frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 26 July 2024
DOI 10.3389/fmicb.2024.1416605

Check for updates

OPEN ACCESS

EDITED BY
Erica Lasek-Nesselquist,
Wadsworth Center, United States

REVIEWED BY
David Couvin,
Institut Pasteur de Guadeloupe, Guadeloupe
Guislaine Refrégier,
Université Paris-Saclay, France

*CORRESPONDENCE
Maria Beatrice Boniotti
✉ mariabeatrice.boniotti@izsler.it

RECEIVED 12 April 2024
ACCEPTED 05 July 2024
PUBLISHED 26 July 2024

CITATION
Scaltriti E, Iyad K, Boniotti MB, Menozzi I, Bolzoni L, Ippolito D, Ciarello FP, Loda D, D'Incau M, Zanoni M, Presti VDML, Mazzone P, Gavaudan S and Pacciarini ML (2024) Inside *Mycobacterium bovis* SB0120 spoligotype circulating in Italy: analysis of the most frequent genotypes by whole genome sequencing. *Front. Microbiol.* 15:1416605. doi: 10.3389/fmicb.2024.1416605

COPYRIGHT
© 2024 Scaltriti, Iyad, Boniotti, Menozzi, Bolzoni, Ippolito, Ciarello, Loda, D'Incau, Zanoni, Presti, Mazzone, Gavaudan and Pacciarini. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Inside *Mycobacterium bovis* SB0120 spoligotype circulating in Italy: analysis of the most frequent genotypes by whole genome sequencing

Erika Scaltriti[1], Karaman Iyad[2], Maria Beatrice Boniotti[2]*, Ilaria Menozzi[1], Luca Bolzoni[1], Dorotea Ippolito[3], Flavia Pruiti Ciarello[3], Daniela Loda[2], Mario D'Incau[2], Mariagrazia Zanoni[2], Vincenzo Di Marco Lo Presti[3], Piera Mazzone[4], Stefano Gavaudan[5] and Maria Lodovica Pacciarini[2]

[1]Risk Analysis and Genomic Epidemiology Unit, Istituto Zooprofilattico Sperimentale Della Lombardia e Dell'Emilia-Romagna (IZSLER), Parma, Italy, [2]National Reference Centre for Bovine Tuberculosis, Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia Romagna – IZSLER, Brescia, Italy, [3]Area Territoriale Barcellona Pozzo di Gotto, Istituto Zooprofilattico Sperimentale della Sicilia "A. Mirri", Barcellona Pozzo di Gotto, Messina, Italy, [4]Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche "Togo Rosati" (IZSUM), Perugia, Italy, [5]Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche "Togo Rosati" (IZSUM), Ancona, Italy

Bovine tuberculosis (bTB) is a chronic inflammatory disease primarily caused by *Mycobacterium bovis*. The infection affects domestic animals and wildlife, posing a zoonotic risk to humans. To understand the dynamics of transmission and genetic diversity in Italy's *M. bovis* population, we conducted whole-genome sequencing (WGS) analysis on two prevalent genotypes, belonging to Spoligotype SB0120, identified in different geographical and temporal contexts. By comparing these genomes with international *M. bovis* isolates, we identified a distinct clade within the lineage La1.2, encompassing the Italian SB0120 isolates, indicating a genomic segregation of Italian *M. bovis* from other European isolates. Within Italy, a significant level of genetic variability emerged across regions, while isolates within epidemiologically linked outbreaks exhibited minimal genetic diversity. Additionally, isolates derived from cattle and wild boars within a tuberculosis hotspot in Central Italy and from cattle and black pigs in Sicily formed unified clonal clusters. This indicates the presence of persistent strains circulating in the examined regions. The genetic diversity within herds was limited, as specific clones endured over time within certain herds. This research enhances our comprehension of the epidemiology and transmission patterns of bTB in Italy, thereby aiding the development of precise control strategies and disease management. Using WGS and implementing standardized protocols and databases will be pivotal in combating bTB and promoting One-Health approaches to address this noteworthy public health concern.

KEYWORDS

whole genome sequencing, bovine tuberculosis, *Mycobacterium bovis*, phylogeny, spoligotypes, VNTR-MIRU genotypes, SNPs

1 Introduction

Bovine tuberculosis (bTB) is a chronic inflammatory disease mainly caused by *Mycobacterium bovis* and, to a lesser extent, by *Mycobacterium caprae*. The infection can affect several domestic animals and wildlife species and has zoonotic relevance owing to its potential transmission to humans.

Mycobacterium bovis is a component of the *Mycobacterium tuberculosis* complex (MTBC) (Huard et al., 2006; Smith et al., 2006), characterized by 99,9% similarity at the nucleotide level and identical 16S rRNA sequences (Huard et al., 2006). MTBC is composed of highly clonal species with no evidence of genetic material transfer between lineages (Smith, 2012). MTBC derived from a common ancestor related to *Mycobacterium canettii* (Blouin et al., 2012), which includes different members: *M. tuberculosis, M. bovis, M. africanum, M. microti, M. caprae, M. canettii, M. pinnipedii, and M. origys* (Brosch et al., 2002; Rodriguez-Campos et al., 2012; van Ingen et al., 2012). It was suggested that members of the MTBC may be considered a series of host-adapted variants designated "ecotypes" of *M. tuberculosis* rather than species (Smith et al., 2006, 2009).

Four clonal complexes of *M. bovis* were defined based on distinct spoligotypes (Kamerbeek et al., 1997), Single Nucleotide Polymorphisms (SNPs) signatures, and/or specific deletions ranging from 806 to 14,094 bp: European 1 (Eu1), European 2 (Eu2), African 1 (Af1), and African 2 (Af2) (Müller et al., 2009; Berg et al., 2011; Rodriguez-Campos et al., 2012; Smith, 2012). However, these clonal complexes do not represent the whole diversity of *M. bovis* genomes (Zimpel et al., 2020). Recently, Branger et al. (2020) proposed the existence of the clonal complex "European 3," corresponding to the "cluster I" described by Hauer et al. (2019), mainly constituted by spoligotype SB0120, which is one of the most common spoligotypes of *M. bovis* responsible for bTB outbreaks in Italy, France, Germany, Zambia, and Algeria (Haddad et al., 2004; Boniotti et al., 2009; Munyeme et al., 2009; Sahraoui et al., 2009; Smith, 2012; Ghavidel et al., 2018). SB0120 is considered the ancestral spoligotype of *M. bovis* since it is generally accepted that spoligotype patterns evolve by loss (deletion) of spacer sequences of the DR locus (Rodríguez et al., 2010; Smith, 2012).

Genetic differentiation of strains has become an indispensable tool to study the evolution, epidemiology, and ecology of this pathogen, providing a means to detect and characterize the spread of pathogens in domestic livestock and establish eventual spread in the wildlife population.

Moreover, molecular typing has been used in many countries as a supporting tool to understand epidemiological links among outbreaks and to conduct efficient bTB control strategies. The most popular approach for a long time has been a combination of spoligotyping and Variable Numbers of Tandem Repeats-Mycobacterial Interspersed Repetitive Unit (VNTR-MIRU) with different markers selected for their high discrimination capacity and based on the MTBC population present in the territory (Boniotti et al., 2009; Gormley et al., 2014). Couvin et al. (2022) performed a global evaluation of the discriminative capacity of 12- and 15-loci VNTR-MIRU panels using data from the SITVITBovis database.[1]

In Italy, a study, performed on 1,560 *M. bovis/M. caprae* strains, allowed to select a panel of 12 VNTR-MIRU loci, which showed the most discriminative ability for VNTR-MIRU typing (Boniotti et al., 2009). Based on this study, molecular typing (spoligotyping/ VNTR-MIRU) has been applied since 2008 on *M. bovis/M. caprae* isolates collected in Italy, providing a national animal database (ITAN-TB) containing information about genotypes, species, and herd of origin, representing an important tool for epidemiological purposes.

Traditional genotyping provided preliminary information on the genetic characteristics of the *M. bovis* population in Italy and other countries, including the presence of *M. caprae* in more than 8% of bTB outbreaks, the low frequency of British Isle spoligotypes corresponding to EU 1 (Boniotti et al., 2009; Smith, 2012), and of SB0121, corresponding to the clonal complex EU 2, common in France, Portugal, and Spain (Duarte et al., 2010; Rodríguez et al., 2010; Hauer et al., 2016), the dominant presence of SB0120, described as the most prevalent (20–25%) in continental European countries (France, Germany, Belgium) (Kubica et al., 2003; Haddad et al., 2004; Hauer et al., 2016), and reported in approximately 50% of bTB outbreaks in Italy (Boniotti et al., 2009).

Although traditional typing performed with spoligotyping/ VNTR-MIRU is a worldwide recognized tool for epidemiological investigation (Hauer et al., 2019), recent studies show the limitations of these methods in tracing the origin of a given infection (Rodriguez-Campos et al., 2011; Guimaraes and Zimpel, 2020). In addition, even if these typing techniques target polymorphic genetic regions, they interrogate less than 1% of the genome and therefore have an intrinsically restricted discriminatory power (Hauer et al., 2019). Furthermore, homoplasy phenomena can occur in DR region (Corner et al., 1995; Smith, 2012). and VNTR loci (Comas et al., 2009) giving rise to identical patterns in separate sub-lineages (Comas et al., 2009; Trewby et al., 2016).

Although the estimated substitution rates of *M. bovis* range from 0.15 to 0.53 single nucleotide variants per genome per year (Trewby et al., 2016; Guimaraes and Zimpel, 2020), the comparison of entire genomes using Whole Genome Sequencing (WGS) has shown a high capacity of differentiation of *M. tuberculosis* and *M. bovis* strains (Biek et al., 2012; Homolka et al., 2012; Joshi et al., 2012; Roetzer et al., 2013; Luo et al., 2014). Accordingly, WGS has been demonstrated to be useful for the differentiation of *M. tuberculosis* strains with identical VNTR-MIRU profiles (Guimaraes and Zimpel, 2020).

