Genomic and temporal analyses of *Mycobacterium bovis* in southern Brazil

Rudielle de Arruda Rodrigues[1], Flábio Ribeiro Araújo[2], Alberto Martín Rivera Dávila[3], Rodrigo Nestor Etges[4], Julian Parkhill[5] and Andries J. van Tonder[5,*]

Abstract

Mycobacterium bovis is a causal agent of bovine tuberculosis (bTB), one of the most important diseases currently facing the cattle industry worldwide. Tracing the source of *M. bovis* infections of livestock is an important tool for understanding the epidemiology of bTB and defining control/eradication strategies. In this study, whole genome sequencing (WGS) of 74 *M. bovis* isolates sourced from naturally infected cattle in the State of Rio Grande do Sul (RS), southern Brazil, was used to evaluate the population structure of *M. bovis* in the region, identify potential transmission events and date the introduction of clonal complex (CC) European 2 (Eu2). *In silico* spoligotyping identified 11 distinct patterns including four new profiles and two CCs, European 1 (Eu1) and Eu2. The analyses revealed a high level of genetic diversity in the majority of herds and identified putative transmission clusters that suggested that within- and between-herd transmission is occurring in RS. In addition, a comparison with other published *M. bovis* isolates from Argentina, Brazil, Paraguay and Uruguay demonstrated some evidence for a possible cross-border transmission of CC Eu1 into RS from Uruguay or Argentina. An estimated date for the introduction of CC Eu2 into RS in the middle of the 19th century correlated with the historical introduction of cattle into RS to improve existing local breeds. These findings contribute to the understanding of the population structure of *M. bovis* in southern Brazil and highlight the potential of WGS in surveillance and helping to identify bTB transmission.

DATA SUMMARY

(1) Supplementary files 1 and 2 (available in the online version of this article).

(2) Raw sequencing data for all sequenced isolates have been deposited in the European Nucleotide Archive (ENA) under project PRJEB39667. Accession numbers are detailed in File S1.

(3) The R code used to analyse the data in this paper is available on Github: https://github.com/avantonder/Brazil2020.

INTRODUCTION

Mycobacterium bovis is a causative agent of bovine tuberculosis (bTB), infecting mainly bovines globally, in addition to a broad host range that includes humans and wild animals. Moreover, the disease in humans constitutes a public health problem and currently is one of the most important diseases faced by the global cattle industry, causing significant economic losses to beef and dairy production [1–3].

In many countries, national control and eradication bTB programmes have been successfully implemented in livestock, based on test and slaughter of infected cattle as a strategy to

manage the disease [1, 2]. In Brazil, which is currently the world's largest exporter of beef [4], the disease is compulsorily reported and regulated by the National Program for the Control and Eradication of Animal Brucellosis and Tuberculosis (PNCEBT), which aims to reduce the prevalence and incidence of bTB in cattle and buffaloes, by intradermal testing and the slaughter of reactive animals, and eventually aims to eradicate bTB in Brazil [5]. In the State of Rio Grande do Sul (RS), since December 2014, the Normative Instruction of the Secretariat of Agriculture, Livestock and Agribusiness no. 002/2014 (IN-SEAPA 002/2014) has worked with the objective of complementing the actions of the PNCEBT, avoiding the dissemination of diseases, as well as promoting the process of detection of outbreaks through the diagnosis of disease by qualified veterinarians from the programme [6].

Official epidemiological surveys of bTB performed in 75% of Brazilian herds show that the prevalence of this disease is heterogeneous among and within states with prevalence of infected herds ranging between 1.0 and 13.9%, prevalence being higher in dairy farms, especially high-production farms [7]. The main risk factors identified were milk production, particularly in farms with some degree of sophistication in the mode of production, and having a larger number of cattle [8–20].

In southern Brazil, RS currently has the seventh largest cattle population, with 12918325 animals (estimated in 2019), which represents 6.05% of the Brazilian national herd [21]. The state stands out in national milk production, being the second largest producer with about 13.2% of national production [22]. Study of the epidemiological situation of bTB in RS reported the prevalence of outbreaks to be 2.8% [95% Confidence Interval (CI): 1.8–4.0%] at the herd level and 0.7% [95% CI: 0.4–1.0%] at the animal level. It was found that there was a concentration of outbreaks in the northern part of the state, characterized by the predominance of milk and mixed beef–dairy farms. The risk factors associated with the outbreak condition were dairy farming (odds ratio, OR=2.90 [95% CI: 1.40–6.13]) and herds with 16 or more cattle at least 24 months old (OR=2.61 [95% CI: 1.20–5.49]) [15]. The PNCEBT-RS statistics recently reported that the prevalence of outbreaks was 1.9% and that of cases was 1.0% [23].

Genetic differentiation of *M. bovis* strains can provide better understanding about the evolution, epidemiology and ecology of bTB [24–27]. The molecular tools currently used worldwide for genotyping *M. bovis* isolates are spoligotyping and variable number tandem repeat (VNTR) analysis [28–31]. However, although these genotyping methods have proven useful in identifying local clusters, there are limitations regarding their ability to discriminate events within clusters involved in persistence and in assessing the spread of bTB [32, 33].

On the other hand, whole genome sequencing (WGS) has been successfully used to study multi-scale bTB epidemiology [26, 34, 35], offering much higher resolution than other established typing methods, which greatly improves the definition of the regional location of *M. bovis* strains [33, 36]. Among *M. bovis* WGS studies, many have focused on investigating

Impact Statement

Mycobacterium bovis is a causal agent of bovine tuberculosis (bTB), one of the most important diseases currently facing the cattle industry worldwide. Tracing the source of *M. bovis* infections of livestock is an important tool for understanding the epidemiology of bTB and defining control/eradication strategies. This is, to date, the largest whole genome sequencing-based study of *M. bovis* in Brazil and South America and will be of interest to many researchers in the bacterial genomics field. Our study highlighted the diversity of *M. bovis* in southern Brazil, identified potential within- and between-herd transmission often over large distances, and estimated a potential date of introduction of clonal complex Eu2 into the region that correlates with the historical introduction of cattle into the State of Rio Grande do Sul to improve existing local breeds. Further, the work showed that even with limited sampling, genome sequencing has the potential to be a useful tool for helping to control bTB transmission in Brazil.

bTB transmission and source within livestock populations [32, 37–40].

The first complete Brazilian *M. bovis* genome, SP38, was published in 2017 and belonged to clonal complex (CC) European 2 (Eu2) [41]. Using WGS-based multi-locus sequencing typing, a recent study showed evidence of *M. bovis* transmission events between farms and multiple introductions into a specific farm in the state of Santa Catarina, southern Brazil [42]. Another recent WGS study in northern Brazil has reported an endemic and unique *M. bovis* clade responsible for ongoing livestock transmission events in both cattle and buffalo [35].

In this study, 74 *M. bovis* isolates, sampled from naturally infected cattle taken from areas of occurrence of bTB in RS, southern Brazil, were sequenced. The aims of the study were to characterize the population structure and diversity of *M. bovis* in RS, attempt to identify potential transmission events over short and long distances, and gain insights into the use of a WGS-based surveillance programme to improve control of bTB in the cattle population in RS. Additionally, a global collection of 316 CC Eu2 isolates, including 64 from this study, was used to date the most recent common ancestor (MRCA) of this lineage and estimate the date of introduction of Eu2 into Brazil.

METHODS

Sample collection

All of the *M. bovis* isolates were obtained from naturally infected dairy and beef cattle, detected using comparative intradermal tuberculin testing (CITT) according to the rules of the PNCEBT [5], from bTB-positive farms (outbreaks) in

the period between 2015 and 2019 by the Secretariat of Agriculture, Livestock and Irrigation (SEAPI). Farms included in this study were chosen if they had a history of bTB, were geographically proximate and/or had a history of animal movements between them. For each infected animal, samples were collected from each lesion suggestive of tuberculosis (LST) at the same time point; for animals with multiple lesions, samples corresponding to each tissue type were cultured. Samples were collected from tissue (lung, liver, carcass) and lymph nodes (submandibular, pulmonary, parotid, mediastinal, mesenteric, retropharyngeal and tracheobronchial; File S1). Positive animals based on CITT were subjected to sanitary slaughter, according to Brazilian legislation. A total of 196 samples were collected from 78 animals from 13 farms.

Isolation and identification of *M. bovis*

All tissue samples with or without LST were macerated with 1.5 ml of sterile distilled water in a Magna Lyser apparatus (Roche Life Science), decontaminated by the Petroff method [43], subjected to cultivation in Stonebrink culture medium and incubated at 37 °C for up to 90 days. Confirmation of the isolates was performed by conventional PCR with primers that flank the differentiation region 4 (RD4), absent in this species but present in all other members of the *M. tuberculosis* Complex (MTBC) [44]. DNA extractions for genomic sequencing from the cultured *M. bovis* isolates were performed according to the protocol of van Embden *et al.* [45], with modifications (File S2). A total of 76 isolates were successfully cultured and sent for sequencing.

