Bovine tuberculosis in a multi-host system in Northern Ireland: spatial distribution, molecular epidemiology and rapid diagnosis

Assel Akhmetova

BSc, MSc

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

School of Biodiversity, One Health & Veterinary Medicine

College of Medical Veterinary and Life Sciences University of Glasgow

February 2023

Abstract

Bovine tuberculosis (bTB) has a complex epidemiology and a wide range of host species. United Kingdom and Ireland are one of the examples where control of bovine tuberculosis (bTB) is complex and very desirable for the farming and economy sectors. Despite the increasing implementation of control measures, such as rigorous farms surveillance, control of animal movements, testing of individual animals, and effective contact tracing are implemented for reducing the bTB spread on a national scale as the prevalence of the disease in cattle herds remains high (9.95% in 2021) (DAERA, 2021). The presence of a wildlife reservoir, the Eurasian badger *Meles meles* and its possible role in bTB persistence and spread in cattle species make bTB epidemiology very complex. Test-and-vaccinate or remove (TVR) wildlife intervention study was implemented in years 2014-2018 to control bTB levels in badgers and prevent disease transmission to livestock in Northern Ireland. Chapter 1 included literature review about bovine tuberculosis, its pathogenesis and diagnosis in animals, summary of previous research in the field of detection and molecular epidemiology as well as recent control system of bTB in Great Britain and Northern Ireland. This Chapter also covered recent issues and research gaps in relation to bTB laboratory diagnosis tests and epidemiology and thesis outline.

Chapter 2 focused on *Mycobacterium bovis* spatial and molecular analysis of data from 1248 cattle and badger individuals within the TVR region (100 km^2 and 2 km buffer zone) south-eastern area in Northern Ireland with high bTB prevalence and high badger density (DAERA, 2018a). The study aimed to estimate the association between the spatial distribution of *M. bovis* multiple locus variable number of tandem repeats analysis (MLVA) types found in cattle and in badgers in the TVR area. Kernel density estimates of the shared MLVA types between cattle and badgers were evaluated to understand the spatial structure of the data. Contours in 95% density levels were estimated to demonstrate the spatial overlap of the main MLVA types, for each of the major strains. Two spatial models were built to assess the spatial distribution associations of *M. bovis* MLVA types in relationship to badger-cattle and cattle-badger transmissions. Thirty-seven MLVA types were found in cattle (n=31) and badger (n=6) species, with four of them (004, 006, 122 and 297) shared between two host species. Genotype 006 was identified as the most frequent and represented >51% *M. bovis* isolates; it was indicated as a

founder MLVA type for other genotypes using goeBURST algorithm. Strong association between spatial distribution of MLVA types in cattle and badgers was identified using kernel discriminant analysis (KDA).

Combining whole genome sequence (WGS) data analysis with the associated epidemiological metadata provided an opportunity for a thorough genomic epidemiological analysis of *M. bovis* transmission in the TVR area. These results were shown in Chapter 3 of current PhD thesis and aimed to explain the relative importance of within and between species transmissions for bTB persistence. In total 619 *M. bovis* isolates collected between years 1986 and 2018 from cattle and badgers were sequenced. From this dataset, previously studied endemic clade (MLVA 006 genotype) comprising of 302 isolates was used to study bTB transmission dynamics using Bayesian coalescent-based methods, Discrete Ancestral Trait Mapping (DATM) approach to reconstruct ancestral states of *M. bovis* collected from cattle and badgers, and *outbreaker2* software to reconstruct *M. bovis* transmission tree and outbreak reconstruction (Campbell et al., 2018).Estimated *M. bovis* substitution rate (mean 0.36-0.37 substitutions per genome per year) and most recent common ancestor (1970-1980s) was similar with other studies for bTB genomic epidemiology (Crispell et al., 2019, Salvador et al., 2019). Results obtained from the transmission trees reconstruction demonstrated high levels of between cattle transmissions and transmissions from cattle to badgers, and within badgers. The evidence of inter-species was also demonstrated in reconstructed phylogenetic trees. Overall, the results of this chapter showed that genetic and genomic *M. bovis* data obtained from historical isolates and the TVR intervention study can provide exceptional resolution for the genomic epidemiology of bTB, shedding light into the role of livestock and wildlife in the transmission of *M. bovis* in the region.

In Chapter 4, I developed a molecular bacterial load assay for rapid quantification of *M. bovis* directly from infected animals' tissues collected in Northern Ireland. Molecular bacterial load assay is currently used as a research method to monitor anti-TB therapy for human tuberculosis and quantifies *M. tuberculosis* bacterial load decline in response to treatment. I optimised this laboratory protocol on bTB-free bovine tissues spiked with *M. bovis* BCG and evaluated the performance of MBLA for bTB detection. This demonstrated that MBLA assay is efficient for bTB diagnostics in animal tissues. MBLA identifying *Mycobacterium tuberculosis*

complex (MTBC) specific ribosomal RNA can quantify viable *M. bovis* in animal tissues with the range of concentrations between 1.59E+09 CFU/ml to 1.68E+02 CFU/ml. MBLA was then applied for the analysis of 214 culture-positive bTB infected bovine tissue samples. Using MBLA, *M. bovis* was detected in 90% of cases and bacterial loads were reliably quantified in 73% of positive samples. Work with hazard group 3 highly contagious bacilli as *M. bovis,* requires the use of facilities with higher biosafety level (containment level 3, CL3) that are not always available in diagnostic and research laboratories. Therefore, I conducted heat inactivation experiments to quantify *M. bovis* BCG using both 16S rRNA and DNA to determine the possibility of working with infected tissues outside a CL3 laboratory, i.e. in a CL2 laboratory after inactivation. This showed that infected samples can be used for RNA-based techniques such as MBLA after heat killing of bacilli without any impact on mycobacterial load in clinical specimens. To enable the MBLA to be used in the field, I also tested the stability of *M. bovis* BCG RNA at room temperature for up to one month. I demonstrated that if tested RNA samples were stored and transported at room temperature within at least one month, it would not affect the quantification of mycobacteria. The results of the MBLA assay in Chapter 4 used for the infected bovine tissue samples collected within the TVR study area suggested that this molecular technique can be used as diagnostic method for rapid (results obtained within the same day as sample collection as opposed to weeks of culture testing) and sensitive detection of bovine tuberculosis directly in animal tissue samples.

Overall, these findings aided to our understanding of bTB transmission within cattle and badgers in an endemic area and offered tools for rapid molecular identification of *M. bovis* in this setting. The discussion of these results was presented in Chapter 5.

To my family, and in memory of my father

Table of Contents

List of Tables

List of Figures

Acknowledgement

The work within this PhD thesis was funded by the JSC “Center for International Programs”, “Bolashak” International Scholarship of the Republic of Kazakhstan.

I would like to express my special thanks to my collaborators Dr Robin Skuce, Dr Adrian Allen, Eleanor Breadon and all the laboratory team in AFBI, Northern Ireland for providing me this amazing data and helping in getting samples for experimental analysis, for their expert advice and encouragement throughout my PhD research project.

I would like to express my special thanks of gratitude to my first supervisor Professor Rowland Kao who gave me the golden opportunity to work on this amazing project and helped me in doing a lot of research and I came to know about so many new things.

I would like to express my gratitude and appreciation for my supervisors, Dr Katarina Oravcova, Dr Theo Pepler and Dr Liliana Salvador whose guidance, support and encouragement has been invaluable throughout this study. I would like to acknowledge the sincere efforts and valuable time given by my supervisors.

Also, I would like to thank my family and friends for their support. Without that support we couldn’t have succeeded in completing this project.

Specific contributions:

Chapter 2 Molecular typing (MLVA) data and spatial data for this analysis was provided by Agri-Food and Biosciences Institute in agreement with the Department of Agriculture, Environment and Rural Affairs (DAERA) of Northern Ireland. The information regarding the cattle herds and households is confidential; therefore, no personal data about farm identification numbers and their owners was shown in this analysis.

Chapter 3 Whole genome sequencing data and metadata used in this Chapter was provided by Agri-Food and Biosciences Institute in agreement with the Department of Agriculture, Environment and Rural Affairs (DAERA) of Northern Ireland.

Chapter 4 Tissue samples collected from cattle in NI was also provided by Agri-Food and Biosciences Institute in agreement with the Department of Agriculture, Environment and Rural Affairs (DAERA) of Northern Ireland. Special thanks to Eleanor Breadon who helped me processing these samples during my visit to AFBI in 2018.

Author's declaration

I declare that this thesis and the research contained within it is my own work unless otherwise stated, and no part of it has been submitted as part of any other degree or qualification. The work was carried from October 2016 to August 2022 at the University of Glasgow under the supervision of Dr Katarina Oravcova, Dr Theo Pepler, Prof Rowland Kao and Dr Liliana Salvador.

Definitions

ADC Albumin dextrose catalase enrichment used for isolation and cultivation of mycobacteria

AFBI Agri-Food and Biosciences Institute, Northern Ireland

BCG Bacille Calmettte Guerin: vaccine for tuberculosis (TB) disease

BEAST Bayesian Evolutionary Analysis Sampling Trees: program for Bayesian analysis of molecular sequences (Drummond and Bouckaert, 2015)

bTB Bovine Tuberculosis

DAERA Department of Agriculture, Environment and Rural Affairs in Northern Ireland

DATM Discrete Ancestral Trait Mapping

ESS Effective Sample Size

GB Great Britain, comprising England, Scotland and Wales

HPD Highest Posterior Density

IC internal control

IFN-γ Interferon-gamma serological test

KDA Kernel discriminant analysis

KDE Kernel density estimation

MBLA Mycobacterial load assay

MCC Maximum Clade Credibility

ML Maximum Likelihood

MLVA-VNTR multiple locus variable number of tandem repeats analysis

MTBC Mycobacterium Tuberculosis Complex

NI Northern Ireland

NND Nearest Neighbour Distance

OADC Oleic albumin dextrose catalase culture growth supplement

PBS Phosphate-buffered saline buffer

PCR Polymerase chain reaction

PPD Purified protein derivative

qRT-PCR Quantitative Reverse Transcription PCR

RBCT Randomised Badger Culling Trial

RFLP Restriction Fragment Length Polymorphism

RTA Road traffic accidents (for badgers killed on roads)

SICCT Single intradermal comparative cervical test

SIT Single intradermal test

SNP Single Nucleotide Polymorphism

tMRCA Time to most recent common ancestor

UK United Kingdom, comprising of Great Britain and Northern Ireland

WGS Whole Genome Sequencing

1 Chapter

Background to bovine tuberculosis in the United Kingdom and Ireland

1 Chapter

1.1 Introduction

Mycobacterium tuberculosis complex (MTBC) consists of a group of closely related bacterial species that can cause tuberculosis (TB) in humans and animals (Wirth et al., 2008). One of them is *Mycobacterium tuberculosis (M. tuberculosis)* that is a global threat for public health (Navin et al., 2002). According to the World Health Organization (WHO) (2020), in 2019 the global number of tuberculosis deaths caused by *M. tuberculosis* were about 1.4 million. Another species, *Mycobacterium bovis (M. bovis),* is responsible for the occurrence of bovine tuberculosis (bTB) in cattle and other animals and has the widest host range from other MTBC species (Karlson and Lessel, 1970, Thoen et al., 2006). Other prominent members of MTBC group, such as *M. canetti, M. africanum*, *M. caprae, M. pinnipedii, M. suricattae* and *M. microti* are known by different phenotypic characteristics and host ranges (Gutierrez et al., 2005). *M. canetti* and *M. africanum* are causing human tuberculosis in Africa, *M. pinnipedii* is causing infection in seals and *M. caprae* in ruminants, *M. suricattae* in meerkats and *M. microti* in voles. All members of MTBC are >95% genetically similar and difficult to differentiate (Wirth et al., 2008).

Bovine tuberculosis is an infectious disease of cattle and many other domestic and wild species that can occasionally be transmitted to humans (O'Reilly and Daborn, 1995). Commercial milk pasteurization helped to successfully eliminate zoonotic tuberculosis in many developed countries (Palmer and Waters, 2011). However, being one of the most serious problems for animal health and the farming sector on a global scale, bovine tuberculosis is also associated with wildlife animals. The most prominent of them are European badgers (*Meles meles*) in the United Kingdom and Ireland (Delahay et al., 2001), Eurasian wild boar (*Sus scrofa*), red deer (*Cervus elaphus*) and fallow deer (*Dama dama*) in Iberian Peninsula and Western Austria (Gortazar et al., 2012, Nigsch et al., 2018), brushtail possums (*Trichosurus vulpecula*) and ferrets (*Mustela furo*) in New Zealand (NZ) (Coleman and Cooke, 2001), white tailed deer (*Odocoileus virginianus*) in Michigan (USA) (Payeur et al., 2002) and African buffalo (*Syncerus caffer*) in South Africa (De Vos et al., 2001). Currently used test-and-slaughter programs for control of tuberculosis infection in cattle are successful in many countries. These measures

combined with post-mortem examination of carcasses at abattoirs, rigorous farm surveillance, animal movement controls, testing of individual animals and effective contact tracing are implemented for reducing bTB spread on a national scale (Humphrey et al., 2014, Livingstone et al., 2015, Godfray et al., 2018). The UK and Ireland are two examples where eradication of bTB is difficult though desirable for the farming sector and economy. It costs the United Kingdom (UK) government about £200 million yearly (Godfray et al., 2018, Northern Ireland Audit Office, 2018) to control *M. bovis* infection in livestock and wildlife animals. UK badger numbers make up about a quarter of the global badger population (Reid et al., 2012, O'Mahony, 2015). However, the role of badgers in bTB transmission is still controversial and challenging for researchers and policy makers. A study by Brooks-Pollock et al. (2014) suggests that current control strategies, together with several factors such as movement of domestic animals, exogenous sources of the infection, and low sensitivity of current diagnostic tools are not enough to reduce annual incidence of bTB.

This literature review describes *M. bovis* pathogenesis and diagnosis in cattle and wildlife species, discusses the current molecular techniques available for fingerprinting of *M. bovis* strains, and methods for investigation of the disease transmission pathways and dynamics used in recently study. Moreover, it outlines the drawbacks and future potential of *Mycobacterium bovis* genome associated studies for molecular epidemiology of bTB.

1.1.1 Pathogenesis of bovine tuberculosis

Mycobacterium bovis has a wide range of domestic hosts such as cattle, pigs, cats and wildlife species such as deer, elk, possums, badgers and others (Fitzgerald and Kaneene, 2013). Infection in animal species usually occur through direct contact of animals, or indirectly by ingestion of contaminated substances (Neill et al., 1994, Menzies and Neill, 2000). However, the number, consistency and size of inhaled organisms play a crucial role in disease occurrence in individuals and a very small amount of aerosol droplets can be enough to initiate lesions (Neill et al., 2001a). First experimental studies by Dean et al. (2005) and Johnson et al. (2007) were conducted to evaluate the minimum dose of *M. bovis* needed for cattle to become infected and symptomatic, and it was shown that only 1 CFU of the bacterium was enough to develop pulmonary tuberculosis in half of the tested

cattle. Infected animals were positive for the tuberculin skin test and the interferon-gamma test, developing granulomas in lymph nodes of the thoracic region. Another study by Buddle et al. (1994) showed that a very low dose (20 CFU) was sufficient for brushtail possums to become infected and develop granulomas in lungs and lymph nodes. These findings suggest that direct contact between animals and respiration of *M. bovis* play important roles in disease spread.

Once *M. bovis* enters a host, it multiplies very slowly and it might take months to years before the host shows any clinical signs (Neill et al., 2001). For this reason, many diagnostic tests cannot detect bTB in the early stages of the disease. This is also because the immune reaction of the organism is minimal initially; however, at a later stage, when granulomas start developing in organs, mainly in lungs and lymph nodes, the immunological response intensifies and disease can be observed in most cases (Neill et al., 1994; Neill et al., 2001b). It is important to point out that, tuberculosis can be spread in animals without showing any symptoms for a long period of time (Pollock et al., 2001, Philips and Ernst, 2012).

Indirect transmission of *M. bovis,* such as ingestion of bacteria from shared contaminated food, water or soil, affects mainly cattle mesenteric lymph nodes (Menzies & Neill, 2000; Domingo et al., 2014). Genital or mammary infections are possible but not very common in most developed countries (Domingo et al., 2014). Some studies that examined cattle carcasses and meat at slaughterhouses and private abattoirs identified lesions distributed across different parts of the body; however, mostly in lungs, head and chest lymph nodes and less in liver and abdominal lymphatic system (Corner et al., 1990). Although many previous studies underscored the importance of understanding infection routes in the past decades (Pritchard, 1988; Neill et al., 1994), there are many factors that can influence the pathogenesis of bTB in host animals, such as immune response, age, breed, sex. The surrounding environment can play an important role for the survival of *M. bovis* bacilli outside the animal bodies and can be a route for indirect transmission of the bacterium.

The mechanisms of immune response to *M. bovis* infection varies in different host species. The latency of *M. bovis* in cattle was previously described in research papers (Alvarez et al., 2009, Ncube et al., 2022), and our understanding of the stages of the infection progression in animals are mostly based on experimental

studies with domestic (Waters et al., 2010) and wild animals (Gormley and Corner, 2017). Therefore, knowledge of these mechanisms is crucial to understand how bTB develops in host species and can help to improve the disease control schemes. Mycobacterium is micro-aerophilic bacterium and requires controlled levels of oxygen for its survival and growth than what is present in the atmosphere (Moore and James, 1982), therefore it can survive in many other organs in the animal bodies (not only in the lungs). The ability of mycobacteria to survive for a long time in macrophages cause chronic tuberculosis or long infections (Flynn et al., 2011). Previous research by Pollock et al. (2001) and Cassidy et al. (2001) suggest that T-cells dominate in early response to intracellular bacteria and cell-mediated immune response (CMI) plays a major in building the immunity to *M. bovis*. Thus T-cells have a protective function but are also responsible for lysis of bacterial macrophages and formation of lesions (Pollock et al., 1996, Cassidy et al., 1998, Liebana et al., 2000). They secrete cytokines such as tumour necrosis factor-alpha (TNF-α) and interferon-gamma (IFN-γ) that are beneficial for the development of granulomas and cause activation of macrophages and their anti-bacterial defence (Wallis et al., 1992). Knowledge of protective processes underlying complex mycobacterial infection immunity can be helpful in vaccine design and may provide opportunities for identification of suitable diagnostic markers.

1.1.2 Practical and laboratory diagnosis

Ante-mortem diagnoses of *M. bovis* in cattle using only clinical or microbiology methods, although challenging, is possible based on the detection of cell-mediated immune response (CMI) initiated by T-cells (Pritchard, 1988). First tuberculin was discovered by Koch (1890) in 1890, he discovered the possibility of using it as a diagnostic test to identify infected TB patients. In some cases, patients induced symptoms like high temperature, sickness and local skin reaction. These findings were useful in developing effective bTB diagnostics such as the tuberculin “skin” test and IFN-γ test that has been successfully used in many developed countries (de la Rua-Domenech et al., 2006) .

