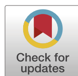
One Health 17 (2023) 100654

Contents lists available at ScienceDirect

One Health

journal homepage: www.elsevier.com/locate/onehlt



Whole genome sequencing improves the discrimination between *Mycobacterium bovis* strains on the southern border of Kruger National Park, South Africa

Eduard O. Roos[a,*,1], Johannes Loubser[a], Tanya J. Kerr[a], Anzaan Dippenaar[a,2], Elizma Streicher[a], Francisco Olea-Popelka[b], Suelee Robbe-Austerman[c], Tod Stuber[c], Peter Buss[d], Lin-Mari de Klerk-Lorist[e], Robin M. Warren[a], Paul D. van Helden[a], Sven D.C. Parsons[a,3], Michele A. Miller[a,*]

[a] *Department of Science and Innovation-National Research Foundation Centre of Excellence for Biomedical TB Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, PO Box 241, Cape Town 8000, South Africa*

[b] *Department of Pathology and Laboratory Medicine, Schulich Medicine & Dentistry, Western University, London, Ontario N6A 5C1, Canada*

[c] *National Veterinary Services Laboratories, Veterinary Services, Animal and Plant Health Inspection Service, United States Department of Agriculture, Ames, IA, USA*

[d] *Veterinary Wildlife Services, South African National Parks, Kruger National Park, Private Bag X402, Skukuza, 1350, South Africa*

[e] *Department of Agriculture Land Reform and Rural Development, Office of the State Veterinarian, Kruger National Park, PO Box 12, Skukuza, 1350, South Africa*

A R T I C L E I N F O

Keywords:
Mycobacterium bovis
Next-generation sequencing
Spoligotyping
Tuberculosis
Whole genome sequencing
Wildlife

A B S T R A C T

Background: Mycobacterium bovis forms part of the *Mycobacterium tuberculosis* complex and has an extensive host range and zoonotic potential. Various genotyping methods (e.g., spoligotyping) have been used to describe the molecular epidemiology of *M. bovis.* Advances in whole genome sequencing (WGS) have increased resolution to enable detection of genomic variants to the level of single nucleotide polymorphisms. This is especially relevant to One Health research on tuberculosis which benefits by being able to use WGS to identify epidemiologically linked cases, especially recent transmission. The use of WGS in molecular epidemiology has been extensively used in humans and cattle but is limited in wildlife. This approach appears to overcome the limitations of conventional genotyping methods due to lack of genetic diversity in *M. bovis*.
Methods: This pilot study investigated the spoligotype and WGS of *M. bovis* isolates ($n = 7$) from wildlife in Marloth Park (MP) and compared these with WGS data from other South African *M. bovis* isolates. In addition, the greater resolution of WGS was used to explore the phylogenetic relatedness of *M. bovis* isolates in neighbouring wildlife populations.
Results: The phylogenetic analyses showed the closest relatives to the seven isolates from MP were isolates from wildlife in Kruger National Park (KNP), which shares a border with MP. However, WGS data indicated that the KNP and MP isolates formed two distinct clades, even though they had similar spoligotypes and identical in silico genetic regions of difference profiles.
Conclusions: Mycobacterium bovis isolates from MP were hypothesized to be directly linked to KNP wildlife, based on spoligotyping. However, WGS indicated more complex epidemiology. The presence of two distinct clades which were genetically distinct (SNP distance of 19–47) and suggested multiple transmission events. Therefore, WGS provided new insight into the molecular epidemiology of the *M. bovis* isolates from MP and their relationship to isolates from KNP. This approach will facilitate greater understanding of *M. bovis* transmission at wildlife-livestock-human interfaces and advances One Health research on tuberculosis, especially across different host species.

* Corresponding authors.
E-mail addresses: dreoroos@gmail.com (E.O. Roos), miller@sun.ac.za (M.A. Miller).
[1] Present address: Clinomics, Uitzich Road, Bainsvlei, Bloemfontein, 9338, South Africa.
[2] Present address: Family Medicine and Population Health, Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium.
[3] Present address: Afrivet, Office Park, Newmark Estate, Silver Lakes Road, Hazeldean, Pretoria, 0081, South Africa

https://doi.org/10.1016/j.onehlt.2023.100654
Received 23 July 2023; Accepted 8 November 2023
Available online 12 November 2023
2352-7714/© 2023 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Mycobacterium bovis is part of the *Mycobacterium tuberculosis* complex (MTBC), a group of pathogenic mycobacteria that infect a wide range of hosts [1,2]. Members of the MTBC have high genetic similarity, with *M. bovis* being >99.95% similar to *M. tuberculosis* at the nucleotide level [3]. *Mycobacterium bovis* infection causes bovine tuberculosis (bTB) in animals, as well as zoonotic tuberculosis in humans who are exposed to infected animals or animal products [4].

In South Africa, previous studies have suggested that bTB in wildlife presents a transmission threat to livestock [2,5,6]. In particular, the study by Musoke et al. [2] concluded that the spillover was from Kruger National Park (KNP) wildlife to cattle on the border based on a shared spoligotypes (SB0121). Spoliogotyping is an assay that detects variability in direct repeat regions of *Mycobacterium* spp. Mycobacterial interspersed repetitive unit-variable number tandem repeat (MIRU-VNTR) patterns have also been used to characterize the prominent infecting strain (KNP VNTR 1), known as the "KNP-strain/cluster". Since there were no reports of bTB in communal cattle during this time (1996–2021), it was concluded that the spillover was most likely from KNP African buffaloes to cattle in the community.

This is a significant concern since bTB is a controlled disease, and spillover could result in socio-economic losses to subsistence farmers living next to game reserves in South Africa. There are further One Health implications in this setting as wildlife could share *M. bovis* with the livestock of the surrounding communities as well as the human population. These communities include subsistence farmers who are in contact with livestock and others who may drink unpasteurised milk. Therefore, the presence of bTB in wildlife presents an under recognized threat to public health in this setting.

The studies reporting spillover have relied on conventional genotyping methods such as spoligotyping and MIRU-VNTR typing for epidemiological investigations [2,6,7]. However, strains that share the same spoligotype or MIRU-VNTR genotype, and are part of the same clade, may be epidemiologically distinct. It has been shown that these conventional methods can overestimate transmission and could result in misleading conclusions that cases are epidemiologically linked [1,8]. The limitations of these methods have been thoroughly documented [1,5–7]. As a monomorphic bacterial pathogen *M. bovis* has little DNA diversity which conventional methods cannot capture as they focus on mobile or repetitive regions rather than a global genome wide comparison [1]. Another drawback of conventional genotyping methods is that there is little to no information on the molecular evolution of these genetic markers. In contrast, WGS interrogates the whole genome and leads to an increase in the sequence diversity based on ability to detect subtle differences between sequences, including single nucleotide polymorphisms [9,10]. This is important for epidemiological investigations since mutations are used to cluster isolates and determine the timeframe of transmission [11].