Sequencing and genomic analysis

The isolates were sequenced using the Illumina NextSeq 500 and HiSeq 2500 platforms. Raw sequence reads were trimmed using TRIMMOMATIC-v0.32 [46] and the species composition of the reads was assessed using Kraken [47] and Bracken [48]. Samples with fewer than 70% *Mycobacterium* reads were excluded (n=2). The trimmed sequence reads were mapped to the *M. bovis* AF2122/97 reference genome (NC0002945) using the Burrows-Wheeler Aligner (BWA; minimum and maximum insert sizes of 50 and 1000 bases) [49]. SNPs were processed with SAMtools mpileup and BCFtools (minimum base call quality of 50 and minimum root squared mapping quality of 30) [50, 51]. Genomic regions consisting of GC-rich sequences such as PPE proteins and PE-PGRS repeats were masked in the resulting alignment using previously published coordinates [52]. Variant sites in the alignment were extracted using snp-sites [53] and a maximum-likelihood phylogenetic tree was reconstructed using IQ-tree [54] accounting for constant sites in the alignment and run for 1000 bootstraps using the extended model selection function. The resulting tree was rooted using the AF2122/97 reference and annotated in iTOL [55]. Genomic clusters were assigned using fastBAPS [56] and pairwise SNP distances were calculated for all pairs of isolates using pairsnp (https://github.com/gtonkinhill/pairsnp).

For the detection of *in silico* spoligotypes of *M. bovis*, SpoTyping-v2.1 [57] was run on the raw sequence reads. The binary code representation of the spoligotypes was compared against the International *Mycobacterium bovis* Spoligotype Database (www.mbovis.org) to assign the spoligotype (SBXXXX) to each isolate [58]. RD-analyzer [59] was used to assign the CCs to the isolates (File S1).

Statistical analysis

The R [60] library ggmap [61] was used to construct a map showing the geographical locations, based on the latitude and longitude of each farm sampled, of each sequenced isolate. The same geographical coordinates were plotted using the R library ggplot [62] and a convex hull, calculated using the getConvexHull function from the R library contoureR [63] and representing the smallest polygon incorporating a particular set of points, was drawn around isolates from each fastBAPS cluster. The distHaversine function from the R library geosphere [64] was used to calculate the pairwise geographical distance in kilometres between each isolate. Mantel tests (1000 permutations to assess significance) were implemented using the R library ade4 [65–68] to test for an association between genetic and spatial distance. Putative transmission clusters were defined using two different pairwise SNP thresholds (15 SNPs to capture older transmission events and five SNPs to identify more recent transmission) using the R library iGRAPH [69]. The networks were then plotted with ggraph [70]. All isolates regardless of which animal they were collected from were treated as independent samples.

Contextual and dating analysis

To provide regional context for the 74 isolates sequenced in this study, sequence data for 35 published isolates from Brazil and Uruguay [35, 39, 41], and assemblies for 18 isolates from Argentina, Brazil and Paraguay [24, 71] (File S1) were downloaded from the European Nucleotide Archive (ENA) and the NCBI respectively. The assemblies were converted to fastq files using fastaq (https://github.com/sanger-pathogens/Fastaq) and the resulting data were processed as above.

In order to estimate a date for the introduction of CC Eu2 into RS, a global collection of 252 published Eu2 isolates was assembled [25, 40, 41, 72–76] (File S1). Quality control and mapping were performed as above and the isolates were combined with the 64 Eu2 isolates sequenced in this study. The presence of temporal signal in the Eu2 dataset was investigated by plotting the root to tip distance for each isolate, calculated using the R library phytools [77], against its sampling date (Fig. S1). The slope, x-intercept (MRCA), correlation coefficient and R^2 value were calculated in R. The Eu2 mapped alignment was used as input for BEAST v1.8.4 [78] and the tip sampling dates were used for calibration. Three runs of 10^8 Markov chain Monte Carlo (MCMC) iterations were performed using an HKY substitution model, strict or constant molecular clock, and constant or exponential population size and growth (12 separate runs). The performance of each model was assessed through the comparison of posterior marginal likelihood estimates (MLEs) and the model with the highest Bayes factor [79] (strict clock/constant population size) was
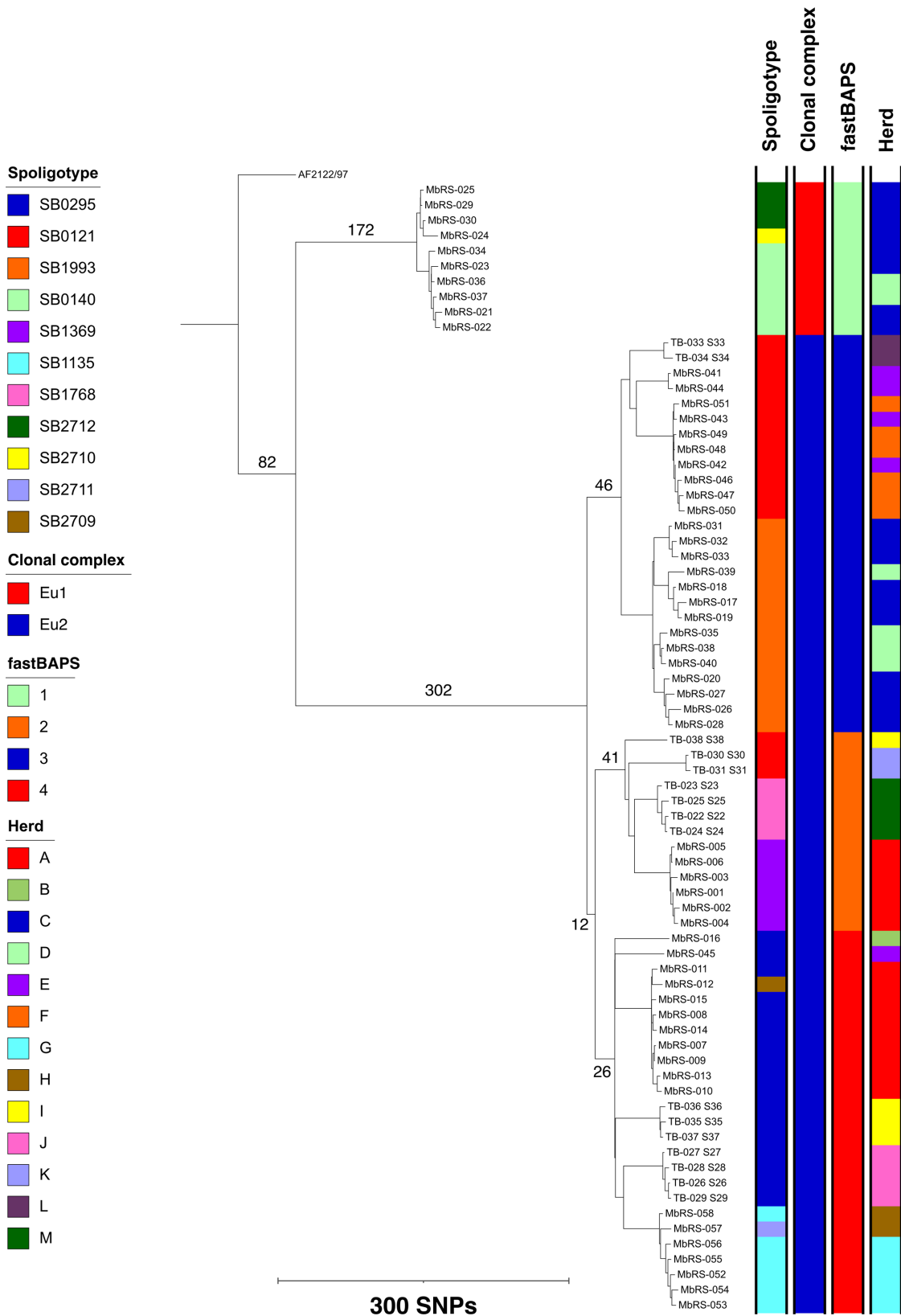
Fig. 1. Maximum-likelihood phylogenetic tree of 74 *M. bovis* isolates from Rio Grande do Sul, Brazil. The tree was rooted using the AF2122/97 reference. Branch lengths for the major branches are given in number of SNPs.
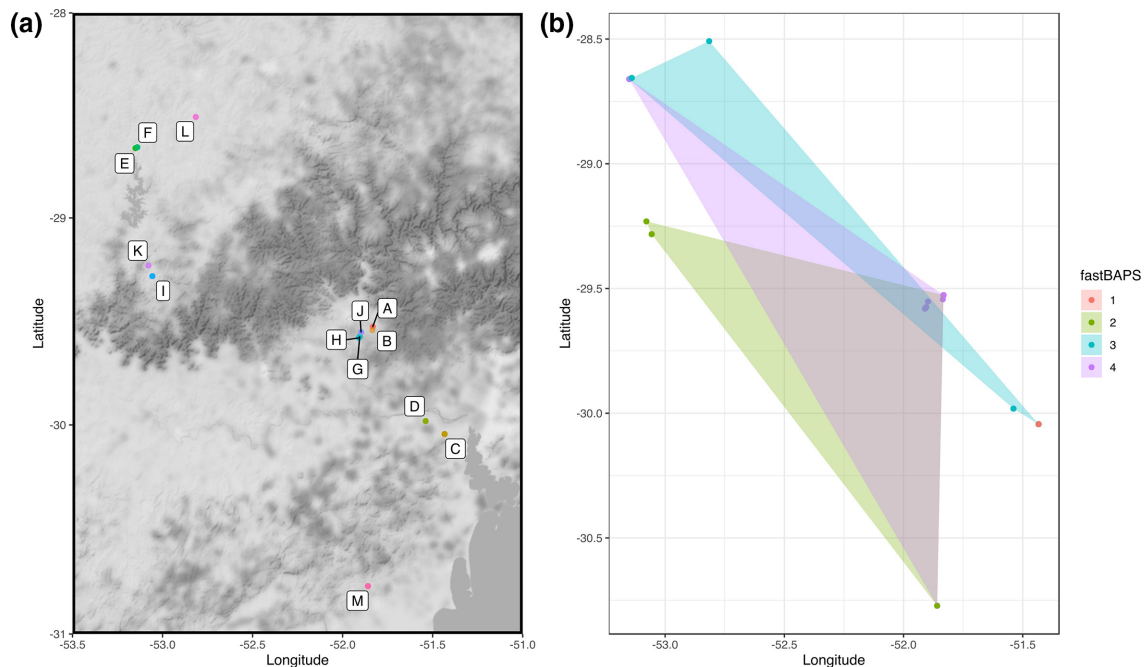
Fig. 2. (a) Location of the 13 farms in Rio Grande do Sul sampled in this study. (b) Spatial analysis of the distribution of isolates coloured by fastBAPS cluster. Each polygon represents the minimum convex polygon of the sampled locations of the isolates from each fastBAPS cluster.