As earlier described by Monaghan et al. (1994) there are two “skin” tests currently used on animals, namely the single intradermal test (SIT) and single intradermal comparative cervical test (SICCT). The SIT screening test is used in mainland Europe, USA and New Zealand as a routine diagnostic test for bTB in individual

animals and herds (de la Rua-Domenech et al., 2006). Both tests include injection of purified protein derivatives (PPDs) of *M. bovis* (and with *M. avium* for SICCT) intradermally in animals' neck. The SIT includes injection of purified protein derivative (PPD) of *M. bovis* intradermally in the animals' neck. The increase in skin injection site after 72 hours will be inspected and if the difference of skin thickness between before and after the tuberculin injection is > 4 mm then the result is positive (Monaghan et al., 1994). SICCT is used in the UK and Republic of Ireland (RoI) for bovine tuberculosis diagnosis in livestock animals and consists of the injection of both *M. bovis* (AN5) and *M. avium* (D4ER or TB56) strains as comparative diagnostics into on the same side of animals' neck. The injection sites (upper for avian and lower for bovine PPD) are usually observed after three days (72 hours) for any inflammatory reaction. If the difference of the reaction (injection site) between bovine tuberculin and avian tuberculin injections is > 4 mm then the animal is considered positive for bovine tuberculosis; in some cases individuals can have 'inconclusive' results when the bovine measure of skin thickness is bigger than the avian one, however when the difference of the skin thickness is < 4 mm, but not negative, animals should be isolated and sent for re-testing (Pollock et al., 2001). Official bTB herd status in NI is given after the annual skin testing and the post-mortem examination (if animals were sent for human consumption) results are known (Northern Ireland Audit Office, 2018). Herds can be assigned officially tuberculosis free (OTF) status, officially tuberculosis free status suspended (OTS), or officially tuberculosis free status withdrawn (OTW).

Sensitivity and specificity are the main measures used to evaluate the tests. Sensitivity is the proportion of infected animals correctly identified by the test, whereas specificity is the proportion of non-infected animals correctly identified as bTB negative. Based on the Great Britain bTB control program, the sensitivity of SICCT was estimated to be between 75.0-95.5% and specificity between 78.8-100% (reviewed in Rua-Domenech et al., 2006). Some studies point out that there are factors that can potentially demonstrate skin test false negative results in cattle, such as when the test was introduced right after the previous tuberculin test or too soon after the animal was infected, *M. bovis* co-infection, vaccination against Johne's disease, if an animal was given some steroids or other drugs and was under the nutritional stress (Buddle et al., 2009, Buddle et al., 2015). Other factors relating to the type of tuberculin used or the conditions of the test application were also listed (de la Rua-Domenech et al., 2006). Therefore, the

tuberculin “skin” test alone is not effective for routine bTB diagnostics, which is why complementary IFN-γ tests and serological tests were introduced for ante-mortem confirmation of tuberculosis in cattle (de la Rua-Domenech et al., 2006).

Post-mortem identification of lesions and further laboratory confirmation are important parts of the diagnosis. However, in some cases animals show no visible lesions (VL) and no *M. bovis* culture growth despite being tuberculin test positive. O'Hagan et al. (2016) evaluated risk factors that were associated with bTB reactor cattle by SICCT in Northern Ireland and presence of visible lesions confirmed in the laboratory and found that 97.5% of diagnostic test results were consistent with the presence or absence of lesions. However, detection of visible lesions and laboratory techniques for *M. bovis* detection can differ between countries/abattoirs. Culturing *M. bovis* is more likely when isolates are extracted from visible granulomas than from non-visible lesions (NVL) (World Health Organization (WHO) et al., 2017). This suggests that a combination of ante-mortem and post-mortem inspections are needed for thorough diagnosis of bovine tuberculosis.

McCallan et al. (2021) described a wide range of serological tests for detection of antibody response against *M. bovis* infection in cattle. However, the low sensitivity and specificity of the tests pointed to the need for a rapid and sensitive diagnostic test such as γ-IFN assay. This *in vitro* immunoassay, based on detection of the host’s CMI response to the *M. bovis* antigen, was introduced as the commercial test kit, Bovigam (Gormley et al., 2006). It measures the amounts of cytokine IFN-γ produced by T-lymphocytes after incubation of blood samples for 16-24 hours with a specific antigen, using an enzyme immunoassay method. After Australia accredited the γ-IFN assay for diagnostic use in early 1990s, it was also successfully used in Ireland, Northern Ireland, USA, Italy, Spain and Brazil for testing bTB in cattle (Wood and Jones, 2001). These studies demonstrated increased sensitivity (median 96.6%) of the new approach and in some cases detecting early infection of bTB (de la Rua-Domenech et al., 2006).

Detecting *Mycobacterium bovis* before animals show any clinical signs is very important for timely isolation of infected animals from the population to limit the potential spread. This requires diagnostic techniques that are robust and cost-effective for routine use on individuals and whole herds. SIT and SICCT are

powerful for early diagnosis of the infection and is beneficial for bTB control, whereas $\gamma$the IFN-γ test is used as ancillary test to identify missed infected animals. This scheme is standard in the UK and many other European countries (de la Rua-Domenech et al., 2006).

Bacteriological culture and isolation of *M. bovis* is the laboratory 'gold standard' method for the confirmation of this bacterium. This method is used to confirm the existence of *M. bovis* from suspect animal tissues found during the post-mortem examination (Gormley et al., 2014). The bacilli are recovered efficiently from granulomas or necrosis in organs, which are mostly found in lungs and pulmonary lymph nodes in cattle (Corner et al., 1990). *M. bovis* are microaerophilic bacteria that grow at 37ºC and have a long generation time. Cultivation and isolation of mycobacterium was studied in various types of growth media, such as in Löwenstein–Jensen egg-based medium, enriched Middlebrook 7H10 and 7H11 agar, tuberculosis blood agar, and liquid Middlebrook 7H9 broth (Corner et al., 2012a, Gormley et al., 2014). The combination of solid and liquid media for *M. bovis* isolation is a widely used approach to increase the sensitivity of its confirmation. However, the sensitivity and specificity of this technique are not always 100%, and it usually takes 3-6 weeks to grow the mycobacteria. In the absence of clinical signs and visible lesions, the detection of mycobacteria can be more challenging.

Although bacteriological culture techniques are used for *M. bovis* confirmation with high sensitivity, there are several limitations to this method. Direct culture confirmation technique requires specific laboratory facilities (BSL3), a long time for the bacteria to grow (up to 2 months), and a relatively high cost. Moreover, successful recovery of *M. bovis* from collected tissues can be influenced by different factors, such as the type of specimens, the conditions of sample storage, transportation, preparation and decontamination, and the selection of the culture medium and growth supplements. In fact, the optimal pre-processing of the samples and decontamination of the tissue specimens is very important for the direct detection of *M. bovis* (Corner et al., 2012a). The growth of contaminating micro-flora can be solved with the use of appropriate detergents, however, some of them may contain toxic chemicals (Corner et al., 2012a).

Less time-consuming and less costly method is smear microscopy, which is used for direct detection of *M. bovis* in tissue sections and smears with acid-fast staining using Ziehl-Neelsen (ZN) staining method (Cook, 1997). However, it is not very sensitive (Gutierrez and Garcia, 1993) and cannot differentiate between viable and non-viable bacteria.

The introduction of automated systems for *M. tuberculosis* growth such as 26 radiometric Bactec TB 460 and MGIT 960 made isolation of mycobacteria less time-consuming but require specific equipment and disposal of hazardous waste (Chien et al., 2000). Therefore cost-effective, reliable tools for *M. bovis* detection and confirmation in a shorter amount of time are crucial for effective bovine tuberculosis control programs.

Mycobacterial Bacterial Load Assay (MBLA) is culture-free molecular technique that was used for detection and quantitation of viable *M. tuberculosis* bacilli in sputum samples from tuberculosis infected patients. It was first introduced in Honeyborne et al. (2011) and described by Gillespie et al. (2017) as an early biomarker for monitoring human TB treatment. This polymerase chain reaction technique (PCR) uses mycobacterial 16S ribosomal RNA (16S rRNA) to detect and quantify the mycobacterial loads directly from human sputum (Honeyborne et al., 2011). The use of mycobacterial 16S rRNA was introduced as a marker of bacterial viability, as present in cells in multiple copies and has a shorter half-life comparing to DNA or other RNA types. MBLA assay was not used for the direct detection and quantitation of *M. bovis* from animal tissue samples before current research study described in Chapter 4 of this thesis.

Detection of *M. bovis* DNA was widely demonstrated in research studies that use different sample types. The IS6110 gene was commonly evaluated as a proposed PCR-target for *M. bovis* diagnostics, as specific for MTBC species (Thacker et al., 2011). Moreover, direct DNA extraction was shown possible from various types of samples, such as animal tissues, nasal swabs and milk. However, extraction efficiencies were identified for detection of mycobacteria in some studies. In a study by de Souza Figueiredo et al. (2010) *M. bovis* DNA was extracted from nasal swabs from cattle and PCR performed targeting the *RvD1-Rv2031c* and *IS6110* sequences. *M. bovis* was detected only in 2 out of 34 samples, with 5.9% sensitivity of the method. These results might be explained by the low mycobacterial loads contained in nasal swabs. A better performance of DNA-based *M. bovis* detections

obtained from animal tissues, such as lungs and lymph nodes, where the bacteria is known to be localized in tuberculous granulomas. Zarden et al. (2013) demonstrated the 78.3% sensitivity of the DNA based PCR.

The use of *M. tuberculosis* RNA as a target for PCR amplification was also used in some studies (Therese et al., 2012, Montenegro et al., 2014), however less is demonstrated for the direct detection of *M. bovis* RNA. Studies based on detection of 16S rRNA were mostly used for differentiation between tuberculosis and non-tuberculosis mycobacteria (Miller et al., 2002, Quan et al., 2017).

Quantitative PCR is a method for estimation of number of copies of a specific DNA/RNA target, which is very useful for application in samples with low copy number of target genes. Little is known about the quantitation of mycobacterial DNA and RNA from animal tissues. Costa et al. (2013) showed the semi-nested PCR method that detected and quantified *M. bovis*, as low as 1×10^0 bacilli/ml, with the overall sensitivity and specificity of 98.2% and 88.7%, respectively. In a different study by Young et al. (2005) 16S rRNA based PCR quantified 1×10^3 to 3.6×10^3 gene copies/g in soil and environmental samples.

The presence of *M. bovis* in environmental samples, such as soils and animal faeces, was previously shown using bacteriological culture methods (Romanowski et al., 1992). However, these methods require decontamination of the samples, which usually are very sensitive, and long incubation time in the culture media (Romanowski et al., 1992). The molecular detection of *M. bovis* directly from clinical specimens is mostly based on DNA targets. Young et al. (2005) performed an experimental study to investigate the presence and survival of mycobacterial DNA and 16S rRNA in the soil microcosms and in environmental samples near known badger setts. The results showed that *M. bovis* can persist in soils for more than 15 months, and that bacterial DNA and 16 S rRNA can be successfully detected using RT-PCR methods. This study also provided further evidence of the presence of viable cells in environmental samples.

1.1.3 Molecular epidemiology of M. bovis

The molecular epidemiology of *M. bovis* has been crucial in understanding the cause of infection and track the infection pathway in the population. Molecular typing techniques can be used to identify the pathogens, reconstruct their

evolutionary history and understand their spread. The understanding of *M. bovis* transmission dynamics and the tracing back of the infection to its ancestor were achieved with the integration of specific molecular typing techniques, some of which will be reviewed in this study (Table 1.1). Reconstructing the evolutionary history of *M. bovis* strains and their transmission patterns is important to gain insights into "who infected whom" and inform control strategies.

1.1.3.1 Spoligotyping

Spacer oligonucleotide typing (spoligotyping) is a polymerase chain reaction (PCR) based method, which includes amplification of a polymorphic direct repeat locus (DR) in mycobacterial chromosome (Hermans et al., 1991, Groenen et al., 1993, Kamerbeek et al., 1997). It is widely used for *M. tuberculosis* and *M. bovis* genotyping, based on the presence/absence of unique 35-41 bp spacer sequences in the DR. For standardization of the method, 43 spacers were selected by Kamerbeek et al. (1997) and introduced as the spoligotype pattern (ordered from 1 to 43 spacers) by Groenen et al. (1993). Later, Smith and Upton (2012) developed a website www.Mbovis.org with a full dataset of available *M. bovis* spoligotypes for convenient use on international level, which allows researchers to conduct inter-laboratory comparison studies.

Spoligotyping requires less mycobacterial DNA and is generally less time-consuming compared to IS*6110* genotyping (Kremer et al., 1999). Combining several typing methods for strain differentiation has greater resolution and reliability (Kremer et al., 1999), such as mixing spoligotyping and Variable-Number Tandem Repeat Typing (VNTR) described below. For efficient control of bTB, molecular methods for rapid and reliable identification and typing are required.

1.1.3.2 Multiple locus variable number of tandem repeats analysis (MLVA-VNTR)

Molecular typing of *M. bovis* strains based on PCR detection of variable-number tandem repeats (VNTRs), a short sequence that vary in number of copies in the bacterial genome, is currently used for *M. bovis* genotyping. Frothingham and Meeker-O'Connel (1998) introduced 11 tandem repeat loci, from which five major polymorphic tandem repeat (MPTR) and six exact tandem repeat (ETR) loci are useful for MLVA-VNTR typing of MTBC strains. Supply et al. (2000) demonstrated the presence of specific elements called mycobacterial interspersed repetitive

units (MIRUs) in *M. tuberculosis* genome that are 40±100 bp in size and show variations in copy numbers. Current *M. tuberculosis* genotyping is based on the use of 12 MIRU-VNTR loci (Supply et al., 2000, Supply et al., 2006). A panel of MLVA-VNTR markers for *M. bovis* isolates typing have been demonstrated by Roring et al. (2002) and other research groups. A comparison of different VNTRs to determine the loci for typing of *M. bovis* isolates were observed by Smittipat and Palittapongarnpim (2000) and Skuce et al. (2002). These studies demonstrated a relatively high discriminating power of the MLVA-VNTR typing method for *M. bovis* and positive correlation with spoligotyping results, where seven loci were determined as suitable for *M. bovis* genotyping in NI (Skuce et al., 2002, Skuce et al., 2005).

The stability of genetic markers and high discriminating power of MTBC typing showed the potential to use MLVA-VNTR typing for inter-laboratory testing, where the same loci are used for analysis (Hilty et al., 2005, Mathema et al., 2006). Homoplasy is known to occur in some MLVA-VNTR types and refers to the evolution of the repetitive regions within the *M. bovis* genome that can occur independently for different genotypes at the same time, therefore makes it difficult to infer phylogenetic relationships between different MLVA-VNTR types (Reyes et al., 2012). Various combinations of tandem repeats were characterized for molecular typing of *M. tuberculosis* complex isolates and varying discriminatory power for different loci. Thus, it is important to select the combination of the VNTR loci that are most relevant for specific geographic locations and MTBC member strains (Skuce et al., 2005).

This high-throughput and highly discriminating technique (Skuce et al., 2005; Allix et al., 2006) is currently widely used for *M. bovis* genotyping and transmission dynamics inference. Spoligotyping and MLVA profiling for *M. bovis* has shown important inferences for bTB epidemiology, where specific genotypes tend to cluster in space in GB and NI indicating local transmission patterns (Skuce et al., 2010, Skuce et al., 2020). Combining these data with surveillance of wildlife reservoirs and contact tracing of cattle can be helpful to identify infection transmission patterns.

1.1.3.3 Whole-genome sequencing

In recent years, diagnostic microbiology and pathogen epidemiology has made increasing use of sequencing technologies. This includes identifying outbreaks, tracing the infection sources and dynamics of the spread. Whole genome sequencing is a complex technique that involves DNA extraction from bacterial culture, library preparation, sequencing of DNA and interpretation of the results using bioinformatics pipelines (Sanger et al., 1977, Illumina, 2017). The first available whole genome sequence of *M. bovis* was collected from cattle in England in 2003 and is currently used as reference genome (AF2122/97) (Garnier et al., 2003).

M. bovis studies mainly used Illumina sequencing, which sequences small areas of the chromosome hundreds of times simultaneously and produces an almost fully covered bacterial genome (Illumina, 2017). The method is called "sequencing by synthesis" where DNA molecules attached to the flow cells are amplified and then sequenced using fluorophore-labelled nucleotides (Illumina, 2017). This produces large amounts of data, which allows in-depth analysis of the given organism in a short amount of time. The raw data can be assembled *de novo* into contiguous sequences (contigs) or aligned to the reference genome. There are many software packages and automated bioinformatic pipelines developed for accurate assembly of the genomes (Sukumar et al., 2021). Points of genetic variations in the genome are called single nucleotide polymorphisms (SNPs). There are also available bioinformatics pipelines that have been developed to extract exact spoligotypes from WGS data (*in silico* spoligotyping) (Bogaerts et al., 2021).

Table 1.1 Molecular typing techniques used for differentiation of *M. bovis* strains (Mathema et al., 2006, Kao et al., 2014)

Typing method	Description	Limitations
Spoligotyping	Based on detection of spacers in direct repeat (DR) region of the mycobacterial chromosome PCR and hybridization-based technique	Lower discriminatory power than MLVA-VNTR typing Targets only single DR region Used for identification and differentiation of *M. bovis* strains

	Can be used to distinguish between some MTBC strains Can be used for cross laboratory comparisons, results shown as binary codes (presence/absence of spacers) An International spoligotyping database is available to use (www.Mbovis.org)	In combination with MLVA-VNTR typing can be used for molecular epidemiology studies, identifying phylogenetic relationships, transmission dynamics
Multiple locus variable number of tandem repeats analysis typing (MLVA-VNTR)	Based on differences in number of tandem repeated regions (~100 bp) in *M. bovis* genome Higher resolution than spoligotyping Good for *M. bovis* strain typing Initially characterized 41 loci for *M. tuberculosis* Only 8 main loci are used for *M. bovis* identification in NI	Discriminatory power depends on number of loci used Various sets of loci can be used for different geographical regions Each locus has specific molecular clock Used for *M. bovis* molecular epidemiology studies and evolutionary investigation studies
SNP analysis	High resolution Currently widely used in *M. bovis* Bioinformatics pipelines developed to extract exact spoligotypes from WGS data (*in silico* spoligotyping)	Utilization for *M. bovis* isolates differentiation and study their phylogenetic relationships Can be used for evolutionary history reconstruction and transmission dynamics studies in various host-pathogen interactions

1.1.4 Use of WGS technologies for M. bovis epidemiology

The *M. bovis* genome is about 4.3 Mb in size and has approximately 4200 genes, including Ribonucleic acid (RNA) molecules. It has high guanine-cytosine (GC) content and a high number of repetitive regions throughout the genome (Garnier et al., 2003). Most MTBC genomes contain proline-glutamate (PE) or proline-proline-glutamate (PPE) gene families and their role is still unclear, although some researchers associate these groups with pathogenesis of TB (Garnier et al., 2003, Delogu et al., 2017).

A series of recent studies have indicated the successful use of WGS in epidemiological investigations (Biek et al., 2012, Roetzer et al., 2013, Kao et al., 2016, Reis and Cunha, 2021), specifically in the course of outbreak. DNA sequencing technologies are becoming cheaper for studying bacterial genomes in a shorter time and higher resolution. Examination of complete genomes of bacterial strains can provide a greater insight into identification of their phylogenetic relationships. Differences in bacterial genomes can occur as point mutations, mostly SNPs or single nucleotide insertions/deletions, and genetic diversity due to homologous recombination (Caugant et al., 1981, Liu et al., 2006, Didelot and Maiden, 2010). However, the rates of bacterial genome evolution and levels of recombination vary from species to species; as an example, *M. bovis* is a slowly evolving pathogen (Hauer et al., 2015, Guimaraes and Zimpel, 2020) with small number of recombination events (Reis and Cunha, 2021), in comparison to other bacteria (Biek et al., 2015). Substitution rates for *M. bovis* (as nucleotide substitutions per genome per year) were described in several studies, in Trewby et al. (2016) as 0.20 (95 % HPD (height posterior density) 0.10-0.30 substitutions per genome per year), in Crispell et al. (2019) as 0.28 (HPD 95% 0.21-0.37 substitutions per genome per year), and in Salvador et al. (2019) as 0.37 (HPD 95% 0.24-0.51). While the mean values of the substitution rates across studies are slightly different, there is a large overlap of the 95% HPD intervals between the estimations from the different studies. Because reverse mutations are rare, the character of occurred mutations should theoretically allow study of the transmission direction throughout the time of the outbreak (Walker et al., 2013).

