Deciphering the evolution of the temporal and geographic distribution of French Mycobacterium bovis genotypes using a high throughput SNP-targeted amplicon sequencing method

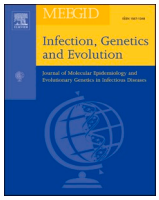
Lorraine Michelet, Jennifer Tambosco, Franck Biet, Patrick Fach, Sabine Delannoy, María Laura Boschiroli

Infection, Genetics and Evolution 114 (2023) 105497

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

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



Deciphering the evolution of the temporal and geographic distribution of French *Mycobacterium bovis* genotypes using a high throughput SNP-targeted amplicon sequencing method

Lorraine Michelet [a,*], Jennifer Tambosco [a], Franck Biet [b], Patrick Fach [c], Sabine Delannoy [c], María Laura Boschiroli [a]

[a] Université Paris-Est, ANSES Animal Health Laboratory, National reference laboratory for Tuberculosis, 94706 Maisons-Alfort, France
[b] ISP, INRAE, Tours, F-37380 Nouzilly, France
[c] IdentyPath Genomics Platform, Laboratory for Food Safety, Anses, 94706 Maisons-Alfort, France

ARTICLE INFO

Keywords:
Mycobacterium bovis
Bovine tuberculosis
France
SNP typing
Genotypic diversity

ABSTRACT

Mycobacterium bovis, which belongs to the *Mycobacterium tuberculosis* complex, is a highly clonal pathogen. However, several lineages of *M. bovis* have been described worldwide and nine different clusters were identified in France. Targeted amplicon sequencing using next-generation sequencing technology of eighty-eight phylogenetically informative single nucleotide polymorphisms (SNPs) were used to infer the phylogenetic relationship of 630 strains of the National Reference Laboratory isolated between 1979 and 2018 from various animal species. This study allowed classifying 618 different genotypic profiles (combination of a spoligotype and 8 loci-MIRU-VNTR profiles) into the nine previously identified clusters. A global analysis of the entire collection of the National Reference Laboratory has made it possible to represent the evolution of clonal complexes and clusters in time and space for better assessing epidemiological changes of bovine tuberculosis in France.

1. Introduction

Mycobacterium bovis (*M. bovis*), a member of the *Mycobacterium tuberculosis* complex (MTBC), is the main causative agent of animal tuberculosis (TB, formerly referred to as bovine tuberculosis (bTB)) (Brosch et al., 2002). *M. caprae* and more recently *M. orygis* has been described as responsible for the disease both in animals and human beings and *M. tuberculosis* as causative agent of reverse zoonosis to animals (Kock et al., 2021; Prodinger et al., 2014). The World Health Organization (WHO), the World Organization for Animal Health (WOAH), the Food and Agriculture Organization of the United Nations (FAO) and the International Union Against Tuberculosis and Lung disease (The Union) have indeed recommended a One Health strategy to promote a common human and veterinary approach against zoonotic tuberculosis. Studying the genetic diversity of *M. bovis* at the human-animal ecosystem interface could help to address the problem of TB (Appegren et al., 2023). In France, organized control campaigns against bovine tuberculosis have been implemented since 1954 (Boschiroli and Bénet, 2014). Systematic screening of TB by skin testing and slaughter of positive reactors led to a rapid decrease of the prevalence (Boschiroli and Bénet, 2014). France has been recognized officially bovine TB free since 2001 with a herd prevalence below 0.1%, albeit the number of outbreaks have been rising steadily since the last ten years. These outbreaks concentrate mainly in endemic areas in Nouvelle-Aquitaine (south western France) (Delavenne et al., 2021) where *M. bovis* not only affects cattle but several wildlife species. Indeed, increasing concern regarding TB in wildlife led to the implementation of a surveillance program in wildlife in 2011 (Reveillaud et al., 2018) that disclosed the presence of *M. bovis* infection in several wildlife species in the most affected TB areas and thus the existence of complex TB multi-host transmission systems. Badgers (*Meles meles*) and wild boar (*Sus scrofa*) are the main species implicated in TB transmission in France (Canini et al., 2022), but other wildlife species have also been found to be infected: red deer (*Cervus elaphus*) (Zanella et al., 2008), roe deer (*Capreolus capreolus*) (Lambert et al., 2017) and fox (*Vulpes vulpes*) (Michelet et al., 2018).

MTBC members are highly homogeneous and clonal organisms that can be differentiated by large sequence polymorphisms (LSPs), genomic insertions or deletions, which are also known as regions of difference (RDs) (Brites et al., 2018; Brosch et al., 2002). Ten years ago, four clonal complexes were described for *M. bovis* based on RDs, single nucleotide

* Corresponding author.
E-mail address: lorraine.michelet@anses.fr (L. Michelet).

https://doi.org/10.1016/j.meegid.2023.105497
Received 13 July 2023; Received in revised form 24 August 2023; Accepted 29 August 2023
Available online 30 August 2023

polymorphisms (SNPs) or particular spoligotype signatures. The European 1 (Eu1) clonal complex, characterized by the deletion of the chromosomal region RDEu1 and the absence of spacer 11, is the major group in the UK and its former trading partners (USA, South Africa, New Zealand, Australia and Canada) (Smith et al., 2011). The European 2 (Eu2) clonal complex, described by the absence of spacer 21 in their spoligotype patterns and by a SNP in *guaA* that is specific to these strains (Rodriguez-Campos et al., 2012), is dominant in the Iberian Peninsula. The African 1 (Af1) clonal complex, defined by a specific chromosomal deletion (RDAf1) and the absence of spacer 30, is found in west-central African countries, such as Mali, Cameroon, Nigeria, and Chad (Muller et al., 2009). The African 2 (Af2) clonal complex is localized in East Africa and is described by a specific chromosomal deletion (RDAf2) and the absence of spacers 3 to 7 (Berg et al., 2011) in their spoligotypes. Recently, a new nomenclature based on whole genome sequencing (WGS) data has been proposed for livestock-associated MTBC members (Zwyer et al., 2021). *M. bovis* has been subdivided into 8 lineages (La1.1 to La1.8).

Based on spoligotype and MIRU-VNTR analysis, in France we showed four predominant types of strains over a collection of 2332 *M. bovis* strains (Hauer et al., 2015). The first one is composed by SB0120 or BCG-like strains, now described as European 3 (Eu 3) clonal complex (Branger et al., 2020). Other representative clusters identified in France are the Eu2 clonal complex (SB0121 strains), SB0134 strains and the F4-family. To further study the genetic diversity of *M. bovis* in France, we performed a WGS analysis of 87 strains that allowed us to describe 9 clusters (Hauer et al., 2019). Six of these clusters have been described previously: cluster A or F4-family, cluster C or SB0134 strains, cluster D or Eu1 clonal complex, cluster F or Eu2 clonal complex, cluster G or F9 family and cluster I or Eu3 clonal complex (Hauer et al., 2015; Rodriguez-Campos et al., 2012; Smith et al., 2011). Clusters B, E and H, which are composed of few strains in this study are not defined by any other characteristics on their spoligotype pattern or MIRU-VNTR markers.