selected (Table S1). The three MCMC runs were combined using LogCombiner v1.8.4 (10% burnin) and convergence was assessed [posterior effective sample size (ESS) >200 for each parameter] using Tracer v1.6. A maximum clade credibility tree summarizing the posterior sample of trees in the combined MCMC runs was produced using TreeAnnotator v1.8.4 and the resulting tree was annotated using ggtree [80]. In order to confirm the temporal signal in the dataset, the R library TIPDATINGBEAST [81] was used to generate 20 new datasets with randomly assigned dates; BEAST was then run on these new datasets using the same strict constant priors. As the estimated substitution rates in the observed data did not overlap with the estimated substitution rates in the randomized data, the temporal signal observed in the observed data was not obtained by chance (Fig. S2).

RESULTS

Dataset characterization

In total, 74 *M. bovis* isolates were sequenced having been obtained from the bacteriological culture of tissues of 56 cattle with or without LST. All colonies indicative of *M. bovis* were confirmed by conventional PCR for the RD4 flanking region. Isolates were obtained from 13 distinct herds (A–M) across three mesoregions in the state of RS, Brazil (Fig. 1). Genotypic characterization by *in silico* spoligotyping revealed seven previously identified profiles: SB0295 ($n=17$), SB0121 ($n=15$), SB1993 ($n=14$), SB1369 ($n=6$), SB1135 ($n=6$), SB0140 ($n=6$) and SB1768 ($n=4$). Four new spoligotypes were defined: SB2709 ($n=1$), SB2710 ($n=1$), SB2711 ($n=1$) and SB2712

($n=3$). Isolates with two different CCs, European 1 (Eu1) and Eu2, were present in the dataset with the majority (64/74, 86.5%) classified as Eu2 (File S1). The phylogenetic tree of the 74 Brazilian isolates, rooted with AF2122/97 as a reference (CC Eu1), showed a distinct phylogenetic structure with the two included CCs clearly segregating (Fig. 1). Sequence-based clustering using fastBAPS showed that there were four distinct clusters within the phylogeny: 1 ($n=10$), 2 ($n=13$), 3 ($n=26$) and 4 ($n=25$), with cluster 1 composed of isolates from Eu1 and clusters 2–4 comprising isolates from Eu2 (Fig. 1).

Spatial distribution of fastBAPS clusters

The locations of the 13 farms sampled in this study are shown in Fig. 2(a). The geographical distribution of the 74 sampled isolates was overlaid with polygons representing the spatial distribution of each fastBAPS cluster (Fig. 2b). These spatial distributions overlapped and could not be separated on the basis of sampling location, showing evidence of multiple *M. bovis* lineages circulating in the region.

Genetic diversity

A multimodal distribution was observed for pairwise SNP distances within the dataset, with the distinct transmission modes representing the population structure of the isolates (Fig. 3a). There were no differences in the distribution of pairwise SNP distances between isolates from the same herd or isolates from different herds, confirming the presence of multiple lineages circulating within and between different herds in the region. Similar levels of pairwise SNP diversity
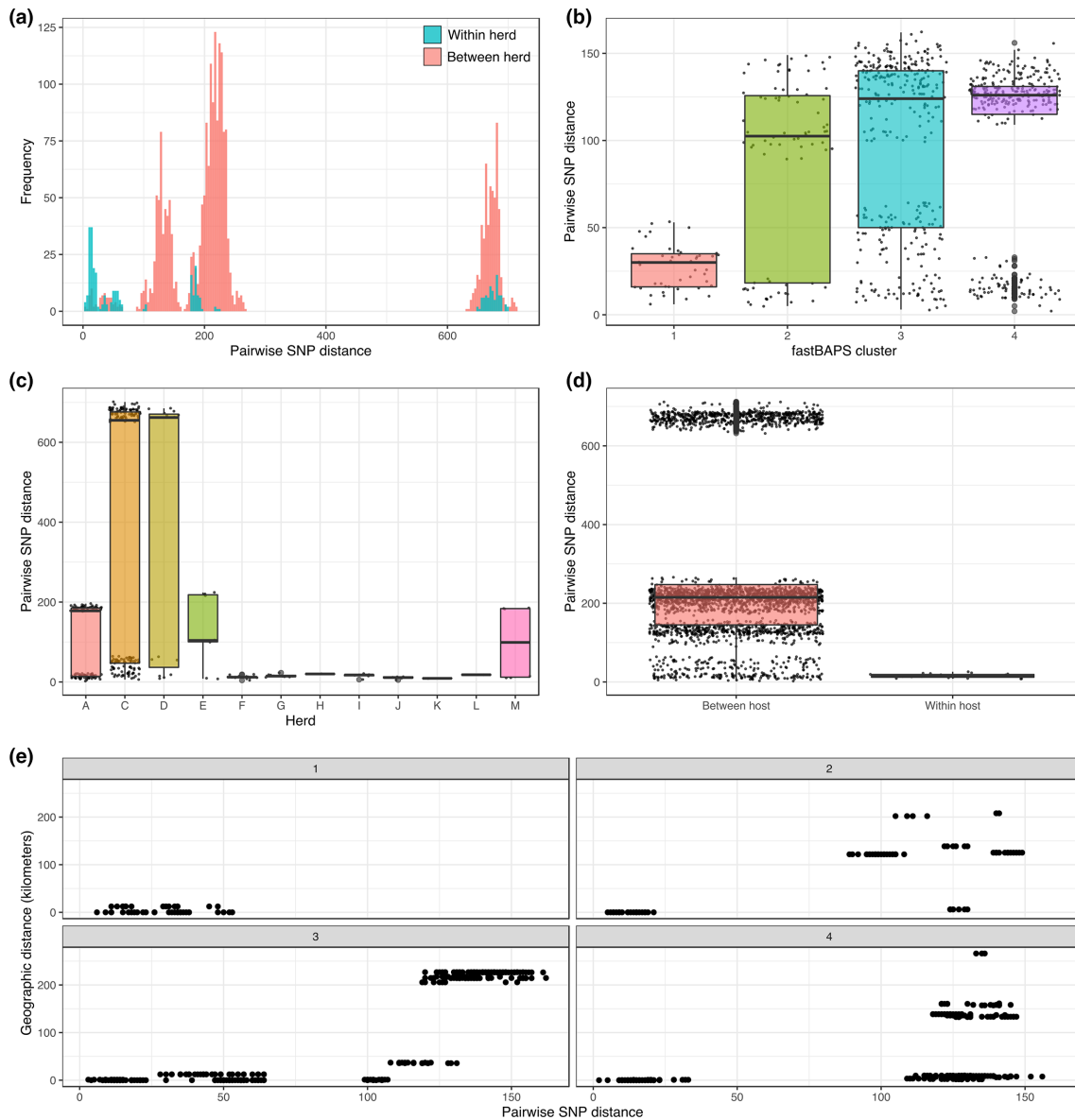
Fig. 3. (a) Histogram of pairwise SNP distances separated by between- and within-herd isolate pairwise distances. (b) Boxplot of all pairwise SNP distances separated by fastBAPS cluster. (c) Boxplot of all pairwise SNP distances separated by herd. (d) Boxplot of all pairwise SNP distances separated by within and between animal. (e) Scatterplots of pairwise SNP distance against geographical distance for all pairs of isolates separated by fastBAPS cluster.

were observed for fastBAPS clusters 2, 3 and 4, with the maximum number of pairwise SNPs between isolates in these clusters ranging between 149 and 162 SNPs (Fig. 3b). By comparison, there was much less diversity within fastBAPS cluster 1, with all isolates being within 53 SNPs of each other. Fig. 3(c) shows the distribution of pairwise SNP distances binned by herd; a high level of diversity was observed for farms C and D, reflecting the presence of both CCs in animals from these herds, whilst a medium level of diversity was seen in herds A, E and M. The remaining herds, F–L, showed much lower levels of diversity, which is probably a reflection of the small number of isolates collected from these herds.