Studies have shown that *M. bovis* strains in naturally infected systems are often geographically localized, rather than host species-specific, with a single dominant strain present across species in the system [1,9,10]. The genetically fixed strain lacks diversity and therefore is a further limitation when using conventional genotyping methods to investigate the epidemiology of *M. bovis* outbreaks, and assigning directionality to transmission events [1,10]. The slow mutation rate of MTBC members contributes to the difficulty of detecting recent transmission events, especially using spoligotyping and MIRU-VNTR [10]. It has been suggested that clusters identified by spoligotyping represent transmission that occurred 200 years ago, whereas MIRU-VNTR clusters suggest transmission over the last 30 years [11].

In order to improve the resolution required to examine the molecular epidemiology of *M. bovis*, whole genome sequencing (WGS) has been employed [9–12]. This technique differentiates isolates on a nucleotide level which allows detection of a broad range of genomic variants, including single nucleotide polymorphisms (SNPs), insertions and deletions [13–15]. High resolution is especially important when exploring recent transmission in multi-host systems, as this allows the identification of epidemiologically linked cases (<5 SNP cut-off) and is used to determine transmission. However, the low mutation rate (mutation rate of 0.5 SNP per genome per year) of *M. bovis* is still a limitation with WGS [5,10,16,17]. Another advantage of WGS is its ability to add a time frame that allows for the estimation of when an event took place, compared to other genotyping techniques, although larger sample sizes are needed for accuracy [10]. The ability of molecular epidemiology to determine when isolates diverged increases our insights into *M. bovis* infections, especially in multi-host systems. Therefore, application of WGS will increase our ability to detect potential epidemiologically linked populations and interspecies transmission at interfaces [3,9,10].

Therefore, the aim of this study was to describe the genetic diversity of *M. bovis* isolates from wildlife in Marloth Park (MP), using high-throughput next-generation WGS to identify an epidemiological link to *M. bovis* infected wildlife in neighbouring KNP, a bTB endemic area. We hypothesized that these two populations would be infected with the same strain of *M. bovis,* based on the available spoligotyping and MIRU-VNTR data [2,5,7,9], which could be confirmed by WGS data.

2. Materials and methods

2.1. Sample collection and mycobacterial cultures

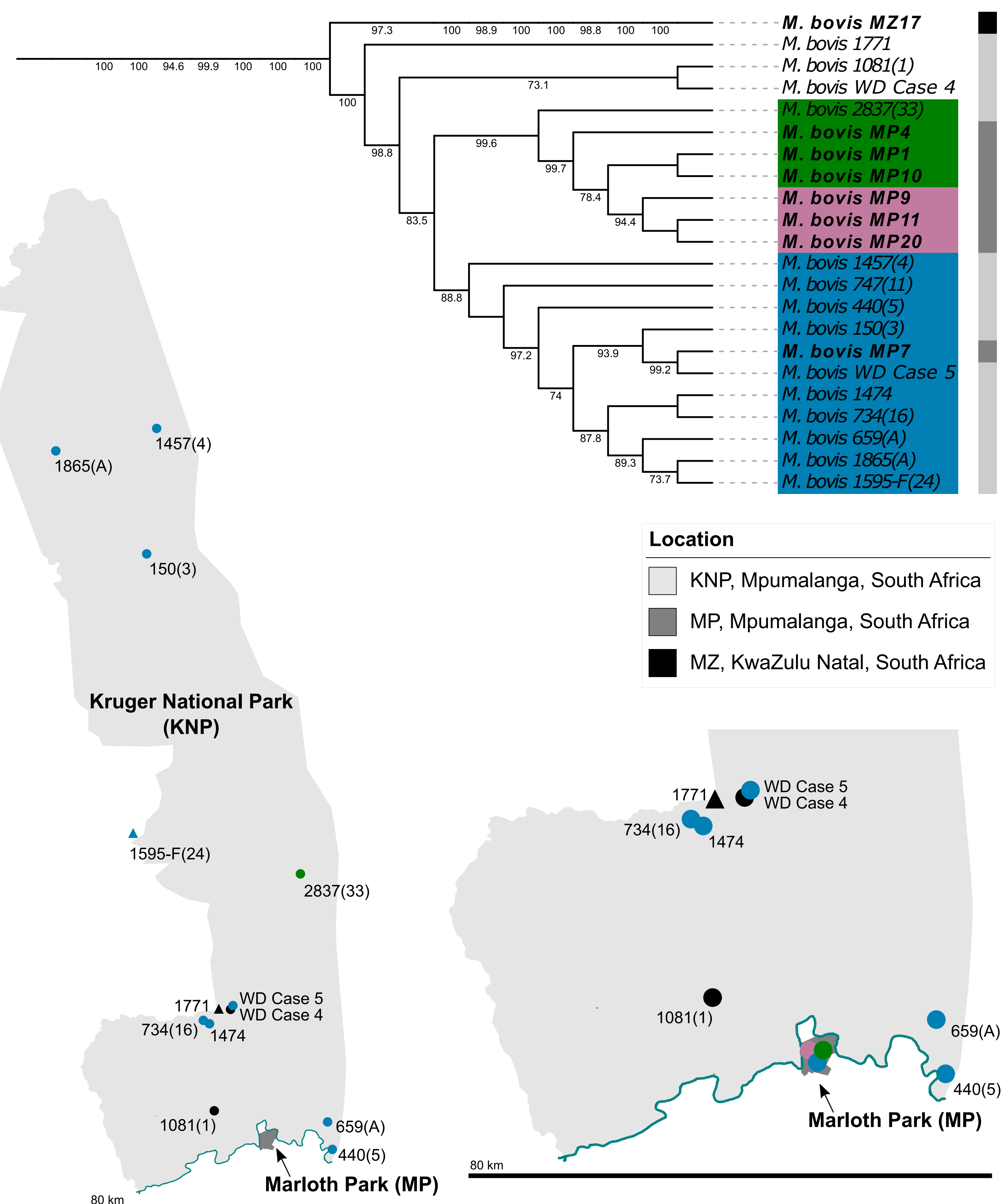
Tissue samples were collected from warthogs (*Phacochoerus africanus*) during routine disease surveillance in MP ($n = 7$), a residential wildlife estate established in 1977 (38 years prior to our sampling in 2015), which is separated from KNP by the Crocodile River (Fig. 1). A single sample from a warthog in a geographically distinct region 255 km to the south of MP, i.e., uMhkuze Nature Reserve (MZ), was also collected in 2015. Lymph nodes were sampled from all animals and stored at −20 °C until processed [18]. Mycobacterial cultures were performed in a BACTEC™ MGIT™ 960 system (BD Biosciences, Franklin Lakes, New Jersey, USA), as described by Goosen et al. [19].