The National Reference Laboratory (NRL) for animal tuberculosis possesses a large collection of *M. bovis* strains isolated from cattle and various wildlife species. To date, this collection is composed of 7438 *M. bovis* strains isolated between 1979 and 2021. Almost 94% (6964 strains) of this collection has been characterized by spoligotyping and/or MIRU-VNTR. Its genetic diversity is represented by 721 different genotypic profiles, which is defined by the combination of a spoligotype profile and a 8 loci MIRU-VNTR profile (6 loci proposed by the Venomyc European consortium plus ETRC and QUB 26 loci) (Hauer et al., 2016). WGS has been shown to be a powerful and informative tool for SNP discovery disclosing genetic variations useful for phylogenetic classification (Hauer et al., 2019; Zimpel et al., 2019; Zwyer et al., 2021). Although performing WGS on bacterial strains is becoming progressively cost effective, a certain level of expertise in bioinformatics is it still required for analyzing the acquired data. Targeted amplicon sequencing of SNPs using next generation sequencing (tNGS) is less expensive and provides a powerful approach for the rapid phylogenetic classification and characterization of large collections (Ison et al., 2016), and is achievable without special bioinformatics skills. We performed such a study based on clade-specific SNPs in order to differentiate *M. bovis* clades, to classify the wide variety of genotype profiles and to analyze the evolution of clusters in time and space for better assessing epidemiological changes of TB in France.

2. Material and methods

2.1. Targeted single-nucleotide polymorphisms

Oligonucleotides were designed to target 88 SNPs. This panel of SNP has been selected based on the phylogenetic analysis of 87 strains representing the French genetic diversity (Hauer et al., 2019). SNP have been chosen on specific branches to classify strains into the different clusters and subclusters (Fig. S1). This classification is based on our previous works ((Hauer et al., 2015; Hauer et al., 2019); main clusters correspond to the previously described F4-family (cluster A), SB0134-group (cluster C), Eu1 clonal complex (cluster D), Eu2 clonal complex (cluster F) and SB0120-group (cluster I/ Eu3 clonal complex) or phylogenetically relevant groups. Sub-clusters correspond to well-defined groups of strains on the phylogenetic tree. A list of SNPs specific to the selected clusters was compiled using an in-house pipeline to identify those present only in all strains belonging to it. If several SNPs were available, we selected those in genomic regions favorable for primer design. SNPs with reference alleles are provided in Table S1. *Mycobacterium bovis* AF2122/97 (Genbank accession LT708304.1) was used as the reference strain.

2.2. Target-specific primer characteristics

The Access Array System for Illumina Sequencing Systems User Guide (Fluidigm, PN144 100–3770 H1) was used with described modifications to perform the 2-primer target-specific PCR amplification on the 48.48 Access Array Integrated Fluidic Circuit (IFC). tNGS with the Fluidigm Access Array amplifies a short genomic region, 100–150 base pairs (bp) in length with a final product of 150–200 bp. Oligonucleotides were designed to target the absolute genome position of the previously identified 88 SNPs with the targeted SNP to be placed in the center of the final amplified product (Table S2). Oligonucleotides were designed with a melting temperature of 59–62 °C, optimum temperature 60 °C. Common sequence 1 (CS1; 5'-ACACTGACGACATGGTTCTACA-3′) and common sequence 2 (CS2; 5'-TACGGTAGCAGAGACTTGGTCT-3′) universal primer sequences, required for the Illumina MiSeq amplicon tagging and indexing, were added to the 5′- end of all target-specific forward (TSF) and reverse (TSR) primers, respectively. The "Access Array generate tagged primers workbook" (Fluidigm, PN 100–3873) was used to add universal sequence tags to the target-specific primers (Table S3). Primers were purchased from Eurofins Genomics (Ebersberg, Germany).

2.3. Primer validation

Target-specific primer validation was performed in a 96-well PeqSTAR Thermocycler. Two *M. bovis* strains thermolysates and water were used as positive and negative controls for primer validation, respectively. The validation master mix solution with final concentrations contained: 1× FastStart High Fidelity Reaction Buffer without MgCl2 (Roche, Meylan, France), 4.5 mM MgCl2 (Roche), 5% DMSO (Roche), 200 μM each PCR grade Nucleotide Mix (Roche), 1× Access Array loading reagent (Fluidigm, San Francisco, USA), tagged TSF and TSR primers (200 nM each), 1 μl of DNA (lysates), 0.05 U/μl FastStart High Fidelity Enzyme Blend (Roche), molecular grade water to 20 μl final volume. The Thermocycler program run with 35 cycles was performed as recommended in the Fluidigm targeted sequencing primer validation protocol (PN 100–3770 H1) and provided in Table S4. PCR products were checked for each PCR reaction on 2% agarose gel to ensure that only one PCR product of the correct size was generated for each target.

Multiplexing primer pairs allowed 88-targeted SNPs to be run on a single 48.48 Access Array IFC for a total of 4224 amplicons generated on each 48.48 Access Array IFC. Primer pairs selected for duplex or triplex were first compared on the ThermoFisher Scientific Multiple Primer Analyzer (https://www.lifetechnologies.com). Seventeen primer pairs were used in simplex, 44 primer pairs were used in duplex and 27 primer pairs were used in triplex (Table S2). Duplex and triplex were validated with a SYBR real-time PCR. The validation master mix solution with final concentrations contained: 1× Power SYBR Green Master Mix (Applied Biosystems), primers (200 nM each), 1 μl of DNA (lysates), molecular grade water to 12 μl final volume.

2.4. Bacterial strains

All the strains included in this study were issued from the French

animal Tuberculosis reference laboratory collection. 630 strains isolated between 1979 and 2018 from various animal species representing 618 different genotypic profiles (spoligotype and MIRU-VNTR) were selected (Table S5). This panel of strains represent 85% of the profile diversity of the whole collection. Strains belonging to the 87 strains of our previous work have been included as control. For the study, Middlebrook 7H9 liquid culture media of these strains heated 1 h at 80 °C, washed with 1 ml NaCl solution and resuspended in 100 µl Tris-EDTA were employed.

2.5. 48.48 Access array IFC workflow

The 2-Primer Target-Specific PCR Amplification was performed according to manufacturer's recommendations (Fluidigm, PN100–1031, Appendix C). The Fluidigm 48.48 Access Array workflow for Illumina sequencing was performed as previously described (Ison et al., 2016). Briefly, a first PCR reaction combining 48 primers assays (as previously described) and 48 samples into 2304 separate microreaction chambers was performed on the 48.48 Access Array IFC. In this step, all the amplicons generated for each sample were collected in the corresponding sample inlet. The tagged target-specific PCR amplicons were then transferred from the Access Array IFC to a 96-well PCR plate; resulting in sample-specific pooled amplicons. Sixteen 48.48 Access Array IFC were used in this study.

