowiak, B., Ritacco, V., et al. (2015). Four decades of transmission of a multidrug-resistant *Mycobacterium tuberculosis* outbreak strain. *Nature Communications*, 6(1), 7119. Menardo, F., Duchêne, S., Brites, D., Gagneux, S. (2019). The molecular clock of *Mycobacterium tuberculosis*. *PLOS Pathogens*, 15(9), e1008067. Rambaut, A., Drummond, A.J., Xie, D. Baele, G., Suchard, M.A. (2018). Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology*, 67(5), 901–904. Rambaut, A., Drummond, A.J. (2010). TreeAnnotator v.1.6.1. [Internet]. [cited 2002 Jan 20]. Available from: <http://beast.bio.ed.ac.uk>

Francisco, A.P., Vaz, C., Monteiro, P.T., Melo-Cristino, J., Ramirez, M., Carriço, J.A. (2012). PHYLO-ViZ: Phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics*, 13(1), 87. Meehan, C.J., Goig, G.A., Kohl, T.A., Verboven, L., Dippenaar, A., Ezewudo, M., et al. (2019). Whole genome sequencing of *Mycobacterium tuberculosis*: current standards and open issues. *Nature Reviews Microbiology*, 9:533–545. Pollock, J.M., Neill, S.D. (2002). *Mycobacterium bovis* infection and tuberculosis in cattle. *Veterinary Journal*, 163(2), 115–127. Queiroz, M.R., Groff, A.C.M., Silva, N.D.S., Filho, J.H.H.G., Amaku, M., Dias, R.A., et al. (2016). Epidemiological status of bovine tuberculosis in the state of Rio Grande do Sul, Brazil. *Semina: Ciências Agrárias*, 37(5), 3647–3658. Murakami, P.S., Monego, F., Ho, J.L., Gibson, A., Javorouski, M.L., Bonat, M., et al. (2012). Detection of RDRIO strain of *Mycobacterium tuberculosis* in tapirs (*Tapirus terrestris*) from a zoo in Brazil. *Journal of Zoo and Wildlife Medicine*, 43(4), 872–875. Valvassoura, T., & Ferreira Neto, J.S. (2014). *Tuberculose em primatas*. São Paulo: Faculdade de Medicina Veterinária e Zootecnia da Universidade de São Paulo, 193 p. Crispell, J., Benton, C.H., Balaz, D., De Maio, N., Akhmetova, A., Allen, A., et al. (2019). Combining genomics and epidemiology to analyse bi-directional transmission of *Mycobacterium bovis* in a multi-host system. *eLife*, 8, e45833. Salvador, L.C.M., O’Brien, D.J., Cosgrove, M.K., Stuber, T.P., Schooley, A.M., Crispell, J., et al. (2019). Disease management at the wildlife-livestock interface: Using whole-genome sequencing to study the role of elk in *Mycobacterium bovis* transmission in Michigan, USA. *Molecular Ecology*, 28(9), 2192–2205. Rutaihwa LK, Menardo F, Stucki D, Gygli SM, Ley SD, Malla B, et al. (2019) Multiple Introductions of *Mycobacterium tuberculosis* Lineage 2–Beijing Into Africa Over Centuries. *Frontiers in Ecology and Evolution*, 7, 112.

**Table 1** . Dating estimates of *Mycobacterium bovis* clusters from the safari park

<i>M. bovis</i> clusters	Dating estimates (years before present)	Approximate year
MRCA of cluster 1	10 – 26	1992 – 2008
MRCA of cluster 2	18 – 39	1979 – 2000
MRCA of cluster 3	15 – 40	1978 – 2003
MRCA of clusters 1 and 2	23 – 47	1971 – 1995
MRCA of clusters 1, 2 and 3	50 – 103	1915 – 1968
MRCA of TB013S13L (cluster 4)	25 – 77	1993 – 1941
MRCA of <i>M. bovis</i> from llamas	10 – 21	1997 – 2008