In Italy, the combined use of spoligotyping and VNTR-MIRU (Boniotti et al., 2009) has highlighted the presence of high frequent profiles isolated in several herds, in different spatiotemporal contexts, apparently without any epidemiological correlation based on the available information.

Furthermore, WGS can improve the accuracy of the phylogenetic and epidemiological analysis of geographically characteristic genotypes over time (Lasserre et al., 2018, Orloski et al., 2018, Price-Carter et al., 2018, Hauer et al., 2019) and provide additional insights into the presence of new clonal complexes or lineages (Hauer et al., 2019; Kohl et al., 2020; Zimpel et al., 2020).

In the last few years, several genomic studies on *M. bovis* national relevant clones and populations have been published, allowing for the description of the phylogenetic and phylodynamic relationships among *M. bovis* isolates circulating in Europe and around the world (Hauer et al., 2019, Loiseau et al., 2020, Zwyer et al., 2021, Pozo et al., 2022). Regarding the *M. bovis* Italian population, very little is known

1 http://www.pasteur-guadeloupe.fr:8081/SITVIT_Bovis

about the association of specific clonal groups (Hauer et al., 2019) or global lineages (Loiseau et al., 2020; Zimpel et al., 2020) with the VNTR-MIRU genotypes circulating in Italy.

The present study aimed to employ SNP analysis based on WGS to identify accurate phylogenetic relationships among isolates belonging to the most frequent *M. bovis* SB0120/VNTR-MIRU profile in Italy, which originates from various epidemiological contexts and distinct geographical regions. Moreover, we compared their sequences with the SB0120 genomes already available and with the main sub-lineages circulating in Europe.

The most frequent SB0120/VNTR-MIRU profile circulating in Italy was selected based on a previous large-scale spoligotyping/VNTR-MIRU genotyping study (Boniotti et al., 2009) (Pacciarini, unpublished results). Our study includes: (1) isolates from outbreaks that were very distant geographically and temporally and (2) isolates that have been circulating for many years in the same restricted geographical area both in cattle and wild species.

Furthermore, this analysis aims to assess the genetic diversity within and between regions. Moreover, it seeks to examine genetic diversity within epidemiologically correlated outbreaks and within individual herds.

This genomic study will contribute to a more accurate description of the phylogenetic relationships between Italian and European *M. bovis* isolates of the SB0120 spoligotype, providing additional insight into the transmission and surveillance of bTB in Italy, defining molecular lineages (Hauer et al., 2019, Zimpel et al., 2020, Zwyer et al., 2021) as well as new clonal targeting profiles circulating in Italy.

2 Materials and methods

2.1 Microbiological and molecular characterization of *Mycobacterium bovis* isolates collected in ITAN-TB

As a national reference Centre for Bovine Tuberculosis (CRN-TB), we collected more than 7,000 *MTBC* isolates from 2000 to 2020 in Italy. The collection, called ITAN-TB, contains *M. bovis/M. caprae* strains mainly isolated from cattle and buffalo bTB outbreaks during the eradication program. The ITAN-TB database is not public, but its data can be requested by Italian authorities or the network of national laboratories for conducting epidemiological investigations.

All the isolates were identified by microbiological methods (Metchock et al., 1999) and by molecular methods described by Kulski et al. (1995). Mycobacteria of the MTBC were identified by PCR/RFLP of the *gyrB* gene, as reported by Boniotti et al. (2014), and by two High Resolution Melting (HRM) assays targeting the *gyrB* gene. The two assays were developed to identify *M. tuberculosis*, *M. microti, M. bovis,* and *M. caprae*: the first reaction allows distinguishing *M. tuberculosis*, *M. bovis/M. caprae*, and *M. microti*, while the second differentiates *M. bovis* from *M. caprae*. The first pair of primers, consisting of myc_HRM_1F (5'-GTCGAGATCAAGCGCGAC-3′) and myc_HRM_1R (5'-TTCGAAAACAGCGGGGTCG-3′), amplifies a fragment of 135 base pairs (bp), while the second pair of primers composed of TBCX_2F (5'-TGACGATATCCGGGAAGGC-3′) and TBCX_2R (5'-TCAAACCAGTGGGTCAGCTG-3′), amplifies a fragment of 150 bp. The reaction was performed in a total volume of 20 µL. One µl of DNA from the sample or control was added to a reaction mixture containing 7 µL of sterile H_2O, 10 µL of 2X mastermix (Bio-Rad), 1 µL of Primer Forward 10 µM, 1 µL of Primer Reverse 10 µM. Reference strains of *M. tuberculosis, M. microti*, *M. bovis,* and *M. caprae* were included as melting curve standards and positive controls in each experiment. The amplification was carried out with an initial activation of the DNA polymerase for 3 min at 98°C, followed by 50 cycles divided into the following two steps: 98°C for 5 s and 60°C for 10 s. A post-amplification cycling step was performed with a denaturation step at 95°C for 1 min and renaturation at 65°C for 1 min followed by an HRM ramping from 65°C to 95°C with an increment of 0,2°C every 10 s. The HRM assay was performed on the CFX96™ Real-Time PCR Detection System (Bio-Rad) with the SsoFast™ EvaGreen® Supermix (Bio-Rad). The software used for the analysis of amplification and melting curves are CFX Maestro (Bio-Rad) v2.2 and Precision Melt Analysis™ (Bio-Rad) v1.3., respectively.

From 2008, *M. bovis/M. caprae* strains were typed systematically by spoligotyping (Kamerbeek et al., 1997) and by VNTR-MIRU with ETRA-E (Frothingham and Meeker-O'Connell, 1998) and 7 additional VNTR/MIRU markers selected in the study of Boniotti et al. (2009), obtaining more accurate genetic profiles. Spoligotyping was performed as described by Kamerbeek et al. (1997). The spacer sequences contained in the direct repeat locus were detected by hybridization onto a spoligotyping membrane (Mapmygenome, Hyderabad, India). Spoligotypes were named according to an agreed international convention by using the www.mbovis.org database (Smith, 2012). VNTR-MIRU typing was performed on 12 VNTR-MIRU loci: ETRA (2165), ETRB (2461), ETRC (577), ETRD (580), ETRE (3192), Qub11a (2163a), Qub11b (2163b), Qub26 (4052), Qub1895 (1895), Qub15 (3155), Qub3232 (3232), and MIRU26 (2996) (Frothingham and Meeker-O'Connell, 1998; Supply et al., 2000; Roring et al., 2002; Skuce et al., 2002). The panel contains the 8 loci recommended by the European consortium SSPE-CT-2004-501903 Veterinary Network of Laboratories Researching into Improved Diagnosis and Epidemiology of Mycobacterial Diseases (VENoMYC).

2.2 Selection of *Mycobacterium bovis* strains for this study

To select the strains for this study, we evaluated the ITAN-TB database collection from 2008 to 2018, which contains mainly *M. bovis/M. caprae* strains isolated from bTB outbreaks during the eradication program and genotypes isolated from other domestic and wild species. Since nearly all the North and Centre of Italy achieved a disease-free status (DFS) in the last 10 years, the ITAN-TB mainly contains MTBC isolates from Southern regions, especially Sicily, Campania, Apulia, and Calabria, where bovine tuberculosis is still present. Therefore, most of the *M. bovis* strains analyzed in this study were isolated from these regions.

In this study, we selected a set of 76 SB0120 isolates, focusing on the most frequent genotype, based on VNTR-MIRU profile (Supplementary Tables S1, S3). In particular, we considered: (i) thirty-six strains belonging to the main VNTR-MIRU genotype circulating in Italy from different regions across Italy, including Campania, Sicily, Apulia, Calabria, Veneto, and Lazio. These strains, characterized by the genetic profile VNTR-MIRU 4,5,5,3,3,10,4,4,4,3,6,5, here named Italian Genotype 1, were

representative of the Italian population (in red Supplementary Figure S1). They were chosen to reflect the temporal and regional distribution observed between 2008 and 2018, (ii) twenty-three strains, also belonging to the Italian Genotype 1, were specifically sampled from cattle and black pigs within a delimited geographic area of Sicily, around the Caronia municipality, which is located in the Nebrodi Park (with an area of nearly 860 km^2 in the Messina province). Over a three-year period (2015–2018), these isolates were obtained from animals with well-documented epidemiological backgrounds (Supplementary Figure S2); Caronia, marked by its extensive farming systems and shared pastures, provided a unique setting for studying disease dynamics, and (iii) seventeen isolates, designated as the Matelica Clade and belonging to SB0120, were identified in cattle and wild boars originating from a geographically area (about 1,000 km^2) around the Matelica municipality in the Macerata province over an eight-year span. These specimens exhibited the genetic profile VNTR-MIRU 3,3,5,3,3,10,4,4,4,3,6,5. They were collected from herds and hunting areas proximal to livestock farms in the Marche region of Central Italy. The epidemiological data associated with these cases of clustered outbreaks included comprehensive information such as outbreak identification records, the timing of the last IDT (intradermal tuberculin) test execution, details regarding shared pastures, animal movements during the interval between the last IDT test and the detection of positive cases, and genotype data. We included the genomes belonging to the Matelica Clade in the analyses with the aim to compare the genomic variability observed between samples originated from a unique introduction in a geographically restricted area (i.e., the Matelica Clade) with respect to the genomic variability observed within clades belonging to the Italian Genotype 1, which were associated, based on the epidemiological and genotypical information, to *M. bovis* introduction in areas with similar territorial extension (such as Caronia).