In the second PCR reaction, performed on a conventional thermocycler, Illumina sequence-specific adaptors with individual sample barcodes on the reverse primer were attached to the amplicons produced in the first PCR reaction (Fluidigm, PN 100–3873). The Fluidigm Access Array Barcode Library for Illumina Sequencers- 384 (Single Direction) (Fluidigm, kit PN 100-4876) barcode numbers 1–384 were used for amplicon tagging in two successive experiments. The Illumina employed specific primer sequences were as follows: forward (PE1_CS1) 5′- AATGATACGGCGACCACCGAGATCTACACTGACGACATGGTTCTACA-3′ and reverse (PE2_Barcode_CS2) 5′-CAAGCAGAAGACGGCATACGAGAT-(BARCODE)- TACGGTAGCAGAGACTTGGTCT-3′. Each DNA sample was assigned a unique barcode sequence, which is used for de-multiplexing samples downstream.

The PCR products generated were controlled on agarose gel. At two different times, 384 PCR products were then pooled to create a single product library. The quality of the pooled library was checked on a LabChip GX (PerkinElmer Inc.) and the library concentration was determined on the QuantiFluor System (Promega).

2.6. Amplicon sequencing using the Illumina MiSeq

Custom Access Array primers (FL1 and FL2) were incorporated to read the amplicons generated from the Access Array IFC. FL1 and FL2 primer mixes were prepared according to the Access Array System for Illumina Sequencing Systems User guide. The FL1 primer mix was prepared by combining CS1 oligonucleotide (5′-A + CA + CTG + ACGACATGGTTCTACA-3′) and CS2 oligonucleotide (5′-T + AC + GGT + AGCAGAGACTTGGTCT-3′). The FL2 primer mix combines the CS1rc oligonucleotide (5′-T + GT + AG + AACCATGTCGTCAGTGT-3′) and CS2rc oligonucleotide (5′-A + GAC + CA + AGTCTCTGCTACCGTA-3′); locked nucleic acid (LNA) nucleotides are preceded by a "+"; oligonucleotides were synthesized by TIB MOLBIOL (Berlin, Germany).

Denaturation and dilution of the Access Array pooled library was performed according to the Illumina Preparing Libraries for Sequencing on the MiSeq protocol (#150369740 Rev. D). The library was first diluted to 2 nM and 4 nM, for the first and the second experiment respectively based on the Quantifluor concentration and the average length (bp) of the libraries. The denatured DNA for 2 or 4 nM library procedure was performed and diluted to a final concentration of 8 pM or 12 pM for the first and the second experiment respectively.

The denatured and diluted sample library was sequenced on an Illumina MiSeq instrument as a 2 × 150 bases paired-end run (MiSeq Flow Cell, v2 reagents) according to the Illumina MiSeq System User Guide (#15027617).

2.7. Bioinformatics analysis

Sequencing reads were de-multiplexed according to the sample specific barcodes on the Illumina MiSeq. A reference sequence was created by concatenating the sequences of the 88 amplicons from the reference AF2122/97 *M. bovis* NCBI NC_002945.4 (Genbank accession LT708304.1) and sequence reads for each *M. bovis* strain were mapped to this reference sequence using the Bionumerics software version 7.6 (Applied Math, Belgium). The wgSNP module in Bionumerics was used to identify and output SNP variants present. A maximum parsimony tree was created by concatenating the 88 SNPs to identify the relationship of 630 *M. bovis* strains.

3. Results

3.1. Classification of M. bovis strains according to phylogenetically informative SNPs

The entire panel studied could be classified despite some missing data. Indeed, only twenty-eight of the targeted loci failed to be amplified in 27 strains (one or two loci per strain), representing 0.05% of missing data. Among the 630 strains, 131 have been identified as belonging to Cluster A, 13 to Cluster B, 94 to Cluster C, 24 to Eu1 clonal complex, 27 to Cluster E, 115 to Eu2 clonal complex, 30 to Cluster G, 9 to Cluster H and 170 to Eu3 clonal complex (Fig. 1 and Table S5). Seventeen strains have not been assigned to a cluster. Indeed, 8 strains are identified at the root of the E + F clusters, 6 strains at the root of the G + H + I clusters and 1 strain identified at the root of the A + B + C + D + E + F clusters. Two strains carry a particular SNP profile that do not allow them to be included in one of the defined groups. These two strains were found to be respectively a *M. caprae* strain (spoligotype SB0416) and a *M. bovis* strain isolated from a monkey (spoligotype SB0912) belonging to the Af2 clonal complex.

Specifically, the strains classified in Cluster A (n = 131) have 33 spoligotypes and 91 VNTR profiles, for a total of 114 genotypes. All spoligotypes are characterized by the absence of spacer 33. The majority of VNTR profiles (except 4) have a truncated repeat at the QUB26 locus. A strain isolated in 1999 shares only SNP Bov4 with the other strains in the cluster and has a spoligotype profile with spacer 33 present. 54 strains belong to the A1 subcluster, representing 45 genotypic profiles. All these strains share a common feature in their spoligotype profile with the absence of spacers 5 and 6. Only 13 strains belong to cluster B with 12 genotypic profiles. These strains appear to be separated into two groups, one of which is characterized by the absence of spacers 4 to 6 and a particular VNTR profile: alleles 7 or 9 in ETR A, 5 in ETR C and 1 in ETR D (Table S5). Cluster C is subdivided into two sub-clusters C1 (n = 36) and C2 (n = 58). However, all the strains belonging to this cluster share the same characteristic on the 25 identified spoligotype profiles, i.e. the absence of spacer 4 and 5. Subcluster C1 and C2 represent respectively 33 and 46 genotypic profiles. The Eu1 clonal complex (cluster D) is represented by 24 strains representing 24 genotypic profiles. Within this cluster, subcluster D1 has been highlighted and characterized by the absence of spacers 6 and 8 to 12 on the spoligotype profile. Cluster E is represented by 27 strains representing 25 genotypic profiles. This cluster does not present any specific characteristic. A third of the strains of this cluster present the SB0120 spoligotype. Strains of the Eu2 clonal complex or cluster F (n = 115) represent 102 genotypes. As expected, all spoligotypes were characterized by the absence of spacer 21. Three different subcluters have been identified F1, F2 and F3, which represent respectively 18, 29 and 19 genotypic profiles. Cluster G is represented by 30 strains with 26 genotypic profiles. These strains are characterized by the absence of spoligotype spacers 1 to 17 and 23–24. Nine strains with 9 genotypic profiles belong to Cluster H. This cluster

does not present any specific characteristic. The Eu3 clonal complex (or cluster I) is the most representative cluster with 170 strains representing 159 genotypic profiles. This cluster has been subdivided according to the phylogenetic tree with a subcluster, named cluster I1, which excludes vaccine strains and other phylogenetically close strains. In our panel, 12 strains have been assigned to cluster I and represent 10 genotypic profiles. The 158 other strains belong to cluster I1, in which 3 sub-clusters have been identified (I1a, I1b and I1c), grouping respectively 6, 23 and 10 genotypic profiles.