The median number of SNPs between isolates from the same animal was 15 and ranged from eight to 26, whilst the number of SNPs between isolates from different animals varied between two and 712 with a median of 215 (Fig. 3d). Fig. 3(e) shows pairwise SNP distances plotted against geographical distances for all pairs of isolates separated by fastBAPS clusters. The result of the Mantel test for each fastBAPS cluster showed that there was no association between genetic and spatial distance for fastBAPS clusters 1 and 4, with a weak association detected for cluster 2 and a stronger association detected for cluster 3 (Table 1).

Table 1. Mantel test association scores for each fastBAPS cluster

Observed correlations close to 1 represent a strong association between spatial and genetic distance

fastBAPS cluster	Observed correlation	Simulated *p*-value
1	−0.01857867	0.615
2	0.5263921	0.001
3	0.7324717	0.001
4	−0.02190707	0.803

Genomic epidemiology and transmission

A total of 14 putative transmission clusters were defined using the conservative pairwise SNP threshold of 15 and varied in size between two and eight isolates with 15 isolates not part of a cluster; two of the clusters comprised solely of isolates from the same animal (Fig. 4a). The minimum within-herd pairwise SNP distance for the clusters varied between two and 14 SNPs. Only three of the putative clusters contained between-herd links, implying longer distance transmission: cluster 1 (herds C and D; 12.3 km apart), cluster 2 (herds F and G; 157.8 km apart) and cluster 3 (herds G and H; <1 km apart), with the remaining clusters all containing isolates from the same herd. A closer inspection showed that the minimum pairwise SNP distance for isolates from different herds in these clusters was 11, three and 14 SNPs for clusters 1, 2 and 3, respectively.

When the more stringent pairwise SNP threshold of five SNPs was applied, the number of putative transmission clusters dropped to four; 65 isolates were not assigned to a cluster (Fig. 4b). Only one of the three clusters contained links between different herds and was made up of isolates found in cluster 2 identified using the 15-SNP threshold. The remaining three clusters were each made up of two isolates from different animals in the same herds (herds A and J).

Comparison with other published *M. bovis* isolates

The 74 isolates from this study were combined with 53 published isolates from Argentina, Brazil, Paraguay and Uruguay to reconstruct a phylogenetic tree (Fig. 5a) that was rooted with *M. caprae* as the outgroup. The three CCs contained in the dataset, Eu1, Eu2 and unknown4, segregated into distinct lineages. The first published Eu2 genome from Brazil, SP38, sits amongst the Eu2 isolates from this study. Eu1 isolates from this study are both ancestral to and derived from Eu1 isolates from Uruguay and Argentina, suggesting two-way transmission between Brazil and neighbouring countries for isolates from this CC in RS.

In order to estimate a date of introduction for CC Eu2 into RS, the 64 sequenced Eu2 isolates from this study were combined with 252 published Eu2 isolates to build a time-scaled phylogenetic tree for the entire CC (Fig. 5b). The estimated age of the MRCA for Eu2 was 1762 AD [95% confidence interval (CI): 1725–1795]. All of the RS Eu2 isolates formed a clade with two recent French isolates; the estimated date of the
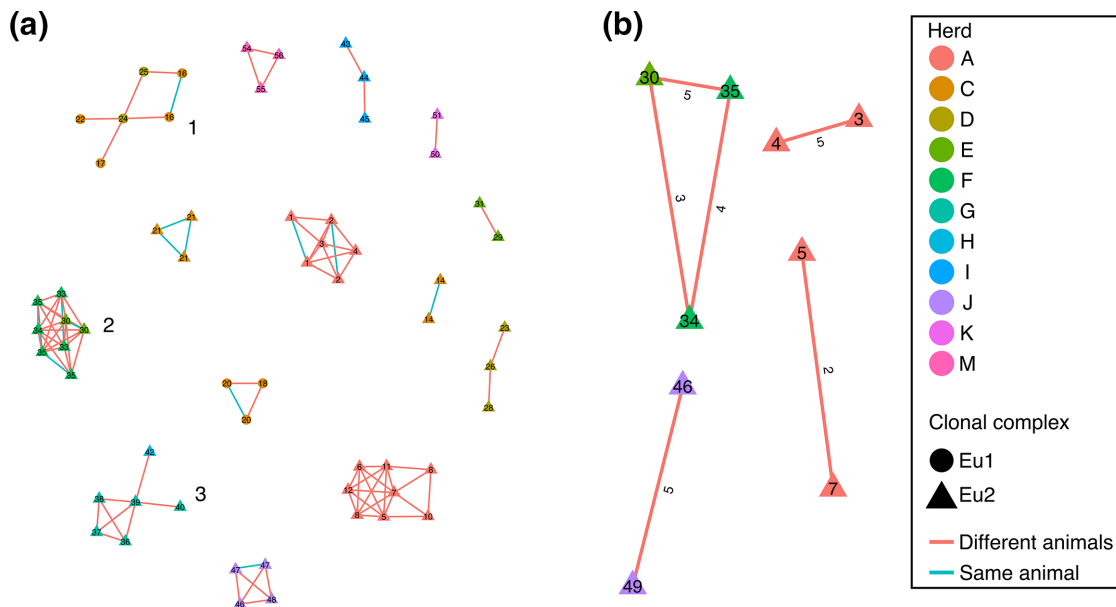
Fig. 4. (a) Putative transmission clusters generated using a pairwise SNP threshold of 15 SNPs and clusters with between-herd transmission events (clusters 1, 2 and 3) are labelled. (b) Putative transmission clusters generated using a pairwise SNP threshold of five SNPs. The pairwise SNP distances between each pair of isolates is shown. Nodes are coloured by herd and labelled with the animal ID. Clonal complexes are shown by different shapes and links between nodes are coloured according to whether the nodes were from the same animal or not.
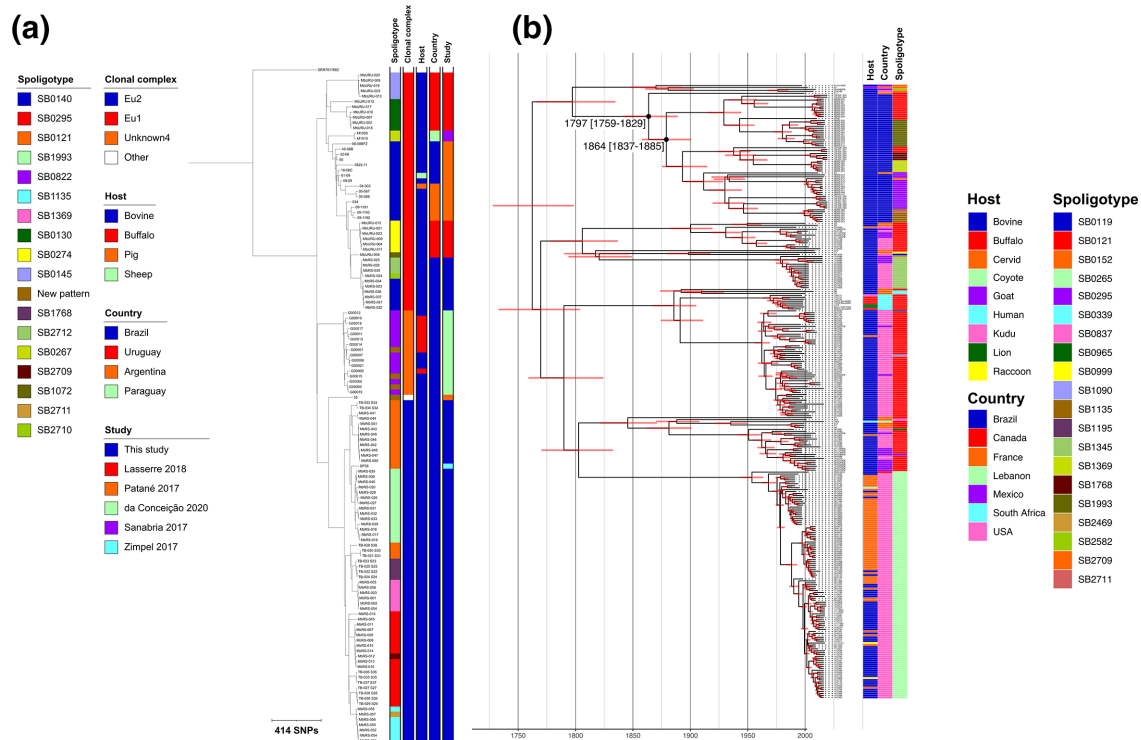
Fig. 5. (a) Maximum-likelihood phylogenetic tree of 127 South American *M. bovis* isolates. The tree was rooted using a *Mycobacterium caprae* isolate (SRR761662) as the outgroup. (b) Time-scaled maximum-likelihood phylogenetic tree of 316 Eu2 isolates.

last common ancestor between the French isolates and the Brazilian Eu2 isolates was 1797 AD [95 % CI: 1759–1829] with the date of the MRCA of the Brazilian isolates being 1864 AD [95 % CI: 1837–1885] (Fig. 5b). These dates imply a possible date of introduction of Eu2 into RS in the middle of the 19th century. The estimated substitution rate for the Eu2 dataset was 0.83 SNPs per genome per year.