Moreover, from a total of 8 herds, 2 to 7 strains from each of these herds were selected and analyzed to assess genetic diversity within herds.

In addition, in the phylogenetic analysis, five *M. bovis* genomes characterized by genotypes different from the Italian Genotype 1 and Matelica Clade were included as outgroups: two strains isolated in Sicily characterized by spoligotype SB0841; one strain isolated in Campania with 2 repeat units in ETRD instead of 3; one strain isolated in Apulia with 3 different markers (Qub11a: 9, Qub1895: 2, Qub15: 2); one strain isolated in Apulia with 2 different markers: Qub11a: 9, Qub3232: 4 (See Supplementary Table S1).

2.3 DNA extraction

Mycobacterium bovis strains frozen in tryptic soy broth with 10% glycerol were thawed and cultured on 5 Stonebrink medium (prepared as solid slants in screw-cap tubes) for 3–4 weeks at 37°C. For each isolate, all the colonies were transferred into 500 µL of Middlebrook 7H9 broth and inactivated for 20 min at 95°C.

DNA was extracted by the Cetyltrimethylammonium bromide (CTAB) protocol as described by Ausubel et al. (2003) with minor modifications. In short, lipase was added to a final concentration of 11,3 µg/µl, and the tube was incubated overnight at 37°C. Fifty µl of proteinase K (20 mg/mL) and 1,5 mL of CTAB extraction solution were added, and the mixture was incubated for 1 h at 65°C. After centrifugation for 10 min at 12.500 g, the pellet was discarded, and an equal volume of chloroform was added to the recovered aqueous phase. The mixture was centrifuged as described above. The aqueous phase was recovered, mixed with 2X volume of CTAB precipitation solution [CTAB 0,5% (w/v), NaCl 40 mM], and incubated for 1 h at room temperature. After centrifugation, the aqueous phase was discarded, and the pellet was resuspended in 350 µL of NaCl 1,2 M. The mixture was extracted with chloroform as described above. The aqueous phase was recovered, and the solution was precipitated by adding 0,6 volume of isopropanol at room temperature. After washing the pellet with 500 µL of Ethanol 70% vol/vol, the pellet was resuspended in 30 µL of TE (Tris–HCl 10 mM, EDTA 1 mM pH8).

2.4 Whole genome sequencing and SNP analysis

Genomic libraries were prepared starting from extracted genomic DNA using the Nextera XT DNA Library Preparation Kit (Illumina), quantified by a Quantus Fluorometer (Promega, Fitchburg, United States), and sequenced on either Miseq or Nextseq 550 systems (Illumina) generating 2 × 250 and 2 × 150 bp paired-end reads, respectively. Raw reads were checked for quality using FastQC (Babraham Bioinformatics, n.d.) and for possible contaminants through Kraken2 software (Wood and Salzberg, 2014). Reads were trimmed with Trimmomatic ver. 0.38 (minimum mapping quality for an alignment to be used of 30 and minimum base quality for a base to be considered of 20) (Bolger et al., 2014) and used as input for SNP analysis.

To contextualize Italian isolates within the international framework, we performed a SNP analysis comprising *M. bovis* isolates sequenced in this study, along with a selection of representative La genomes from international databases. This selection includes a broad representation of La 1.2, to which the Italian isolates from the study of Zwyer et al. (2021) appear to belong (Supplementary Table S2). The SNP analysis was performed starting from trimmed reads with the CFSAN SNP pipeline ver. 0.8 (Davis et al., 2015) using the *M. bovis* AF2122/97_BX24833 genome (Accession N. AF2122/97_BX248333) *M. bovis* lineages/sub-lineages of Italian isolates were checked running KvarQ (Steiner et al., 2014) with the new KvarQ testsuite (MTBC_animals) indicated in Zwyer et al. (2021). A second analysis based on a sub-dataset (only *M. bovis* Italian isolates sequenced in this study plus outgroups) was also performed using the CFSAN SNP pipeline ver. 0.8 (Davis et al., 2015) with default settings except for the following parameters: minimum mapping quality for an alignment = 30 minimum base quality for a base = 20. More deeply, trimmed reads were mapped to the *M. bovis* Mb3601 genome, representative of a new clonal complex EU3 isolated in France and characterized by SB0120 spoligotype (Assembly Accession Number: LR699570.1) (Branger et al., 2020) using Bowtie2 (Langmead and Salzberg, 2012), and SNPs were identified using VarScan (Koboldt et al., 2012). Abnormal SNPs (i.e., SNPs from the 500 nt-edge regions of contigs and recombinant and repetitive regions with more than 3 SNPs per 1,000 bases) were removed. The high-quality core SNP matrixes obtained from both SNP analyses were used to infer phylogeny through a Maximum Likelihood (ML) algorithm using

TABLE 1 *Mycobacterium bovis/Mycobacterium caprae* population in Italy in 2008–2020 (*N* = 3,353 outbreaks).

Spoligotypes	%	VNTR-MIRU	%	Spoligo_VNTR-MIRU	%
SB0120	44,4	4553310444365	7	SB0120_4553310444365*	5,6
SB0134	12,1	5453410454365	3,9	SB0134_5453410454365	3,5
SB0841	10,7	5453310444365	3,9	SB0841_5453310444365	3
SB0418	6,8	3353310444365	3,2	SB0120_3353310444365**	2,6
SB0872	2,2	5353310444365	2,7	SB0120_5553410454365	1,5
SB0121	2,2	5553310444365	2,3	SB0841_555335444365	1,5
Other Spoligos	23	Other VNTR-MIRUs	77	Other Spoligo_VNTR-MIRUs	82,3

VNTR-MIRU order of markers: ETRA, ETRB, ETRC, ETRD, ETRE, Qub11a, Qub11b, Qub26, Qub1895, Qub15, Qub3232, MIRU26; *Italian genotype; **Matelica Clade.

RaxML software with 100 bootstrap replicates (Stamatakis, 2014). Phylogenetic trees were edited using iTOL ver.6.[2]

3 Results

3.1 Spoligotypes and VNTR-MIRU genotypes in Italy

Since 2008, all the isolates collected in Italy have been genotyped by spoligotyping and VNTR-MIRU. Spoligotyping differentiated the 4,845 MTBC strains from 3,353 herd outbreaks collected between 2008 and 2020 into 136 spoligotypes (Table 1). Fourteen spoligotypes belonged to *M. caprae,* while 122 to *M. bovis*. The most frequent spoligotypes were SB0120 (44,4%), SB0134 (12,1%), SB0841 (10,7%) for *M. bovis,* and SB0418 (6,8%) for *M. caprae*. These spoligotypes represent 77% of bTB strains circulating in Italy. Seventy-one spoligotypes were found in only one outbreak.

The combination of spoligotyping/VNTR-MIRU identified a total of 1,035 genotypes, including 318 clusters of 2–190 isolates and 717 unique profiles. Of these, 421 belonged to SB0120, 103 to SB0134, and 61 to SB0841. Despite the high degree of resolution (H index >0.98) (Boniotti et al., 2009), twelve genotypes belonging to SB0120 were found in more than 56 herd outbreaks, mainly located in Sicily, Campania, and Apulia since 2008 (Supplementary Tables S1, S3).

Among them, the most prevalent was the genetic profile VNTR-MIRU 4,5,5,3,3,10,4,4,4,3,6,5, as described in M&M section (Italian genotype 1), which was identified between 2008–2020 in 190 cattle/buffalo herds (5,6% of outbreaks). This genotype was widespread in different regions (Veneto, Campania, Marche, Lazio, Apulia, Calabria, Sicily), and was the cause of several herd outbreaks localized in a specific geographical area of Sicily (Caronia, province of Messina). The second most frequent profile belonging to SB0120 was VNTR-MIRU 3,3,5,3,3,10,4,4,4,3,6,5 (Matelica Clade), which differs from the Italian genotype 1 in the ETRA and ETRB loci (3 repeats instead of 4 and 5, respectively). The Matelica Clade has been isolated in 111 cases of bTB in cattle herds and wild boars (77), mainly in the hunting district of Monte San Vicino, within the municipality of Matelica (Macerata, Marche region, Central Italy). Since the first isolate of *M. bovis* was detected in a hunted wild boar in 2002, the same genetic strain has been circulating in this area. Several bTB outbreaks were detected in cattle herds in the following years, as well as in wild boars, raising suspicion of a wildlife-livestock interface. Out of the total cases, we were able to sequence only 17 samples with high-quality DNA; retrieving DNA from samples collected during the 2010–2017 period, especially before 2010, proved to be difficult.

3.2 Italian SB0120 *Mycobacterium bovis* population in the international context

Since little is known about the global lineages of Italian *M. bovis* genotypes, a comparative SNP analysis was performed, including 81 sequences from Italy, 76 SB0120 isolates of the Italian Genotype 1 and Matelica Clade, and 5 outgroups (see Materials and Methods section). Additionally, a selection of 100 sequences covering the European genetic diversity of *M. bovis* (see Supplementary Table S2) (Hauer et al., 2019, Kohl et al., 2020, Loiseau et al., 2020, Zwyer et al., 2021) was added to the SNP analysis. In particular, we included the well-known *M. bovis* clonal complexes EU1 (La 1.8.1) and EU2 (La 1.7.1), the new *M. bovis* sub-lineages Unknown 7 (La 1.8.2) and Unknown 3 (La 1.4) described by Zwyer et al. (2021) and Loiseau et al. (2020), which showed a broad distribution in Europe and Asia, South America and Asia, respectively. Furthermore, based on the studies of Zwyer et al. (2021), and Loiseau et al. (2020) we selected several genomes of different geographical origins characterized by spoligotype SB0120 and belonging to the La 1.2 clade, which corresponds to the Unknown 2 described by Loiseau et al. (2020) and to the EU3 described by Branger et al. (2020). *M. caprae* isolates (La2) were included in the analysis as additional outgroups. Among the international genomes selected for this study, we included a unique Italian isolate from human (SRR7131117, Loiseau et al., 2020) that did not cluster with newly sequenced Italian isolates, but with French and Swiss isolates (minimum pairwise SNPs with SRR7131117 = 104 SNPs) (Supplementary Table S5).