Walker et al. (2013) determined possible thresholds to assess transmission links between isolates based on their genetic distances, where five or less, but no more

than twelve SNPs was found likely. A large number of existing studies in the broader literature examined various approaches to study genomic data for molecular epidemiology purposes. Phylogenetic analyses use bacterial sequencing data to study interrelationships of isolates based on their evolutionary history (Brosch et al., 2002, Rodriguez-Campos et al., 2012, Stucki and Gagneux, 2013) and make assumptions about transmission patterns among isolates. It is challenging to reconstruct transmission links based only on genomic data, but combining metadata about host species, dates of infection and contact tracing is helpful for accurate epidemiological investigation studies (Kao et al., 2014, Kao et al., 2016).

The use of WGS for *M. bovis* isolates was introduced by Biek et al. (2012) showing important spatio-temporal links between isolates from infected cattle and badgers of the same VNTR type in NI. This study was the first to demonstrate the evidence of the genetic relationships of *M. bovis* in multi-host system in NI and the estimated mutation rate was 3.40 (CI: 0.87–5.93) × 10^{-8} substitutions per site per year (Biek et al., 2012). Trewby et al. (2016) then targeted the single VNTR-10 type to infer transmission dynamics within the cattle population of NI and investigate the correlation of *M. bovis* genetic and spatial distances in the presence of cattle movements. The genetic variance was low within VNTR-10 isolates, however both cattle and badger isolates were found in close spatially (<1.5 km). The evolutionary rate estimated using Bayesian analysis was 0.2 substitutions per genome per year and most recent common ancestor (MRCA) identified from the 1970s (Trewby et al., 2016). These studies suggest that WGS data add significant insights for understanding bTB epidemiology in a complex multi-host system.

Glaser et al. (2016) analysed *M. bovis* isolates collected from infected cattle and white-tailed deer in Minnesota to investigate the association of *M. bovis* genetic diversity and its spatial distribution in both host species. The authors found that although similar genotypes (spoligotyping and VNTR) were circulating in both cattle and deer, WGS of *M. bovis* strains isolated from cattle and deer did not indicate strong association between the genetic and spatial distribution of *M. bovis* among host species.

Crispell et al. (2017) studied the use of WGS to understand *M. bovis* transmission in cattle and brushtail possums in NZ. The study addresses the role of the brushtail

possums in the persistence of bTB in cattle and estimated the substitution rate of the *M. bovis* population as 0.53 substitutions per genome per year. Using WGS and spatio-temporal data, Salvador et al. (2019) evaluated the role of wildlife species as reservoirs of infection. The *M. bovis* evolutionary rate was estimated as 0.37 substitutions per genome per year, and evidence of *M. bovis* inter-species transmissions among cattle, deer and elk in Michigan was shown. Later Crispell et al. (2019) studied bTB genomic epidemiology in a densely sampled Woodchester park area using phylogenetic Beast2 approach and demonstrated the presence of bi-directional *M. bovis* transmissions among studied cattle and badgers, with much higher badger-cattle transition rates than cattle-badgers transitions. In a similar study by (Rossi et al., 2022) investigated bTB phylodynamics in Cumbria, and estimated higher rates of transmission from badgers to cattle than vice-versa. More recent study by van Tonder et al. (2021) inferred *M. bovis* transmission links between cattle and badgers from randomised badger culling trial area and showed the importance of both cross-species and within species bTB transmissions.

Reconstructing phylogenetic relationships of the isolates in a sampled population, following estimation of evolutionary rates, can indicate the evolutionary history of the pathogen, explain introduction of the infection in time and indicate inter-species transmissions (Pybus et al., 2012, Kao et al., 2014, Patane et al., 2017). The substitution rates of *M. bovis* and *M. tuberculosis* vary across different research studies from 0.15-0.5 events per genome per year for different countries and host species (Trewby et al., 2016; Biek, O'Hare, Wright, Mallon, McCormick, Orton, McDowell, Twerby, et al., 2012; Bryant et al., 2013; Walker et al., 2013; Crispell et al., 2017). Despite the broad range of substitution rate estimates, it is important that WGS increase the resolution of the studied data and allows for greater understanding of bTB dynamics, and investigation of the role of inter-species transmission (Kao et al., 2014, Hatherell et al., 2016).

1.1.5 Control of bTB in Great Britain and Northern Ireland

Bovine tuberculosis caused by *M. bovis* is recognized as a serious problem in many countries, particularly in Europe, North, Central and South America, Africa, and Oceania (Humblet et al., 2009, World Health Organization (WHO) et al., 2017, Garcia et al., 2021, Perea et al., 2023). In the United Kingdom, every year about £100 million is spent by the government to control bTB infection. Moreover, these

expenses are increasing every year and it is crucial to achieve an efficient bTB control plan. Although the UK has developed a strategy for bTB control, the infection still occurs and is continuing to increase (Godfray et al., 2018). Current strategies include disease monitoring of domestic animals with skin tests and movement restrictions, followed by slaughterhouse inspection. Furthermore, in Britain, failure of the bTB control program has been linked to the disease prevalence in badgers, and transmission to local cattle (Delahay et al., 2001, Donnelly and Nouvellet, 2013).

The national control policies in Great Britain and Northern Ireland from 1935 onwards were based on a test-and-slaughter system, which reduced bTB incidents in cattle to a very low level. However, disease prevalence in cattle increased again in the mid-1980s (Woodroffe et al., 2009, Robinson, 2015) . One of the confirmed factors of bTB presence in animals in Britain can be explained by contacts between cattle and wild animals (Humblet et al., 2009). Further examinations confirmed a high prevalence of the disease in badgers compared to other species of wild animals (Delahay et al., 2001). Therefore, infection control through badger culling had been carried out in the UK under different schemes between 1973 and 1998 (Jenkins et al., 2007). Badger and cattle movements are another major risk of disease transmission and played an important role in *M. bovis* infection spread in the area. Adjacent cattle grazing lands and unrecorded movements of cattle are not in particular related to interspecies transmission between badger and cattle. The Randomised Badger Culling Trial, or RBCT, was introduced in Great Britain as a badger culling field trial to lower the risk of bTB incidence in cattle and assess the effectiveness of the approach (Donnelly et al., 2006). Later researchers showed that reduction of badger populations using the culling strategy, which covered large territories (100 km^2), was effective within the culling areas and decreased the infection presence in cattle, but a “perturbation effect” was seen in 2 km surrounding areas, where infection increased in badgers and the risk of contacts between badger and cattle increased (Woodroffe et al., 2009). Hence, such extensive culling tactics were considered time-consuming, expensive and counterproductive.

In Northern Ireland, annual monitoring of bTB using the SICCT plays an essential role in disease control. Here control is particularly important, due to the high value of beef and dairy farms for the agricultural sector (Abernethy et al., 2006).

Animals showing positive reaction to the skin test are called reactors, and these animals are subsequently slaughtered (Northern Ireland Audit Office, 2009). In some cases, the antemortem examination is conducted using the γ-IFN test; slaughtered animal carcasses (after having skin test positive results) are visually examined in the abattoir, after which tissue samples are taken for laboratory confirmation using bacteriology methods. Here disease control methods aim to eradicate the infection agent through testing and culling, control of transmission routes and the establishment of double fencing boundaries between neighbouring herds.

According to the report of the Northern Ireland Audit Office (NIAO) (2009), there is a higher risk of disease spread from local sources compared to imported animals, which include markets and illegal movement of animals. Good biosecurity measures in individual herds can be effective in preventing disease transmission between neighbouring farms, which can be cost-effective for herd holders. However, movement of cattle is still considered as one of the major risks of bTB spread within and outside of the herds where the movement of undetected infected animals also takes place (Milne et al., 2020, Skuce et al., 2020). In 2004 the EU therefore recommended pre-movement testing for cattle that for some reason missed the annual testing, to mitigate infection transmitted through animal movements (Northern Ireland Audit Office, 2009).

The role of European badgers in disease transmission in Northern Ireland has been described in several independent research studies (Biek et al., 2012, Trewby et al., 2016). Investigations showed that about 20% of badgers killed by vehicles were infected by *M. bovis* (Abernethy et al., 2006). According to the Department of Agriculture, Environment and Rural Affairs statistics (DAERA, 2021) in Northern Ireland (NI) bTB herd incidence in 2021-2022 was 9.95%. In order to obtain more evidence of bTb distribution in wildlife species, a new study called “test and vaccinate or remove” (TVR) was conducted in NI (DAERA, 2018a, Menzies et al., 2021). This intervention was introduced in 2014 to 2018, in the Banbridge area of County Down. The aim of the study was to test badgers for bTB, vaccinate if negative and remove those that tested positive. This approach provided an opportunity to examine *M. bovis* in badgers and potentially address the issue of the disease spread in cattle.

1.2 Thesis outline

This thesis studies epidemiology of *Mycobacterium bovis* in a multi-host system in NI within the intensely sampled wildlife intervention area. It consists of three main parts, the first one investigates the association of spatial distribution of *M. bovis* molecular types within a 100 km^2 area to understand the underlying processes of bTB transmission between cattle and badgers. The introduction of test and vaccinate or remove intervention to study the infection in badgers in Northern Ireland (DAERA, 2018a) provided an opportunity to obtain a unique dataset about *M. bovis* in the local badger population. In current study 1248 *M. bovis* isolates and 36 MLVA types from cattle and badgers and their spatial locations were analysed to evaluate the role of each host in disease transmission in the TVR area in Northern Ireland. By studying these data, it was possible to determine the association of *M. bovis* molecular types in cattle and badger species in the area.

With recent progress in using whole genome sequencing (WGS) it was widely used in studies for *M. bovis* transmission dynamics in different countries. WGS data was used to understand *M. bovis* evolution and intra-species transmission patterns in cattle and badgers; therefore, reconstructed phylogenetic and transmission trees of *M. bovis* isolates using WGS and epidemiological data of the historically known endemic *M. bovis* lineage in the TVR area were compared. In this study, an exceptional data set of 619 *M. bovis* isolates collected from badgers and cattle in a 100 km^2 area in Northern Ireland within the TVR intervention study were analysed. Historical molecular subtyping data permitted the targeting of an endemic pathogen lineage, whose long-term persistence provided a unique opportunity to study disease transmission dynamics in unparalleled detail. Graph transmission tree methods and coalescent analyses indicated most of the *M. bovis* diversity was mostly focused on one endemic lineage. Results pointed that bTB transmissions from cattle to badger and from badger to cattle play almost equal important role in disease dynamics. These assumptions indicate that cattle to cattle transmissions were more frequent in comparing to other inter and within species bTB transmissions.

The experimental part of the current thesis was focused on the development of a novel polymerase chain reaction (PCR) based technique for the detection and

quantitation of *M. bovis* bacterial load directly from animal tissue samples. Currently used "gold standard" for laboratory confirmation of *M. bovis* in cattle and other animals is bacteriological culture. However, it is a very labour intensive and time-consuming technique and there is a high demand for more accurate and sensitive molecular techniques that can also produce the results in a shorter amount of time. Molecular bacterial load assay is a culture-free technique that measures mycobacterial load directly from specimens and was widely used to quantify bacterial load decline in response to TB treatment in human samples (Honeyborne et al., 2011, Honeyborne et al., 2014). Several laboratory experiments were conducted to optimize the use of MBLA for detection and quantitation of *M. bovis* 16S rRNA directly from animal tissue samples. This technique has already been used for *M. tuberculosis* detection (Honeyborne et al., 2011, Honeyborne et al., 2014), and using it for the detection of *M. bovis* will improve current diagnostic methods, confirm bTB cases in a much shorter amount of time, and provide a reliable molecular method that can be used directly from infected material.

2 Chapter

Spatial distribution of *Mycobacterium bovis* multiple locus variable number of tandem repeats analysis (MLVA) types in cattle and badgers

2 Chapter

2.1 Introduction

Bovine tuberculosis (bTB) is a chronic bacterial disease of cattle and other animals, caused primarily by *Mycobacterium bovis,* a member of *Mycobacterium tuberculosis* complex (MTBC) (O'Reilly & Daborn, 1995; Grange, 2001). While cattle are the main reservoir of *M. bovis,* the disease has also been found in humans and many other domestic and wildlife animals (Grange, 2001, Thoen et al., 2006). Many developed countries managed to eliminate the disease in livestock, however it is still significantly affecting the farming sector and economy in some countries, especially where infection persists in wildlife species.

Pathogenesis of the disease is complex and can be different in individual animals. *Mycobacterium bovis* droplets can be inhaled by animals through direct contact with infected species or, less often, indirectly when ingested from the environment or contaminated surfaces (Neill et al., 1994). The bacteria grow slowly, and infected animals may show clinical signs only after several months or years, or sometimes can be dormant with no symptoms for long period of time (Pollock & Neill, 2002). Therefore, it is important to detect bTB before it spreads to other animals within herds.

Where disease persists in wildlife populations, applying effective control strategies becomes more challenging. Current control measures in Great Britain cost government about £150 million (Godfray et al., 2018), and £44 million in Northern Ireland every year (Northern Ireland Audit Office, 2018). BTB testing, using a variant of the tuberculin skin test, has been compulsory in the UK since the 1950s, with diagnostic positive 'reactors' removed from affected farms. According to the Department of Agriculture, Environment and Rural Affairs (DAERA), bTB herd incidence in Northern Ireland (NI) in 2019 was 7.85% and increased to 8.40% in 2021 (DAERA, 2019, 2021). This indicates that bTB is still a major animal health problem in NI, likely exacerbated by the persistence of the disease in the surrounding wildlife (Allen et al., 2018).

Badgers' social behaviour is one of the important aspects of bTB transmission dynamics. In the UK, in areas with high badger densities, badgers are mostly territorial and do not move far between social groups (Woodroffe et al., 1995).

The distribution of *M. bovis* infection in badgers was identified as highly clustered within their social groups (Olea-Popelka et al., 2003; Woodroffe et al., 2005).

Strategies to control bovine tuberculosis in cattle through badger culling were effectively implemented in Republic of Ireland (Eves, 1999; Griffin et al., 2005) and Great Britain (Clifton-Hadley et al., 1993), with significant reductions in number of incidences in cattle. However, during randomized badger culling trial (RBCT) conducted in GB within the "reactive" badger culling areas there was an increase in bTB cases in cattle (Donnelly et al., 2003, Donnelly et al., 2007). The badger culling strategies in high-risk areas of England were an important evidence of alterations in geographical distribution of *M. bovis* infection in badgers in response to the repeated badgers' removal (Jenkins et al., 2007). Previous studies showed significant clustering of the bTB infection in both cattle and badgers, and strong association of their *M. bovis* genotypes (Woodroffe et al., 2005), however after culling was implemented in wide areas clustering levels were reduced and increased the infection numbers in cattle (Donnelly et al., 2003, Donnelly et al., 2006, Donnelly et al., 2007). Field culls potentially affected spatial organization of badger setts and led to the negative outcomes of the trial in cattle, as movement of badgers for longer distances allow more risk for the disease transmission (Cheeseman et al., 1993, Roper and Lüps, 1993, Tuyttens et al., 2000). Therefore, the development of efficient bTB control policies in both domestic and wildlife animals can be specific for each area and should be implemented considering potential benefits and drawbacks.

A "test and vaccinate or remove" (TVR) (Menzies et al., 2021) intervention study on badgers, involving selective badger culling, was undertaken in a 100 km^2 area (and a 2 km buffer area) in County Down in NI, an area with the highest bTB prevalence and badger population density in the country from 2014 to 2018 (Agri-Food and Biosciences Institute, 2014, DAERA, 2018b). Animals were captured, sett-side tested for infection (Ashford et al., 2020) and released if test positive or vaccinated if test negative (Menzies et al., 2021). No badgers were culled during first year of study. The aim of the research project was to evaluate the impact of the TVR protocol on badger social structure, ranging behaviour and test positivity, to monitor any effects on bTB in local cattle and gain expertise in badger field-craft (Courcier et al., 2020, Arnold et al., 2021, Menzies et al., 2021). Before the introduction of this study, province-wide passive surveillance of bTB in badgers

included testing of animals killed in road-traffic accidents (RTA) (Courcier et al., 2018) and laboratory confirmation of *M. bovis* in sampled animals (Courcier et al., 2018, Milne et al., 2020).

Spoligotyping and Multiple-Locus Variable number tandem repeat analysis (MLVA) typing methods are used for *M. bovis* molecular epidemiological inference in both cattle and badger isolates (Milne et al., 2020; Skuce et al., 2020). The relative frequency and distribution of different MLVA types in NI has changed over recent decades (Skuce et al., 2010). The herd-level surveillance studies showed the most common *M. bovis* genotypes by their spatial structure in years 2003-2008, where significant changes in their geographical localization can be explained by so-called clonal expansion of *M. bovis* in GB (Smith et al., 2003). This indicated that although *M. bovis* population has a highly "clonal" genetic structure and some MVLA types were predominant in specific areas, their frequency distributions in the region remain diverse within cattle populations (Skuce, 2010).

Previous work has also shown spatial clustering of *M. bovis* molecular types in cattle herds and neighbouring badger setts, indicating that both species are involved in the disease's epidemiology, likely through inter- and intra-species transmission. More recently however, the potential for indirect transmission between species has received support from the findings that cattle and badgers tend to not come into close proximity with each other (O'Mahony, 2015, Woodroffe et al., 2016, Campbell et al., 2019). And this has renewed interest in the potential for localised environmental persistence of *M. bovis* and how it may constitute an epidemiological risk (Allen et al., 2021). *M. bovis* molecular types and their geographical localization in cattle and badgers in NI has been studied previously at different levels (Skuce and Neill, 2001, Milne et al., 2020). However, very limited data was available on bTB in badgers; the main source of *M. bovis* data in badgers came from RTA animals, which lacked the kind of systematic sampling necessary to make robust epidemiological inferences. Conversely, the systematic badger sampling employed in the TVR study enabled the collection of a densely-sampled representative dataset, with which to evaluate the associations between distribution of *M. bovis* genetic types in cattle and badgers and investigate important transmission links.

In this study the investigation of whether cattle and badgers share similar *M. bovis* MLVA types in the TVR area, and if they were spatially localized was investigated.

These two processes (*M. bovis* isolate genetic similarity and spatial overlap between host species) will give a broader understanding of the disease occurrence in a multi-host system and provide insights into how much the presence of the disease in badgers influences the presence of the disease in cattle (and/or vice-versa) in this well-sampled 100 km^2 region. Evaluation of the spatial distribution of *M. bovis* MLVA types to inform understanding of the persistence and association of shared pathogen MLVA types in cattle and badgers, and what that may mean for the transmission dynamics in the study region was performed.

2.1.1 Objectives

1. The distribution of MLVA types and their genetic similarity among cattle and badger isolates was determined using the goeBURST algorithm. This might explain the relative frequency and the prevalence of some MLVA types in the area.

2. Investigation of the spatial clustering of *M. bovis* in cattle and badger was done within the TVR area to understand whether there were any patterns of local transmission of bTB in the area as introduced in previous studies throughout NI (Milne et al., 2020).