DISCUSSION

To date, this is the largest WGS-based study of *M. bovis* in Brazil and South America and included 74 isolates sourced from infected cattle in RS, southern Brazil. WGS genotyping identified 11 distinct spoligotypes including four new profiles and two CCs, Eu1 and Eu2. Four genetic lineages were assigned using fastBAPS. The overlap of the spatial distributions of these four fastBAPS clusters and the overlap of the within-herd and between-herd pairwise SNP distances, as well as the lack of association identified in two of the fastBAPS clusters using Mantel tests, showed that there is little spatial clustering of the genetic lineages circulating in this region. This was further confirmed by the presence of multiple genetic lineages, shown by high levels of genetic diversity in the majority of herds in the study dataset. Using thresholds of five and 15 SNPs, both putative within- and between-herd transmission events could be identified, indicating relatively recent long-distance transmission that is probably due to cattle movements. The inclusion of previously published isolates from Argentina,

Brazil, Paraguay and Uruguay showed some evidence for the exchange of Eu1 between RS and Uruguay or Argentina. Dating analysis using a global Eu2 dataset suggested that Eu2 was introduced into southern Brazil sometime in the middle of the 19th century.

SB0295 was the most commonly observed spoligotype in this study, and is considered the second most frequently observed in Brazil [82]. Recently, this spoligotype was reported in northern Brazil [30], and had already been described in the states of Paraíba [83], Mato Grosso do Sul, Santa Catarina [84, 85], São Paulo [86], Bahia [87], Mato Grosso and Goiás [88]. SB0121 was the second most predominant spoligotype, and was also prevalent in previous studies in states of different Brazilian regions, including RS in southern Brazil [89]; São Paulo and Minas Gerais in south-eastern Brazil [84, 86]; Mato Grosso do Sul, Mato Grosso and Goiás in the midwest [84, 88]; and Bahia and Paraíba in the north-east [84, 87].

In a previous study, Eu1 was shown to be distributed across the Americas, with a high frequency in Argentina, Chile, Ecuador, Mexico and North America. In this study, Eu1 was found in 13.5 % (10/74) of samples, which is similar to previous estimates of prevalence in Brazil that ranged between 14.4 % (13/90) and 16 % (6/36) [83, 84]. The prevalence of Eu2, found in 64/74 (86.5 %) of isolates in this study, was similar to the prevalence identified in the states of Minas Gerais, São

Paulo, Mato Grosso, Mato Grosso do Sul, Goiás, Tocantins and Pará, where the prevalence was 81.1% (73/90) [90].

Given the presence of two more distantly related CCs within the dataset, the high level of genetic diversity observed between isolates from different animals was not unexpected. This high level of diversity was also reflected in the pairwise SNP distances observed in isolates from the same herds, in particular herds C and D (Fig. 3c). The analysis of within-host diversity identified isolates from the same animal that were up to 26 SNPs apart, highlighting the likelihood of these animals suffering from mixed infections, higher rates of within-host evolution or long-term infections that are not being detected. This level of within-host diversity is similar to that observed in cattle in Ethiopia [91], although further analyses of multiple isolates from the same animals in other settings will be required to see if this is common to all animals infected with bTB. Examination of the pairwise SNP diversity amongst the defined fastBAPS clusters showed that there was much less diversity amongst cluster 1, which comprised Eu1 isolates, when compared to the other three clusters comprising Eu2 isolates which had similar levels of diversity (Fig. 3b). Given the limited sampling and the close geographical proximity of the farms from which the cluster 1 isolates were collected, it is difficult to draw any strong conclusions but these results suggest that the local population of Eu1 has had less time to evolve and diversify and has possibly been circulating for a shorter time than Eu2.

We observed no correlation between genetic and spatial distance for fastBAPS clusters 1 and 4 but there was a stronger, significant correlation in clusters 2 and 3. Cluster 2 is almost entirely formed by six Jersey cows from a herd in the municipality of Teutônia (herd A) and six Holstein cows from two herds in Arroio do Tigre, approximately 177 km away, and belonging to the same mesoregion (Centro Oriental Rio-Grandense). The six Jersey cows from herd A were moved from the city of Candelária, 57 km from Arroio do Meio, 8 months previously. This suggests that there was a possible long-distance transmission of a lineage or lineages persisting in herds surrounding Arroio do Tigre and Candelária. The strong correlation in cluster 3 is probably due to the close geographical proximity of herds C and D (12 km apart) comprising beef cattle, that also share the same owner (see below), and herds E and F (1.1 km apart) which are dairy herds and share a stream between them. Whilst there are no official records of cattle movements between the dairy herds E, F and L, there is a strong local culture of animal replacements in dairy herds from nearby farms which are often unreported.

As there is currently no consensus as to the most appropriate SNP threshold to apply to defining transmission clusters in either bTB or *M. tuberculosis sensu stricto* [92–94], we used two different thresholds. Given the very slow mutation rate of *M. bovis* (0.5–1 SNP per genome every 2 years) [95], a threshold of 15 SNPs was chosen as it would potentially identify older transmission events and provide an indication of how likely transmission, recent or otherwise, could be identified using WGS. Using this threshold, we identified

a limited number of transmission events including both within- and between-herd transmission (Fig. 4a). Assuming recent transmission (within the last 5 years) is indicated by a pairwise SNP threshold of five SNPs or fewer, we could find evidence for recent within-herd transmission in three of the 13 herds (A, F, J) and evidence of between herd-transmission between herds C and D (Fig. 4b). Farms C and D belong to the same owner: farm C is for cow/calf operations, and provides weaned calves for fattening operations on farm D.

The spatial overlap of the fastBAPS clusters, high within-herd genetic diversity, and the general lack of an association between genetic and spatial distance gives strong evidence for the *M. bovis* population in this region being well mixed and probably maintained by short- and long-distance cattle movements. In the majority of countries where bTB is prevalent, there tends to be strong geographical localization of *M. bovis* genotypes such as VNTR and spoligotypes [95–97]. This localization is typically attributed to infection from contiguous herds, the presence of a wildlife reservoir such as the Eurasian badger in Great Britain and the wild boar in mainland Europe, and short-range movements of infected animals between herds [96, 98, 99]. This pattern is consistent across countries with very different *M. bovis* populations such as Great Britain, New Zealand and the USA [95–97]. One notable exception to this is Uruguay where clusters of infection change location from year to year, suggesting long-distance transmission [100]. Given the small sample size in our study, we can only speculate on the differences observed between RS and other settings. The state of RS is located in the extreme south of Brazil, and has environmental and cultural peculiarities that make it distinct from other states in Brazil; it has severe winters and the areas for livestock production are not vast, with consequent specificities in livestock production [101]. This means that there are localized areas of livestock production and dairy farming rather than a contiguous area across the state with the majority of infected herds concentrated in the north. We were able to identify processes associated with geographical localization, such as transmission occurring between contiguous herds, herds with the same owner and herds within the same region, but given the distribution of our fastBAPs clusters a higher level of long-distance transmission is probably helping to homogenize the *M. bovis* population across the state with little or no geographical clustering being observed. Additionally, whilst there are confirmed cases of bTB in wild boar in RS [102], to date no WGS has been done on these isolates to assess their similarity to local bovine isolates, so it would be premature to speculate on the role wild boar may play in maintaining bTB reservoirs in RS.

Our temporal analysis of the global Eu2 collection using BEAST estimated the substitution rate of Eu2 as 0.83 SNPs per genome per year, which is in line with previous estimates of *M. bovis* substitution rates [37, 95, 103]. Two previous studies have estimated the Eu2 MRCA date as part of wider *M. bovis* lineage dating analyses using different software and evolutionary models. Using a relaxed clock and constant population size model in BEAST, Zimpel *et al.* [27] estimated the MRCA of 183 Eu2 isolates and five closely related isolates

to fall between 355 AD and 1673 AD, although the exact date of the divergence between the Eu2 isolates and the five related isolates was not provided. Loiseau *et al.* [104] used a combination of various models in BEAST as well as LSD to estimate the date of the Eu2 MRCA as being between 1416 and 1705, although the dataset they used only included 17 Eu2 isolates due to the computational restraints typical to using BEAST. Our analysis estimated the date for the Eu2 MRCA to be 1762 with narrow 95% CIs of approximately 30 years either side of this date. Although this date falls outside the upper range of the previous two estimates, our larger dataset of 316 isolates, the evidence of strong temporal signal (Figs S2 and S3) and comparatively narrow CIs mean that we can be confident about the date we obtained.