All Italian sequenced isolates belonged to La1.2 lineage and exhibited SNPs specific to La1.2, as defined as by Zwyer et al. (2021) except for 232080–2010 sample as resulted in KvarQ analysis (Supplementary Table S4). In this study, the maximum pairwise genetic distance among Italian isolates of the La1.2 lineage was 319 pairwise SNPs (Supplementary Table S5), while the highest pairwise SNPs difference with international isolates of different La lineages was 1,080 pairwise SNPs.

2 https://itol.embl.de/

In the context of La1.2 lineage, Italian isolates of La1.2 clustered separately from other European strains included in the study, except for one isolate from Germany (ERR2815526, Kohl et al., 2020) that shared a minimum of 82 pairwise SNPs with the nearest Italian genomes. In particular, apart from the isolates of the Matelica Clade, all the Italian isolates of Genotype 1 clustered in a single polyphyletic clade that also comprises outgroup isolates with different spoligotype and VNTR-MIRU profiles (Figure 1). Two isolates (232080–1-2010 and 217495–8-2008) lay outside this polyphyletic clade: 232080–1-2010 was not classified as a La 1.2 strains, while 217495/8 diverged from the other isolates in this study (pairwise SNPs range = 1–140 with an average number of SNPs of 72), despite their belonging to the same Spoligotype, VNTR-MIRU profile and La1.2 lineage (Supplementary Table S5); this could be possibly explained with convergent evolution events of VNTR markers, already observed in *Mycobacterium tuberculosis* (Comas et al., 2009). This last strain will be excluded from SNP metrics associated with the Italian La1.2 isolates, ensuring that genetic variability is assessed solely based on point mutations.

3.3 WGS analysis of *Mycobacterium bovis* Italian genotype 1 and Matelica Clade

A second phylogenetic analysis was performed on *M. bovis* Italian isolates sequenced in this study and outgroups described in the Materials and Methods section (Figure 2). This analysis aims to evaluate the genetic diversity within and between Italian regions (corresponding to the Nomenclature of Territorial Unit for Statistics 2 level), as well as within individual herds and between epidemiologically correlated herd outbreaks.

As expected, isolates belonging to Italian Genotype 1 and Matelica Clade, which differed in ETRA and B loci, clustered separately with an average of 254 pairwise SNPs (min 239, max 334, median 259) (Supplementary Table S6). Within the Italian genotype 1, where the pairwise SNP differences ranged from 1 to 242 SNPs, sub-clusters were distinguishable, and they were generally associated with different Italian regions (shown with different colors in Figure 2). Clusterization by regions was more evident for Apulia (Blue) with almost all the isolates grouped in clade A (pairwise SNPs range = 2–50 SNPs) and for Sicily (dark green) where the majority of isolates grouped in clade B and C (max pairwise SNPs = 77 SNPs), which derived from the same evolutionary branch. The pairwise SNP differences between clade A (Apulia) and clade B + C (Sicily) ranged from 102 to 129 SNPs.

To evaluate the genetic diversity within epidemiologically correlated herd outbreaks, we considered 3 cases where the observed epidemiological links (i.e., animal movements between herds and geographical proximity), were supported by the genetic relationships among isolates: (1) Isolates of RM1 and RI1 herds from Lazio region (light green) showed few pairwise SNPs (10 SNPs), consistent with the transmission of bTB by the trading of an infected animal from RM1 to RI1 and with the evolution of the outbreak strain in five years; (2) FG2 isolate corresponded to an animal originating from the FG1 herd (SNPs = 7); and (3) TV1 isolate corresponded to an animal originating from the RG1 herd (SNPs = 3).

Within-herds genetic diversity was evaluated in 8 herds (C12, MC1, CE1, FG1, C6, C1, C3, C10) with 2–7 strains from each herd. Variation detected in isolates of the same herd (CE1, FG1) was, respectively, of 4 and 5 SNPs. Two isolates of C10 herd showed a strict genomic correlation (pairwise SNPs range = 0–1 SNPs) with a black swine isolate of the same herd, suggesting contamination among different animals of the same herd. Also, in C3 herd the spread of a single clone over years (2015–2017) was observed (pairwise SNPs range = 0–5 SNPs).

Additionally, we evaluated the genetic diversity observed between bTB isolates originating from herds linked by geographical proximity. Specifically, we examined two scenarios:

1. Clade B analysis: this clade includes isolates from epidemiologically linked herds (C1, C2, C6, C7, C8, C9, C12, C13) located in Caronia municipality. In this clade, the pairwise SNPs ranged from 8 to 15 SNPs (see SNP heatmap in Figure 3). Notably, black swine isolates clustered alongside bovine isolates with a pairwise SNPs range of 3 to 12 SNPs (Figure 3). These herds, located within the Nebrodi Park, were epidemiologically correlated due to the sharing of grazing areas (Supplementary Figure S2);
2. Matelica clade analysis: the Matelica Clade consisted of 17 isolates collected between 2010 and 2017 (see Supplementary Table S1). This genotype included 9 isolates from three bovine herds (MC1, MC2, MC3) and 8 wild boars. All isolates grouped within a single cluster (Figure 2) with low SNPs diversity (pairwise SNPs range = 0–9 SNPs), thus supporting a geographical epidemiological correlation. The maximum SNP distance among isolates of the MC1 herd was three pairwise SNPs, while the distance ranged from one to seven SNPs among the three herds (MC1, MC2, and MC3), showing high genomic similarity in this geographic area. Moreover, wild boar isolates clustered together with bovine isolates (pairwise SNPs range among wild boar isolates = 0–7 SNPs; pairwise SNPs range among wild boar and bovine isolates = 0–9 SNPs). In particular, some wild boar isolates (namely 169505, 207778-20, 23230 and 231240) had no (0) SNPs diversity concerning MC1 isolates (105517–20, 105486–20,105493–20,105505,105513). Isolates of the herd MC3 were more distant from wild boar isolates (pairwise SNPs range = 5–9 SNPs) and MC1 and MC2 bovine isolates (pairwise SNPs range = 4–7 SNPs) than the other herds. In general, despite the collection of these strains over time (seven years), the majority of wild boar isolates showed only a few SNPs of difference (median 2.5 SNPs).

Despite the longer time elapsed among the sampling date of the Matelica Clade strains with respect to the Clade B ones (seven vs. two years), the Matelica Clade showed a significantly smaller pairwise genetic variability than Clade B (generalized linear model, GLM, with Poisson distributed error, $p < 0.0001$). In addition, the strains in the Matelica Clade showed a significant increase in the pairwise genetic variability when the difference elapsed between the sampling dates of strains increased (GLM with Poisson distributed error, $p < 0.0001$), see Figure 4.

4 Discussion

In Italy, despite recent progress in the eradication campaign, bovine tuberculosis (bTB) continues to have a significant economic impact due to high control and eradication costs, compensation for

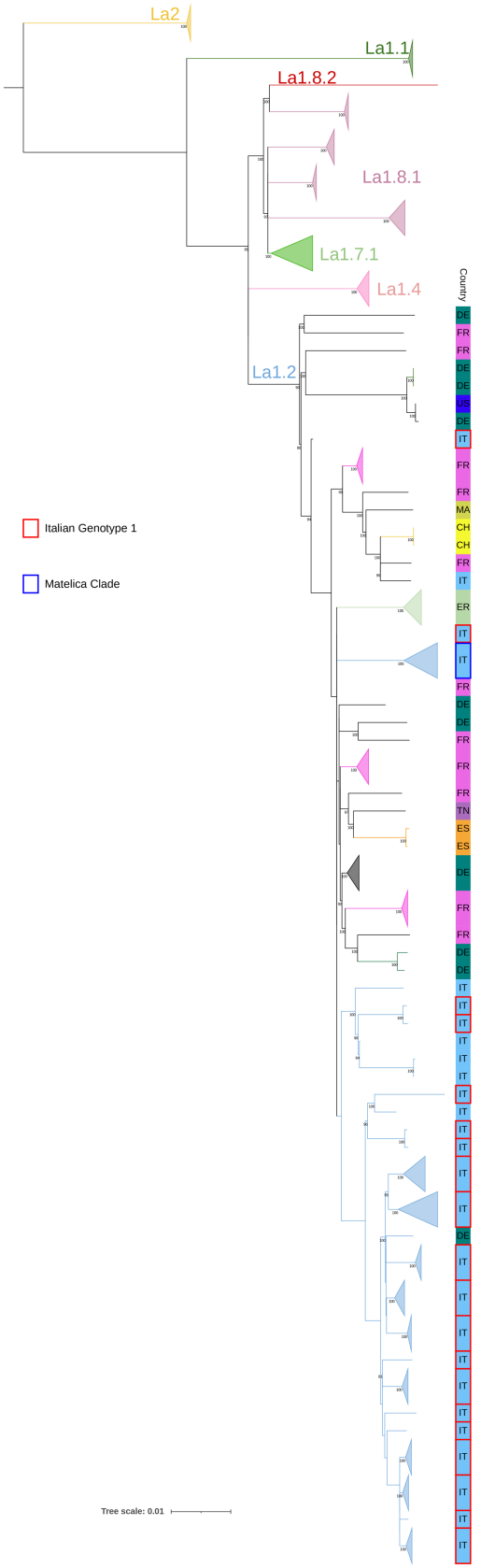

FIGURE 1
SNP based phylogenetic analysis of the *M. bovis* population. Lineages, previously described by Zwyer et al. (2021), are reported. Isolates belonging to the Matelica Clade and the Italian Genotype 1 are indicated in blue and red, respectively (see legend). Isolates of La1.2 lineage of same Country and type of Clade with 100% bootstrap values (>3 isolates) are collapsed. Other European strains are made explicit in the figure and their international acronym is reported. Bootstrap values higher than 80% were shown near their relative nodes.