3. Investigation of whether there were any MLVA types that were shared between cattle and badgers, as an indicative of the interspecies transmission. And studied the spatial association of MLVA types identified in cattle and badgers; whether using these data it is possible to predict the distribution of MLVA types in one another based on their kernel density estimations. These findings might be useful in understanding the roles of cattle and badgers in bTB spread within the TVR area

2.2 Materials and methods

2.2.1 Study area

The TVR area is a 100 km^2 territory with a 2 km “buffer” zone in County Down, Northern Ireland. During the TVR wildlife intervention study (DAERA, 2018) a total of 1,248 *M. bovis* isolates were collected from cattle and badgers. The dataset included five different sources of isolates shown in Figure 2.1. A total of 1,073 *M. bovis* cattle isolates collected through the national bTB herd surveillance scheme was used in this study. From these, historical samples were obtained during 1986-2013 and samples that were collected during the 2014-2018 TVR study project.

Then, 142 *M. bovis* isolates from 49 badgers sampled during the TVR programme in 2015-2018. It is important to note that in the first year (2014) of the wildlife intervention study no badgers were culled; they were merely recorded, tested and released. This data was used to understand the initial prevalence of the bTB infection within badger population and collect the epidemiological information about badgers. As described in Menzies et al. (2021) all captured badgers were first anaesthetised, then inspected for any physical conditions, parasites or injures. After recording these characteristics, the animals were microchipped for further identification, the spatial locations where they were found were documented, and samples (blood, tracheal aspirates and swabs) were taken for *M. bovis* laboratory testing. Before badgers being released back to the areas where they were captured, GPS collars were put in each individual to track their movement behaviour. In subsequent years of the TVR study (2015-2018), bTB positive badgers were culled; and negative badgers vaccinated and released. Additional to the intervention study data, 33 isolates were collected from 14 badgers killed in road traffic accidents (RTA) and identified as positive for *M. bovis* during 1999-2018 (Figure 2.1). Therefore, in total we analysed multi-locus variable analysis (MLVA) data and spatial data from 1073 unique cattle and 63 badgers.

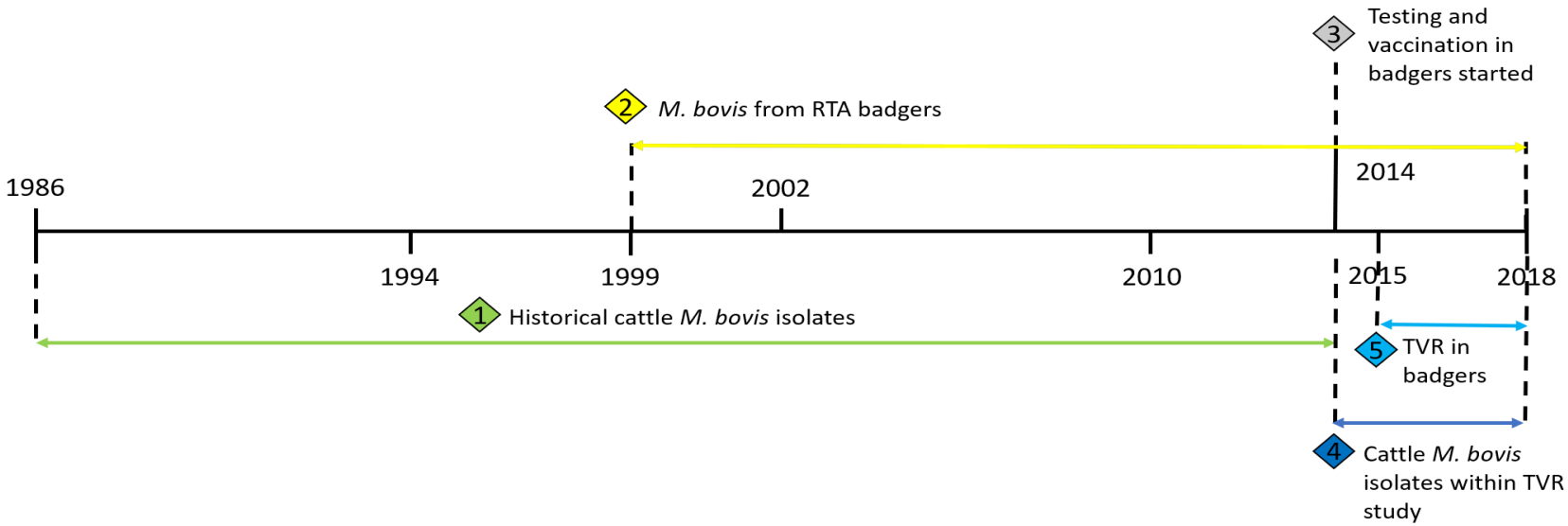


Figure 2.1 Description of the *M. bovis* data used in current analysis

1. Historical cattle *M. bovis* isolates sampled from the TVR area and the buffer zone from previous years 1986-2013
2. *M. bovis* isolates collected from badgers killed in traffic accidents (RTA), 1999-2018
3. Start of the TVR intervention study, no badgers were culled during the first year of the program, 2014
4. *M. bovis* isolates detected in cattle herds within the TVR area and the buffer zone, 2014-2018
5. Sampling of bTB infected badgers, 2015-2018

2.2.2 Cattle data

All cattle herds are subject to routine bTB surveillance according to the TB control programme in NI (TB Strategic Partnership Group, 2016). Cattle data for this analysis were obtained from cattle herds confirmed with *M. bovis* infection. Animals that were tested positive for bTB during the annual skin testing surveillance scheme were slaughtered and sent for post-mortem examination. After inspection, tissue samples (whether or not have tuberculosis lesions) were collected and sent for *M. bovis* confirmation by bacteriology culture (Northern Ireland Audit Office, 2018). Moreover, under the current eradication scheme animals sent to slaughter for meat consumption are also inspected for bTB, if suspected positive they are collected and cultured (Northern Ireland Audit Office, 2018). All tissue samples taken from any infected animals are cultured and sent for molecular typing using spacer oligonucleotide typing ('spoligotyping', Kamerbeek et al. 1997) and VNTR-MLVA typing (Supply et al., 2001; Roring et al., 2002).

For this analysis MLVA data from infected cattle herds that were sampled particularly within the TVR intervention study and buffer area was used. Data was chosen in years during the programme and from previous years from statutory testing in cattle.

2.2.3 Badger data

During the five-year TVR intervention study period, badgers in adjoining locations to cattle herds were trapped and captured (DAERA, 2018). This analysis did not include any animal movement data about badgers captured in the first year of the study. From the second year, badgers were trapped in cages and tested using Dual Path Platform (DPP) VetTB test in the field (Courcier et al., 2020), with further gamma interferon testing of blood samples and swabs sent for bacteriological cultures and molecular typing.

Data from badgers killed in road traffic accidents (RTA) in NI was included. These data, about *M. bovis* in badgers from years 1999 to 2018, were obtained from the RTA Survey (TB Strategic Partnership Group, 2016). When animals were found dead on roads they were reported and sent for laboratory and post-mortem analyses; bTB positive tissues were sent for *M. bovis* culture and molecular typing.

This survey does not however, provide a thorough investigation of the situation about bTB in badgers.

2.2.4 Molecular typing

Multi Locus VNTR Analysis (MLVA) is a PCR-based *M. bovis* molecular typing method (Supply et al., 2000, Skuce et al., 2010). This typing technique is based on the identification of copy numbers of tandem repeated DNA sequences in mycobacterial genomes (Frothingham and Meeker-O'Connell, 1998, Roring et al., 2002). All the procedures for the *M. bovis* cultures, DNA extraction and molecular typing were undertaken in accordance with the internal standard operational procedures (SOPs) at the Bacteriology branch of the Veterinary Science Division (VSD) in Agri-Food and Biosciences Institute.

After the *M. bovis* cultures were confirmed and isolated (Bactec/MGIT, Lowenstain-Jensen slopes, Middlebrook broth and solid culture), genomic DNAs were extracted by incubation in distilled and boiling water for 7 minutes and then centrifuged for further use in PCR amplification using specific primers. With known sizes of each VNTR at different loci, the sizes of amplification products were evaluated using gel-electrophoresis and correlated to the number of tandem repeats. Eight loci (2163b, 4052, 2461, 1955, 1895, 2165, 2163a and 3232) were selected systematically for molecular typing of *M. bovis* isolates in NI (Skuce et al., 2005) and their profiles represented as a concatenated string of numbers, which was simplified to a laboratory code based on their frequency (Table 2.1).

Table 2.1 ***M. bovis*** **VNTR-MLVA types found in cattle and badgers with corresponding spoligotypes (Mbovis.org) (as described in (Skuce et al., 2010)**

#	MLVA type (lab use)	MIRU-VNTR loci (copy numbers)								Spoligotype
		MV 2163b	MV 4052	MV 2461	MV 1955	MV 1895	MV 2165	MV 2163a	MV 3232	
1	001	4	4	5	3	4	7	11	9	140
2	002	4	4	5	3	3	7	11	7	142
3	003	4	4	5	1	4	6	10	7	140
4	004	4	4	5	3	4	7	11	7	140
5	005	4	2	5	3	4	7	11	7	140
6	006	4	4	4	1	4	7	10	7	263
7	007	4	4	5	1	4	6	10	9	140
8	008	4	2	5	3	4	6	11	7	140
9	009	4	4	4	3	2	7	6	8	273
10	010	3	4	5	3	4	7	11	9	140
11	011	3	3	3	3	4	7	11	8	145
12	019	4	4	5	3	4	7	11	4	140
13	020	4	4	5	3	4	5	11	9	131
14	025	3	4	5	1	4	6	10	7	140
15	027	4	4	5	2	3	5	10	7	140
16	049	4	2	5	3	4	7	11	8	140
17	068	3	4	5	3	4	6	11	7	978
18	073	4	4	5	3	4	6	11	7	140
19	114	4	4	7	3	4	7	11	9	140
20	117	4	4	5	3	4	6	10	9	140
21	122	4	4	4	1	4	6	10	7	263
22	145	4	4	5	3	4	7	11	5	140
23	146	4	4	5	1	4	5	10	9	140
24	149	5	4	5	3	4	7	3	7	140
25	158	4	4	4	1	4	7	9	7	263
26	169	4	3	5	3	4	7	9	7	140
27	255	4	4	5	3	3	2	11	7	140
28	266	4	3	4	1	4	7	10	7	263
29	293	4	4	5	1	4	5	10	10	140
30	297	3	4	4	1	4	7	10	7	263
31	421	4	4	4	1	4	5	10	7	263
32	423	4	4	4	1	4	7	10	8	263
33	464	4	4	4	1	2	7	10	7	263
34	471	4	4	5	3	4	1	11	9	140
35	543	4	4	5	1	4	5	11	9	131
36	997	3	5	5	1	4	5	11	5	120

2.2.5 Genetic relatedness of M. bovis isolates

M. bovis MLVA types were separated into three groups: isolates that were isolated only from badgers, MLVA types isolated only from cattle, and MLVA types isolated from both cattle and badgers. Genetic relatedness of MLVA types was assessed using globally optimized eBURST software (Francisco et al., 2009), which groups similar multiple locus sequence typing (MLST) profiles into clonal complexes, and identifies potential founding genotypes based on differences in their allelic profiles. The initial eBURST algorithm (Feil et al., 2004) was designed to reconstruct hypothetical relationships between different MLST (multi locus sequence typing) (Baker et al., 2004) profiles of bacteria. It focused on building a spanning tree where two connected nodes were only one locus different from each other. The implemented goeBURST algorithm included an option to connect these single trees together depending on the number of differences in locus variants if more than one. The optimized goeBURST approach (Francisco et al., 2009) first identifies a probable "founder" genotype, then connects profiles that are different at one locus (single locus variants (SLVs)); two loci (double locus variants (DLVs)); three loci (triple locus variants (TVLs)) and so on. GoeBURST grouped MLVA profiles in clonal complexes without necessarily linking all isolates together; types not linked were identified as singletons separated from the tree.

2.2.6 Spatial analyses of M. bovis MLVA types

Spatial statistical techniques can be used to investigate relationships between observed entities using their geographical characteristics, such as positions, distances and/or patterns. In this study, spatial locations of cattle herds were represented by farm home building locations; badger positions were associated with badger trapping locations and with the positions of animals killed on roads. All spatial points were characterized by their X (UK Easting) and Y (UK Northing) coordinates. Spatial clustering methods were used to explore the distribution of *M. bovis* MLVA types isolated from each host species.

The *k*-means algorithm (MacQueen, 1967) (using "cluster" package implemented in R (v. 3.5.2)) (Maechler et al., 2021, R Core Team, 2021), was used to examine the clustering of *M. bovis* in cattle and badger species in a given area. This algorithm partitions the dataset into a predefined *k* number of clusters (Jain et

al., 1999). The significant number of clusters and their fit to the spatial data were examined using two clustering mechanisms, such as elbow curve method and silhouette analysis (using the 'factoextra' and 'NbClust' packages implemented in R v.3.5.2) (Charrad et al., 2014, Kassambara and Mundt, 2020, R Core Team, 2021). These methods determined the optimal number of clusters for the *k*-means clustering.

Due to the large number of *M. bovis* isolates that were densely collected within the TVR area, the density-based clustering approach was also applied (Ester, 1996, Hahsler et al., 2019, R Core Team, 2021). This method investigates the spatial structure of geographical points, based on the density around each object; and groups points with nearby neighbours, leaving spatial outliers in the lower density areas. For this approach users are required to select two parameters, such as the minimum number of neighbours ('MinPts') and epsilon ('Eps'), the radius of the neighbourhood around each spatial point (Kriegel et al., 2011). We defined these parameters for each data set (cattle and badgers) and selected the 'MinPts' based on *k*-nearest neighbours algorithm (kNN) (within the 'dbscan' package).

2.2.7 Association between cattle and badger MVLA types

Kernel density estimation (KDE) (Silverman, 1986) is a non-parametric method that we used for visualization and analysis of spatial data for different MLVA types. A smooth density distribution was estimated for geographical points that belong to each MLVA type, to show the existing spatial patterns of the observed locations and the potential to estimate areas with higher densities for particular MLVA types or groups of MLVA types (Kloog et al., 2009). Smoothed kernel distributions provide visualisations of the spatial "shapes" of *M. bovis* MLVA types found in both cattle and badgers and help to understand the extent to which they are related to each other.

To investigate potential spatial association between disease in cattle and disease in badgers, a modelling approach based on Kernel Discriminant Analysis (KDA) (Bouveyron et al., 2015) was applied for the shared *M. bovis* genotypes 004, 006, and 297. To estimate the kernel densities for these genotypes the minimum required number of spatial points was three, therefore, MLVA types that had fewer samples (for example, MLVA type 122) were excluded from this analysis. Permutation test was used to determine whether the MVLA distributions of cattle

and badgers are unrelated, against the alternative that there is a spatial association between the two MLVA distributions in the feature space. Based on the kernel density estimates of the *M. bovis* MLVA types that belong to cattle and badgers, the model predicts spatial locations of cattle MLVA types using badger data and *vice-versa*. For each model the misclassification rate was calculated, which is the proportion of incorrect MLVA type predictions that varies on a scale from zero to one. For each direction, a permutation test (10,000 replicates) was used to assess the evidence against the null hypothesis (H_0: No spatial association between the MVLA distributions of cattle and badgers). The distributions of the predicted values were calculated and compared with the observed misclassification rates.

2.3 Results

2.3.1 Study area

M. bovis MLVA data in this analysis represented by 1248 isolates obtained from bTB infected cattle and badgers from 1986-2018 years (Figure 2.3a) period collected within the TVR intervention study area in County Down, South-East of Northern Ireland. Total 175 *M. bovis* isolates were collected from 63 badger species, this is because multiple samples were taken from one animal during examination. From the 1073 bTB positive cattle in this study, 53% were homebred in local farms and 47% bought-in to the area in different years from 1986 to 2018.

Figure 2.2 shows the map of NI and the approximate location of the TVR study area and the buffer zone (2 km) with *M. bovis* positive isolates collected from cattle herds and badgers used in this analysis.

All isolates obtained for this study were *M. bovis* culture confirmed and MLVA typed, however metadata regarding the animals was not included, such as age, breed, any data regarding the non-infected animals and whether any visible or non-visible lesions were detected in infected cattle or badgers.

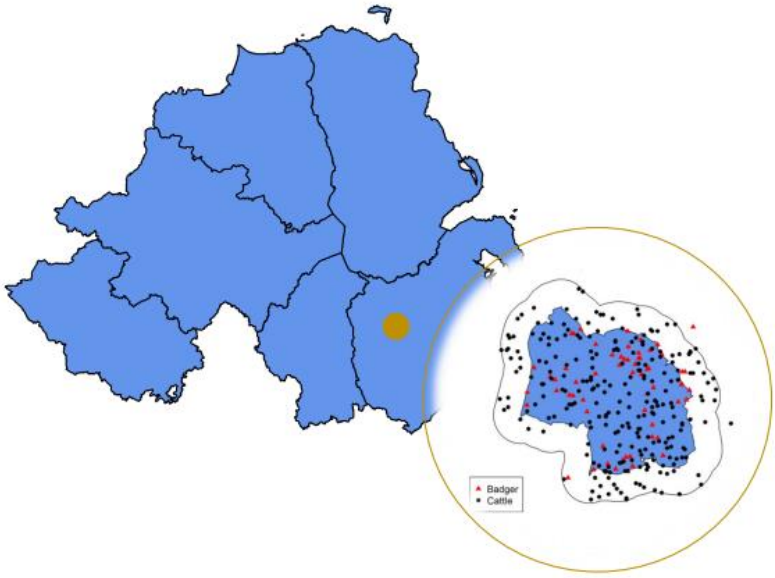


Figure 2.2 Map of Northern Ireland with the TVR intervention area

TVR as shown as brown point, and an inset of the TVR area with zoomed locations of *M. bovis* isolates from cattle (black dots) and badgers (red triangles) within the TVR study area (shown in blue) and the buffer zone (shown in white). Some spatial points fall outside the buffer zone; these represent relevant historical M. bovis isolates that were added to the dataset.

The sampling of M. bovis isolates has increased during the TVR study period due to intensive testing within 2015-2018 when compared to the sampling performed during the historical period (1986-2013). Table 2.2 shows the distribution of M. bovis isolates collected from cattle and badgers during the years used for the current study.

Table 2.1 Distribution across time of *M. bovis* isolates used in this study

Years of sampling	Host	Number of isolates	Study
1986-2013	Cattle	263	Historical
1999-2013	Badgers	33	RTA
2015-2018	Cattle	810	BTB Surveillance (within TVR area)
2015-2018	Badgers	142	TVR

2.3.2 Genetic diversity and relatedness of MLVA types

In total, thirty-six MLVA types were identified among the 1248 *M. bovis* isolates (Figure 2.3b), of which thirty-three were found in cattle and seven in badgers. Four genotypes were shared among cattle and badgers (MLVA types 004, 006, 122 and 297), of which type 006 was the most prevalent among overall genotypes in

cattle and badgers (Figure 2.3c). MLVA types identified in cattle on herd-level were associated with homebred and bought-in animals, where number of MLVA types found in purchased cattle were thirty-one, comparing to eighteen types found in homebred animals. This diversity might be explained by the farm to farm movement of cattle infected animals undetected during compulsory bTB testing.