2.2. Speciation and culture of *M. bovis* isolates

Samples with positive growth in MGIT were genetically speciated using 16S rDNA sequencing and genetic regions of difference (RD) analysis [20,21]. Isolates identified as *M. bovis* were inoculated (100 μl) onto Middlebrook 7H11 medium (BD Biosciences) agar plates (50 ml) supplemented with 0.5% sodium pyruvate (Sigma-Aldrich, St. Louis, Missouri, USA) and incubated for 6–8 weeks until confluent growth was observed. Thereafter, colonies were harvested from agar plates and DNA was extracted using two methods, the phenol/chloroform method described by Warren et al. [22] as well as boiling some of the scraped pure colonies for 30 min at 100 °C in phosphate-buffered saline (PBS). Bacterial isolates were subjected to spoligotyping twice, once as a crude boiled culture from the MGIT after *M. bovis* RD confirmation, and after DNA extraction (phenol/chloroform method and boiled pure colonies) from pure colonies on agar plates [23]. All spoligotypes from this study were compared to the *Mycobacterium bovis* Spoligotype database (https://www.mbovis.org/database) [24].

2.3. Whole genome sequencing and bioinformatic analyses

Extraction of DNA and whole genome sequencing (WGS) were performed on the 8 samples (MP = 7, MZ = 1) sent to United States Department of Agriculture National Veterinary Services Laboratories (NVSL) in Ames, Iowa. Sequencing was performed using the Illumina MiSeq platform (Illumina, San Diego, California, USA) using a paired-end approach (2 × 250 bp) with the 500-cycle MiSeq Reagent Kit v2 (Illumina). Library preparation was done with the Nextera XT DNA Library Preparation Kit (Illumina), following the manufacturer's instructions. Sequences were submitted to the European Nucleotide



(caption on next page)

Fig. 1. Map of Kruger National Park (KNP, light grey) and Marloth Park (MP, dark grey and arrow) bordering KNP. The enlarged map on the bottom right indicates the three different *Mycobacterium bovis* strains that were identified in MP. The coloured circles represent the approximate location of each isolate and black triangles represent isolates outside KNP and MP. The Crocodile River (in blue) acts as a geographical barrier between the two parks on the western, northern, and eastern sides of MP. The top right insert is the maximum likelihood molecular phylogeny inferred from a bootstrap consensus tree with 1000 replicates of the 31,640 variable positions between the concatenated sequences. The figure represents the relationship of the study isolates to its closest relative ancestors for visualisation purposes. All study isolates are highlighted within the cluster in bold. The full molecular phylogeny can be view in Supplementary Fig. 1. The analysis represents the evolutionary history of the taxa (Felsenstein 1989). The phylogenetic tree was produced by IQTree, based on the variable positions that were identified with respect to the *M. tuberculosis* H37Rv reference sequence (Stamatakis 2006, 2014) and visualised using the Interactive Tree of Life v6 online tool (Letunic & Bork 2019) (available at https://itol.embl.de/). All *M. bovis* strains are annotated with the location of origin. In summary: MZ17 – warthog, uMkuze, 2015; 1771 – leopard, KNP, 2000; 1081(1) – greater kudu, KNP, 1997; WDCase 4 – African wild dog, KNP, 2016; 2837(35) – buffalo, KNP, 2001; MP 1,4,7,9,10,11,20 – warthog, MP, 2015; 1457 (4) – buffalo, KNP, 1999; 747(11) – buffalo, KNP, NA; 440(5) – buffalo, KNP, 1998; 150(3) – buffalo, KNP, 1998; WDCase 5 – African wild dog, KNP, 2016; 1474 – lion, KNP, 1998; 731(16) – baboon, KNP, 1996; 659(A) – buffalo, KNP, 1996; 1865(A) – buffalo, KNP, 2000; 1595-F(24) – lion, KNP, 1999. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Archive (Supplementary Table S1), project accession: PRJEB27859.

Mycobacterium bovis sequences from this study ($n = 8$), along with previously published WGS for *M. bovis* isolates (from KNP wildlife, including lion, leopard, African buffalo, baboon, and greater kudu sampled 1997–2001, $n = 12$, and two African wild dogs sampled in 2016; other South African wildlife and cattle, $n = 5$, and other available sequences, $n = 15$) and a representative set of the MTBC sequences ($n = 41$, Supplementary Table S2; Supplementary Fig. 1) were analysed using various open source software, listed below [9,25]. Briefly, reads were trimmed with Trimmomatic version-0.32 [26]. Sequence reads were aligned to the reference genome of *M. tuberculosis* H37Rv (GenBank NC000962.3) with Novoalign (Novocraft, Selangor, Malaysia; version-3.02.13), Burrows-Wheeler Aligner (BWA) (version-0.6.2) [27], and SMALT (version-0.7.5) [28] (Supplementary Table S1). The Genome Analysis Tool Kit (GATK) and SAMTools (version-1.3) were used to identify single nucleotide variants, small insertions and deletions from the alignment files of the three different mapping algorithms [29]. Only variants that were present in all three alignments, according to GATK and SAMToools, overlapping in their position and base identity with an allele frequency of 95% were used. Additionally, a depth of coverage cut-off of 30%, relative to the average coverage across the genome, was applied to only include high-confidence variants. Variants in *pe/ppe* family genes, repeat regions, insertion sequences, and phages were excluded [30].

High confidence variable sites ($n = 31{,}640$), including coding and non-coding SNPs, were concatenated to generate a multi-FASTA file and used to construct a maximum likelihood phylogeny of the isolates included in this analysis with IQ-TREE with 1000 bootstrap pseudo-replicates, using an ultrafast and automatic model selection method [31]. The resulting tree file was visualised and annotated using iTOL v6 [32].

The SNPs identified in the study isolates were compared with one another to determine their pairwise distance using customised, in-house Python scripts. Isolates from various host species within KNP (African buffalo, greater kudu, baboon, African wild dog, lion, leopard; $n = 14$) were also selected and compared to isolates from MP to determine the number of variants between these isolates [9]. We used the proposed 5 SNP cut-off to identify "household" contacts and 15 SNP cut-off as a cluster identification cut-off and 2 SNPs as a highly related isolate cut-off value [4–6,33].