Broadly speaking, the lineages observed in the Eu2 phylogeny are specific to a single country or neighbouring countries such as Mexico and the USA that have extensive cattle trade. With the caveat of limited sampling from countries such as Canada, this suggests that inter-country trade is not a major source of infection, particularly in North America. For example, the clade composed of SB0265 isolates, with the exception of the outgroup from Canada, is exclusively composed of isolates from the USA. The estimated date of the MRCA of the entire clade is 1954 [95% CI: 1944–1963] and of the American isolates is 1968 [95% CI: 1961–1974; Fig. 5b]. This lineage is associated with outbreaks in farmed cervids and cattle in Canada and states such as Nebraska, Indiana and Michigan [25].

The state of RS was originally located in Spanish territory, to the west of the area covered by the Treaty of Tordesillas. The first Europeans to colonize the state were Jesuit priests, who founded missions and imported cattle from the Iberian Peninsula where Eu2 is the predominant genotype [105]. After the expulsion of the Jesuits in 1759, the cattle became wild and spread throughout the south of Brazil, Uruguay and Argentina. These cattle, known as the creole breed, formed the basis of the beef industry in RS in the 18th century. To improve this industry, the government began to import British breeds in the middle of the 19th century, directly from the UK, or via Argentina and Uruguay [70]. The estimated date of introduction of Eu2 into southern Brazil sometime in the middle of the 19th century is consistent with this re-stocking rather than it already being present in the older cattle populations. The MRCA date for the entire CC, 1762, would also support this. Whilst we did not perform any temporal analysis on the Brazilian Eu1 isolates, it is possible this CC, found at a high frequency in former colonies and trading partners of the UK, including Argentina and Uruguay, was also introduced into southern Brazil at this time [106]. However, the much lower diversity observed in the Eu1 isolates compared to those from Eu2 suggest a much more recent date of introduction (Fig. 3b). Further sampling and sequencing alongside a robust temporal analysis of the global Eu1 population would be required to confirm this.

A limitation of this study was the small sample size compared to the number of infected cattle in the region, estimated to be in the region of 84000 head in 2013. Although this was due to limited available resources, clearly this hampered the detection of transmission events. However, we were still able to characterize significant diversity within the herds sampled and provide the first WGS-based description of the population structure of *M. bovis* in southern Brazil. Even with the limited sampling, we were able to detect within- and between-herd transmission at different SNP thresholds and found supporting evidence for one of these transmission events in the form of known cattle movements between farms with the same owner (herds C and D). Future studies could focus on collecting a larger number isolates from more herds with a focus on adjacent farms where the cattle movement data are available and bTB prevalence is consistently high.

Recently, the Brazilian bTB control programme (PNCEBT) [5] strategy has been revised: sanitation of outbreaks will be mandatory based on the classification of the federative units in relation to the degree of risk for bTB, and this was established in Normative Instruction 10/17 [5, 107]. With this new regulation, bTB state programmes can define strategies for eradicating the disease, based on a system for detecting bTB lesions in slaughterhouses and tracing back to the farm of origin, for sanitation of outbreaks. Evidently, surveillance for eradication must be supported by tools that allow an accurate assessment of the diversity and local persistence of outbreaks of bTB, as well as their dissemination, and WGS can be a useful and more reliable tool than traditional genotyping methods, as spoligotyping and MIRU-VNTR tend to be influenced by homoplasy [39, 108]. The results of this study show that genome sequencing has the potential to provide novel, detailed insights into local bTB transmission dynamics and should be considered as a useful tool for helping to control bTB transmission in RS.

Author contributions
R.A.R. designed the study, performed the culture of *M. bovis* isolates, DNA extraction, PCR, sequencing of some of the RS isolates and participated in part of the bioinformatics analysis. F.R.A. provided funds for the study, designed the study, performed the culture of *M. bovis* isolates, DNA extraction, PCR and sequencing of some of the RS isolates. A.M.R.D. performed the sequencing of some of the RS isolates. R.N.E. and J.P. helped write the manuscript. A.J.vT. designed the study, performed the analysis and helped write the manuscript. All authors provided comments on the draft of the manuscript.

Conflicts of interest
The authors declare that there are no conflicts of interest.

Ethical statement
All the animals in this study were culled following the guidelines of the Brazilian National Program for the Control and Eradication of Bovine Brucellosis and Tuberculosis [5]. No animals were slaughtered specifically for research.

References

1. **WHO, OIE, FAO, The Union**. In: *Roadmap for Zoonotic tuberculosis*. Geneva: WHO Press, 2017. https://www.oie.int/fileadmin/Home/eng/Our_scientific_expertise/docs/pdf/Tuberculosis/Roadmap_zoonotic_TB.pdf

2. **OIE**. In: *Chapter.3.4.6: Bovine tuberculosis*. OIE: Terrestrial manual, 2018. https://www.oie.int/fileadmin/Home/eng/Health_standards/tahm/3.04.06_BOVINE_TB.pdf

3. **Caminiti A**. In: *The Socio-economic Costs of Bovine tuberculosis*. Bulletin: World Organisation for Animal Health (OIE), 2019. https://oiebulletin.com/index.php?panorama=3-05-tb-costs-en&edition=7591&pdf=panorama&article=8480

4. **USDA**. In: *Livestock and Poultry: World Markets and Trade*. United States Department of Agriculture, Foreign Agricultural Service, 2020. https://apps.fas.usda.gov/psdonline/circulars/livestock_poultry.pdf

5. **Secretary of Agricultural Defense, Ministry of Agriculture, Livestock and Supply, Brazil**. In: *SDA normative instruction N° 10, March 3, 2017*. Brazil: Ministry of Agriculture, Livestock and Supply, Secretary of Agricultural Defense, 2017. pp. 1–4. http://www.in.gov.br/materia/-/asset_publisher/Kujrw0TZC2Mb/content/id/19124587/do1-2017%E2%80%9306-20-instrucao-normativa-n-10-de-3-de-marco-de-2017%E2%80%9319124353

6. **Secretariat of Agriculture, Livestock and Agribusiness, Department of Agricultural Defense, State of Rio Grande do Sul, Brazil**. In: *Normative Instruction SEAPA 002/2014*. Brazil: Secretariat of Agriculture, Livestock and Agribusiness, Department of Agricultural Defense, State of Rio Grande do Sul, 2014. https://www.agricultura.rs.gov.br/upload/arquivos/201611/01155402-pncebt-instrucao-normativa-seapi-002-com-alteracoes.pdf

7. **Ferreira Neto JS, da Silveira GB, Rosa BM, Gonçalves VSP, Grisi-Filho JHH**. Analysis of 15 years of the National program for the control and eradication of animal brucellosis and tuberculosis, Brazil. *Semina Ciênc Agrár* 2016;37:3385.

8. **Bahiense L, de Ávila LN, Bavia ME, Amaku M, Dias RA**. Prevalence and risk factors for bovine tuberculosis in the state of Bahia, Brazil. *Semina Ciênc Agrár* 2016;37:3549.

9. **de Barbieri JM, de Oliveira LF, Dorneles EMS, de Alencar Mota A, Gonçalves VSP**. Epidemiological status of bovine tuberculosis in the state of Minas Gerais, Brazil, 2013. *Semina Ciênc Agrár* 2016;37:3531–3548.

10. **Dias RA, Ulloa-Stanojlovic FM, Belchior APC, Ferreira RDS, Gonçalves RC**. Prevalence and risk factors for bovine tuberculosis in the state of São Paulo, Brazil. *Semina Ciênc Agrár* 2016;37:3673.

11. **Galvis JOA, Grisi-Filho JHH, Costa DD, Said A, Amaku M**. Epidemiologic characterization of bovine tuberculosis in the state of Espírito Santo, Brazil. *Semina Ciênc Agrár* 2016;37:3567.

12. **Guedes IB, Bottene IFN, Monteiro L, Leal Filho JM, Heinemann MB**. Prevalence and risk factors for bovine tuberculosis in the state of Mato Grosso do Sul, Brazil. *Semina Ciênc Agrár* 2016;37:3579.

13. **Lima PRB, do Nascimento DL, de Almeida EC, Pontual KAQ, Amaku M**. Epidemiological situation of bovine tuberculosis in the state of Pernambuco, Brazil. *Semina Ciênc Agrár* 2016;37:3601.

14. **Néspoli JMB, Negreiros RL, Amaku M, Dias RA, Ferreira F**. Epidemiological situation of bovine tuberculosis in the state of Mato Grosso, Brazil. *Semina Ciênc Agrár* 2016;37:3589.

15. **Queiroz MR, Groff ACM, Silva NDS, Grisi-Filho JHH, Amaku M**. Epidemiological status of bovine tuberculosis in the state of Rio grande do Sul, Brazil. *Semina Ciênc Agrár* 2016;37:3647.

16. **Ribeiro LA, Gonçalves VSP, Francisco PFC, de Alencar Mota ALA, do Nascimento GT *et al.*** Epidemiological status of bovine tuberculosis in the federal district of Brazil. *Semina* 2016;37:3561.

17. **Rocha WV, Jayme VDS, de Alencar Mota ALA, de Brito WMED, de Castro Pires GR *et al.*** Prevalence and herd-level risk factors of bovine tuberculosis in the state of Goiás, Brazil. *Semina* 2016;37:3625.