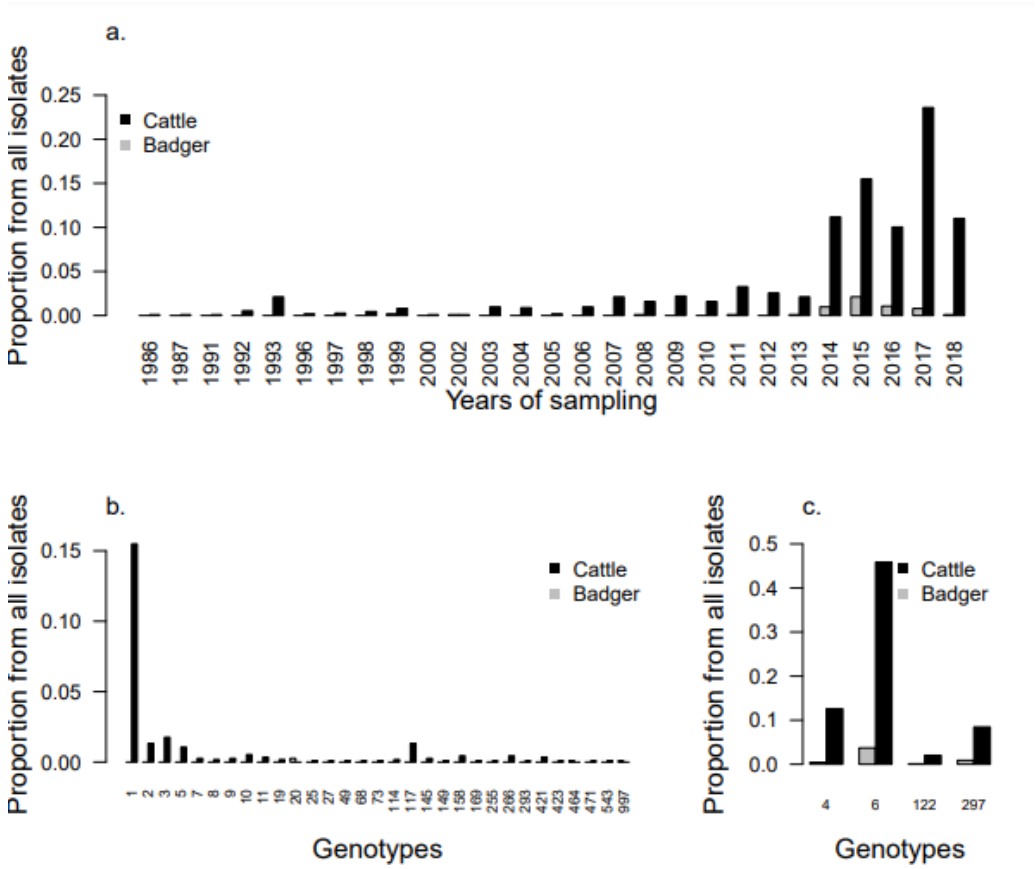


Figure 2.3 (a, b, c) Distribution of *M. bovis* isolations from cattle and badgers

a. Proportion from all *M. bovis* isolates collected from cattle (n=1073) in black and badgers (n=63) in grey by sampling years. b. *M. bovis* MLVA-VNTR types in cattle only (black) and badger only (grey); c. *M. bovis* MLVA-VNTR types shared between cattle (black) and badgers (grey).

From seven MLVA types that were obtained in badger species, types 020, 464 and 997 were only badger specific and not found in cattle in this dataset. Some MLVA types were represented by only single isolates. The goeBURST (Francisco et al., 2009) algorithm visualized the genetic relatedness of different MLVA types (Figure 4). Groups were defined based on differences at one, two or more VNTR loci from the predicted founding MLVA genotype. Input data used for goeBURST algorithm included only 'strain number' for each MLVA type without accounting for the frequency of the samples per type. This software was used to understand the relatedness of MLVA types and their patterns of descent, which together with the

spatial localization of the types might explain some transmission links of the infection in the area.

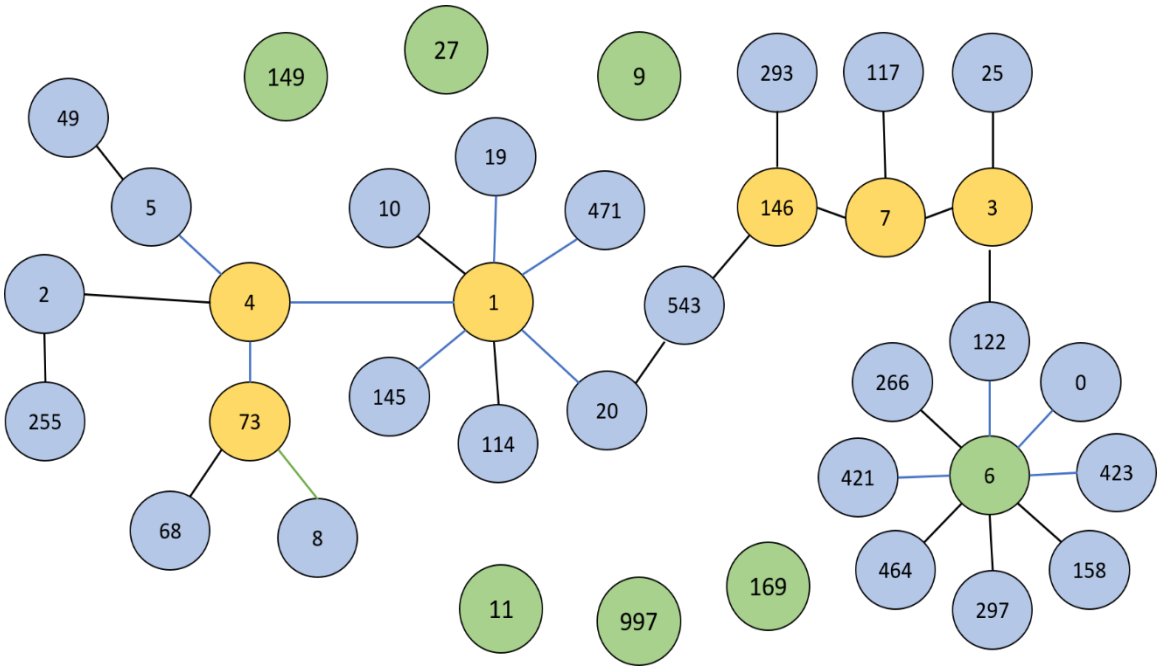


Figure 2.4 The GoeBURST tree of the thirty *M. bovis* MLVA types and six singletons

Type 006 was predicted as the founder of the clonal group (shown in green), MLVA types 001, 004 and 073 were identified as subgroup founders of linked clusters, whereas MLVA types 003, 007 and 146 were identified as subgroup founders with a unique link (in yellow). MLVA types 009, 011, 027, 149, 169 and 997 were shown as singletons (in green).

The majority of MLVA types (30/36) were grouped in one clonal complex with the MLVA type 006 the predicted as founder (shown in green in Figure 2.4). Genotype 006 in this analysis comprised of 52% of total *M. bovis* isolates, which was 42% from total bTB infected cattle; 72% from total isolates obtained from badgers sampled in various years. Branch colours between MLVA types (nodes) on the graph demonstrated links between profiles according to tiebreak rules implemented in the goeBURST algorithm (Francisco et al., 2009), where blue and green colours represented differences in one and two VNTR loci, respectively, and black showed links without ties between profiles. MLVA molecular types highlighted in yellow (001, 003, 004, 007, 073 and 146) indicate the probable subgroup founder MLVA types; other connected types were one or two loci different from these sub-founders. Some MLVA types were not included in the clonal group and were identified as singletons, such as MLVA types 009, 011, 027, 149, 169 and 997. The majority of MLVA types that are shared between cattle and badgers tend to group together (MLVA types 006, 122 and 297), and some host-

specific types were single locus variants (SLVs) of the founder and sub-founder types; for example, cattle types 421, 423 are SLVs of the putative founding MLVA type 006; MLVA type 020 (found in badgers sample) is an SLV of sub-founder MLVA 001 (found only in cattle) as well as MLVA types 145, 019 and 471.

It is also important to note that MLVA typing has relatively low discriminatory power for *M. bovis* typing, compared to whole-genome sequencing (Allix et al., 2006; Kao et al., 2014); genetic difference between genotypes would tend to be overestimated.

2.3.3 BTB spatial distribution

The co-localization and clustering of different MLVA types in cattle and wildlife hosts has been identified spatially in different years at a province-wide and regional scale throughout NI (Skuce et al., 2010; Milne et al., 2020; Skuce et al., 2020). Here, the spatial distribution of the *M. bovis* infection in cattle and badgers sampled within 30 years period in a densely-sampled but small TVR area was examined. Figure 2.5a demonstrated the geographical distribution of the bTB infections isolated from cattle and badgers across TVR area. The distribution of the distances to the nearest infected neighbours among all cattle and badgers, between just cattle and just badgers were shown in Figure 2.5b. These plots clearly showed that *M. bovis* isolations from both host species were identified within spatially close areas. And the distribution of the distances between different groups suggested that distances between *M. bovis* identified in badgers were shorter than in cattle.

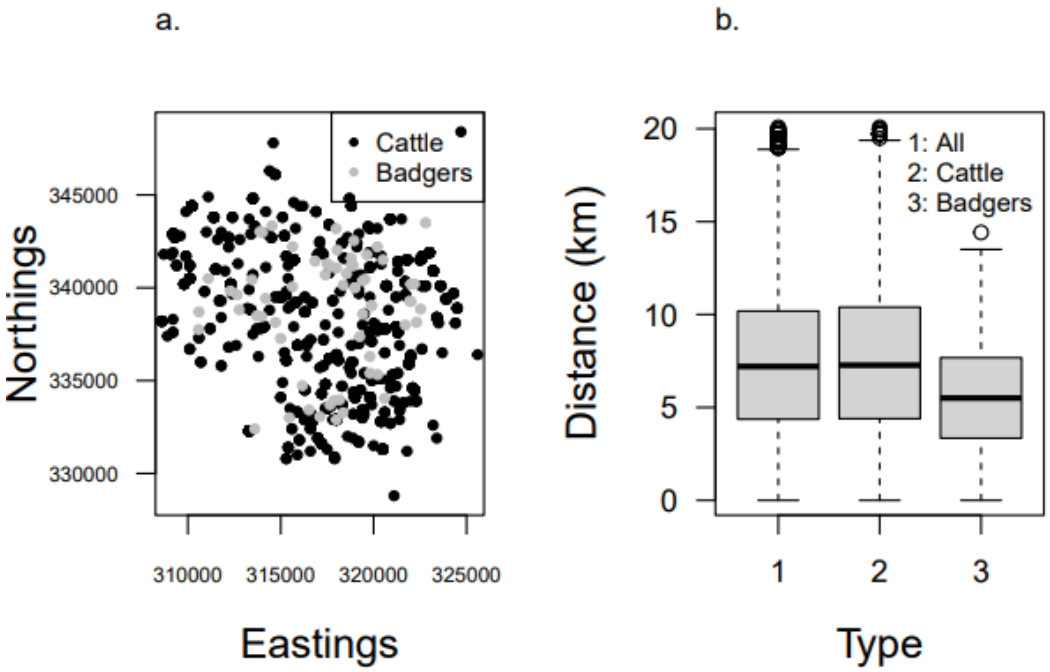


Figure 2.5 (a, b) Spatial distribution of *M. bovis* in cattle and badgers

a. Spatial distribution of M. bovis in cattle and badgers shown as black (cattle) and grey (badgers) dots. b. Distribution of Euclidean distances between bTB infections within all data (1), cattle (2), badgers (3).

To understand the spatial clustering of *M. bovis* in cattle and in badgers we used two clustering algorithms, *k*-means (Steinley and Brusco, 2007) analysis and density-based clustering (Hahsler et al., 2019). These methods were used to indicate any evidence of the geographical co-localization of the bTB infection in cattle and badgers separately, which might indicate whether the bTB spread is driven by local transmission.

2.3.4 K-means clustering in cattle M. bovis isolates

Exploring the optimal number of clusters for the *k*-means algorithm using elbow curve plot suggested that three clusters were optimal for *M. bovis* in cattle data. The most suitable number of clusters for this dataset was identified as three. Another clustering method was a density-based clustering of *M. bovis* in cattle. For this analysis we first computed the distance matrix between the bTB infections in cattle, then identified the minimum distances within the *k* number of neighbouring infections (*kNN* distance) and then used these values to run the *dbscan* (Hahsler et al., 2019) model. *MinPts* (minimum number of neighbours) was indicted as 5 with the *eps* (neighbourhood radius) of 1 km. Within the suggested

parameters the density-based clustering approach divided the 1073 cattle *M. bovis* isolates into 13 clusters and 18 noise points (not included in any cluster).

2.3.5 Spatial clustering of M. bovis in badgers

The same techniques were used to detect statistically significant clustering of *M. bovis* infection in badgers. As previously described we tested the optimal number of clusters in the range of ten for the *k*-means clustering approach, both methods (Elbow curve, Silhouette values) suggested that the *k*=3 number of clusters were the most suitable for this data with the highest statistical values and average silhouette value of 0.53 (results not shown). When applied the *dbscan* statistics for the badgers *M. bovis* isolates we used the default number of *MinPts*=5 and identified the required *eps*=2 km plotting the *kNN* distances. *M. bovis* badger isolates partitioned into 3 spatial clusters and 6 noise points. Both clustering approaches indicated significant clustering of the bTB infection in badgers which suggested that the distribution of the infection in badgers was different from random. Although the clustering algorithms were not applied to the MLVA types in this analysis, when visualizing which MLVA types were included in each of the cluster in badger *M. bovis* isolates we identified that the biggest cluster represented by MLVA types 004, 006, 020, 122 and 297. Whereas other two clusters included only genotypes 004 and 006; and 004, 006 and 464. From looking at the Figure 2.4 (goeBURST) it is seen that the MLVA types 006, 297, 122 were also grouped together. MLVA types 122, 464 and 997 were found only as single isolates.

Previous studies by Skuce et al. (2010) have indicated that the bTB infection was spatially clustered in cattle throughout NI, and the investigation of whether clustering of bTB in badgers was associated with the same in cattle and if this specifically associated with the MLVA types that were shared between species. This could determine if the inter species transmission influenced the spatial distribution of *M. bovis* in the area.

For investigation of the *M. bovis* spread in the TVR area both cattle-cattle and badger-badger transmissions are important. However, what is more important is to understand the inter-species transmission pathways and underlying drivers of these events. In current study the spatial distribution of the MLVA types that were found in both cattle and badgers and the relative contribution of each was examined.

2.3.6 Relationships between genetic and spatial diversity of M. bovis MLVA types between hosts.

Kernel density estimates for the three shared *M. bovis* MLVA types 004, 006 and 297 were calculated and shown in Figure 2.6. MLVA type 122 was excluded from these estimations as it required minimum of 3 number of *M. bovis* isolates per species which was not enough in our data. The heatmaps for each of the types 004, 006 and 297, were shaded in orange and demonstrated the spatial distribution of MLVA type locations with darker areas showing the "hotspots" with larger numbers of points for each type. When combining these kernel density estimates for all cattle-badger shared MLVA types (Figure 2.6), with contours that include 95% of isolates in each type, there was a high degree of overlap between the different types, even at the TVR area scale. As shown on a graph, MLVA type 006 included more than a half of all cattle and badgers *M. bovis* isolates and belongs to the central density kernel. MLVA types 006 and 297 were genetically similar and the core density kernels were mostly overlapping.

Therefore, these findings may suggest that the spatial distribution of the bTB infections in cattle and badgers are linked and there is a higher chance of the transmissions to be local in the TVR area as the density "hotspots" of the infection were similar in the shared MLVA types.

All these assumptions may lead to the conclusion that *M. bovis* transmission in the area is strongly associated with the MLVA types shared among cattle and badgers, therefore inter-species transmission. It is interesting to see if there were any historical or environmental underlying processes that might affect the spatial clustering of the infection or the localization of the density kernels of the particular MLVA types.

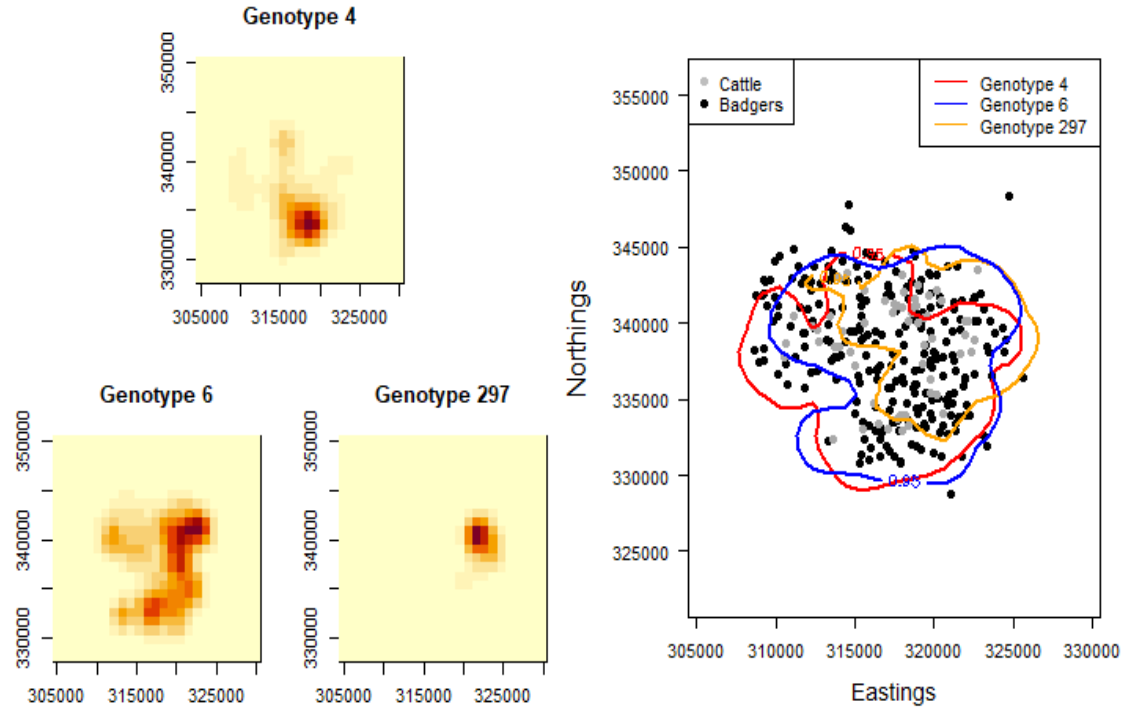


Figure 2.6 Kernel density estimate (KDE) heatmaps for MLVA types 004, 006 and 297

KDE map with 95% confidence contours for the three shared M. bovis MLVA molecular types (red, blue and orange lines). Geographical positions of bTB-affected herd-level cattle (black dots) and culture-confirmed bTB TVR/RTA badgers (grey dots) are displayed.

According to the study by Skuce et al. (2020) some MLVA types were spatially associated with particular Divisional Veterinary Office (DVO) regions with higher probabilities of occurrence in the area at the herd-level. MLVA type 006 was strongly associated with Newry DVO which includes the TVR study area. Therefore, as the most prevalent MLVA type in the region it might highly influence the spatial distribution of other genotypes.

MLVA type 006 was the most prevalent in this dataset and was indicated in 48.6% of bTB cattle and 66.7% of bTB badger data from the total number comparing to other types. The kernel discriminant analysis (KDA) used in current analysis was introduced to determine the association between the distribution of the MLVA types in cattle and badger species. Based on the spatial density estimates for each shared genotype in the two groups two models were compared: cattle-badgers and badgers-cattle; and aimed to predict their spread respectively. Firstly, we input the spatial locations of *M. bovis* in badgers and cattle separately as independent data, then assign their MLVA types as input genotypes data. Then run the kernel discriminant analysis (Duong, 2007) function to which allocate every spatial point to one of the groups (MLVA types) based on estimated densities.