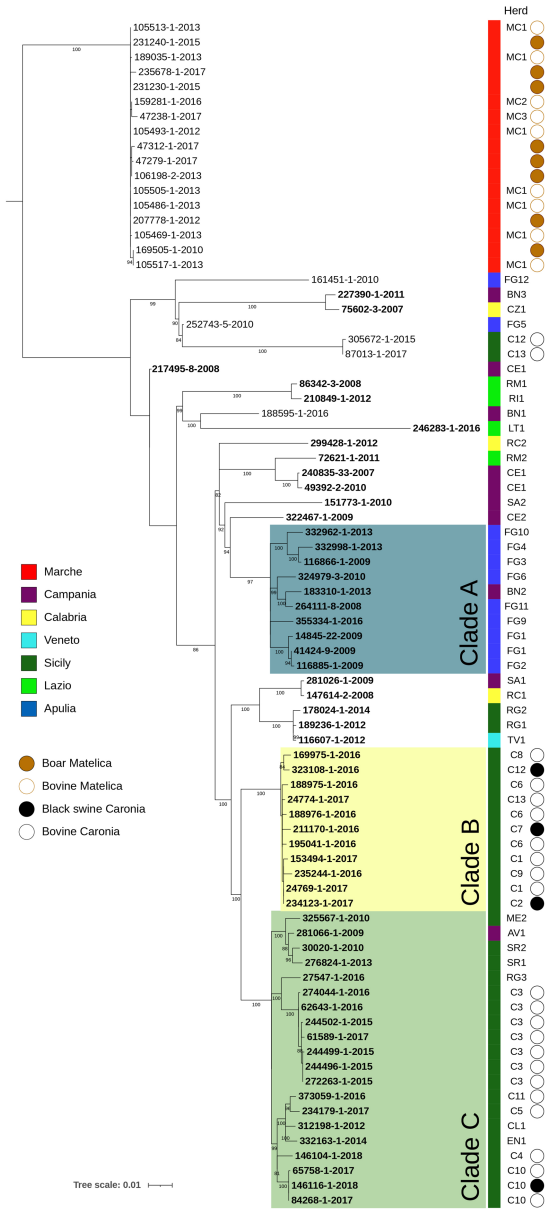

FIGURE 2
SNP based phylogenetic analysis of the Italian *M. bovis* strains considered in this study. Colored squares indicate the belonging to different Italian regions, as reported in the legends. The herd column highlights the herd to which each isolate comes from. Isolates from black swine collected from Caronia are marked with a bold black circle, while bovine isolates from the same area are reported with an empty circle. The Matelica column contains a bold brown circle figuring out boar isolates, while an empty circle indicates bovine isolates. Bootstrap values higher than 80% were shown near their relative nodes.

slaughtered TB-positive animals, and trade restrictions on live animals and animal products. The new Animal Health Law (Regulation (EU) 2016/429, 2016) has introduced changes in disease management, including updated diagnostic protocols, risk assessment, and epidemiological investigations (European Food Safety Authority and European Centre for Disease Prevention and Control, 2022).

Molecular genotyping is essential for tracking infectious agents in epidemiological investigations. In bovine tuberculosis (bTB), techniques like whole-genome sequencing (WGS) offer insights into the pathogen's evolution, transmission, and distribution. WGS surpasses traditional methods (Spoligotyping and VNTR-MIRU) by enabling rapid sequencing of entire bacterial genomes, thus improving epidemiological accuracy, especially in cases with frequent Spoligo-VNTR-MIRU profiles. Few Italian *M. bovis* genomes have been published, and there's limited knowledge on the association of specific clonal groups with VNTR-MIRU genotypes in Italy (Hauer et al., 2019; Kohl et al., 2020; Loiseau et al., 2020; Zimpel et al., 2020). This study is the first large-scale genomic analysis of *M. bovis* in Italy, aiming to use SNP analysis from WGS to determine the phylogenetic relationships of the two most common SB0120-VNTR-MIRU profiles in Italy (Genotype 1 and Matelica Clade) and compare them with European sub-lineages. The study analyzed 81 *M. bovis* isolates, primarily from regions where bTB persists: Sicily, Campania, and Apulia.

Despite the presence of SB0120 in other lineages (Branger et al., 2020), all sequenced Italian isolates (Matelica Clade and Italian genotype 1) belong to La1.2 lineage (maximum pairwise genetic distance of 319 pairwise SNPs among Italian isolates of La1.2 lineage; Figure 1). These isolates clustered separately from European isolates, suggesting an Italian genomic segregation. Our SNP analysis confirms the high genetic variability within SB0120. For instance, the French reference strain Mb3601 (ERR3825346 in Figure 1; Branger et al., 2020) differs from 116 to 291 SNPs from isolates of the Italian Genotype 1. This variability aligns with findings observed by Hauer et al. (2019).

Another purpose of our study was to evaluate the genetic diversity within and between Italian regions, within epidemiological correlated outbreaks, and within herds. To assess genetic diversity within and between Italian regions, we considered 36 isolates from Campania, Sicily, Apulia, Calabria, Veneto, Lazio collected between 2008 and 2018. As expected, the phylogenetic analysis showed that isolates belonging to Italian Genotype 1 and Matelica Clade, which differ in ETR-A and -B loci, clustered separately (Figure 2), with an average genetic distance of 254 pairwise SNPs (min 239, max 334, median 259). Within the Italian genotype 1, which is the prevalent genotype in Italy, there is a genetic diversity ranging from 1 to 242 SNPs, indicating prolonged circulation across multiple regions. Sub-clusters were generally region-specific (i.e., Apulia in blue and Sicily in dark green in Figure 2), with high genetic variability within regions (pairwise SNPs range = 2–70 SNPs).

A recent study in Spain (Pozo et al., 2022), on *M. bovis* isolates from different host species, including cattle, wild boar, and red deer, revealed the genetic distance was primarily geographic rather than host-specific. Isolates from both cattle and wildlife from a given region were more closely related than isolates from the same species but geographically distant. Moreover, the genetic diversity among the sequenced isolates, all belonging to the SB0339 spoligotype, ranged from 49 to 88 SNPs. These values are lower than the ones observed for the Italian genotype 1 (pairwise SNPs range = 1–242), all belonging to the same VNTR-MIRU profile, suggesting longer circulation of Italian genotype 1 in Italy than Spanish isolates.

To asses genetic diversity within outbreaks in geographically close areas, two main groups of isolates were considered: (i) 24 isolates from cattle and black pigs in a delimited *M. bovis* hot-spot area in Sicily (Caronia, Messina province) with known epidemiological information over three years (2015–2018) and (ii) 17 isolates, belonging to Matelica Clade, from cattle and wild boar collected from 2010 to 2017.

The isolates located in the Caronia municipality clustered into clades Clades B and C (Figure 3), with Clade C showing three sub-clusters (c.a, c.b and c.c) originating from a common ancestor, and including isolates from other areas of Sicily, thus suggesting a large-scale circulation of these lineages in Sicily. While the co-circulation in the Caronia municipality of different Spoligotype-VNTR-MIRU profiles has been already documented, suggesting multiple introductions of bTB in the area (Marianelli et al., 2019, 2023), here we showed that strains belonging to the same Spoligotype-VNTR-MIRU profile originated from distinct introductions.

On the other hand, Clade B was associated with peculiar geo-epidemiological and managerial characteristics of the infected herds. Eight cattle herds fall within Clade B (C1, C2, C6, C7, C8, C9, C12, C13), all with established epidemiological connections due to direct and indirect mutual contacts. These farms are located within the Nebrodi Park, a large nature reserve of about 86,000 hectares, where bTB is historically endemic. The park's diverse wildlife, including foxes, wild boars, martens, and fallow deer, shares feeding and watering points with livestock, facilitating disease transmission. Notably, the feral and free-roaming pigs dominate in terms of population density among these species.

Herds in Clade B are located within a 20 km^2 area in 2 geographic areas: the first (C8, C12, C1, C6, C13) with an average inter-herd distance of about 1.5 km, while the second (C2, C6, C7, C9), 5 km away, with an average inter-herd distance of 2.1 km. Although these farms have distinct registration codes, they constitute a single epidemiological unit due to consistent direct and indirect interactions over time and space. Operating as mixed farms, hosting cattle, pigs, sheep, goats, donkeys, and horses, these farms practiced extensive grazing, with minimal shelter and fencing, leading to constant animal-wildlife contact, especially at shared water and foraging sites (Fine et al., 2011). Marianelli et al. (2023) highlighted that TB spread is only limited by natural geographic barriers without appropriate biosecurity measures. Moreover, studies in the Nebrodi Park suggested that the Black Nebrodi pig, typically raised in a free-roaming system, could act as a bTB reservoir (Di Marco et al., 2012).