The targeted SNP sequence data combined with the WGS of the 87 strains was able to classify 618 genotypic profiles into the 9 clusters. These profiles represent 85% of those in the strain collection. Profiles showing only minor variations (one repeat on one allele) were considered to belong to the same cluster as the one characterized by tNGS. This information allowed us to assign each of the 6843 strains in the NRL collection (for which a genotypic profile had previously been determined) into a specific cluster.

3.2. Temporal evolution of M. bovis cluster diversity

The above results permitted us to follow the evolution of the number of strains in each cluster from 1979 to 2021 (Fig. 2). Clusters B and G can be considered absent in the country, given that they have not been detected since the last 10 years. These clusters are respectively composed of 39 and 125 strains of the collection and have not been observed since 2014 and 2012 respectively. The Eu1 clonal complex and the cluster H are observed sporadically throughout the period. They comprise 93 and 31 strains in the collection respectively. Clusters C, E and Eu2 clonal complex are recurrently observed and include respectively 1105, 214 and 694 strains. Cluster E was not observed between 2003 and 2010, but did not really disappeared since it has been responsible for several outbreaks in the last 13 years. Finally, the two major clusters in France are cluster A and the Eu3 clonal complex with respectively 2006 and 2536 strains. These two clusters have been present in the collection since the 1980s and became dominant in the last fifteen years.

3.3. Phylogeography of M. bovis clusters in France between 2000 and 2021

In addition to the temporal study, our aim was to study the evolution of the geographical distribution of strains of the whole collection belonging to each cluster by “department” (French administrative regional division). The data associated with strains collected before 2000 were too incomplete for this geographical analysis. We therefore represented the strain geographic origin (5592 strains) by 5-year periods from 2001 to 2021 (Fig. 3).

The first period (2001–2005) concerns 541 strains of the collection, which are distributed throughout France. No Cluster B strain was observed during this period. Clusters A, C and the Eu3 clonal complex represent each 148, 165 and 132 strains distributed in different French departments. In the second period (2006–2010), an increase in the number of strains was observed in 3 departments. In particular, cluster C (n = 446) emerged in the north of France in Normandy, which corresponds to the first outbreak of bovine tuberculosis in wildlife. Eu3 clonal complex strains (n = 365) emerge in two departments in Burgundy (Côte

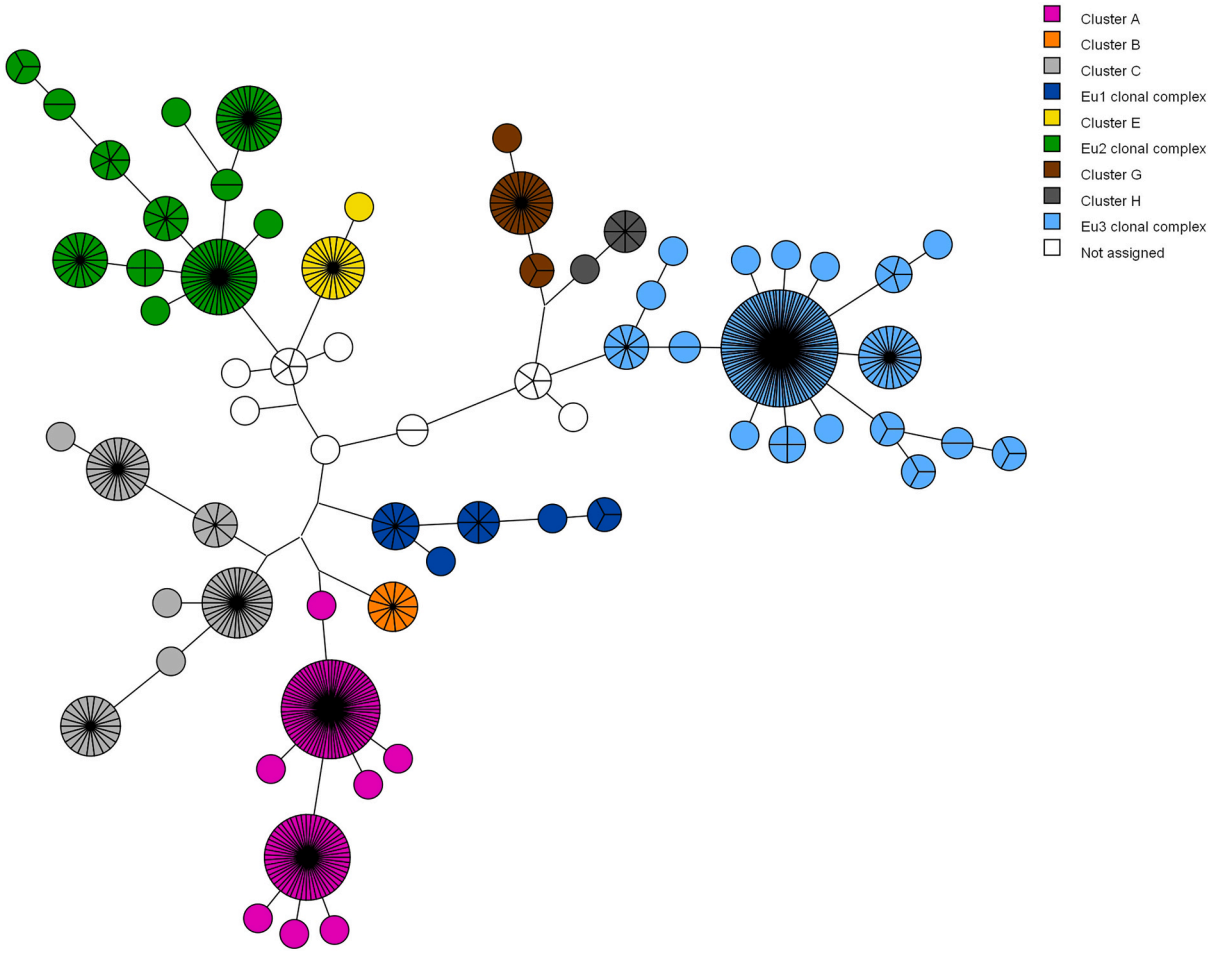


Fig. 1. Maximum parsimony tree of 630 *M. bovis* strains based on the 88 SNPs.