Therefore, to explore the association between bTB distribution in cattle and badgers we used the cattle data to identify MLVA types in badgers and vice versa. After this step, the results obtained by the models were compared with the real dataset and the misclassification rates estimated: 0,38 and 0,26 for cattle MLVA types predicted using badgers' data and for badgers MLVA types using cattle data. Following this analysis, the permutation test (randomization test, x10.000 repeats) was carried out to test the hypothesis, which provided the distributions of the misclassification errors for both KDA models. When comparing the observed density model results with permuted it was identified that model mostly predicts MLVA type 006 correctly in both hosts. However, it was more likely that the model will misclassify MLVA type 297 or 004 to MLVA 006 isolates.

From the total 1136 MLVA typed isolates, MLVA types that were shared between cattle and badgers in 817 isolates, from which 18% belonged to type 004, 69% to MLVA type 006 and only 13% to type 297. Taking this into account it was attempted to demonstrate the relative importance of the MLVA type 006 among other types in the area and if there was any effect for our models' results. For this reason, the analysis was run on two types of *M. bovis* data: including and excluding the MLVA type 006. The results of these tests were shown in Figures 2.7 and 2.8.

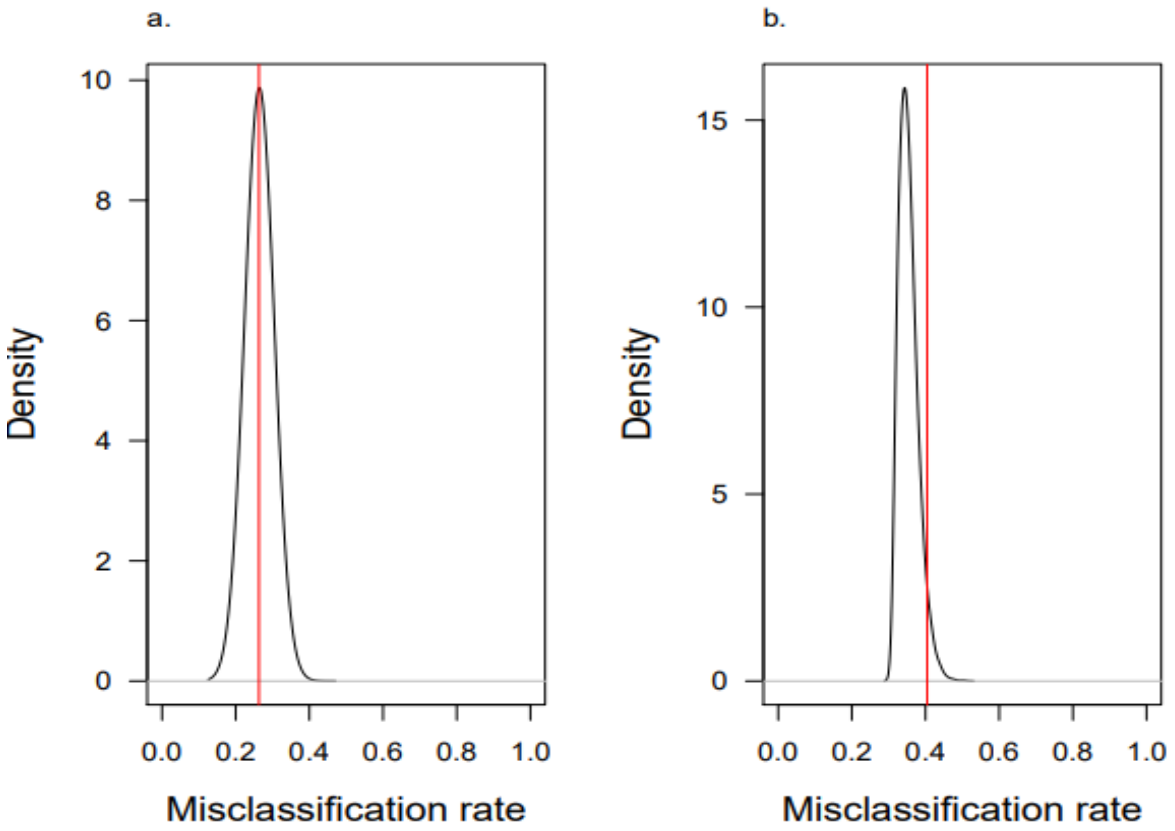


Figure 2.7 (a, b) Misclassification rates for permutation tests for a cattle/badger model predicting badger/cattle *M. bovis* (MLVA types 004, 006, 297)

a. MLVA types in cattle predicting types in badgers; b. MLVA types in badgers predicting types in cattle. The red lines represent the observed misclassification rates.

The exclusion of MLVA type 006 produced significantly different results: misclassification rates for cattle-badger predictions were 0.23; and 0.22 when predicting *M. bovis* types in badgers from cattle types (Figure 2.8 (a, b)). The p-values for the observed with expected misclassification rates within each comparison type after permutation were statistically significant (cattle-badger: p = 0,0006; badger-cattle: p = 0). These findings indicated that using the KDA model for MLVA types 004 and 297, the model successfully predicted the types that were present in cattle using the data of the spatially close badgers and vice-versa.

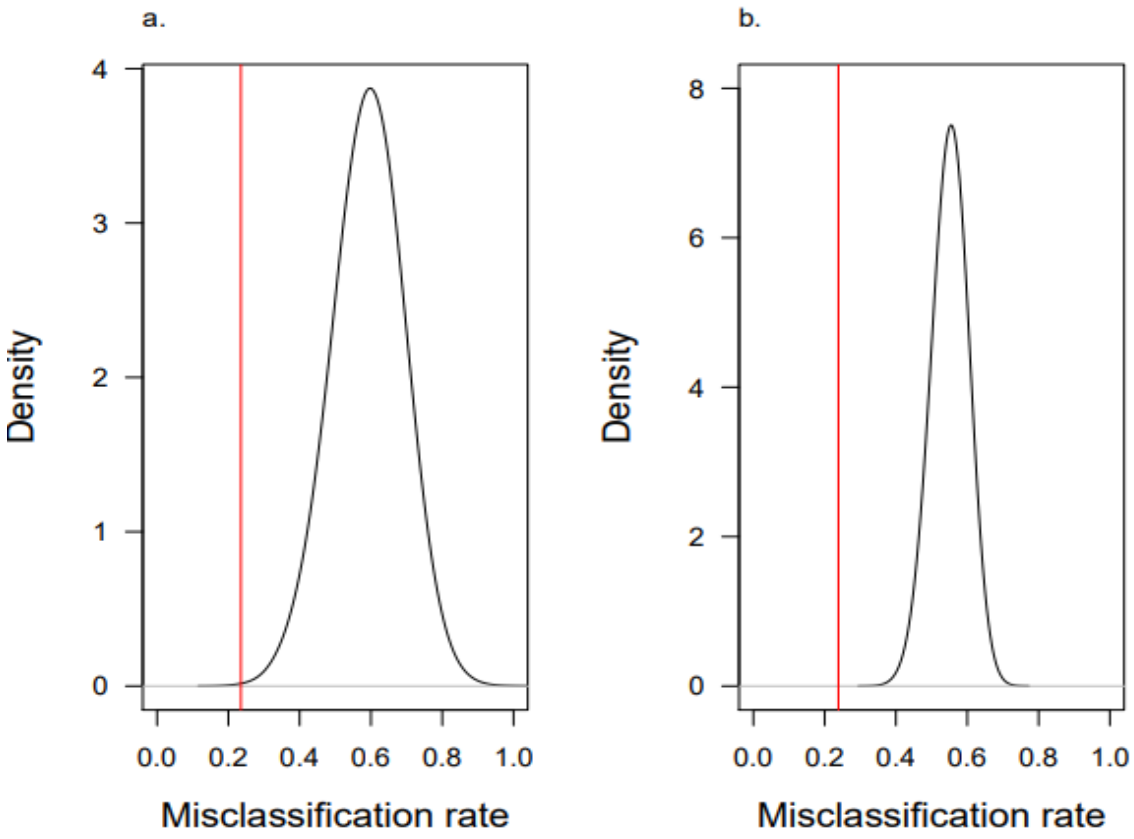


Figure 2.8 (a, b) Distribution of misclassification rates for the permutation tests (badger KDE model predicting cattle *M. bovis* MLVA types, and vice-versa, for MLVA types 004 and 297 only)

a. Cattle M bovis type predicting badger *M. bovis* type; b. Badger *M. bovis* type predicting cattle *M. bovis* type. The red lines indicate observed misclassification rates for each comparison.

2.4 Discussion

This study examined the genotypic and spatial distribution of *M. bovis* MLVA types collected from Northern Irish herd-level cattle and badgers within a novel intervention study in badgers undertaken in 100 km^2 area within 2014-2018 years period. Previously, within the UK and Ireland, badger culling was one of the major approaches to control bTB infection in badgers (Griffin et al., 2005, Jenkins et al., 2010, Sheridan, 2011, Godfray et al., 2013). The study area was selected based on the evidence of high bTB prevalence, high cattle densities and comparatively high densities of badger setts (Reid et al., 2012, Menzies et al., 2021). This was the first study in NI where selective culling of infected badgers and vaccination of the non-infected badgers were undertaken over a relatively large area (Menzies et al., 2021). The outputs of the approach provided detailed data on badger population characteristics in the area and the presence of *M. bovis* in badgers which is essential in understanding the role of badgers in inter/intra species transmission of bTB. Surveillance data about *M. bovis* in badgers pre-TVR were available only from RTA cases (Courcier et al., 2018).

Genetic variability of the *M. bovis* MLVA types was evaluated using goeBURST algorithm (Francisco et al., 2009) where we distinguished the most likely ancestor type and patterns of descent according to allelic profiles. Where local control measures are insufficient to break chains of transmission, *M. bovis* continues to spread and generate new variants, in this case indexed as VNTR copy number variants. Although MLVA typing has lower discriminatory power than whole genome sequencing, the accurate selection of VNTR loci helps to determine the main clonal complexes in the region or country level (Skuce et al., 2020). Results of current work suggest that MLVA type 006 was the founder of the main clonal complex (Francisco et al., 2009) and the six other MLVA types were sub-founders (001, 003, 004, 007, 073, 146) (Figure 2.4). However, the relative frequency and density of occurrence of some MLVA types varied significantly. Some *M. bovis* MLVA types, such as types 006, 001, 004 and 297 were more prevalent, whereas some molecular types were found only in single animals or single species. Although cattle *M. bovis* data was mostly represented by herd-level animals sampled in different years, the distribution of the MLVA types was heterogeneous across farms. The types of cattle either homebred or imported might be affected by the cattle movements and/or herd types as described in Milne et al. (2019). The

occurrence of newly introduced MLVA types within same herds can be explained by the long-distance movements of cattle between different areas with known history of bTB. MLVA 006 was known to be historically endemic in the TVR area (Skuce et al., 2010; Skuce et al., 2020) which was consistent from the findings of current studied demonstrated the prevalence of MLVA 006 in used data; MLVA types 001 and 004, mainly associated with the Newtownards DVO (Skuce et al., 2020) , might have been brought to the area when animals were moved. Badgers' *M. bovis* data showed lesser diversity of MLVA types, and more than 90% of the types found in badgers were also found in cattle. Where among RTA badgers only shared *M. bovis* MLVA types were identified, similar as in previous studies throughout NI (Trewby, 2016; Milne et al., 2020). These findings might be influenced by various factors, as cattle population sizes being much higher than badgers across whole NI (Reid et al., 2012), and therefore acquiring more MLVA types. According to NI statistics, the cattle population in 2020 was about 1.6 million which is about 40 times higher than the badger population (Reid et al., 2012). Badgers are also known to be territorial, therefore not moving for long distances (Woodroffe et al., 1995). This implies that the spatial distribution of *M. bovis* in badgers will be more clustered in the area (Olea-Popelka et al., 2003, Woodroffe et al., 2005) and have less diversity in MLVA types.

Significant co-localization and spatial structure of *M. bovis* molecular types in livestock and wildlife populations was described in many research studies across the UK (Olea-Popelka et al., 2005, Woodroffe et al., 2005, Goodchild et al., 2012). Herd-level cattle surveillance studies in NI showed evidence of spatial clustering of specific MLVA types (Skuce et al., 2010; Skuce et al., 2020) and their associations with the geographic regions. Another study by Milne et al. (2020) was therefore consistent with the previous evidence that local spread was an important driver of the bTB epidemic in farm cattle and RTA badgers in NI. Current study findings also indicated that the spatial distribution of *M. bovis* in cattle and badgers was not random and the spatial clustering suggests that the local factors of bTB spread might be more important for the TVR area. However, the study occupied only a 100 km^2 territory and we cannot make conclusions about the whole NI, and these results must be interpreted with caution. But with regards to the *M. bovis* sampling densities in wildlife, this study data provided exceptional badgers' *M. bovis* information.

Many studies found significant associations of *M. bovis* distributions in cattle and badgers in Britain (Olea-Popelka et al., 2005, Woodroffe et al., 2005, Balseiro et al., 2013). However, identification of *M. bovis* transmission patterns in both host species may improve our understanding of the persistence and spread of the infection in the area. Genetic diversity of *M. bovis* in badgers and cattle showed some similarity, as seen from the four genotypes found in both species (MLVA 004, 006, 122 and 297), demonstrating evidence of intra- and inter-species transmission events. The use of KDE for the three shared MLVA types 004, 006 and 297 allowed to estimate the distributions of these *M. bovis* types.

Kernel density models were used to study associations between MLVA type distributions in cattle and badgers. Based on KDEs estimated for the shared MLVA types, MLVA types in cattle were predicted from the badger MVLA type distributions, and vice versa, within the TVR area. Using this approach, the level of inter-species transmissions was explored where the MLVA types shared between the two hosts was assumed to indicate such transmissions. There was some evidence of spatial clustering of *M. bovis* in both hosts; and the density kernels of MLVA types 004, 006 and 297 were largely overlapping, suggesting that transmission between cattle and badgers might account for some infection spread in the area. However, these cannot infer the direction of these transmissions. The results of the two prediction models in the absence of MLVA type 006 demonstrated therefore that there were some important patterns underlying the transmission of *M. bovis* in the TVR area. The spread of MLVA type 006 was studied by Skuce et al. (2020) and associated with the TVR region in cattle herds, therefore, it might be mostly driven by the cattle-to-cattle transmissions, and thereafter spilling over to the badger population. Whereas the results for MLVA types 004 and 297 might be explained by the stronger genetic differentiation between these types. While MLVA type 297 is only one VNTR tandem repeat different from MLVA type 006 (Table 2.1), MLVA type 004 is two tandem repeats different (Table 2.1) and was a sub-founder of a different group (Figure 2.4); it was likely brought in to the area (Skuce et al., 2020).

Many recent studies on *M. bovis* epidemiology in multi-host systems demonstrated the value of whole genome sequencing in detecting important links in transmission dynamics between cattle herds and badgers (Biek et al., 2012, Trewby et al., 2016, Patane et al., 2017) and quantification of the role of each host transmissions

(Crispell et al., 2019). Trewby et al. (2016) has shown that, the application of WGS is useful for understanding bTB epidemiology even when applied to *M. bovis* isolates from the same and/or closely related MLVA groups. Phylogenetic methods to reconstruct the infection transmission chains and understand the "who infected whom" question can be used in together with thorough epidemiological data (Kao et al., 2014). For example, study by Crispell et al. (2019) provided the evidence that badgers transmit *M. bovis* more often to cattle than vice-versa in Woodchester park area, with estimated transmission rates. However, due to the slow evolutionary rate of *M. bovis*, inferences about transmission between hosts must be interpreted with care and might be highly influenced by the rates at which these transmissions occur Kao et al. (2014). Considering the relatively large TVR area, 100 km^2, obtaining WGS data for the same isolates used in current study combined with spatio-temporal analyses might be effective in further investigations of *M. bovis* transmission patterns and processes within intervention area.

Overall, there was a strong association found between the distributions of *M. bovis* in cattle and badgers in the TVR area. These findings suggest the occurrence of both within and between species transmission, highlighting the need for improved control measures for both livestock and wildlife hosts. Based on experience in other areas, such as Great Britain and Ireland, badger culling can have positive or negative effects on bTB persistence and spread in both species (Donnelly and Nouvellet, 2013, van Tonder et al., 2021). Therefore, a deeper analysis of the obtained *M. bovis* data from badgers together with long term cattle movement data must be undertaken. Data on bTB infection in badgers in the first year (before culling) was important for understanding the distribution of *M. bovis* in wildlife pre-TVR. It will be interesting to explore the changes in badgers *M. bovis* persistence and movement dynamics if affected by badger-removal activities and vaccination of non-infected animals in years after the TVR study. Two of the aims of the spatial analysis described in this Chapter were to explore if *M. bovis* MLVA types found in badgers were spatially clustered together and whether the intervention study increased the distances badgers move from their setts, which could spread new MLVA types into other areas.

2.4.1 Limitations of the study

It is recognized that the TVR study implemented for a relatively short period of 5 years for *M. bovis* surveillance in wildlife, for bTB as an endemic disease (Menzies et al., 2021). However, this was the first wildlife intervention study in NI to obtain such thorough data for the under-sampled *M. bovis* badger population. Besides this, *M. bovis* isolates collected from cattle herds were essential in this analysis to demonstrate the occurrence of between species bTB transmission, which should be taken into account when designing future control strategies in cattle. While any temporal trends were not analysed for the bTB transmissions in cattle, this study included cattle *M. bovis* data from historical isolates. However, some infections might still be missing in cattle during the routine surveillance of bTB. Cattle location data represent farm locations (same for all cattle within farm), therefore, using more accurate locations for each cattle would help to track bTB transmission patterns.

Even though *M. bovis* molecular typing techniques such as MLVA typing has low resolution compared to WGS, they are useful for discrimination of the isolates at genetic and spatial scale. VNTR MLVA typing in NI, based on a specific loci combination for the region (Roring et al., 2002; Skuce et al., 2005), combined with spoligotyping (Aranaz et al., 1996) provides higher resolution data for typing of *M. bovis* strains. However, this is still not optimal for contact tracing or identification of the infection source and therefore using WGS analysis of these samples can improve the understanding of *M. bovis* epidemiology in the area.

3 Chapter

Evolutionary dynamics of *M. bovis* in a multi-host system in the TVR area

Parts of this analysis were performed in collaboration with researchers at the following institutions: University of Glasgow, Glasgow, UK; Centro de Investigation en Alimentacion y Desarrollo A.C., Hermosillo, Sonora, Mexico; Fios Genomics, Edinburgh, UK; Department of Infectious Diseases, College of Veterinary Medicine, University of Georgia, Athens, GA, USA; Foreign, Commonwealth and Development Office, Glasgow, UK; Agri-Food and Biosciences Institute, AFBI Stormont, Belfast, UK; University of Edinburgh, Roslin Institute, Edinburgh, UK; Department of Agriculture, Environment and Rural Affairs (DAERA), Belfast, UK; Department of Zoology, University of Oxford, Oxford, UK and Department of Agriculture Food and the Marine (DAFM), Dublin, Ireland.

This work has been published in bioRxiv pre-print doi:
https://doi.org/10.1101/2021.03.12.435101

3 Chapter

3.1 Introduction

Bovine tuberculosis (bTB), caused by *Mycobacterium bovis (M. bovis),* has a long history in the United Kingdom (UK), with known persistence in European badger (*Meles meles*) populations (Clifton-Hadley et al., 1993, Delahay et al., 2001). Bovine TB surveillance of the disease in cattle populations in Northern Ireland (NI) has contributed with valuable insights about the prevalence and spread of *M. bovis* strains in the region (Skuce et al., 2010; Skuce et al., 2020). Current bTB control strategy in NI is focused on disease surveillance in cattle through annual testing and post-mortem examination of positive animals, and further confirmation by molecular strain typing from identified lesions (Abernethy et al., 2006, TB Strategic Partnership Group, 2016). Badgers are protected by the law throughout the United Kingdom (1992) and historically only limited data are available regarding bTB in wildlife, mostly from animals killed in road traffic accidents (RTA). The test and vaccinate or remove (TVR) research study proposed by the Department of Agriculture Environment and Rural Affairs of NI started in 2014 and was aimed at controlling bTB in livestock by removing bTB positive badgers and protecting the uninfected ones (DAERA, 2018; Menzies et al., 2021).