18. **Silva MDCP, Gonçalves VSP, de Alencar Mota ALA, Koloda M, Ferreira Neto JS *et al.*** Prevalence and herd-level risk factors for bovine tuberculosis in the state of Paraná, Brazil. *Semina* 2016;37:3611.

19. **Veloso FP, Baumgarten KD, de Alencar Mota ALA, Ferreira F, Ferreira Neto JS *et al.*** Prevalence and herd-level risk factors of bovine tuberculosis in the state of SANTA Catarina. *Semina* 2016;37:3659.

20. **Vendrame FB, Amaku M, Ferreira F, Telles EO, Grisi-Filho JHH**. Epidemiologic characterization of bovine tuberculosis in the state of Rondônia, Brazil. *Semina Ciênc Agrár* 2016;37:3639.

21. **ABIEC**. *Beef Report, Brazilian Livestock Profile. (*http://abiec.com.br/en/publicacoes/beef-report-2020-2/). Brazil: Brazilian Beef Exporters Association; 2020.

22. **IBGE**. *IBGE Indicators, Livestock Production Statistics 2019. (*https://biblioteca.ibge.gov.br/visualizacao/periodicos/2380/epp_2019_4tri.pdf). Brazil: Brazilian Institute of Geography and Statistics; 2020.

23. **SEAPDR**. *PNCEBT Statistics in RS: review 2019 (*https://www.agricultura.rs.gov.br/secao-de-vigilancia-zoosanitaria-pncebt). Secretariat of Agriculture, Livestock and Rural Development, Brazil: Porto Alegre; 2020.

24. **Patané JSL, Martins J, Castelão AB, Nishibe C, Montera L**. Patterns and processes of *Mycobacterium bovis* evolution revealed by phylogenomic analyses. *Genome Biol Evol* 2017;9:521–535.

25. **Orloski K, Robbe-Austerman S, Stuber T, Hench B, Schoenbaum M**. Whole genome sequencing of *Mycobacterium bovis* isolated from Livestock in the United States, 1989–2018. *Front Vet Sci* 2018;5:253.

26. **Crispell J, Cassidy S, Kenny K, McGrath G, Warde S**. *Mycobacterium bovis* genomics reveals transmission of infection between cattle and deer in Ireland. *Microb Genomics*.

27. **Zimpel CK, Patané JSL, Guedes ACP, de Souza RF, Silva-Pereira TT *et al.*** Global distribution and evolution of *Mycobacterium bovis* lineages. *Front Microbiol* 2020;11:843.

28. **Ghavidel M, Mansury D, Nourian K, Ghazvini K**. The most common spoligotype of *Mycobacterium bovis* isolated in the world and the recommended loci for VNTR typing; A systematic review. *Microb Pathog* 2018;118:310–315.

29. **Weerasekera D, Pathirane H, Madegedara D, Dissanayake N, Thevanesam V**. Evaluation of the 15 and 24-*loci* MIRU-VNTR genotyping tools with spoligotyping in the identification of *Mycobacterium tuberculosis* strains and their genetic diversity in molecular epidemiology studies. *Infect Dis* 2019;51:206–215.

30. **Carneiro PAM, Pasquatti TN, Takatani H, Zumárraga MJ, Marfil MJ**. Molecular characterization of *Mycobacterium bovis* infection in cattle and buffalo in Amazon Region, Brazil. *Vet Med Sci* 2020;6:133–141.

31. **Figueiredo Rocha VC, de Souza-Filho AF, Ikuta CY, Hildebrand E Grisi Filho JH, de Azevedo Issa M *et al.*** High discrimination of *Mycobacterium bovis* isolates in Brazilian herds by spoligotyping. *Prev Vet Med* 2020;179:104976.

32. **Trewby H, Wright D, Breadon EL, Lycett SJ, Mallon TR**. Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis. *Epidemics* 2016;14:26–35.

33. **Escárcega DAV, Razo CAP, Ruíz SG, Gallegos SLS, Suazo FM**. Analysis of bovine tuberculosis transmission in Jalisco, Mexico through whole-genome sequencing. *J Vet Res* 2020;64:51–61.

34. **Salvador LCM, O'Brien DJ, Cosgrove MK, Stuber TP, Schooley AM**. Disease management at the wildlife-livestock interface: Using whole-genome sequencing to study the role of elk in *Mycobacterium bovis* transmission in Michigan, USA. *Mol Ecol* 2019;28:2192–2205.

35. **da Conceição ML, Conceição EC, Furlaneto IP, da Silva SP, dos Santos GuimarãesAE**. Phylogenomic perspective on a unique *Mycobacterium bovis* clade dominating bovine tuberculosis infections among cattle and buffalos in Northern Brazil. *Sci Rep* 1747;2020:10.

36. Kao RR, Price-Carter M, Robbe-Austerman S. Use of genomics to track bovine tuberculosis transmission. *Rev Sci Tech* 2016;35:241–268.

37. Biek R, O'Hare A, Wright D, Mallon T, McCormick C. Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. *PLoS Pathog* 2012;8:e1003008.

38. Ghebremariam MK, Hlokwe T, Rutten V, Allepuz A, Cadmus S. Genetic profiling of *Mycobacterium bovis* strains from slaughtered cattle in Eritrea. *PLoS Negl Trop Dis* 2018;12:e0006406.

39. Lasserre M, Fresia P, Greif G, Iraola G, Castro-Ramos M. Whole genome sequencing of the monomorphic pathogen *Mycobacterium bovis* reveals local differentiation of cattle clinical isolates. *BMC Genomics* 2018;19:2.

40. Hauer A, Michelet L, Cochard T, Branger M, Nunez J. Accurate phylogenetic relationships among *Mycobacterium bovis* strains circulating in France based on whole genome sequencing and single nucleotide polymorphism analysis. *Front Microbiol* 2019;10:955.

41. Zimpel CK, Brandão PE, de Souza Filho AF, de Souza RF, Ikuta CY. Complete genome sequencing of *Mycobacterium bovis* SP38 and comparative genomics of *Mycobacterium bovis* and *M. tuberculosis* strains. *Front Microbiol* 2017;8:2389.

42. Anzai EK. *Whole Genome Sequencing of Mycobacterium bovis as an Instrument of Surveillance System in the State of Santa Catarina*. São Paulo: Faculty of Veterinary Medicine and Animal Science; 2019.

43. Makovcova J, Babak V, Slany M, Slana I. Comparison of methods for the isolation of mycobacteria from water treatment plant sludge. *Antonie van Leeuwenhoek* 2015;107:1165–1179.

44. Sales ML, Fonseca AA, ÉB S, Cottorello ACP, Issa MA. Evaluation of molecular markers for the diagnosis of *Mycobacterium bovis*. *Folia Microbiol* 2014;59:433–438.

45. van Embden JDA, van Soolingen D, Small PM, Hermans PWM. Genetic markers for the epidemiology of tuberculosis. *Res Microbiol* 1992;143:385–391.

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

47. Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol* 2014;15:R46.

48. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput Sci* 2017;3:e104.

49. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009;25:1754–1760.

50. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J. The sequence Alignment/Map format and SAMtools. *Bioinformatics* 2009;25:2078–2079.

51. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* 2011;27:2987–2993.

52. Price-Carter M, Brauning R, de Lisle GW, Livingstone P, Neill M *et al*. Whole Genome Sequencing for Determining the Source of *Mycobacterium bovis* Infections in Livestock Herds and Wildlife in New Zealand. *Frontiers in Veterinary Science* 2018;5:272.

53. Page AJ, Taylor B, Delaney AJ, Soares J, Seemann T. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genomics* 2016;4:e000056.

54. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.

55. Letunic I, Bork P. Interactive tree of life (iTOL) V3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 2016;44:W242–W245.

56. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res* 2019;47:5539–5549.

57. Xia E, Teo Y-Y, Ong RT-H. SpoTyping: fast and accurate in silico Mycobacterium spoligotyping from sequence reads. *Genome Med* 2016;8:19.

58. Smith NH, Upton P. Naming spoligotype patterns for the RD9-deleted lineage of the *Mycobacterium tuberculosis* complex; www.Mbovis.org. *Infect Genet Evol* 2012;12:873–876.

59. Faksri K, Xia E, Tan JH, Teo Y-Y, Ong RT-H. In silico region of difference (RD) analysis of Mycobacterium tuberculosis complex from sequence reads using RD-Analyzer. *BMC Genomics* 2016;17:847.

60. Team RC. *R: A Language and Environment for Statistical Computing*. Vienna, Austria; 2013.

61. Kahle D, Wickham H. ggmap: spatial visualization with ggplot2. *R J* 2013;5:144.

62. Wickham H. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York; 2016.

63. HamiltonN. contourer: Contouring of non-regular three-dimensional data. 2015. https://CRAN.R-project.org/package=contoureR

64. Hijmans RJ, Williams E, Vennes C. 2019. Geosphere: spherical trigonometry. https://CRAN.R-project.org/package=geosphere

65. Chessel D, Dufour AB, Thioulouse J. The ade4 package - I : One-table methods. *R news* 2004;4:5–10.

66. Dray S, Dufour A-B. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J Stat Softw* 2007;22.

67. Dray S, Dufour AB, Chessel D. The ADE4 package — II: Two-table and K-table methods. *R news* 2007;7:47–52.

68. Bougeard S, Dray S. Supervised Multiblock Analysis in *R* with the ade4 Package. *J Stat Softw* 2018;86.

69. Csardi G, Nepusz T. The igraph software package for complex network research. InterJournal. *Complex Systems* 2006;1695.