For in silico spoligotyping, SpoTyping (v2.0) was implemented. The paired fastQ files were used and input and the octal code output verified on the SITVIT database (http://www.pasteur-guadeloupe.fr:8081/SITVIT_Bovis/). Furthermore, in silico RD-analysis was performed by viewing the NOVO-align alignment files against the reference H37Rv *M. tuberculosis* genome in Artemis [34] and identifying the presence or absence of previously published RDs [9].

3. Results

All eight *M. bovis* isolates from warthogs that were newly sequenced had an average depth of coverage >50×, with >98.5% of reads mapping to the H37Rv reference genome. Spoligotyping identified two different genotypes (SB1275, $n = 3$; and SB0121, $n = 4$) from the seven MP *M. bovis* isolates and SB0140 for the MZ isolate (Table 1). The spoligotype patterns for the crude boiled MGIT cultures matched those of the extracted DNA samples. All *M. bovis* isolates from MP had the RDbovis(c)_Kruger region (23,877 bp) deleted but RD17 was intact (in silico analysis), while RD17 was deleted in the MZ isolate (Supplementary Table S3). Still, the *M. bovis* isolates from MP had the same in silico RD deletion profile as isolates from KNP (Supplementary Table S3) [9].

The phylogenetic analysis showed that the MP and KNP isolates formed two distinct clades (Fig. 1). This subclade was extracted from the comprehensive phylogenetic analyses, using the entire dataset of 31,640 loci. The paired inter-isolate SNP distances between 6 isolates from MP ranged between 0 and 15 variants, with 3 isolates sharing <5 SNPs indicative of close contacts (Fig. 2). The two MP subclades consisted of sequences from three warthog isolates each, with 2837(35) also clustering with the MP clade. This as an isolate from a KNP buffalo sampled in 2001 and had only 7–12 unique SNPs compared to the 6 closely related MP warthog isolates (Fig. 1 and Fig. 2). The MP7 isolate was an outlier with distances of 25–46 SNPs when compared to each of the other MP isolates (Fig. 2). The SNP distances between MP7 and KNP isolates from lion, leopard, baboon, buffalo, African wild dog, and greater kudu, were between 21 and 49 SNPs. Similarly, the other 6 MP warthogs varied from the KNP sequences by 21–49 SNPs (Fig. 1 and Fig. 2). Interestingly, the single warthog isolate from Mkuze (MZ17) was genetically distinct from the MP warthog sequences with SNP distances of 627–636 (Fig. 2).

4. Discussion

Using WGS, *M. bovis* isolates from KNP and MP were differentiated into two distinct clades, a "KNP-cluster/strain" and "MP-strain" which, by conventional genotyping methods, would have been missed. Sequences of isolates 2837 (KNP buffalo) and MP7 (warthog) appeared to occur outside their home range, which provides some evidence of lineages being transmitted from one location to another, although the findings suggest the absence of recent transmission between KNP and MP. The *M. bovis* isolates from KNP and MP shared a most recent common ancestor, although more isolates are required to determine the direction of transmission. Considering the evolutionary rate of 0.5 SNPs per year, the SNP distance between the MP and KNP isolates (21–49 SNPs) suggests divergence approximately 38 years ago, the time period between establishment of MP and sampling of warthogs [5,10,16,17]. In contrast, spoligotype analysis suggested continuous transmission between animals in KNP and MP, based on shared strain type (SB0121). Similarly, a recent study on *M. bovis* showed that WGS identified various clades from similar spoligotypes [35]. This was also true when comparing MIRU-VNTR data, a conventional genotyping method, with WGS for *M. tuberculosis* in a study by Gardy et al. [36], in which WGS could differentiate isolates into different clades with the same MIRU-VNTR genotype. Consequently, this study demonstrates that WGS provided greater resolution for understanding the molecular epidemiology

Table 1
Spoligotype pattern and number of *Mycobacterium bovis* isolates from warthogs (*Phacochoerus africanus*) in Marloth Park (MP) and uMhkuze (MZ), South Africa.

Pattern number	N	Location	Spoligotype pattern
SB0121	4	MP	
SB1275	3	MP	
SB0140	1	MZ	

of *M. bovis* than conventional genotyping methods, which may miss or overestimate transmission events, in the South African (SA) wildlife-livestock-human context.

The use of WGS, but not spoligotyping, was able to infer potential epidemiological links between isolates from warthogs in MP. The close genetic distance (<5 SNPs) between some of the MP isolates (MP9, MP11 and MP20) supports the notion that infection was the result of either exposure to a shared source (i.e., contaminated grazing) or potential transmission between individuals in close contact (i.e., burrowing together, using shared wallows or mother feeding young) [14,16,17]. If there was direct transmission, it would support the possibility that warthogs could serve as a reservoir host of *M. bovis* within MP. However, a larger WGS dataset would be required for the construction of spatial network models to confirm transmission within MP, since the slow mutation rate of *M. bovis,* at 0.5 SNPs per genome per year, is a caveat when determining transmission events [14,33]. However, our observations merit further investigation of *M. bovis* transmission occurring in warthogs in the MP setting. A greater effort to acquire more samples from MP will support development of a molecular clock and transmission network to assist in assigning transmission direction between the isolates [11,14,33].

Only the WGS results allowed for the differentiation of *M. bovis* isolates from KNP and MP into distinct clades. The WGS data further identified potential clonal expansion among MP isolates as WGS identified a sub-clade within MP, with three isolates forming the sub-clade (<5 SNPs), although this is speculative due to the small sample size. These three isolates further shared a unique spoligotype compared to the other MP spoligotypes. There was only a single spacer difference between SB0121 and SB1275 (loss of spacer 30), which suggests that SB1275 was derived from SB0121 and that it could become fixed in the MP population [37]. The phylogenetic analysis suggests that the MP isolates shared a more recent common ancestor with one another, compared to isolates from KNP [37,38]. Although we were not able to determine transmission direction, it seems likely that transmission between the two parks has occurred on more than one occasion. These events could have been followed by evolution and clonal expansion within the MP population after being geographically separated from KNP since MP was established in 1977 (expected SNP distance between MP and KNP 19 SNPs) [9]. These potential links would likely have been missed by conventional genotyping methods [1,9,14] since SNPs tend to mutate on a shorter time scale than changes in spoligotypes or MIRU-VNTR profiles [5]. However, evidence for clonal expansion requires additional sampling, including of other species in MP to rule-out having missed other lineages.

The higher diversity of *M. bovis* in KNP (average inter-isolate SNP distance 24.5 SD 8.5) compared to MP (average inter-isolate SNP distance 18.2 SD 14.9) indicates genetic diversity in both populations. This is not surprising since the *M. bovis* isolates were from diverse wildlife species in KNP, and likely have evolved from the initial introduction of *M. bovis* that occurred in the 1950s or 1960s [9]. It is interesting that the diversity in MP was in a single host species. This could be explained by repeated introductions and/or rapid mutation within the MP population, although more samples are needed to investigate these hypotheses.