Compared to currently used *M. bovis* molecular typing techniques such as spoligotyping (Kamerbeek et al., 1997) and VNTR-MLVA typing (Skuce et al., 2002), discrimination of strains based on whole genome sequencing (WGS) enables a more comprehensive genetic understanding of the pathogen (Kao et al., 2014). Garnier and colleagues (2003) were the first ones to sequence the whole genome of *M. bovis*, which was isolated from a virulent strain in Great Britain. Comparative genomic studies found that there is >99.95% similarity between *M. bovis* and *M. tuberculosis* at the nucleotide level, which showed a very close evolutionary relatedness of the two Mycobacterium Tuberculosis Complex (MTBC) members (Galagan, 2014, Guimaraes and Zimpel, 2020). There was also a high level of variation in gene sequences responsible for different functions, such as virulence, host adaptation, and antimicrobial resistance (Brosch et al., 2002, Brites and Gagneux, 2017, Guimaraes and Zimpel, 2020). The evolution of *M. bovis* has been shown to be clonal, with absence of recombination events (Smith et al., 2003, Smith et al., 2006, Ceres et al., 2022), however recent work has suggested the

existence of recombination sites and has recommended that recombination events in *M. bovis* should be taken in consideration and that further investigation is needed (Reis and Cunha, 2021).

The global distribution of *M. bovis* was studied in different research papers and it was grouped into for four main clonal complexes (CC), namely African 1, African 2, European 1 (Smith, 2012) and European 2 (Rodriguez-Campos et al., 2012) (Rodriguez-Campos et al., 2012). Moreover, a study by Zimpel et al. (2020) evaluated over 2,000 publicly available genomes and showed that the global *M. bovis* distribution depended more on geographic location than on host species and four *M. bovis* main lineages were identified different from the CC described above. With higher discriminatory power, WGS has the ability to distinguish between different strains, potentially allowing identification of infection sources, especially for pathogens like *M. bovis* with little genetic diversity (Kao et al., 2014, Hatherell et al., 2016). There are various publicly available and commercial tools for the analysis of raw sequencing data and SNP calling (Faksri et al., 2016, Sukumar et al., 2021). With the availability of advanced phylogenetic tools, it became possible to use SNP data to study the evolutionary dynamics of the pathogens and monitor disease outbreaks (Bentley and Parkhill, 2015, Kao et al., 2016, Trewby et al., 2016). Previous studies have used whole genome sequence (WGS) data to identify single nucleotide polymorphisms (SNPs) in the *M. bovis* genome, determine the evolutionary dynamics of *M. bovis*, and infer inter-species transmission events in particular geographical locations (Biek et al., 2012; Trewby et al., 2016; Crispell et al., 2017; Crispell et al., 2019; Salvador et al., 2019).

The association of bTB incidence in cattle and badgers has been widely assessed, and it has been demonstrated that there are high genetic and spatial distribution similarities of *M. bovis* molecular types sampled from cattle and badgers (Olea-Popelka et al., 2005; Woodroffe et al., 2005; Milne et al., 2020). Research on *M. bovis* epidemiology in multi-hosts systems with the use of WGS technologies has shown little *M. bovis* genetic diversity obtained from the available cattle and badger data, suggesting possible transmission events between the two species, but with the small number of wildlife hosts, it was not possible to show the directionality of these transmissions (Biek et al., 2012, Trewby et al., 2016). The first study that could estimate such direction was in Woodchester park area in Great Britain, where transmission from badgers to cattle played a bigger role than

from cattle to badgers (Crispell et al., 2019). In contrast, a study by Rossi et al. (2022) in East Cumbria demonstrated the lesser role of badgers in the inter-species dynamics of *M. bovis* compared to cattle hosts. A study by van Tonder et al. (2021) in the South West of England showed that in the identified transmission clusters, badger to cattle *M. bovis* transmissions occurred more often than from cattle to badgers.

Other research has used genomic data and contact tracing to identify common factors responsible for the spread of other infection diseases such as human tuberculosis (Gardy et al., 2011). In a study by Colijn and Gardy (2014), it was demonstrated that transmission patterns can be explained based on phylogenetic tree topologies estimated from genomic data. Ideally, sampling of isolates must represent the pathogen population, but as this is not always possible and many important ancestors can be missed, the interpretation of transmission events must be given with caution. Previous studies have estimated the *M. bovis* minimum genetic distances between isolates found in two host-species from which direct transmissions can be inferred (Bryant et al., 2013; Bentley & Parkhill, 2015). Walker et al. (2013) suggested that the divergence for the epidemiologically linked *M. tuberculosis* isolates can occur within 5-12 SNPs difference, while divergence in more than 12 SNPs between isolates demonstrate that they cannot be considered as a direct transmission event. Walker et al. (2013) suggested that the divergence for the epidemiologically linked *M. tuberculosis* isolates can occur within 5-12 SNPs difference, while divergence in more than 12 SNPs between isolates demonstrate that they cannot be considered as a direct transmission event. These findings can be useful in inferring the transmission patterns of *M. bovis* between host species, where isolates that are genetically closer to each other are associated with the same bTB infection source. Furthermore, the combination of pathogen genomic data with epidemiological data, animal movements and/or other potential contact information can give a more thorough understanding of *M. bovis* dynamics and epidemiology.

The reconstruction of pathogen transmission trees using graph methods to understand “who infected whom” uses sequencing data and dates of sampling (Jombart et al., 2014, Hall et al., 2015). Inference of *M. bovis* transmission networks is complicated by the bacterium slow generation times and slow

mutation rates (Skuce & Neill, 2001). Although the topology of phylogenetic and transmission trees can be similar, in phylogenetic trees the internal nodes represent the most recent common ancestors and are considered potential sources of transmission; in transmission trees, the links connect isolates that are available, thereby assuming that there are no unsampled nodes (isolates) (Ypma et al., 2012).

In the current analysis, I investigated the evolutionary dynamics of *M. bovis* amongst cattle and badgers in a densely sampled area with known bTB history through the TVR study. Specifically, I studied the evolutionary history of the *M. bovis* endemic lineage present in the TVR area using both phylogenetic and transmission tree tools on an extensive *M. bovis* WGS data collected from infected cattle and badgers during 2014-2017, together with historically sampled *M. bovis* isolates collected from cattle and RTA badgers during 1986-2013 (Akhmetova et al., 2021).

The current analysis was aimed to understand the following objectives:
1. Determine *M. bovis* population genetic structure and relatedness of the strains isolated from cattle and badgers
2. Estimate the evolutionary rate of *M. bovis*, identify the time of the most recent common ancestor based on the *M. bovis* reconstructed phylogenetic trees, and to model *M. bovis* population size changes through time in the TVR area.
3. Infer important inter-species transmission dynamics of *M. bovis* using different methods.

3.2 Materials and methods

3.2.1 Sampling of M. bovis from cattle and badgers

Data for this analysis were provided by Agri-Food and Biosciences Institute (AFBI) in Northern Ireland (NI). The *Mycobacterium bovis* isolates were sampled from tuberculosis infected cattle and badgers within the TVR study undertaken by the Department of Agriculture, Environment and Rural Affairs (DAERA) of NI during the years 2014-2017. This intervention study was implemented to study the dynamics of bovine tuberculosis (bTB) in cattle and badgers in Down County, which was

known to have high prevalence of bTB in previous years and with a relatively high badger density (~3.88 animal/km^2) (Reid et al., 2012).

In order to obtain the necessary information regarding badger setts and bTB prevalence in badgers in the area, no badgers were culled in the first year of the TVR study. Badgers were trapped, captured and tested for *M. bovis* using dual path platform (DPP) (Ashford et al., 2020) serological tests. Test-negative badgers were vaccinated in all years using Bacillus Calmette Guerin (BCG) vaccine. If badgers tested positive in years 2015-2018, they were culled, and carcasses sent for post-mortem examination, samples sent for laboratory confirmation using bacteriological culture and molecular typing (Menzies et al., 2021).

In parallel, cattle underwent bTB compulsory testing within the TVR intervention area and within a 2 km buffer area in the same years 2014-2017. Bacteriological culturing of historical *M. bovis* cattle isolates from previous years, 1986-2013, were undertaken from the same geographic area and added to the analysis. Badger *M. bovis* historical isolates were selected from the RTA samples (Courcier et al., 2018) and also included in the study.

Overall, for the current study, a total of 642 *M. bovis* isolates were collected from the bTB infected cattle and badgers and sequenced within the TVR intervention study with an additional (31) from the neighbouring zone. These also included 15 re-sequenced (duplicates) and the reference genome AF2122/97 (Malone et al., 2017, Farrell et al., 2020) as controls. *M. bovis* isolates collected from cattle and badgers used for WGS in different years are described in Table 3.1.

Table 3.1 Distribution of *M. bovis* isolates collected from cattle and badgers, and whole genome sequenced within the TVR study area in different years

Years of sampling	Host	Number of isolates	Study
1986-2013	Cattle	233	Historical
1986-2013	Badgers	10	RTA
2014-2017	Cattle	282	BTB Surveillance (within TVR area)
2015-2018	Badgers	117	TVR

3.2.2 M. bovis isolates preparation and molecular typing

M. bovis isolation from infected cattle and badger tissues was undertaken in a biosafety category 3 laboratory (BSL3) in the Bacteriology branch of the Veterinary Science Division (VSD), AFBI. All bacteriological cultures were isolated according to the internal laboratory standard operational procedures (SOPs). Initial cultures were grown on three different selective media: BACTEC MGIT vials, egg-based Stonebrinks medium and Löwenstein-Jensen (LJ) slants. *M. bovis* single colonies were isolated and heat inactivated (30 mins at 80^0C), and the DNAs were extracted using the ionic detergent cetyltrimethylammonium bromide (CTAB) method (van Helden et al., 2001). Extracted DNA were sent for molecular typing using spoligotyping (Kamerbeek et al., 1997) and MLVA-VNTR typing techniques using 8 loci specific for NI (Skuce et al., 2010).

Figure 3.1 (a, b) shows the distribution of *M. bovis* isolates sampled from cattle and badgers in different years and the distribution of *M. bovis* MLVA types and spoligotypes identified in isolates in both species. Further in this study, *M. bovis* lineages were shown as MLVA/spoligotype groups, e.g. isolates belonging to MLVA type 6 within spoligotype 263, as 6.263 *M. bovis* lineage.

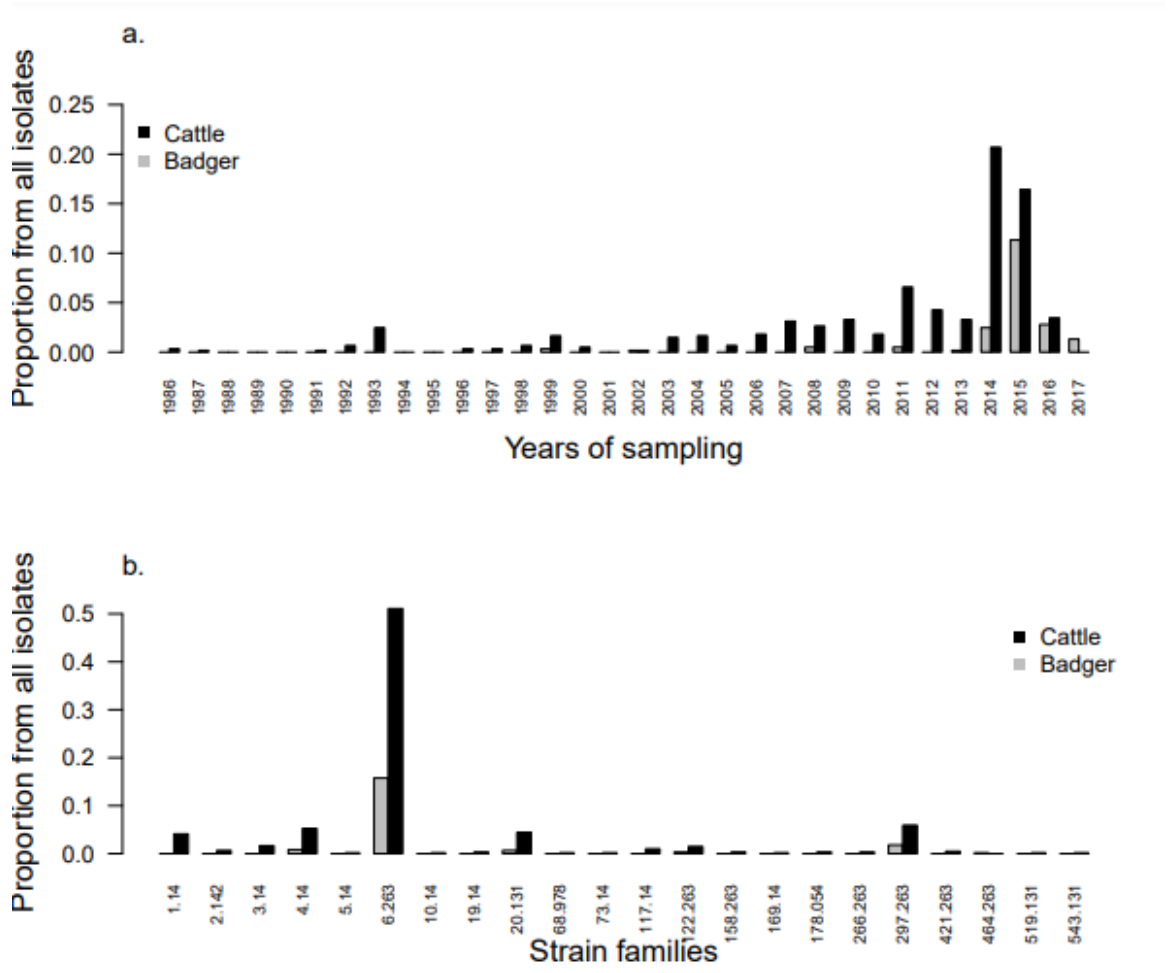


Figure 3.1 (a, b) *M. bovis* isolates sampled from cattle and badgers

Distribution of *M. bovis* isolates collected from bTB infected cattle (black bars) and badgers (grey bars) as a proportion from all *M. bovis* isolates by a. years of sampling; b. MLVA type/spoligotype groups.

Most of the *M. bovis* badger isolates were identified to belong to the spoligotype 263, MLVA types 006, 122, 297 and 464. However, more than 80% of *M. bovis* found in badgers were found to be part of the endemic lineage 6.263 in years 1999-2017. Very little number of isolates belonged to strain families 20.131 and 4.140 where multiple samples were taken from two and five unique badgers respectively in years 2015-2017.

3.2.3 M. bovis whole genome sequencing and bioinformatics analysis

Illumina Nextera XT kits were used for preparation of *M. bovis* DNA libraries according to the protocol, and isolates were sent for sequencing in three laboratories: AFBI, University of Glasgow Polyomics facility and Eurofins Scientific (Akhmetova et al., 2021). A total of 100 isolates were sequenced using Illumina

Miseq with paired-end reads of 250 bp. Another 100 isolates were sequenced using the same platform but producing paired-end reads of 2x300bp. The remaining isolates were sequenced using the Illumina HiSeq platform and producing 2 x250 bp paired-end reads. Additional analysis was undertaken in AFBI, where 15 (random) samples were chosen to be re-sequenced and compared with initial isolates. Bioinformatic analysis of sequencing data was performed using RedDog V1beta.10.3 (Edwards et al., 2015) and reads were mapped to the *M. bovis* AF2122/97 reference genome (GenBank record LT708304.1) (Malone et al., 2017, Farrell et al., 2020). Alignment and mapping were performed in Bowtie2 v2.2.9 (Langmead and Salzberg, 2012) (Langmead and Salzberg, 2012), and SAMtools and BCFtools were used for SNP calling (Li et al., 2009, Danecek et al., 2021). Filtering parameters were shown in Figure 3.2.

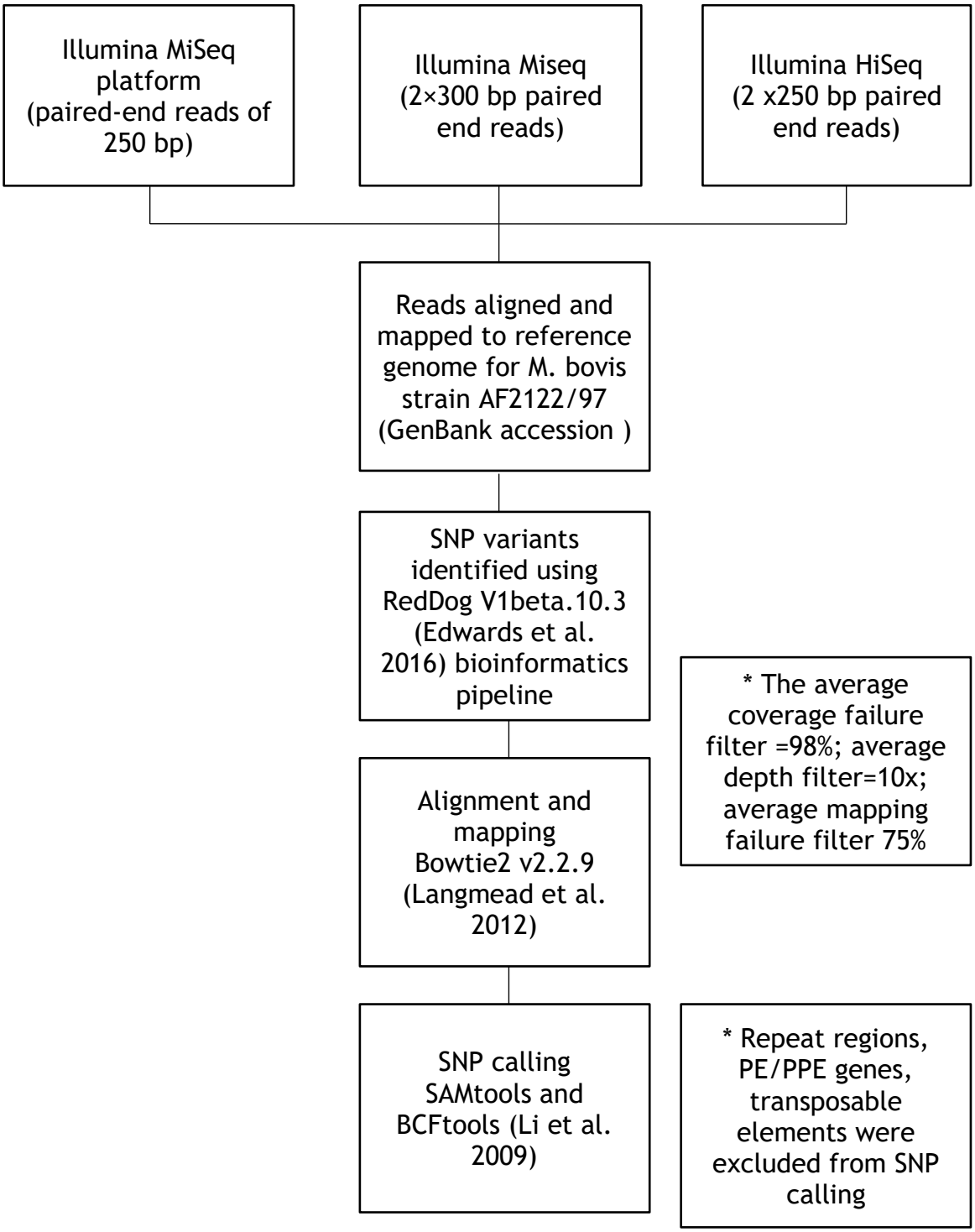


Figure 3.2 Workflow for genome sequencing and bioinformatic analyses performed on all the *M. bovis* isolates

3.2.4 Phylogenetic reconstruction of Mycobacterium bovis

The historical endemic *M. bovis* lineage MLVA type 006 and spoligotype 263 described in Skuce et al. (2020) was the focus of the phylogenetic analyses, which included 302 isolates collected from cattle (n=248) and badgers (n=54). To

compare different nucleotide substitution models, a FASTA alignment with the detected informative SNPs was used as input data into the 'modelTest' function from the package 'Phangorn' (version 2.8.1) (Schliep, 2010) in R (R Core Team, 2021). First, the alignment was imported and transformed the data into the required "phyDat" format. According to the 'modelTest' results, the best fitting models for the dataset were the General Time Reversible (GTR) (Tavaré, 1986) and Hasegawa Kishino Yano (HKY) (Hasegawa et al., 1985), with the lowest AIC: 26875.16 for the GTR substitution model. This model was used to build the Maximum Likelihood (ML) tree. The ML phylogenetic tree model was optimized using the 'Phangorn' (version 2.8.1) package (Schliep, 2010) in R (R Core Team, 2021) and visualized in FigTree 1.4.4 (Rambaut, 2010).