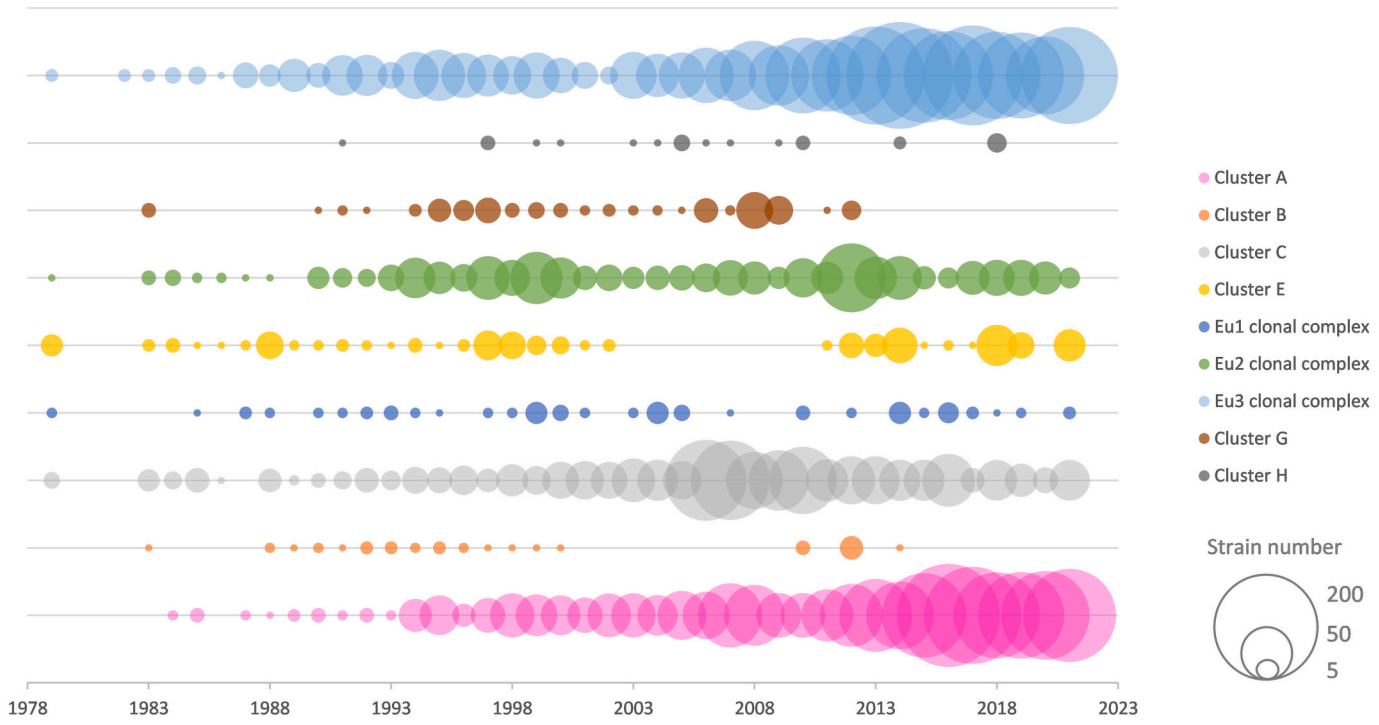


Fig. 2. Graphic representation of strain number per year for each cluster.

d'Or) and Nouvelle Aquitaine (Dordogne). Few strains of clusters B, H and Eu1 clonal complex are observed over this period (less than 10 each) and none of Cluster E. In the third period (2011–2015), all clusters are observed albeit less frequently for Clusters B ($n = 11$), E ($n = 48$), G ($n = 8$), H (n = 3) and Eu1 clonal complex ($n = 13$). Cluster C and Eu2 clonal complex include respectively 179 and 184 strains over this period isolated in several departments. Three areas of high endemicity appeared during this period, with the emergence of strains of cluster A ($n = 434$) notably in the south of Nouvelle-Aquitaine (Pyrénées-Atlantiques and Landes) and strains of the Eu3 clonal complex ($n = 759$) mainly in Côte d'Or and Dordogne. Over the last period (2016–2021), the strong regionalization of TB with major clusters per department is confirmed. The Eu3 clonal complex ($n = 886$) has dominated all over the north of Nouvelle-Aquitaine and remains present in Côte d'Or. Cluster A ($n = 943$) is emerging in the south of France, mainly in the south of Nouvelle-Aquitaine and in Corsica.

4. Discussion

In this study, a high-throughput tNGS approach was applied to 630 strains of *M. bovis* in order to obtain a phylogenic inference. We characterized 618 of the 721 genetic profiles identified in the whole collection. Some genotypes differing only by one allele at a MIRU-VNTR locus are thus considered quasi identical to simplify our study in particular since some loci are hypervariable in the different clusters, i.e. Qub3232 for cluster A, Qub26 for cluster C, Qub11a for Eu2 clonal complex and ETR B for Eu3 clonal complex (Hauer et al., 2016). We focused on the most frequently identified genotypic profiles in the collection. The 88 phylogenetically informative SNPs investigated in this study allowed classification of the strains into 9 previously described clusters (Hauer et al., 2019), highlighting a greater diversity within cluster A and the Eu3 clonal complex. The three minor clusters previously described (clusters B, E and H) were identified but remain largely in minority in terms of genotypic profiles. The Eu1 clonal complex, which is predominant in the United Kingdom (UK) and Ireland (Smith et al., 2011), is present in France but in a more limited way. The Eu2 clonal complex, which is dominant in the Iberian Peninsula (Rodriguez-Campos et al., 2012), is frequently observed with a certain genotypic diversity. The SNPs identified here can be used more widely in Europe and worldwide to classify *M. bovis* strains into different clusters or lineages. Indeed, some are specific to lineages 1.7.1 (Eu2), 1.8.1 (Eu1), 1.8.2 (A + B + C) or more widely 1.8 (Zwyer et al., 2021).

Our results highlight some genetic features specific to each cluster. Up to a certain limit, spoligotype and MIRU-VNTR profiles were good indicators for classifying strains into Clusters A, C, G or Eu1 and Eu2 clonal complexes as previously described (Hauer et al., 2015; Hauer et al., 2016). This study disclosed sub-cluster D1, characterized by the absence of spacers 6 and 8 to 12 in their spoligotypes. However, SNP genotyping clearly brings a better insight in the genetic differentiation for SB0120 spoligotype *M. bovis* strains, which are distributed in cluster B, E or the Eu3 clonal complex. The results also highlighted the limit of classical genotyping method, especially spoligotyping, which presents some homoplasia (Guimaraes and Zimpel, 2020).

Some of the targeted loci failed to be amplified in 27 strains for one locus and two loci in one strain. However, this missing SNP information had no negative consequences in strain classification on the final panel for determining phylogenetic relationships of strains within and between clonal complexes as several SNP were investigated for each sub cluster. A comparable failure rate was observed in a similar study on *Escherichia coli* O26:H11 cattle strains (Ison et al., 2016).