The Matelica Clade isolates, including 9 isolates from three bovine herds (MC1, MC2, MC3) and 8 wild boars from 2010 to 2017, formed a single clonal cluster (pairwise SNPs range = 0–9 SNPs) (Figure 2), supporting the geographical epidemiological correlation. This indicates an eight-year contamination period in this area, with MC1 and 4 wild boars isolates showing no genetic differences. Epidemiological investigations revealed numerous epidemiological connections between MC1 and MC2, including shared pastures for several years, while MC3 introduced bTB via a bovine bought from a herd located in the same province, but geographically more distant from MC1 and MC2. However, all herds had possible contact with infected wild boars either on the pasture (MC1 and MC2) or at the cattle shed.

In this area, the same genetic strain has been circulating since 2002, when the first isolation of *M. bovis* from a hunted wild boar occurred (unpublished data). Since then, many bTB outbreaks have been detected in cattle herds accompanied by a notable rise in *M. bovis* isolations within wild boar populations, suggesting a wildlife-livestock interface (Figure 2). In this region, native beef cattle practice seasonal migration to shared pastures and are maintained in semi-free stabling, fostering interactions with wild boar. Moreover, the lack of biosecurity measures to protect herds from contact with wild animals has escalated the spread of bTB outbreaks among cattle herds and spread the disease within the wild boar population across the Marche region.

Notably, both Matelica and Caronia cases showed clustering of isolates from different species with bovine isolates (Figures 2, 3), suggesting repeated transmission between cattle and wild boar or black pig. As reported by other studies (Johnston et al., 2011;

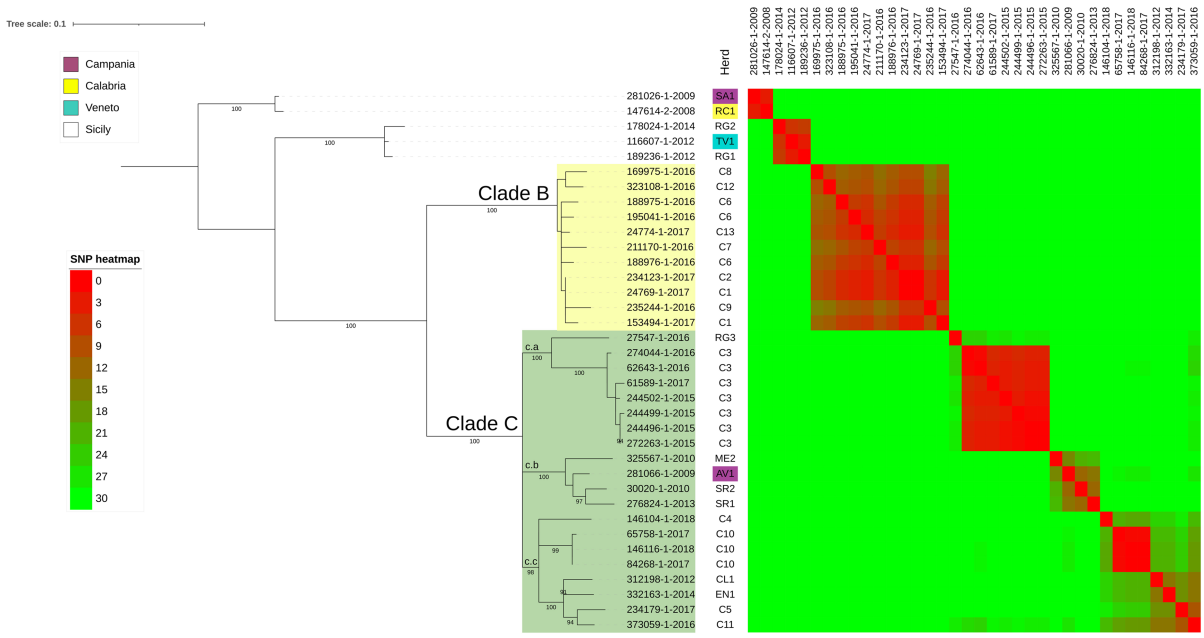

FIGURE 3
SNP based phylogenetic analysis of the node including all the Sicilian isolates. Herds of each isolate are shown in the first column and herds of region different from Sicily are colored following the legend. The heatmap represents the number of SNPs between each couple of isolates; heatmap colors are described in the legend. Bootstrap values higher than 89% were shown near their relative nodes.

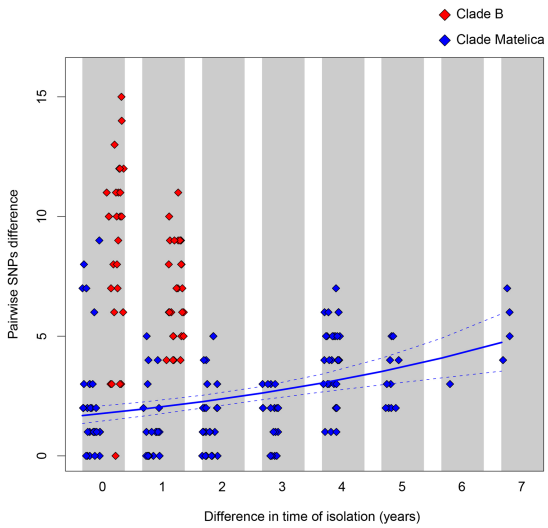

FIGURE 4
Pairwise SNPs differences between strains belonging to Clade B (red diamonds) and to the Matelica Clade (blue diamond) as a function of the time difference elapsed between the sampling dates. The solid blue line represents the pairwise SNPs difference as a function of the sampling date difference within the Matelica Clade estimated through a generalized linear model with Poisson distributed error. The dashed lines represent the confidence interval of the estimate.

Martínez-López et al., 2014) the transmission, diffusion, and persistence of bovine tuberculosis within and between species are closely related to risk factors such as shared pasture, presence of feeding sites, and the absence of physical barriers. The transmission of bTB in wildlife raises public health concerns, including risks for hunters handling infected game meat and consumers of potentially contaminated wild boar products. Our NGS findings highlight the need for veterinary and public health authorities to develop specific and dedicated bTB control plans for geographical areas where bTB is present in wild animals or in multi-host context.

By comparing the genetic findings in Clade B and Matelica Clade, we observed a larger variability in pairwise SNP differences between isolates belonging to Clade B with respect to the variability observed between isolates belonging to the Matelica Clade (see Figure 4). While the genetic variability observed within the Matelica Clade is similar to those observed in mycobacteria outbreaks investigated through WGS (see Biek et al., 2012; Trewby et al., 2016; Bolzoni et al., 2021), the genetic variability observed within Clade B in Caronia municipality is higher than expected. The higher genetic divergence between isolates in Clade B can be explained by hypothesizing an older common ancestor of Clade B, suggesting a really old introduction of this bTB clade in the Caronia area.

Finally, the genetic diversity within 7 herds (MC1, CE1, FG1, C6, C1, C3, C10) showed persistence of clones within herds, with SNP pairwise differences ranging from 0 to 7 SNPs. There was a strict genomic correlation among isolates of different species within these herds, such as black swine and bovine isolates from C10 (0–1 SNPs), and among clones over several years, such as C3 with isolates from 2015 to 2017 (0–5 SNPs). Our data differ from those published by Pozo et al. (2022) in Spain, where genetic distance within herds ranged between 0 and 44 SNPs. However, in our study, we considered only isolates with the same Spoligo-VNTR-MIRU profile, resulting in a reduced genetic distance.

Next, Generation Sequencing is rapidly becoming the gold standard for molecular genotyping of human and bovine tuberculosis. This powerful technology enables the sequencing of entire bacterial genomes, providing unprecedented resolution in the identification and tracking of infectious agents to their source and transmission routes. However, to fully leverage the potential of NGS in tuberculosis genotyping for a One Health approach, it is essential to ensure comparability of data across different studies and regions. This requires the implementation of standardized protocols and the establishment of a global database, allowing for cross-species comparisons and contributing to a more integrated and comprehensive approach to disease surveillance, prevention, and control. The use of NGS and the establishment of standardized protocols and databases will be key in the fight against tuberculosis, enabling a more effective One-Health approach to tackle this disease that affects both animals and humans.

Data availability statement

Raw reads of isolates described in this article were submitted to EBI-ENA (https://www.ebi.ac.uk/ena/browser/home) under study accession n. PRJEB46575 (https://www.ebi.ac.uk/ena/browser/text-search?query=PRJEB46575).

Ethics statement

Ethical approval was not required for the studies on animals in accordance with the local legislation and institutional requirements because only commercially available established cell lines were used.

Author contributions

ES: Conceptualization, Data curation, Methodology, Writing – original draft. KI: Investigation, Methodology, Writing – original draft. MB: Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. IM: Data curation, Investigation, Methodology, Writing – original draft. LB: Data curation, Writing – review & editing. DI: Investigation, Methodology, Writing – original draft. FC: Investigation, Methodology, Data curation, Writing – review & editing. DL: Investigation, Methodology, Writing – original draft. MD'I: Investigation, Methodology, Writing – original draft. MZ: Investigation, Writing – review & editing, Data curation, Methodology. VP: Investigation, Writing – review & editing, Conceptualization, Supervision. PM: Investigation, Writing – review & editing, Methodology. SG: Writing – review & editing, Data curation, Investigation. MP: Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Supervision, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the Italian Ministry of Health (IZSLER PRC2014013-E88F15000140001, IZSLER PRC 08/2021-E55F21002760006). The work of LB was supported by EU funding within the NextGenerationEU

– MUR PNRR Extended Partnership initiative on Emerging Infectious Diseases (project no. PE00000007, PE13 INFACT, Spoke 4).

Acknowledgments

We gratefully acknowledge Marcello Fin, Ivana Barraco, and Anna Mangeli for their technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1416605/full#supplementary-material

References

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., et al. (eds.) (2003). *Current protocols in molecular biology.* John Wiley & Sons, Inc.