70. Pedersen TL. ggraph: An implementation of grammar of graphics for graphs and networks. 2020. https://CRAN.R-project.org/package=ggraph

71. Sanabria L, Lagrave L, Nishibe C, Ribas ACA, Zumárraga MJ. Draft genome sequences of two *Mycobacterium bovis* strains isolated from beef cattle in paraguay. *Genome Announc* 2017;5.

72. Dippenaar A, Parsons SDC, Miller MA, Hlokwe T, Gey van Pittius NC. Progenitor strain introduction of *Mycobacterium bovis* at the wildlife-livestock interface can lead to clonal expansion of the disease in a single ecosystem. *Infect Genet Evol* 2017;51:235–238.

73. Sandoval-Azuara SE, Muñiz-Salazar R, Perea-Jacobo R, Robbe-Austerman S, Perera-Ortiz A. Whole genome sequencing of *Mycobacterium bovis* to obtain molecular fingerprints in human and cattle isolates from Baja California, Mexico. *Int J Infect Dis* 2017;63:48–56.

74. Andrievskaia O, Duceppe M-O, Lloyd D. Genome sequences of five *Mycobacterium bovis* strains isolated from farmed animals and wildlife in Canada. *Genome Announc* 2018;6:e00258–18.

75. Panossian B, Salloum T, Araj GF, Khazen G, Tokajian S. First insights on the genetic diversity of MDR *Mycobacterium tuberculosis* in Lebanon. *BMC Infect Dis* 2018;18:710.

76. Perea Razo CA, Rodríguez Hernández E, Ponce SIR, Milián Suazo F, Robbe-Austerman S. Molecular epidemiology of cattle tuberculosis in Mexico through whole-genome sequencing and spoligotyping. *Plos One* 2018;13:e0201981.

77. Revell LJ. Phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 2012;3:217–223.

78. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the beast 1.7. *Mol Biol Evol* 2012;29:1969–1973.

79. Kass RE, Raftery AE. Bayes factors. *J Am Stat Assoc* 1995;90:773–795.

80. Yu G, Smith DK, Zhu H, Guan Y, Lam TT. ggtree : an r package for visualization and annotation of phylogenetic trees with

their covariates and other associated data. *Methods Ecol Evol* 2017;8:28–36.

81. Rieux A, Khatchikian CE. tipdatingbeast: an R package to assist the implementation of phylogenetic tip-dating tests using beast. *Mol Ecol Resour* 2017;17:608–613.

82. Zumárraga MJ, Arriaga C, Barandiaran S, Cobos-Marín L, de Waard J. Understanding the relationship between *Mycobacterium bovis* spoligotypes from cattle in Latin American Countries. *Res Vet Sci* 2013;94:9–21.

83. Higino SSDS, Pinheiro SR, de Souza GO, Dib CC, do Rosário TR. *Mycobacterium bovis* infection in goats from the Northeast region of Brazil. *Braz J Microbiol* 2011;42:1437–1439.

84. Parreiras PM, Andrade GI, do Nascimento TDF, Oelemann MC, Gomes HM *et al*. Spoligotyping and variable number tandem repeat analysis of *Mycobacterium bovis* isolates from cattle in Brazil. *Mem Inst Oswaldo Cruz* 2012;107:64–73.

85. Cazola DdeO, Jorge KdosSG, Zumárraga MJ, Souza-Filho AF, Araújo FR *et al*. Identificação e genotipagem de *Mycobacterium bovis* em bovinos positivos no teste intradérmico para tuberculose em Mato Grosso do Sul. *Pesq Vet Bras* 2015;35:141–147.

86. Rocha VCF, de Figueiredo SC, Rosales CAR, de Hildebrand e Grisi Filho JH, Keid LB *et al*. Molecular discrimination of *Mycobacterium bovis* in São Paulo, Brazil. *Vector Borne Zoonotic Dis* 2013;13:17–21.

87. Alzamora Filho F, Vasconcellos SEG, Gomes HM, Cavalcante MP, Suffys PN. Múltiplas estirpes de isolados de *Mycobacterium bovis* identificados por tipagem molecular em bovinos abatidos em matadouros-frigoríficos. *Pesqui Veterinária Bras* 2014;34:103–108.

88. Carvalho RCT, Vasconcellos SEG, Issa M de A, Soares Filho PM, Mota P. Molecular typing of *Mycobacterium bovis* from cattle reared in Midwest Brazil. *Plos One* 2016;11:e0162459.

89. Ramos DF, Silva ABS, Fagundes MQ, von Groll A, da SPEA. Molecular typing of *Mycobacterium bovis* isolated in the south of Brazil. *Braz J Microbiol* 2014;45:657–660.

90. ÉB S, de Alencar AP, Hodon MA, Soares Filho PM, de Souza-Filho AF. Identification of clonal complexes of *Mycobacterium bovis* in Brazil. *Arch Microbiol* 2019;201:1047–1051.

91. Almaw G, Mekonnen GA, Mihret A, Aseffa A, Taye H, *et al*. Population structure and transmission of *Mycobacterium bovis* in Ethiopia. *Microb Genom* 2021;7:000539.

92. Walker TM, CL I, Harrell RH, Evans JT, Kapatai G *et al*. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *Lancet Infect Dis* 2013;13:137–146.

93. Xu Y, Cancino-Munoz I, Torres-Puente M, Villamayor LM, Borras R *et al*. High-resolution mapping of tuberculosis transmission: whole genome sequencing and phylogenetic modelling of a cohort from Valencia region, Spain. *PLoS Med* 2019;16:e1002961.

94. Hatherell H-A, Colijn C, Stagg HR, Jackson C, Winter JR, *et al*. Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. *BMC Med* 2016;14:21.

95. Crispell J, Zadoks RN, Harris SR, Paterson B, Collins DM. Using whole genome sequencing to investigate transmission in a multi-host system: bovine tuberculosis in New Zealand. *BMC Genomics* 2017;18:180.

96. Smith NH, Gordon SV, de la Rua-Domenech R, Clifton-Hadley RS, Hewinson RG. Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*. *Nat Rev Microbiol* 2006;4:670–681.

97. Kao RR, Price-Carter M, Robbe-Austerman S. Use of genomics to track bovine tuberculosis transmission Rev. *Sci Tech* 2016;35:241–258.

98. O'Hagan MJ, Matthews DI, Laird C, McDowell SW. Herd-level risk factors for bovine tuberculosis and adoption of related biosecurity measures in Northern Ireland: a case-control study. *Vet J* 2016;213:26–32.

99. Milne G, Allen A, Graham J, Kirke R, McCormick C *et al*. *Mycobacterium bovis* population structure in cattle and local badgers: co-localisation and variation by farm type. *Pathogens* 2020;9:592.

100. Picasso C, Alvarez J, VanderWaal KL, Fernandez F, Gil A *et al*. Epidemiological investigation of bovine tuberculosis outbreaks in Uruguay (2011–2013). *Prev Vet Med* 2017;138:156–161.

101. Queiroz MR, Groff ACM, Silva NDS, Grisi-Filho JHH, Amaku M *et al*. Epidemiological status of bovine tuberculosis in the state of Rio grande do Sul, Brazil. *Semina* 2016;37:3647–3658.

102. Maciel ALG, Loiko MR, Bueno TS, Moreira JG, Coppola M *et al*. Tuberculosis in Southern Brazilian wild boars (*Sus scrofa*): First epidemiological findings. *Transbound Emerg Dis* 2018;65:518–526.

103. Menardo F, Duchêne S, Brites D, Gagneux S. The molecular clock of *Mycobacterium tuberculosis*. *PLoS Pathog* 2019;15:e1008067.

104. Loiseau C, Menardo F, Aseffa A, Hailu E, Gumi B *et al*. An African origin for *Mycobacterium bovis*. *Evol Med Public Health* 2020;2020:49–59.

105. Rodriguez-Campos S, Schürch AC, Dale J, Lohan AJ, Cunha MV *et al*. European 2 – a clonal complex of *Mycobacterium bovis* dominant in the Iberian Peninsula. *Infect Genet Evol* 2012;12:866–872.

106. Smith NH, Berg S, Dale J, Allen A, Rodriguez S *et al*. European 1: A globally important clonal complex of *Mycobacterium bovis*. *Infect Genet Evol* 2011;11:1340–1351.

107. Carneiro PAM, Kaneene JB. *Bovine tuberculosis* control and eradication in Brazil: lessons to learn from the US and Australia. *Food Control* 2018;93:61–69.

108. Comas I, Homolka S, Niemann S, Gagneux S. Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS ONE* 2009;4:e7815.