MRCA = most recent common ancestor

**Table 2**. Possible *Mycobacterium bovis* transmission links observed between cattle and their spoligotype patterns.

Transmission links	<i>M. bovis</i> genomes	SNP-distance	Spoligotypes
Link 1 (deer)	TB005S5L and TB004S4L	4	SB1401
Link 2 (deer)	TB006S6L and TB007S7L	4	SB1401
Link 3 (deer)	TB10S10L and TB0076S7L	10	SB1401
Link 4 (deer)	TB004S4L and TB007S7L	8	SB1401
Link 5 (cattle)	ERR4450971 and ERR4450970	11	SB1135
Link 6 (cattle)	ERR4450970 and ERR4450974	11	SB1135
Link 7 (cattle)	ERR4450974 and ERR4450972	8	SB1135
Link 8 (cattle)	ERR4450974 and ERR4450973	11	SB1135
Link 9 (cattle)	ERR4451184 and ERR4451185	3	SB095
Link 10 (cattle)	ERR4450923 and ERR4450920	7	SB1369
Link 11 (cattle)	ERR4450923 and ERR4450922	10	SB1369
Link 12 (cattle)	ERR4450923 and ERR4450924	10	SB1369
Link 13 (cattle)	ERR4450919 and ERR4450923	9	SB1369
Link 14 (cattle)	ERR4450919 and ERR4450925	11	SB1369, SB0295
Link 15 (cattle)	ERR4450932 and ERR4450928	10	SB0295

Link 16 (cattle)	ERR4450934 and ERR4450927	7	SB0295
Link 17 (cattle)	ERR4450932 and ERR4450926	11	SB0295
Link 18 (cattle)	ERR4450932 and ERR4450929	8	SB0295
Link 19 (cattle)	ERR4450929 and ERR4450930	7	SB0295, SB2709
Link 20 (cattle)	ERR4450953 and ERR4450944	12	SB1993
Link 21 (cattle)	ERR4450944 and ERR4450948	6	SB1993, SB2712
Link 22 (cattle)	ERR4450944 and ERR4450949	5	SB1993
Link 23 (cattle)	ERR4450955 and ERR4450941	6	SB0140
Link 24 (cattle)	ERR4450941 and ERR4450940	6	SB0140
Link 25 (cattle)	ERR4450955 and ERR4450954	7	SB0140
Link 26 (cattle)	ERR4450938 and ERR4450937	12	SB1993
Link 27 (cattle)	ERR4450962 and ERR4450950	9	SB1993, SB0121
Link 28 (cattle)	ERR4450969 and ERR4450964	6	SB0121
Link 29 (cattle)	ERR4450964 and ERR4450966	9	SB0121, SB1135
Link 30 (cattle)	ERR4450969 and ERR4450968	7	SB0121
Link 31 (cattle)	ERR4450969 and ERR4450965	9	SB0121
Link 32 (cattle)	ERR4450965 and ERR4450961	11	SB0121

## Figure Legends

**Figure 1.** Location of the safari park (red drop) and the cities of the cattle farms are represented by different colors on the map of the Rio Grande do Sul state, Southern Brazil. The safari park is located in the vicinity of Porto Alegre, the capital city of Rio Grande do Sul. Pink: cattle herd from Não-Me-Toque; Midnight Green: Arroio do Tigre; Zomp: Teutônia; Green: Charqueadas; Coral Pink: Estrela; Orange: Fazenda Vila Nova; Cyan: Bom Retiro do Sul; Purple: Eldorado do Sul; Black: Camaquã.

**Figure 2.** Timeline of the tuberculosis outbreak in the safari park according to data collection of government authorities from the state of Rio Grande do Sul, Brazil. Years in red represent zoo closure (for visitors), following a court order in 2013. \**Mycobacterium bovis* isolation performed in two llamas, and positive PCR in a third animal. Isolation was not performed in the red deer. \*\* T+ S in deer, ND + N in a camel and an antelope. PPD: Purified protein derivative of *M. bovis* – indicates that animals were subjected to tuberculin test (comparative cervical test).