Babraham Bioinformatics—FastQC a quality control tool for high throughput sequence data (internet). [cited Jul 20, 2023]. Available at: https://www.bioinformatics.babraham.ac.uk/projects/fastqc/

Berg, S., Garcia-Pelayo, M. C., Müller, B., Hailu, E., Asiimwe, B., Kremer, K., et al. (2011). African 2, a clonal complex of *Mycobacterium bovis* epidemiologically important in East Africa. *J. Bacteriol.* 193, 670–678. doi: 10.1128/JB.00750-10

Biek, R., O'Hare, A., Wright, D., Mallon, T., McCormick, C., Orton, R. J., et al. (2012). Whole genome sequencing reveals local transmission patterns of *Mycobacterium bovis* in sympatric cattle and badger populations. *PLoS Pathog.* 8:e1003008. doi: 10.1371/journal.ppat.1003008

Blouin, Y., Hauck, Y., Soler, C., Fabre, M., Vong, R., Dehan, C., et al. (2012). Significance of the identification in the horn of Africa of an exceptionally deep branching *Mycobacterium tuberculosis* clade. *PLoS One* 7:e52841. doi: 10.1371/journal.pone.0052841

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

Bolzoni, L., Scaltriti, E., Russo, S., Pongolini, S., Garbarino, C., Leo, S., et al. (2021). Transmission patterns of a *Mycobacterium avium* subsp. paratuberculosis clone within a single heard investigated by whole genome sequencing. *Vet. Microbiol.* 263:109272. doi: 10.1016/j.vetmic.2021.109272

Boniotti, M. B., Gaffuri, A., Gelmetti, D., Tagliabue, S., Chiari, M., Mangeli, A., et al. (2014). Detection and molecular characterization of *Mycobacterium microti* isolates in wild boar from northern Italy. *J. Clin. Microbiol.* 52, 2834–2843. doi: 10.1128/JCM.00440-14

Boniotti, M. B., Goria, M., Loda, D., Garrone, A., Benedetto, A., Mondo, A., et al. (2009). Molecular typing of *Mycobacterium bovis* strains isolated in Italy from 2000 to 2006 and evaluation of variable-number tandem repeats for geographically optimized genotyping. *J. Clin. Microbiol.* 47, 636–644. doi: 10.1128/JCM.01192-08

Branger, M., Loux, V., Cochard, T., Boschiroli, M. L., Biet, F., and Michelet, L. (2020). The complete genome sequence of *Mycobacterium bovis* Mb3601, a SB0120 spoligotype strain representative of a new clonal group. *Infect. Genet. Evol.* 82:104309. doi: 10.1016/j.meegid.2020.104309

Brosch, R., Gordon, S. V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., et al. (2002). A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. *Proc. Natl. Acad. Sci. USA* 99, 3684–3689. doi: 10.1073/pnas.052548299

Comas, I., Homolka, S., Niemann, S., and Gagneux, S. (2009). Genotyping of genetically monomorphic bacteria: DNA sequencing in *Mycobacterium tuberculosis* highlights the limitations of current methodologies. *PLoS One* 4:e7815. doi: 10.1371/journal.pone.0007815

Corner, L. A., Trajstman, A. C., and Lund, K. (1995). Determination of the optimum concentration of decontaminants for the primary isolation of *Mycobacterium bovis*. *N. Z. Vet. J.* 43, 129–133. doi: 10.1080/00480169.1995.35871

Couvin, D., Cervera-Marzal, I., David, A., Reynaud, Y., and Rastogi, N. (2022). SITVITBovis—a publicly available database and mapping tool to get an improved overview of animal and human cases caused by *Mycobacterium bovis*. *Database (Oxford)* 2022:baab081. doi: 10.1093/database/baab081

Davis, S., Pettengill, J. B., Luo, Y., Payne, J., Shpuntoff, A., Rand, H., et al. (2015). CFSAN SNP pipeline: an automated method for constructing SNP matrices from next-generation sequence data. *PeerJ Comput Sci* 1:e20. doi: 10.7717/peerj-cs.20

Di Marco, V., Mazzone, P., Capucchio, M. T., Boniotti, M. B., Aronica, V., Russo, M., et al. (2012). Epidemiological significance of the domestic black pig (*Sus scrofa*) in maintenance of bovine tuberculosis in Sicily. *J. Clin. Microbiol.* 50, 1209–1218. doi: 10.1128/JCM.06544-11

Duarte, E. L., Domingos, M., Amado, A., Cunha, M. V., and Botelho, A. (2010). MIRU-VNTR typing adds discriminatory value to groups of Mycobacterium bovis and *Mycobacterium caprae* strains defined by spoligotyping. *Vet. Microbiol.* 143, 299–306. doi: 10.1016/j.vetmic.2009.11.027

European Food Safety Authority and European Centre for Disease Prevention and Control (2022). The European Union one health 2021 Zoonoses report. *EFSA J.* 20:e07666. doi: 10.2903/j.efsa.2022.7666

Fine, A. E., O'Brien, D. J., Winterstein, S. R., and Kaneene, J. B. (2011). An effort to isolate *Mycobacterium bovis* from environmental substrates during investigations of bovine tuberculosis transmission sites (cattle farms and wildlife areas) in Michigan, USA. *ISRN Vet Sci* 2011:787181. doi: 10.5402/2011/787181

Frothingham, R., and Meeker-O'Connell, W. A. (1998). Genetic diversity in the *Mycobacterium tuberculosis* complex based on variable numbers of tandem DNA repeats. *Microbiology (Reading)* 144, 1189–1196. doi: 10.1099/00221287-144-5-1189

Ghavidel, M., Mansury, D., Nourian, K., and Ghazvini, K. (2018). The most common spoligotype of *Mycobacterium bovis* isolated in the world and the recommended loci for VNTR typing; a systematic review. *Microb Pathog* 118, 310–315. doi: 10.1016/j.micpath.2018.03.036

Gormley, E., LaL, C., Costello, E., and Rodriguez-Campos, S. (2014). Bacteriological diagnosis and molecular strain typing of Mycobacterium bovis and *Mycobacterium caprae*. *Res. Vet. Sci.* 97 Suppl:30 97, S30–S43. doi: 10.1016/j.rvsc.2014.04.010

Guimaraes, A. M. S., and Zimpel, C. K. (2020). *Mycobacterium bovis*: from genotyping to genome sequencing. *Microorganisms* 8:667. doi: 10.3390/microorganisms8050667

Haddad, N., Masselot, M., and Durand, B. (2004). Molecular differentiation of *Mycobacterium bovis* isolates. Review of main techniques and applications. *Res. Vet. Sci.* 76, 1–18. doi: 10.1016/s0034-5288(03)00078-x

Hauer, A., Michelet, L., Cochard, T., Branger, M., Nunez, J., Boschiroli, M., et al. (2019). Accurate phylogenetic relationships among *Mycobacterium bovis* strains circulating in France based on whole genome sequencing and single nucleotide polymorphism analysis. *Front. Microbiol.* 10:955. doi: 10.3389/fmicb.2019.00955

Hauer, A., Michelet, L., De Cruz, K., Cochard, T., Branger, M., Karoui, C., et al. (2016). MIRU-VNTR allelic variability depends on *Mycobacterium bovis* clonal group identity. *Infect. Genet. Evol.* 45, 165–169. doi: 10.1016/j.meegid.2016.08.038

Homolka, S., Projahn, M., Feuerriegel, S., Ubben, T., Diel, R., Nübel, U., et al. (2012). High resolution discrimination of clinical *Mycobacterium tuberculosis* complex strains based on single nucleotide polymorphisms. *PLoS One* 7:e39855. doi: 10.1371/journal.pone.0039855

Huard, R. C., Fabre, M., de Haas, P., Lazzarini, L. C. O., van Soolingen, D., Cousins, D., et al. (2006). Novel genetic polymorphisms that further delineate the phylogeny of the *Mycobacterium tuberculosis* complex. *J. Bacteriol.* 188, 4271–4287. doi: 10.1128/JB.01783-05

Johnston, W. T., Vial, F., Gettinby, G., Bourne, F. J., Clifton-Hadley, R. S., Cox, D. R., et al. (2011). Herd-level risk factors of bovine tuberculosis in England and Wales after the 2001 foot-and-mouth disease epidemic. *Int. J. Infect. Dis.* 15, e833–e840. doi: 10.1016/j.ijid.2011.08.004

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

Kamerbeek, J., Schouls, L., Kolk, A., van Agterveld, M., van Soolingen, D., Kuijper, S., et al. (1997). Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J. Clin. Microbiol.* 35, 907–914. doi: 10.1128/jcm.35.4.907-914.1997

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. doi: 10.1101/gr.129684.111

Kohl, T. A., Kranzer, K., Andres, S., Wirth, T., Niemann, S., and Moser, I. (2020). Population structure of *Mycobacterium bovis* in Germany: a long-term study using whole-genome sequencing combined with conventional molecular typing methods. *J. Clin. Microbiol.* 58:1573. doi: 10.1128/JCM.01573-20

Kubica, T., Rüsch-Gerdes, S., and Niemann, S. (2003). *Mycobacterium bovis* subsp. caprae caused one-third of human *M. bovis*-associated tuberculosis cases reported in Germany between 1999 and 2001. *J. Clin. Microbiol.* 41, 3070–3077. doi: 10.1128/JCM.41.7.3070-3077.2003

Kulski, J. K., Khinsoe, C., Pryce, T., and Christiansen, K. (1995). Use of a multiplex PCR to detect and identify Mycobacterium avium and *M. intracellulare* in blood culture fluids of AIDS patients. *J. Clin. Microbiol.* 33, 668–674. doi: 10.1128/jcm.33.3.668-674.1995