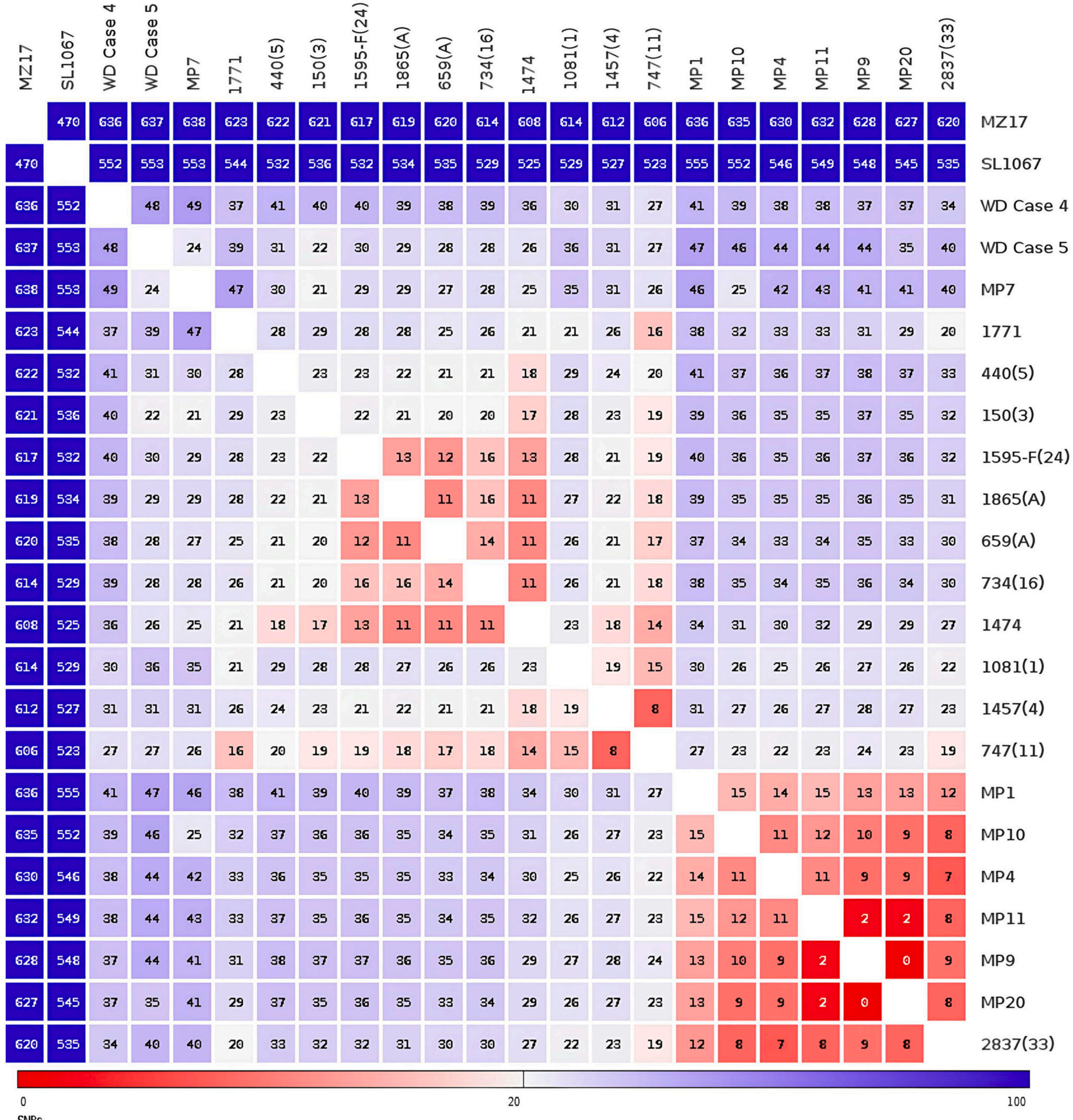
The KNP and MP *M. bovis* isolates were from animals on opposite sides of a man-made (i.e., fence) and geographical barrier (a river between KNP and MP) (Fig. 1). The WGS results suggest that transmission across these barriers rarely occur. This is in contrast to previous studies, using conventional genotyping methods, that suggest inter-species transmission frequently occurs across man-made barriers like fences [2,6–8]. Our results serve as a benchmark for future disease transmission studies in KNP and its surrounding ecosystems. Marloth Park is an example of scenarios where humans and wildlife are in proximity as well as surrounding communities with cattle that graze with the neighbouring wildlife. Characterization of genetic relatedness of *M. bovis* isolates from multiple species highlights the value of a One Health approach to improve understanding of transmission in this complex system.

This study was limited by the small number of samples and having isolates from only a single species for WGS from MP. Therefore, it is possible that an unsampled lineage from MP could have been missed. Therefore, our hypotheses need to be tested using an increased sample size. An additional limitation was the inability to culture additional isolates for WGS, due to overgrowth of non-tuberculous mycobacteria. Furthermore, no demographic data or observations were collected that could be used to determine potential contact between animals. Thus, future research needs to focus on increased sample numbers from multiple species to increase the WGS dataset. In addition, information to support spatial network analysis and modelling of *M. bovis* transmission, using Bayesian phylogenetic approaches, will improve elucidation of routes of spread. This will allow timestamping of the phylogenetic trees and enhance our understanding of transmission in multi-host systems.

5. Conclusions

This study determined that *M. bovis* isolates from multiple wildlife species in MP and KNP form two distinct clades, even though they share a recent common ancestor. This would have been missed by conventional genotyping methods and led to misinterpreting that isolates were epidemiologically linked, as has been suggested by previous data. The higher resolution of WGS, compared to conventional genotyping, has shown that there was significant genetic distance between the KNP and MP *M. bovis* isolates (21–49 SNPs), using a SNP cutoff of 5 for recent transmission. In addition, there was evidence of transmission outside the isolate home ranges. However, more data are needed to confirm the directionality of transmission. The WGS analysis further highlighted potential recent transmission among warthogs in MP, as some isolates had limited genetic distances (0–2 SNPs).