This study allows us to decipher the genetic diversity of a large strain collection isolated between 1978 and 2021 as 92% of the *M. bovis* strains has been assigned to a cluster. It provides an overall view of the evolution of TB in France, which can be related to certain events linked to the surveillance programs or the evolution of breeding systems. Over the past 50 years, a bottleneck in the genetic diversity of strains has been observed. This may be related to the introduction of total slaughter of infected herds since the late 1990s (Boschiroli and Bénet, 2014). This bottleneck may also be linked to changes in breeding practices following the introduction of milk quotas (MAAPAR, 1990) and the encouragement of beef cattle breeding by the European Union (Agreste, 2004). This period was marked by a fall in the prevalence of bovine tuberculosis, from 25% of infected herds in 1965 to less than 0.1% in 2000 (Boschiroli and Bénet, 2014). The number of strains in the collection

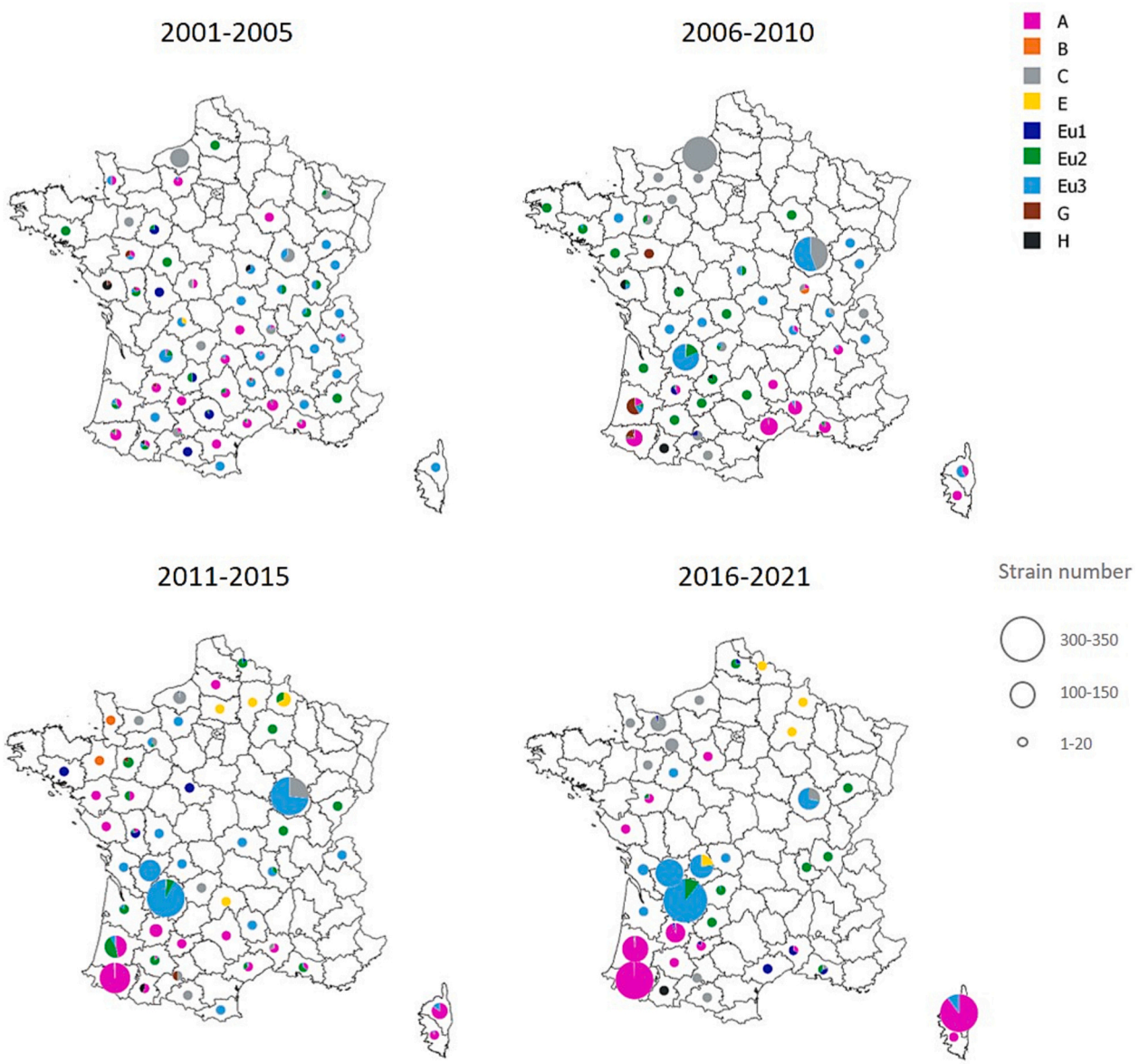


Fig. 3. Geographic distribution of *M. bovis* clusters in France. Pie charts represent the percentage of each cluster per department and are proportional to the number of strains.

varies considerably depending on the period: approximately 400 in the 1980–90s, 1200 in the 1991–2000s, 1900 in the 2001–2010s and 3800 in the 2011–2020s. This increase in strain collection, which does not reflect an increase in TB incidence in cattle, is the consequence an improvement in the control program within the infected herds, the progressive introduction of bacterial culture for confirmation of infection and outbreak declaration, and to a stronger surveillance pressure both in cattle and in wildlife in the recent years. Over the last 15 years, the distribution of TB in France has suggested a strong relationship with suckler herds (Delavenne et al., 2021), particularly in Nouvelle Aquitaine, which has the largest number of beef herds in France. Changes in breeding practices have led to the decline of certain clones in favor of the current clones, which are sometimes linked to particular breeds (Limousine or Blonde d'Aquitaine in Nouvelle Aquitaine, or Charolaise in Burgundy). A decrease in the number of strains and genotypes was observed around the 2000s, concomitant with obtaining the officially TB free (OTF) status in 2001. Bayesian phylodynamic studies could be used to confirm these different hypotheses by comparing models incorporating genomic data from the strains and different epidemiological parameters.

For some clusters, particular events may explain a coincident increase in strains' number. For example, the increase in cluster C strains in Normandy in 2006 followed the detection of *M. bovis* in wildlife, and an extensive epidemiological investigation in the region combined with total depopulation of the red deer population and an strong decrease of wild boar numbers (Zanella et al., 2012). An increase of cluster A and Eu3 complex strains since 2010 is also explained by the implementation of the national wildlife surveillance program Sylvatub that disclosed infection in several other highly affected regions in the South west of France as well as in Burgundy (Reveillaud et al., 2018). For other clusters, the evolution of the number of strains is linked to the trading context. Strains belonging to the Eu1 clonal complex are mainly associated with introductions of animals from the United Kingdom while for the Eu2 clonal complex, some genotypes are linked to introductions of animals from Spain but some other are specific to certain French regions (Boschiroli et al., 2015).

The use of the Access Array in the targeted amplicon sequencing procedure has many advantages in terms of cost, massive multiplexing capabilities and reduced handling time (Lange et al., 2014). Furthermore, the primers designed to target the published SNPs investigated in

this work can also be used in a high-resolution melting (HRM) assay or PCR followed by Sanger sequencing to rapidly identify a cluster. The application of HRM has previously been applied for rapid differentiation of MTBC members or pyrazinamide susceptibility testing (Landolt et al., 2019; Pholwat et al., 2014).