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lasserre, M., Fresia, P., Greif, G., Iraola, G., Castro-Ramos, M., Juambeltz, A., et al. (2018). Whole genome sequencing of the monomorphic pathogen *Mycobacterium bovis* reveals local differentiation of cattle clinical isolates. *BMC Genomics* 19:2. doi: 10.1186/s12864-017-4249-6

Loiseau, C., Menardo, F., Aseffa, A., Hailu, E., Gumi, B., Ameni, G., et al. (2020). An African origin for *Mycobacterium bovis*. *Evol Med Public Health* 2020, 49–59. doi: 10.1093/emph/eoaa005

Luo, T., Yang, C., Peng, Y., Lu, L., Sun, G., Wu, J., et al. (2014). Whole-genome sequencing to detect recent transmission of *Mycobacterium tuberculosis* in settings with a high burden of tuberculosis. *Tuberculosis (Edinb.)* 94, 434–440. doi: 10.1016/j.tube.2014.04.005

Marianelli, C., Amato, B., Boniotti, M. B., Vitale, M., Pruiti Ciarello, F., Pacciarini, M. L., et al. (2019). Genotype diversity and distribution of Mycobacterium bovis from livestock in a small, high-risk area in northeastern Sicily, Italy. *PLoS Negl. Trop. Dis.* 13:e0007546. doi: 10.1371/journal.pntd.0007546

Marianelli, C., Verrubbi, V., Pruiti Ciarello, F., Ippolito, D., Pacciarini, M. L., Presti, D. M. L., et al. (2023). Geo-epidemiology of animal tuberculosis and *Mycobacterium bovis* genotypes in livestock in a small, high-incidence area in Sicily, Italy. *Front. Microbiol.* 14:1107396. doi: 10.3389/fmicb.2023.1107396

Martínez-López, B., Barasona, J. A., Gortázar, C., Rodríguez-Prieto, V., Sánchez-Vizcaíno, J. M., and Vicente, J. (2014). Farm-level risk factors for the occurrence, new infection or persistence of tuberculosis in cattle herds from south-Central Spain. *Prev. Vet. Med.* 116, 268–278. doi: 10.1016/j.prevetmed.2013.11.002

Metchock, B. G., Nolte, F. S., and Wallace, R. J. (1999). "Mycobacterium" in Manual of clinical microbiology. eds. E. Murray, J. Baron, M. A. Pfaller, F. C. Tenover and R. H. Yolken (Washington, D.C.: American Society for Microbiology), 399–437.

Müller, B., Hilty, M., Berg, S., Garcia-Pelayo, M. C., Dale, J., Boschiroli, M. L., et al. (2009). African 1, an epidemiologically important clonal complex of *Mycobacterium bovis* dominant in Mali, Nigeria, Cameroon, and Chad. *J. Bacteriol.* 191, 1951–1960. doi: 10.1128/JB.01590-08

Munyeme, M., Rigouts, L., Shamputa, I. C., Muma, J. B., Tryland, M., Skjerve, E., et al. (2009). Isolation and characterization of *Mycobacterium bovis* strains from indigenous Zambian cattle using spacer oligonucleotide typing technique. *BMC Microbiol.* 9:144. doi: 10.1186/1471-2180-9-144

Orloski, K., Robbe-Austerman, S., Stuber, T., Hench, B., and Schoenbaum, M. (2018). Whole genome sequencing of *Mycobacterium bovis* isolated from livestock in the United States, 1989-2018. *Front Vet Sci* 5:253. doi: 10.3389/fvets.2018.00253

Pozo, P., Lorente-Leal, V., Robbe-Austerman, S., Hicks, J., Stuber, T., Bezos, J., et al. (2022). Use of whole-genome sequencing to unravel the genetic diversity of a prevalent *Mycobacterium bovis* Spoligotype in a multi-host scenario in Spain. *Front. Microbiol.* 13:915843. doi: 10.3389/fmicb.2022.915843

Price-Carter, M., Brauning, R., de Lisle, G. W., Livingstone, P., Neill, M., Sinclair, J., et al. (2018). Whole genome sequencing for determining the source of *Mycobacterium bovis* infections in livestock herds and wildlife in New Zealand. *Front Vet Sci* 5:272. doi: 10.3389/fvets.2018.00272

Regulation (EU) 2016/429 of the European Parliament and of the Council of 9 March 2016 on transmissible animal diseases and amending and repealing certain acts in the area of animal health ('Animal Health Law'). Eur-Lex. Available at: http://data.europa.eu/eli/reg/2016/429/oj

Rodríguez, S., Romero, B., Bezos, J., de Juan, L., Alvarez, J., Castellanos, E., et al. (2010). High spoligotype diversity within a *Mycobacterium bovis* population: clues to understanding the demography of the pathogen in Europe. *Vet. Microbiol.* 141, 89–95. doi: 10.1016/j.vetmic.2009.08.007

Rodriguez-Campos, S., Aranaz, A., de Juan, L., Sáez-Llorente, J. L., Romero, B., Bezos, J., et al. (2011). Limitations of spoligotyping and variable-number tandem-repeat typing for molecular tracing of *Mycobacterium bovis* in a high-diversity setting. *J. Clin. Microbiol.* 49, 3361–3364. doi: 10.1128/JCM.00301-11

Rodriguez-Campos, S., Schürch, A. C., Dale, J., Lohan, A. J., Cunha, M. V., Botelho, A., et al. (2012). European 2--a clonal complex of *Mycobacterium bovis* dominant in the Iberian Peninsula. *Infect. Genet. Evol.* 12, 866–872. doi: 10.1016/j.meegid.2011.09.004

Roetzer, A., Diel, R., Kohl, T. A., Rückert, C., Nübel, U., Blom, J., et al. (2013). Whole genome sequencing versus traditional genotyping for investigation of a *Mycobacterium tuberculosis* outbreak: a longitudinal molecular epidemiological study. *PLoS Med.* 10:e1001387. doi: 10.1371/journal.pmed.1001387

Roring, S., Scott, A., Brittain, D., Walker, I., Hewinson, G., Neill, S., et al. (2002). Development of variable-number tandem repeat typing of *Mycobacterium bovis*: comparison of results with those obtained by using existing exact tandem repeats and spoligotyping. *J. Clin. Microbiol.* 40, 2126–2133. doi: 10.1128/JCM.40.6.2126-2133.2002

Sahraoui, N., Müller, B., Guetarni, D., Boulahbal, F., Yala, D., Ouzrout, R., et al. (2009). Molecular characterization of *Mycobacterium bovis* strains isolated from cattle slaughtered at two abattoirs in Algeria. *BMC Vet. Res.* 5:4. doi: 10.1186/1746-6148-5-4

Skuce, R. A., McCorry, T. P., McCarroll, J. F., Roring, S. M. M., Scott, A. N., Brittain, D., et al. (2002). Discrimination of *Mycobacterium tuberculosis* complex bacteria using novel VNTR-PCR targets. *Microbiology (Reading)* 148, 519–528. doi: 10.1099/00221287-148-2-519

Smith, N. H. (2012). The global distribution and phylogeography of *Mycobacterium bovis* clonal complexes. *Infect. Genet. Evol.* 12, 857–865. doi: 10.1016/j.meegid.2011.09.007

Smith, N. H., Hewinson, R. G., Kremer, K., Brosch, R., and Gordon, S. V. (2009). Myths and misconceptions: the origin and evolution of *Mycobacterium tuberculosis*. *Nat. Rev. Microbiol.* 7, 537–544. doi: 10.1038/nrmicro2165

Smith, N. H., Kremer, K., Inwald, J., Dale, J., Driscoll, J. R., Gordon, S. V., et al. (2006). Ecotypes of the *Mycobacterium tuberculosis* complex. *J. Theor. Biol.* 239, 220–225. doi: 10.1016/j.jtbi.2005.08.036

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Steiner, A., Stucki, D., Coscolla, M., Borrell, S., and Gagneux, S. (2014). KvarQ: targeted and direct variant calling from fastq reads of bacterial genomes. *BMC Genomics* 15:881. doi: 10.1186/1471-2164-15-881

Supply, P., Mazars, E., Lesjean, S., Vincent, V., Gicquel, B., and Locht, C. (2000). Variable human minisatellite-like regions in the *Mycobacterium tuberculosis* genome. *Mol. Microbiol.* 36, 762–771. doi: 10.1046/j.1365-2958.2000.01905.x

Trewby, H., Wright, D., Breadon, E. L., Lycett, S. J., Mallon, T. R., McCormick, C., et al. (2016). Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis. *Epidemics* 14, 26–35. doi: 10.1016/j.epidem.2015.08.003

van Ingen, J., Rahim, Z., Mulder, A., Boeree, M. J., Simeone, R., Brosch, R., et al. (2012). Characterization of Mycobacterium orygis as *M. tuberculosis* complex subspecies. *Emerg. Infect. Dis.* 18, 653–655. doi: 10.3201/eid1804.110888

Wood, D. E., and Salzberg, S. L. (2014). Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* 15:R46. doi: 10.1186/gb-2014-15-3-r46

Zimpel, C. K., Patané, J. S. L., Guedes, A. C. P., de Souza, R. F., Silva-Pereira, T. T., Camargo, N. C. S., et al. (2020). Global distribution and evolution of *Mycobacterium bovis* lineages. *Front. Microbiol.* 11:843. doi: 10.3389/fmicb.2020.00843

Zwyer, M., Çavusoglu, C., Ghielmetti, G., Pacciarini, M. L., Scaltriti, E., Van Soolingen, D., et al. (2021). A new nomenclature for the livestock-associated *Mycobacterium tuberculosis* complex based on phylogenomics [version 2; peer review: 2 approved]. *Open Research Europe* 1, 1–25. doi: 10.12688/openreseurope.14029.2