**Figure 3.** Phylogenetic reconstruction and principal component analysis (PCA) of 18 *Mycobacterium bovis* genomes obtained from the safari park. *Mycobacterium bovis* genomes starting with “TB0” were obtained from deer ( $n=16$ ) and genomes starting with “L” were obtained from llamas ( $n=2$ ). Phylogenetic clusters are shown in the PCA using the same colors. NC\_00295.4 is the reference genome of *M. bovis* AF2122/97. Phylogenetic tree was constructed from a core-SNP matrix using maximum likelihood in IQ-Tree with 1,000 bootstrap replicas. Strains of *Mycobacterium caprae* were used as an outgroup. Lineages were identified according to Zimpel et al. (14). Horizontal bar shows substitutions per nucleotide.

**Figure 4.** Bayesian phylogenetic tree of *Mycobacterium bovis* genomes from the safari park ( $n=18$ ) and cattle herds ( $n=71$ ) of Rio Grande do Sul. Phylogenetic tree was generated in BEAST v1.10.4, using a core SNP-matrix. Horizontal bar indicates years before present. *Mycobacterium caprae* strains are used as outgroup.

**Figure 5.** Minimum spanning tree of 18 *Mycobacterium bovis* genomes from the safari park and 71 *M. bovis* from cattle herds of Rio Grande do Sul state. Nodes are *M. bovis* genomes and edges represent the number of SNPs between two genomes (number of SNPs indicated on each edge). Nodes in orange represent possible transmission links using a SNP cutoff of 12. Genomes of *M. bovis* starting with TB0 (deer) and L (llama) are from the safari park, while genomes of *M. bovis* starting with ERR are from cattle herds.

**Appendix Table 1.** List of *Mycobacterium* spp genomes and respective accession numbers used in this study.

**Appendix Table 2.** Pairwise Single Nucleotide Polymorphism (SNP) distance-matrix of *Mycobacterium bovis* genomes used in this study.

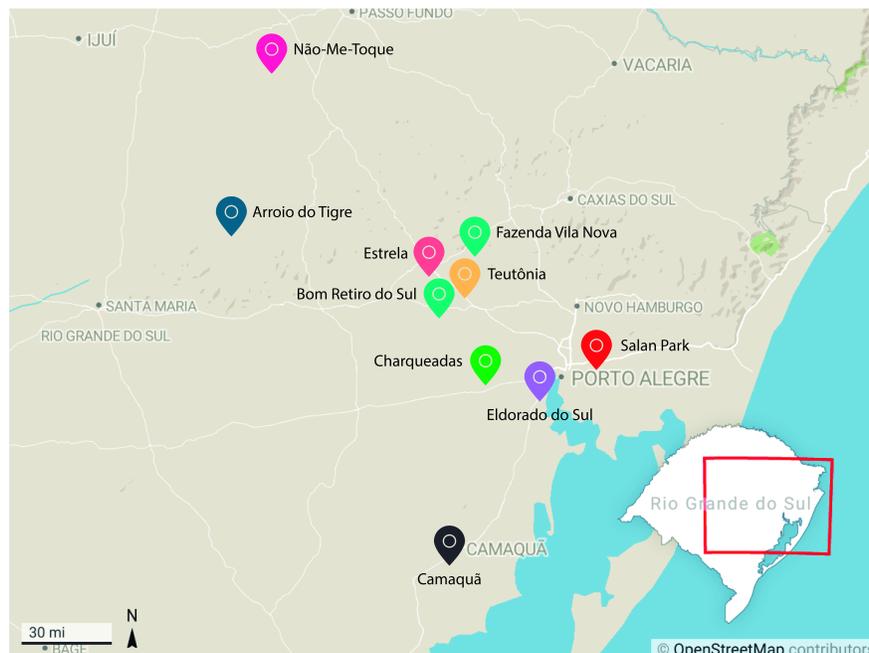
**Appendix Figure 1.** Three-dimensional principal component analysis (PCA) of *Mycobacterium bovis* genomes from the safari park. Colors correspond to clusters depicted in Figure 1 of the article. Component 1 = 66.58%, Component 2 = 24.36%, Component 3 = 3.88%. Red: *Mycobacterium bovis* AF2122/97 reference genome.