In conclusion, this study, based on a new high-throughput technology, has enabled a better description of the genetic diversity of French *M. bovis* strains. This rapid genotyping of a very large number of strains has made it possible to explore the evolution of tuberculosis and *M. bovis* lineages in France over the last 45 years. This new technology offers many advantages in terms of cost, avoiding complete sequencing and in terms of bioinformatics analysis and skills required. SNP screening could be useful for preliminary epidemiological tracing, saving resources for future more in-depth analyses.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2023.105497.

CRediT authorship contribution statement

Lorraine Michelet: Conceptualization, Methodology, Formal analysis, Writing – original draft. **Jennifer Tambosco:** Formal analysis, Writing – review & editing. **Franck Biet:** Funding acquisition, Writing – review & editing. **Patrick Fach:** Resources, Writing – review & editing. **Sabine Delannoy:** Formal analysis, Resources, Writing – review & editing. **María Laura Boschiroli:** Conceptualization, Funding acquisition, Writing – review & editing.

Declaration of Competing Interest

None.

Data availability

The 87 genomes are available at GenBank under the Bioproject Accession No. PRJNA485121. Meanwhile, the datasets used and analyzed in this study are available at the corresponding author.

Acknowledgment

The French Ministry of Agriculture mainly financed this study in the framework of the RFSA (Réseau Français pour la Santé Animale) call on TB projects.

References

Agreste, 2004. L'agriculture, la forêt et l'industrie agroalimentaire en 2004. MAAPAR-SCEES.

Appegren, A., Boschiroli, M.L., De Cruz, K., Michelet, L., Hery-Arnaud, G., Kempf, M., Lanotte, P., Bemer, P., Peuchant, O., Pestel-Caron, M., Skalli, S., Brasme, L., Martin, C., Enault, C., Carricajo, A., Guet-Revillet, H., Ponsoda, M., Jacomo, V., Bourgoin, A., Trombert-Paolantoni, S., Carriere, C., Dupont, C., Conquet, G., Galal, L., Banuls, A.L., Godreuil, S., 2023. Genetic diversity and population structure of *Mycobacterium bovis* at the human-animal-ecosystem interface in France: "A one health approach". Pathogens 12.

Berg, S., Garcia-Pelayo, M.C., Muller, B., Hailu, E., Asiimwe, B., Kremer, K., Dale, J., Boniotti, M.B., Rodriguez, S., Hilty, M., Rigouts, L., Firdessa, R., Machado, A., Mucavele, C., Ngandolo, B.N., Bruchfeld, J., Boschiroli, L., Muller, A., Sahraoui, N., Pacciarini, M., Cadmus, S., Joloba, M., van Soolingen, D., Michel, A.L., Djonne, B., Aranaz, A., Zinsstag, J., van Helden, P., Portaels, F., Kazwala, R., Kallenius, G., Hewinson, R.G., Aseffa, A., Gordon, S.V., Smith, N.H., 2011. African 2, a clonal complex of *Mycobacterium bovis* epidemiologically important in East Africa. J. Bacteriol. 193, 670–678.

Boschiroli, M.-L., Bénet, J.-J., 2014. Bovine tuberculosis eradication in France. In: Thoen, C.O., Steele, J.H., Kaneene, J.B. (Eds.), Zoonotic Tuberculosis: Mycobacterium Bovis and Other Pathogenic Mycobacteria -Third Edition. Wiley Blackwell, West Sussex, UK, pp. 341–347.

Boschiroli, M.L., Michelet, L., Hauer, A., De Cruz, K., Courcoul, A., Hénault, S., Palisson, A., Karoui, C., Biet, F., Zanella, G., 2015. Tuberculose bovine en France : cartographie des souches de *Mycobacterium bovis* entre 2000-2013. Bull. Epid. Santé. Anim. Alim. 70, 2–8.

Branger, M., Loux, V., Cochard, T., Boschiroli, M.L., Biet, F., 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.

Brites, D., Loiseau, C., Menardo, F., Borrell, S., Boniotti, M.B., Warren, R., Dippenaar, A., Parsons, S.D.C., Beisel, C., Behr, M.A., Fyfe, J.A., Coscolla, M., Gagneux, S., 2018. A new phylogenetic framework for the animal-adapted *mycobacterium tuberculosis* complex. Front. Microbiol. 9, 2820.

Brosch, R., Gordon, S.V., Marmiesse, M., Brodin, P., Buchrieser, C., Eiglmeier, K., Garnier, T., Gutierrez, C., Hewinson, G., Kremer, K., Parsons, L.M., Pym, A.S., Samper, S., van Soolingen, D., Cole, S.T., 2002. A new evolutionary scenario for the *Mycobacterium tuberculosis* complex. Proc. Natl. Acad. Sci. U. S. A. 99, 3684–3689.

Canini, L., Modenesi, G., Courcoul, A., Boschiroli, M.L., Durand, B., Michelet, L., 2022. Deciphering the role of host species for two *Mycobacterium bovis* genotypes from the European 3 clonal complex circulation within a cattle-badger-wild boar multihost system. Microbiol. Open. 12.

Delavenne, C., Desvaux, S., Boschiroli, M., Carles, S., Durand, B., Forfait, C., Gache, K., Garapin, F., Girard, S., Keck, N., Moyen, J.L., Pieus, A., Réveillaud, E., Richomme, C., Rivière, J., Dupuy, C., Chevalier, F., 2021. Surveillance de la tuberculose due à *Mycobacterium bovis* en France métropolitaine en 2019 : résultats et indicateurs de fonctionnement. Bull. Epidémiol. Santé Anim. Aliment. 94, 1–23.

Guimaraes, A.M.S., Zimpel, C.K., 2020. *Mycobacterium bovis*: from genotyping to genome sequencing. Microorganisms 8.

Hauer, A., De Cruz, K., Cochard, T., Godreuil, S., Karoui, C., Henault, S., Bulach, T., Banuls, A.L., Biet, F., Boschiroli, M.L., 2015. Genetic evolution of *Mycobacterium bovis* causing tuberculosis in livestock and wildlife in France since 1978. PLoS One 10, e0117103.

Hauer, A., Michelet, L., De Cruz, K., Cochard, T., Branger, M., Karoui, C., Henault, S., Biet, F., Boschiroli, M.L., 2016. MIRU-VNTR allelic variability depends on *Mycobacterium bovis* clonal group identity. Infect. Genet. Evol. 45, 165–169.

Hauer, A., Michelet, L., Cochard, T., Branger, M., Nunez, J., Boschiroli, M.L., Biet, F., 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.

Ison, S.A., Delannoy, S., Bugarel, M., Nagaraja, T.G., Renter, D.G., den Bakker, H.C., Nightingale, K.K., Fach, P., Loneragan, G.H., 2016. Targeted amplicon sequencing for single-nucleotide-polymorphism genotyping of attaching and effacing *Escherichia coli* O26:H11 cattle strains via a high-throughput library preparation technique. Appl. Environ. Microbiol. 82, 640–649.