As WGS becomes more accessible and affordable, it will be invaluable for investigating the molecular epidemiology of *M. bovis* at wildlife-livestock-human interfaces [17,36,39]. This study demonstrated how WGS data can change our understanding of *M. bovis* transmission in the South African wildlife context [2,6,7], since it can detect and confirm epidemiologically linked cases with greater resolution than



	MZ17	SL1067	WD Case 4	WD Case 5	MP7	1771	440(5)	150(3)	1595-F(24)	1865(A)	659(A)	734(16)	1474	1081(1)	1457(4)	747(11)	MP1	MP10	MP4	MP11	MP9	MP20	2837(33)
MZ17		470	636	637	638	623	622	621	617	619	620	614	608	614	612	606	636	635	630	632	628	627	620
SL1067	470		552	553	553	544	532	536	532	534	535	529	525	529	527	523	555	552	546	549	548	545	535
WD Case 4	636	552		48	49	37	41	40	40	39	38	39	36	30	31	27	41	39	38	38	37	37	34
WD Case 5	637	553	48		24	39	31	22	30	29	28	28	26	36	31	27	47	46	44	44	44	35	40
MP7	638	553	49	24		47	30	21	29	29	27	28	25	35	31	26	46	25	42	43	41	41	40
1771	623	544	37	39	47		28	29	28	28	25	26	21	21	26	16	38	32	33	33	31	29	20
440(5)	622	532	41	31	30	28		23	23	22	21	21	18	29	24	20	41	37	36	37	38	37	33
150(3)	621	536	40	22	21	29	23		22	21	20	20	17	28	23	19	39	36	35	35	37	35	32
1595-F(24)	617	532	40	30	29	28	23	22		13	12	16	13	28	21	19	40	36	35	36	37	36	32
1865(A)	619	534	39	29	29	28	22	21	13		11	16	11	27	22	18	39	35	35	35	36	35	31
659(A)	620	535	38	28	27	25	21	20	12	11		14	11	26	21	17	37	34	33	34	35	33	30
734(16)	614	529	39	28	28	26	21	20	16	16	14		11	26	21	18	38	35	34	35	36	34	30
1474	608	525	36	26	25	21	18	17	13	11	11	11		23	18	14	34	31	30	32	29	29	27
1081(1)	614	529	30	36	35	21	29	28	28	27	26	26	23		19	15	30	26	25	26	27	26	22
1457(4)	612	527	31	31	31	26	24	23	21	22	21	21	18	19		8	31	27	26	27	28	27	23
747(11)	606	523	27	27	26	16	20	19	19	18	17	18	14	15	8		27	23	22	23	24	23	19
MP1	636	555	41	47	46	38	41	39	40	39	37	38	34	30	31	27		15	14	15	13	13	12
MP10	635	552	39	46	25	32	37	36	36	35	34	35	31	26	27	23	15		11	12	10	9	8
MP4	630	546	38	44	42	33	36	35	35	35	33	34	30	25	26	22	14	11		11	9	9	7
MP11	632	549	38	44	43	33	37	35	36	35	34	35	32	26	27	23	15	12	11		2	2	8
MP9	628	548	37	44	41	31	38	37	37	36	35	36	29	27	28	24	13	10	9	2		0	9
MP20	627	545	37	35	41	29	37	35	36	35	33	34	29	26	27	23	13	9	9	2	0		8
2837(33)	620	535	34	40	40	20	33	32	32	31	30	30	27	22	23	19	12	8	7	8	9	8	

Fig. 2. Single nucleotide polymorphism (SNP) distance matrix (red shading indicates less 20 SNPs and blue >100 SNPs distance between isolates) between *Mycobacterium bovis* isolates from Marloth Park (MP), Kruger National Park (annotated as isolate identifiers only), uMhkuze (MZ) and St. Lucia (SL), South Africa. All the isolate identifiers also correspond to those in Fig. 1. The SNP analysis was done in a pairwise manner. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

conventional genotyping methods [38]. New genomic approaches to improve detection of recent transmission events are crucial for supporting One Health research to elucidate intra- and inter-species spread of *M. bovis* at livestock-wildlife-human interfaces as well as understanding the contribution of environmental contamination to infection burden in complex ecosystems.

Funding source

This study was supported by the South African government through the South African Medical Research Council and NRF South African Research Initiative (SARChI, grant 86949), and the American Association of Zoological Medicine Wild Animal Health Fund (#7–2016). The content of this publication is the sole responsibility of the authors and none of the funding agencies were involved in any of the research at any

stage. The publication does not necessarily represent the views of the funders.

Ethical approval statement

The Stellenbosch University Animal Care and Use committee approved this study (SU-ACUD15–00029). Section 20 approval was granted by the South African Department of Agriculture, Land Reform and Rural Development (DALRRD 12/11/1/7/2 on 23 December 2014).

Declaration of Competing Interest

None to declare.

Data availability

Data will be made available on request.

Acknowledgements

The authors would like to extend their gratitude to the following people who contributed to the project: Dr. Jennie Hewlett, Marius Kruger, Leana Freese, Guy Hausler, Alicia McCall, Eduard Goosen, and the Veterinary Wildlife Services capture team from Kruger National Park, for their assistance during the sample collection and preparation in the lab.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.onehlt.2023.100654.

References

[1] N.H. Smith, S.V. Gordon, R. de la Rua-Domenech, R.S. Clifton-Hadley, R. G. Hewinson, Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*, Nat. Rev. Microbiol. 4 (2006) 670–681, https://doi.org/10.1038/nrmicro1472.

[2] J. Musoke, T. Hlokwe, T. Marcotty, B.J.A. du Plessis, A.L. Michel, Spillover of *Mycobacterium bovis* from wildlife to livestock, South Africa, Emerg. Infect. Dis. 21 (2015) 448–451, https://doi.org/10.3201/eid2103.131690.

[3] T. Garnier, K. Eiglmeier, J.-C. Camus, N. Medina, H. Mansoor, M. Pryor, S. Duthoy, S. Grondin, C. Lacroix, C. Monsempe, S. Simon, B. Harris, R. Atkin, J. Doggett, R. Mayes, L. Keating, P.R. Wheeler, J. Parkhill, B.G. Barrell, S.T. Cole, S.V. Gordon, R.G. Hewinson, The complete genome sequence of *Mycobacterium bovis*, Proc. Natl. Acad. Sci. U. S. A. 100 (2003) 7877–7882, https://doi.org/10.1073/pnas.1130426100.

[4] M.C. Hlavsa, P.K. Moonan, L.S. Cowan, T.R. Navin, J.S. Kammerer, G.P. Morlock, J. T. Crawford, P.A. LoBue, Human tuberculosis due to *Mycobacterium bovis* in the United States, 1995–2005, Clin. Infect. Dis. 47 (2008) 168–175, https://doi.org/10.1086/589240.

[5] T.M. Hlokwe, P. van Helden, A.L. Michel, Evidence of increasing intra and inter-species transmission of *Mycobacterium bovis* in South Africa: are we losing the battle? Prev. Vet. Med. 115 (2014) 10–17, https://doi.org/10.1016/j.prevetmed.2014.03.011.