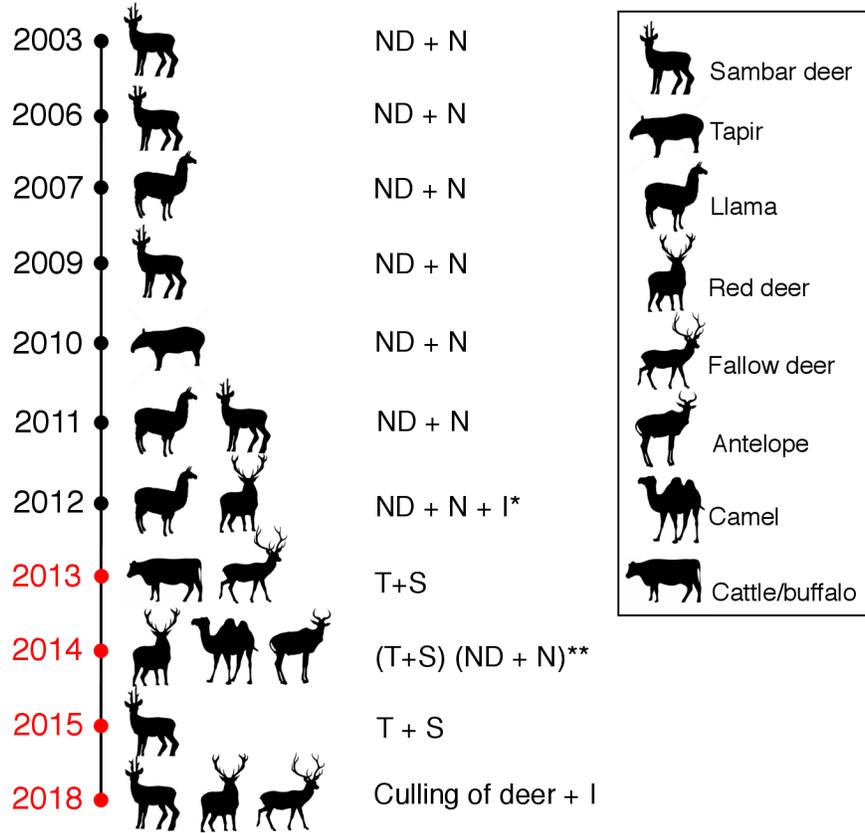
**Appendix Figure 2.** Single nucleotide polymorphism (SNP)-distance between *Mycobacterium bovis* genomes of the safari park (n=18) and of cattle (n=71) from the state of Rio Grande do Sul. This matrix was generated from core-SNP matrix of all genomes using the software PHYLOViZ 2.0 with default parameters. Appendix Table 2 shows individual values. *Mycobacterium bovis* genomes starting with “TB0” were obtained from cervids and genomes starting with “L” were obtained from llamas. In green it is possible to see comparisons between *M. bovis* genomes of different lineages (Lb3 and Lb4) detected in the dataset.

**Appendix Figure 3 .** Single nucleotide polymorphism (SNP)-distance between *Mycobacterium bovis* genomes of the safari park (n=18). This matrix was generated from core-SNP matrix of all genomes using the software PHYLOViZ 2.0 with default parameters. *Mycobacterium bovis* genomes starting with “TB0” were obtained from cervids and genomes starting with “L” were obtained from llamas.

**Appendix Figure 4 .** Comparison of the spoligotype patterns found in cervids (SB1401) and llamas (SB0295) of the safari park. Spoligotype patterns were detected using sequencing reads (fastq files) in SpoTyping software.

**Appendix Figure 5.** Discordant spoligotype patterns of the recent transmission links of *Mycobacterium bovis* detected using genomics.





**ND:** Natural death or humane euthanasia    **I:** *Mycobacterium bovis* isolation in culture  
**N:** Necropsy with positive Ziehl Neelsen    **T + S:** PPD and euthanasia of positive animals