Kock, R., Michel, A.L., Yeboah-Manu, D., Azhar, E.I., Torrelles, J.B., Cadmus, S.I., Brunton, L., Chakaya, J.M., Marais, B., Mboera, L., Rahim, Z., Haider, N., Zumla, A., 2021. Zoonotic tuberculosis - the changing landscape. Int. J. Infect. Dis. 113 (Suppl. 1), S68–S72.

Lambert, S., Hars, J., Réveillaud, E., Moyen, J.L., Gares, H., Rambaud, T., Gueneau, E., Faure, E., Boschiroli, M., Richomme, C., 2017. Host status of wild roe deer in bovine tuberculosis endemic areas. Eur. J. Wildl. Res. 63, 15.

Landolt, P., Stephan, R., Scherrer, S., 2019. Development of a new high resolution melting (HRM) assay for identification and differentiation of *Mycobacterium tuberculosis* complex samples. Sci. Rep. 9, 1850.

Lange, V., Bohme, I., Hofmann, J., Lang, K., Sauter, J., Schone, B., Paul, P., Albrecht, V., Andreas, J.M., Baier, D.M., Nething, J., Ehninger, U., Schwarzelt, C., Pingel, J., Ehninger, G., Schmidt, A.H., 2014. Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. BMC Genomics 15, 63.

MAAPAR, 1990. La réforme de la PAC.

Michelet, L., De Cruz, K., Henault, S., Tambosco, J., Richomme, C., Reveillaud, E., Gares, H., Moyen, J.L., Boschiroli, M.L., 2018. *Mycobacterium bovis* infection of red fox, France. Emerg. Infect. Dis. 24, 1150–1153.

Muller, B., Hilty, M., Berg, S., Garcia-Pelayo, M.C., Dale, J., Boschiroli, M.L., Cadmus, S., Ngandolo, B.N., Godreuil, S., Diguimbaye-Djaibe, C., Kazwala, R., Bonfoh, B., Njanpop-Lafourcade, B.M., Sahraoui, N., Guetarni, D., Aseffa, A., Mekonnen, M.H., Razanamparany, V.R., Ramarokoto, H., Djonne, B., Oloya, J., Machado, A., Mucavele, C., Skjerve, E., Portaels, F., Rigouts, L., Michel, A., Muller, A., Kallenius, G., van Helden, P.D., Hewinson, R.G., Zinsstag, J., Gordon, S.V., Smith, N. H., 2009. African 1, an epidemiologically important clonal complex of *Mycobacterium bovis* dominant in Mali, Nigeria, Cameroon, and Chad. J. Bacteriol. 191, 1951–1960.

Pholwat, S., Stroup, S., Gratz, J., Trangan, V., Foongladda, S., Kumburu, H., Juma, S.P., Kibiki, G., Houpt, E., 2014. Pyrazinamide susceptibility testing of *Mycobacterium tuberculosis* by high resolution melt analysis. Tuberculosis (Edinb) 94, 20–25.

Prodinger, W.M., Indra, A., Koksalan, O.K., Kilicaslan, Z., Richter, E., 2014. *Mycobacterium caprae* infection in humans. Expert Rev. Anti-Infect. Ther. 12, 1501–1513.

Reveillaud, E., Desvaux, S., Boschiroli, M.L., Hars, J., Faure, E., Fediaevsky, A., Cavalerie, L., Chevalier, F., Jabert, P., Poliak, S., Tourette, I., Hendrikx, P., Richomme, C., 2018. Infection of wildlife by *Mycobacterium bovis* in France assessment through a National Surveillance System. Sylvatub. Front. Vet. Sci. 5, 262.

Rodriguez-Campos, S., Schurch, A.C., Dale, J., Lohan, A.J., Cunha, M.V., Botelho, A., De Cruz, K., Boschiroli, M.L., Boniotti, M.B., Pacciarini, M., Garcia-Pelayo, M.C., Romero, B., de Juan, L., Dominguez, L., Gordon, S.V., van Soolingen, D., Loftus, B., Berg, S., Hewinson, R.G., Aranaz, A., Smith, N.H., 2012. European 2–a clonal complex of *Mycobacterium bovis* dominant in the Iberian Peninsula. Infect. Genet. Evol. 12, 866–872.

Smith, N.H., Berg, S., Dale, J., Allen, A., Rodriguez, S., Romero, B., Matos, F., Ghebremichael, S., Karoui, C., Donati, C., Machado Ada, C., Mucavele, C., Kazwala, R.R., Hilty, M., Cadmus, S., Ngandolo, B.N., Habtamu, M., Oloya, J., Muller, A., Milian-Suazo, F., Andrievskaia, O., Projahn, M., Barandiaran, S., Macias, A., Muller, B., Zanini, M.S., Ikuta, C.Y., Rodriguez, C.A., Pinheiro, S.R.,

Figueroa, A., Cho, S.N., Mosavari, N., Chuang, P.C., Jou, R., Zinsstag, J., van Soolingen, D., Costello, E., Aseffa, A., Proano-Perez, F., Portaels, F., Rigouts, L., Cataldi, A.A., Collins, D.M., Boschiroli, M.L., Hewinson, R.G., Ferreira Neto, J.S., Surujballi, O., Tadyon, K., Botelho, A., Zarraga, A.M., Buller, N., Skuce, R., Michel, A., Aranaz, A., Gordon, S.V., Jeon, B.Y., Kallenius, G., Niemann, S., Boniotti, M.B., van Helden, P.D., Harris, B., Zumarraga, M.J., Kremer, K., 2011. European 1: a globally important clonal complex of *Mycobacterium bovis*. Infect. Genet. Evol. 11, 1340–1351.

Zanella, G., Durand, B., Hars, J., Moutou, F., Garin-Bastuji, B., Duvauchelle, A., Ferme, M., Karoui, C., Boschiroli, M.L., 2008. *Mycobacterium bovis* in wildlife in France. J. Wildl. Dis. 44, 99–108.

Zanella, G., Bar-Hen, A., Boschiroli, M.L., Hars, J., Moutou, F., Garin-Bastuji, B., Durand, B., 2012. Modelling transmission of bovine tuberculosis in red deer and wild boar in Normandy, France. Zoonoses Public Health 59 (Suppl. 2), 170–178.

Zimpel, C.K., Patané, J.S.L., Proença Guedes, A.C., de Souza, R.F., Pereira-Silva, T.T., Soler Camargo, N.C., de Souza Filho, A.F., Ikuta, C.Y., Ferreira Neto, J.S., Setubal, J. C., Heinemann, M.B., Guimarães, A.M.D.S., 2019. Global distribution and evolution of *Mycobacterium bovis* lineages. bioRxiv 642066.

Zwyer, M., Çavusoglu, C., Ghielmetti, G., Pacciarini, M., Scaltriti, E., Van Soolingen, D., Dötsch, A., Reinhard, M., Gagneux, S., Brites, D., 2021. A new nomenclature for the livestock-associated *Mycobacterium tuberculosis* complex based on phylogenomics [version 2; peer review: 2 approved]. Open. Res. Europe 1.