[6] P.R. Sichewo, T.M. Hlokwe, E.M.C. Etter, A.L. Michel, Tracing cross species transmission of *Mycobacterium bovis* at the wildlife/livestock interface in South Africa, BMC Microbiol. 20 (2020) 49, https://doi.org/10.1186/s12866-020-01736-4.

[7] T.M. Hlokwe, P. van Helden, A. Michel, Evaluation of the discriminatory power of variable number of tandem repeat typing of *Mycobacterium bovis* isolates from southern Africa., Transbound, Emerg. Dis. 60 (Suppl. 1) (2013) 111–120, https://doi.org/10.1111/tbed.12096.

[8] D. Stucki, M. Ballif, M. Egger, H. Furrer, E. Altpeter, M. Battegay, S. Droz, T. Bruderer, M. Coscolla, S. Borrell, K. Zürcher, J.-P. Janssens, A. Calmy, J. Mazza Stalder, K. Jaton, H.L. Rieder, G.E. Pfyffer, H.H. Siegrist, M. Hoffmann, J. Fehr, M. Dolina, R. Frei, J. Schrenzel, E.C. Böttger, S. Gagneux, L. Fenner, Standard genotyping overestimates transmission of *Mycobacterium tuberculosis* among immigrants in a low-incidence country, J. Clin. Microbiol. 54 (2016) 1862–1870, https://doi.org/10.1128/JCM.00126-16.

[9] A. Dippenaar, S.D.C. Parsons, M.A. Miller, T. Hlokwe, N.C. Gey van Pittius, S. A. Adroub, A.M. Abdallah, A. Pain, R.M. Warren, A.L. Michel, P.D. van Helden, 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. 51 (2017) 235–238, https://doi.org/10.1016/j.meegid.2017.04.012.

[10] J. Crispell, R.N. Zadoks, S.R. Harris, B. Paterson, D.M. Collins, G.W. De-Lisle, P. Livingstone, M.A. Neill, R. Biek, S.J. Lycett, R.R. Kao, M. Price-Carter, Using whole genome sequencing to investigate transmission in a multi-host system: bovine tuberculosis in New Zealand, BMC Genomics 18 (2017) 180, https://doi.org/10.1186/s12864-017-3569-x.

[11] C.J. Meehan, P. Moris, T.A. Kohl, J. Pečerska, S. Akter, M. Merker, C. Utpatel, P. Beckert, F. Gehre, P. Lempens, T. Stadler, M.K. Kaswa, D. Kühnert, S. Niemann, B.C. de Jong, The relationship between transmission time and clustering methods in *Mycobacterium tuberculosis* epidemiology, EBioMedicine. 37 (2018) 410–416, https://doi.org/10.1016/j.ebiom.2018.10.013.

[12] D. Joshi, N.B. Harris, R. Waters, T. Thacker, B. Mathema, B. Krieswirth, S. Sreevatsana, Single nucleotide polymorphisms in the *Mycobacterium bovis* genome resolve phylogenetic relationships, J. Clin. Microbiol. 50 (2012) 3853–3861, https://doi.org/10.1128/JCM.01499-12.

[13] T. Rausch, T. Zichner, A. Schlattl, A.M. Stutz, V. Benes, J.O. Korbel, DELLY: structural variant discovery by integrated paired-end and split-read analysis, Bioinformatics. 28 (2012) i333–i339, https://doi.org/10.1093/bioinformatics/bts378.

[14] T.M. Walker, C.L.C. Ip, R.H. Harrell, J.T. Evans, G. Kapatai, M.J. Dedicoat, D. W. Eyre, D.J. Wilson, P.M. Hawkey, D.W. Crook, J. Parkhill, D. Harris, A.S. Walker, R. Bowden, P. Monk, E.G. Smith, T.E.A. Peto, Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study, Lancet Infect. Dis. 13 (2013) 137–146, https://doi.org/10.1016/S1473-3099(12)70277-3.

[15] A. Dippenaar, S.D.C. Parsons, S.L. Sampson, R.G. Van Der Merwe, J.A. Drewe, A. M. Abdallah, K.K. Siame, N.C. Gey Van Pittius, P.D. Van Helden, A. Pain, R. M. Warren, Whole genome sequence analysis of *Mycobacterium suricattae*, Tuberculosis. 95 (2015) 682–688, https://doi.org/10.1016/j.tube.2015.10.001.

[16] V. Nikolayevskyy, S. Niemann, R. Anthony, D. van Soolingen, E. Tagliani, C. Ködmön, M.J. van der Werf, D.M. Cirillo, Role and value of whole genome sequencing in studying tuberculosis transmission, Clin. Microbiol. Infect. 25 (2019) 1377–1382, https://doi.org/10.1016/j.cmi.2019.03.022.

[17] F. Zakham, S. Laurent, A.L. Esteves Carreira, A. Corbaz, C. Bertelli, E. Masserey, L. Nicod, G. Greub, K. Jaton, J. Mazza-Stalder, O. Opota, Whole-genome sequencing for rapid, reliable and routine investigation of *Mycobacterium tuberculosis* transmission in local communities, New Microbes New Infect. 31 (2019), 100582, https://doi.org/10.1016/j.nmni.2019.100582.

[18] E.O. Roos, P. Buss, L.-M. de Klerk-Lorist, J. Hewlett, G.A. Hausler, L. Rossouw, A. J. McCall, D. Cooper, P.D. van Helden, S.D.C. Parsons, M.A. Miller, Test performance of three serological assays for the detection of *Mycobacterium bovis* infection in common warthogs (*Phacochoerus africanus*), Vet. Immunol. Immunopathol. 182 (2016) 79–84, https://doi.org/10.1016/j.vetimm.2016.10.006.

[19] W.J. Goosen, M.A. Miller, N.N. Chegou, D. Cooper, R.M. Warren, P.D. van Helden, S.D.C. Parsons, Agreement between assays of cell-mediated immunity utilizing *Mycobacterium bovis*-specific antigens for the diagnosis of tuberculosis in African buffaloes (*Syncerus caffer*), Vet. Immunol. Immunopathol. 160 (2014) 133–138, https://doi.org/10.1016/j.vetimm.2014.03.015.

[20] M.-C. Leclerc, N. Haddad, R. Moreau, M.-F. Thorel, Molecular characterization of environmental *Mycobacterium* strains by PCR-restriction fragment length polymorphism of hsp65 and by sequencing of hsp65, and of 16S and ITS1 rDNA, Res. Microbiol. 151 (2000) 629–638, https://doi.org/10.1016/S0923-2508(00)90129-3.

[21] R.M. Warren, N.C. Gey van Pittius, M. Barnard, A. Hesseling, E. Engelke, M. de Kock, M.C. Gutierrez, G.K. Chege, T.C. Victor, E.G. Hoal, P.D. van Helden, Differentiation of *Mycobacterium tuberculosis* complex by PCR amplification of genomic regions of difference, Int. J. Tuberc. Lung Dis. 10 (2006) 818–822. http://www.ncbi.nlm.nih.gov/pubmed/16850559.

[22] R. Warren, M. de Kock, E. Engelke, R. Myburgh, N. Gey van Pittius, T. Victor, P. van Helden, Safe *Mycobacterium tuberculosis* DNA extraction method that does not compromise integrity, J. Clin. Microbiol. 44 (2006) 254–256, https://doi.org/10.1128/JCM.44.1.254-256.2006.

[23] J. Kamerbeek, L. Schouls, A. Kolk, M. van Agterveld, D. van Soolingen, S. Kuijper, A. Bunschoten, H. Molhuizen, R. Shaw, M. Goyal, J. van Embden, Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology, J. Clin. Microbiol. 35 (1997) 907–914, https://doi.org/10.1128/jcm.35.4.907-914.1997.

[24] N.H. Smith, P. Upton, Naming spoligotype patterns for the RD9-deleted lineage of the *Mycobacterium tuberculosis* complex; www.Mbovis.org, Infect. Genet. Evol. 12 (2012) 873–876, https://doi.org/10.1016/j.meegid.2011.08.002.

[25] P.A. Black, M. de Vos, G.E. Louw, R.G. van der Merwe, A. Dippenaar, E. M. Streicher, A.M. Abdallah, S.L. Sampson, T.C. Victor, T. Dolby, J.A. Simpson, P. D. van Helden, R.M. Warren, A. Pain, Whole genome sequencing reveals genomic heterogeneity and antibiotic purification in *Mycobacterium tuberculosis* isolates, BMC Genomics 16 (2015) 1–14, https://doi.org/10.1186/s12864-015-2067-2.

[26] A.M. Bolger, M. Lohse, B. Usadel, Trimmomatic: a flexible trimmer for Illumina sequence data, Bioinformatics. 30 (2014) 2114–2120, https://doi.org/10.1093/bioinformatics/btu170.

[27] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows-Wheeler transform, Bioinformatics. 25 (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.

[28] H. Ponsting, Z. Ning, SMALT - A New Mapper for DNA Sequencing Reads, in: F1000Posters, 2010, p. 1, https://doi.org/10.7490/F1000RESEARCH.327.1.

[29] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernytsky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, M.A. DePristo, The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing

data, Genome Res. 20 (2010) 1297–1303, https://doi.org/10.1101/gr.107524.110.
[30] M. Coscolla, A. Lewin, S. Metzger, K. Maetz-Rennsing, S. Calvignac-Spencer, A. Nitsche, P.W. Dabrowski, A. Radonic, S. Niemann, J. Parkhill, E. Couacy-Hymann, J. Feldman, I. Comas, C. Boesch, S. Gagneux, F.H. Leendertz, Novel *Mycobacterium tuberculosis* complex isolate from a wild chimpanzee, Emerg. Infect. Dis. 19 (2013) 969–976, https://doi.org/10.3201/eid1906.121012.
[31] B.Q. Minh, H.A. Schmidt, O. Chernomor, D. Schrempf, M.D. Woodhams, A. von Haeseler, R. Lanfear, IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era, Mol. Biol. Evol. 37 (2020) 1530–1534, https://doi.org/10.1093/molbev/msaa015.
[32] I. Letunic, P. Bork, Interactive tree of life (iTOL) v4: recent updates and new developments, Nucleic Acids Res. 47 (2019) W256–W259, https://doi.org/10.1093/nar/gkz239.
[33] A.J. Van Tonder, M.J. Thornton, A.J.K. Conlan, K.A. Jolley, L. Goolding, A. P. Mitchell, J. Dale, E. Palkopoulou, P.J. Hogarth, R.G. Hewinson, J.L.N. Wood, J. Parkhill, Inferring *Mycobacterium bovis* transmission between cattle and badgers using isolates from the randomised badger culling trial, PLoS Pathog. 17 (2021) 1–24, https://doi.org/10.1371/journal.ppat.1010075.
[34] T.J. Carver, K.M. Rutherford, M. Berriman, M.-A. Rajandream, B.G. Barrell, J. Parkhill, ACT: the Artemis comparison tool, Bioinformatics. 21 (2005) 3422–3423, https://doi.org/10.1093/bioinformatics/bti553.
[35] A. Hauer, L. Michelet, T. Cochard, M. Branger, J. Nunez, M.-L. Boschiroli, F. Biet, Accurate phylogenetic relationships among *Mycobacterium bovis* strains circulating in France based on whole genome sequencing and single nucleotide polymorphism analysis, Front. Microbiol. 10 (2019) 1–14, https://doi.org/10.3389/fmicb.2019.00955.
[36] J.L. Gardy, J.C. Johnston, S.J.H. Sui, V.J. Cook, L. Shah, E. Brodkin, S. Rempel, R. Moore, Y. Zhao, R. Holt, R. Varhol, I. Birol, M. Lem, M.K. Sharma, K. Elwood, S. J.M. Jones, F.S.L. Brinkman, R.C. Brunham, P. Tang, S.J. Ho Sui, V.J. Cook, L. Shah, E. Brodkin, S. Rempel, R. Moore, Y. Zhao, R. Holt, R. Varhol, I. Birol, M. Lem, M.K. Sharma, K. Elwood, S.J.M. Jones, F.S.L. Brinkman, R.C. Brunham, P. Tang, Whole-genome sequencing and social-network analysis of a tuberculosis outbreak, N, Engl. J. Med. 364 (2011) 730–739, https://doi.org/10.1056/NEJMoa1003176.
[37] N.H. Smith, The global distribution and phylogeography of *Mycobacterium bovis* clonal complexes, Infect. Genet. Evol. 12 (2012) 857–865, https://doi.org/10.1016/j.meegid.2011.09.007.
[38] A.M.S. Guimaraes, C.K. Zimpel, *Mycobacterium bovis*: from genotyping to genome sequencing, Microorganisms. 8 (2020) 667, https://doi.org/10.3390/microorganisms8050667.
[39] V. Nikolayevskyy, K. Kranzer, S. Niemann, F. Drobniewski, Whole genome sequencing of *Mycobacterium tuberculosis* for detection of recent transmission and tracing outbreaks: a systematic review, Tuberculosis. 98 (2016) 77–85, https://doi.org/10.1016/j.tube.2016.02.009.