To assess whether the molecular data had sufficient temporal signal to perform an evolutionary dynamics analysis, the TempEst 1.5.1 software was used (Rambaut et al., 2016). The existence of temporal signal in the data represents a strong association between mutation rate and the time of sampling. Evolutionary relationships among *M. bovis* isolates were generated using a Bayesian Coalescent analysis performed using Beast2 (Bayesian Evolutionary Analysis by Sampling Trees) (Bouckaert et al., 2019) software using a GTR nucleotide substitution model, a coalescent constant population model and a relaxed log normal clock model. Three independent Markov chain Monte Carlo (MCMC) analyses were run for 100,000,000 iterations with a 10% burn-in. The results were combined in LogCombiner (Bouckaert et al., 2019) and visualized in Tracer 1.7.1 (Rambaut et al., 2018). Tracer software provides statistical mean, median and standard deviation values, highest posterior density (HPD) intervals for the overall performance of the phylogenetic analyses, traces for posterior, prior and likelihood values. Model parameters were assessed for convergence and sufficient Effective Sample Sizes (ESSs) for the categorical and continuous parameters (ESS >200), which means that the number of independent samples from the posterior distribution was sufficient for the analyses.

M. bovis clades were identified with a posterior probability (PP) value for the internal nodes > 0.95. The *M. bovis* substitution rate as well as the times (and correspondent 95% HPD intervals) for the most recent common ancestors (TMRCA) were estimated for the data and compared to the values identified in previous research studies.

The past population dynamics of the *M. bovis* endemic lineage, were examined using the coalescent Bayesian Skyline analysis in Beast2 (Drummond et al., 2005). For this analysis GTR nucleotide substitution model with relaxed clock model with default number of dimensions to estimate the effective population size (Ne) were used. In 100,000,000 MCMC chain length and 10% burn-in in three replicates. The obtained log files were first combined in LogCombiner (Bouckaert et al., 2019) and assessed in Tracer 1.7.1 (Rambaut et al., 2018).

3.2.5 Using host-species as discrete traits for ancestral state reconstruction

The Discrete Ancestral Trait Mapping (DATM) approach implemented in Beast2 (Bouckaert et al., 2019) software in the Beast-Classic package was used to reconstruct the *M. bovis* phylogenetic tree, using host-species (cattle and badgers) as discrete traits. The host-state associated posterior probability (PP) were reported for each trait associated to the internal nodes and branches of the tree. For the current analysis, two different sets of specifications were used for the BEAUti files. The alignment was imported to BEAUti, and dates of sampling were specified. For the first model, GTR nucleotide substitution model, relaxed molecular clock and coalescent constant population models were used. The two state host species (badger and cattle) analysis estimated over time the posterior probability that *M. bovis* transitions between the two states (badger-> cattle or cattle->badger). If the probabilities are high, then the data strongly support (evaluated by Bayes's factor values) a model of asymmetric transitions between the two host-species. The second model analysis was performed using the same GTR substitution model and relaxed clock, but with a Bayesian Skyline population model. Three independent simulations were run for both models, with a 500,000,000 MCMC chain length, discarding 10% burn-in and storing every 50,000 samples. Beast2 outputs (.log and .trees) were combined in LogCombiner (Bouckaert et al., 2019) and analysed in Tracer 1.7.1 (Rambaut et al., 2018) (ESS > 200). The .trees files were combined, and the Maximum Clade Credibility tree was estimated and annotated using TreeAnnotator and visualized in Figtree 1.4.4 (Rambaut, 2010). The posterior probability values higher than 50% were shown for the host associated PP support for the nodes. After analysing the Bayesian skyline plot for any possible changes in the demographics of *M. bovis* population over time, the initial dataset was split into two subsets based on times of population

expansion (2011). Beast2 analyses were performed on both datasets of *M. bovis* before the population expansion. Subset 1 included *M. bovis* WGS data collected from cattle and badgers in years 1986-2011 and the second subset from 2012 to 2017. The bTB incidence rates increased in those years and therefore, the DATM analyses were repeated using two models: GTR substitute model, relaxed clock model, Coalescent constant population and second using GTR substitute model, relaxed clock model, Bayesian skyline population model, both for 500,000,000 MCMC and 10% burn-in. Both models were run in three repetitions.

3.2.6 Transmission tree reconstruction

Different approaches can be used to reconstruct transmission trees using genomic and epidemiological data (Ypma et al., 2012). One of the available algorithms is SeqTrack (Jombart et al., 2011), which aims to reconstruct the most likely genealogy directly from a sampled dataset. This method uses genetic information and sampling dates to identify the most plausible ancestors between the sampled isolates. SeqTrack is based on graph theory and assumes that no recombination has occurred. For each isolate, only one ancestor can be assigned, which has evolved earlier in time with weighting minimum pairwise genetic distances. SeqTrack was run using the *adegenet* package (Jombart, 2008) in R (R Core Team, 2021) and provided as output a network with isolates as nodes and genetic distance (in SNPs) as branches.

To compare if the relationship between isolates remains the same, I also used another method to reconstruct transmission trees - the R package (R Core Team, 2021) *outbreaker2* (Campbell et al., 2018) tool. This method reconstructs outbreaks using a Bayesian framework based on genomic data, pathogen generation time and sampling times of isolates. The *M. bovis* generation time prior distribution (i.e. the time interval of infections between the primary and secondary cases) and the *M. bovis* incubation period prior (i.e. the time interval between infection and the first symptoms), were taken from the systematic review by Ma et al. (2018) and from articles that described the average time periods of animals being infected with bTB before being detected (Pollock and Neill, 2002, van Tonder et al., 2021). The lowest and highest generation time values were described to be 0.57 and 3.5 years, respectively (Ma et al., 2018). Based on the proposed generation time prior distribution, the values used for the

current analysis were from 0.5 to 3 years. The incubation time period distributions differed between studies, and were based on specific features of the studied bTB outbreaks, with means of 0.56 years (van Tonder et al., 2021) and 1.4 years (Didelot et al., 2017). Based on these studies, the current analysis used an incubation time period from 0.5 to 2 years. Default values were used for the remaining prior parameters such as the values for the prior distributions of the reporting probability and of the mutation rate (between 0.8 and 1, and between 0 and 1, respectively). The MCMC chain length of 500,000 iterations were run with sampling from the posterior distribution every 1,000 iterations. Estimations of the mutation rate (per site per generation), the number of generations between the infected case and its ancestors (assumed to be the number of unobserved cases, 50%), and the number of mutations among cases were derived under the *outbreaker2* genetic likelihood model (Campbell et al., 2019). As outputs, I generated the transmission tree with the highest likelihood values, the posterior probability support for each node, and the inferred date of infection.

3.3 Results

3.3.1 M. bovis phylogenetic reconstruction

From the 619 *M. bovis* WGS isolates sampled between 1986 and 2017 from badgers and cattle, we identified a total of 1562 SNPs. The ML tree reconstructed from the full dataset of 619 isolates is shown in Figure 3.3, where the strain families are represented for each clade. We identified five major clades that include all the *M. bovis* isolates and their corresponding MLVA types, having the reference genome AF2122/97 (Malone et al., 2017, Farrell et al., 2020) as an outgroup. The biggest clade (shown in blue) represents the *M. bovis* spoligotype 263 and family MLVA types 006, 122, 158, 297, 421, 464. The second biggest clades (shown in pink) includes three major spoligotypes 140, 142 and 978; and MLVA types 001, 002, 004, 005, 010, 068, 073, 169, 117. The 20.131 strain family was indicated as a separate clade (shown in red), which has the longest branch lengths, and according to a previous study by Allen et al. (2013) it was derived from a different common ancestor. Two smaller clades shown in green and in dark blue represent, respectively, strain families 3.140 and 19.140.

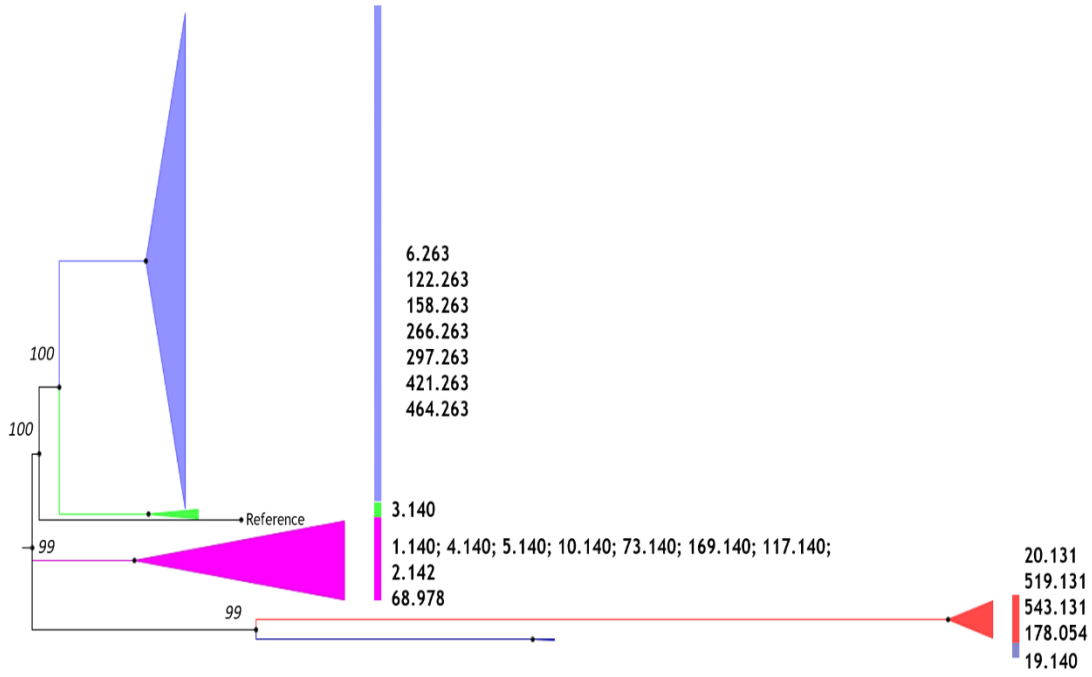


Figure 3.3 Maximum Likelihood tree reconstruction for 619 *M. bovis* isolates collected from badgers and cattle in Northern Ireland between 1986 and 2017

Different colours represent different clades and their correspondent strain families. The *M. bovis* reference genome AF2122/97 (Malone et al., 2017, Farrell et al., 2020) was used as an outgroup. Bootstrap values are labelled for the internal nodes of each clade.

The distribution of pairwise genetic distances (Figure 3.4) illustrated that most of the isolates differ only in 5-10 SNPs (mean 7.5 SNPs) from each other in the biggest clade (blue) of spoligotype 263. This endemic *M. bovis* lineage (part of the biggest clade, blue in Figure 3.3) has been found to be prevalent within the TVR region for the past decades (Skuce et al., 2020). The long history of this lineage in the study area, and its presence in both resident cattle and badgers made it as an ideal dataset to to assess the transmission dynamics using phylogenetic based methods. Consequently, phylogenetic analyses were conducted for the *M. bovis* 302 isolates (248 from cattle, 54 from badgers). Using these isolates another ML phylogenetic tree was constructed and five major clades were identified (Figure 3.5).

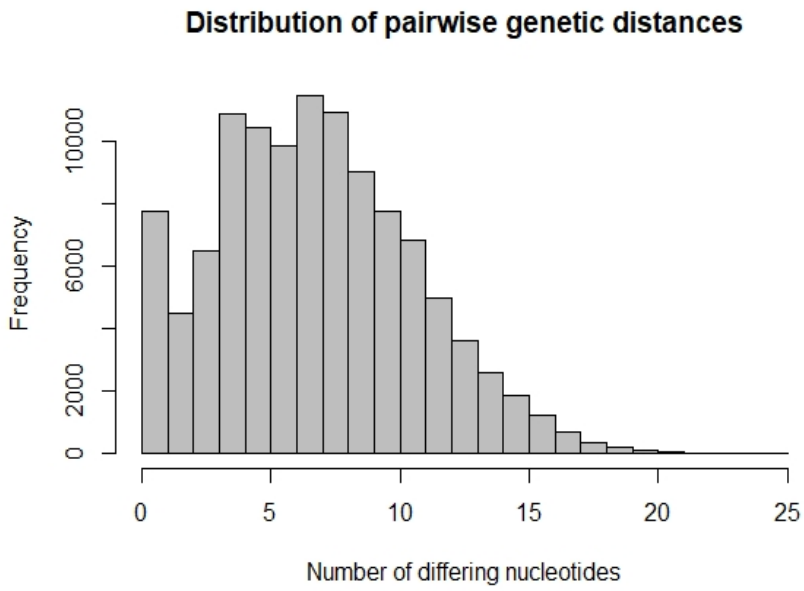


Figure 3.4 Distribution of genetic distances between sampled isolates of *M. bovis* spoligotype 263 (total of 421 isolates sampled from cattle and badgers).

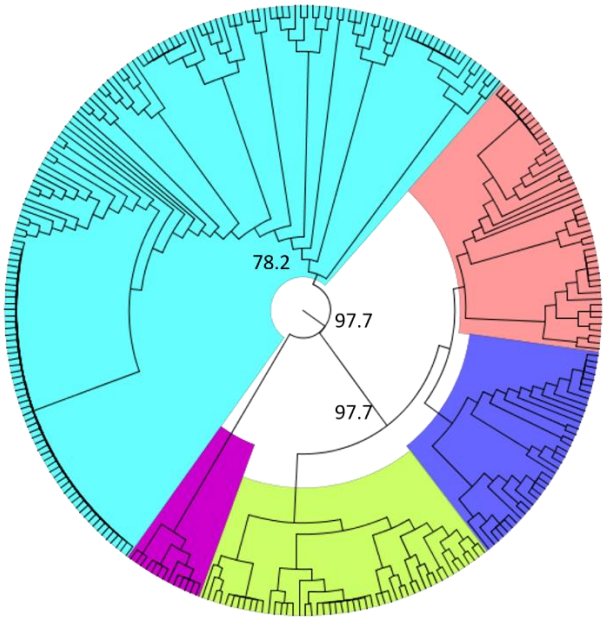


Figure 3.5 Maximum likelihood tree of *M. bovis* isolates belonging to the 6.263 strain family
302 isolates sampled from cattle and badgers during 1986-2017 in the TVR area – one isolate per affected farm per year, and one isolate per affected badger.

The investigation of the presence of temporal signal in the *M. bovis* endemic subset (302 isolates) using TempEst and receiving as input the maximum likelihood phylogeny (including the reference genome) and the isolates correspondent collection dates showed a linear relationship between genetic divergence of the

M. bovis data and sampling times. The TempEst results for the full dataset (n=302) (Table 3.2, Figure 3.6) support the existence of temporal signal in the endemic 6.263 lineage (correlation coefficient >50%, R^2=0.27).

Table 3.2 TempEst results for the 302 *M. bovis* isolates (endemic clade) to determine the existence of temporal signal in the data

Date range	31
Slope (rate)	1.8446E-4
X-Intercept (TMRCA)	1980.6381
Correlation Coefficient	0.5232
R squared	0.2737
Residual Mean Squared	3.2351E-6

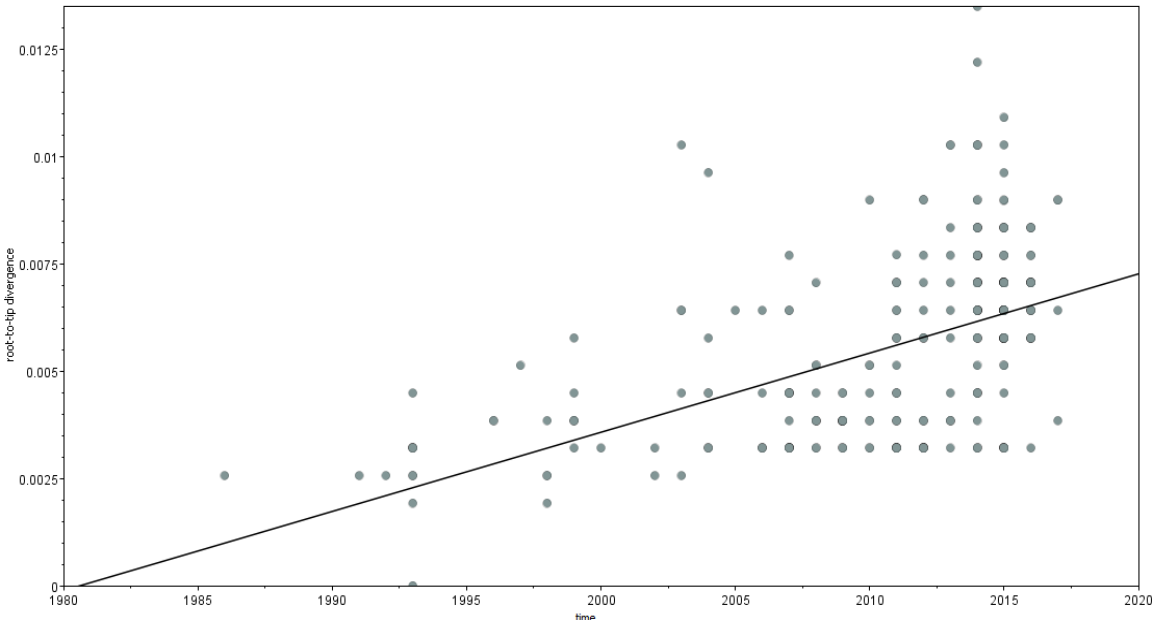


Figure 3.6 TempEst plot root-to-tip divergence over time for ML tree

302 *M. bovis* isolates, time period between 1986 and 2017. TempEst plot showing the correlation between the evolutionary rate of the data and the time of sampling.

The time-measured phylogenies estimated under a GTR substitution model, and relaxed log normal clock and the constant population size demographic model estimated the mean evolutionary rates of 0.37 substitutions per genome per year (highest density interval (HPD) 95%: 0.27-0.47). Similar substitution rate was estimated using Bayesian skyline demographic model, with 0.36 substitution per genome per year (HPD 95%: 0.28-0.45). These results are consistent with the

results of previous studies of *M. bovis* in different regions and host systems (Table 3.3) (Trewby et al., 2016, Crispell et al., 2017, Crispell et al., 2019, Salvador et al., 2019).

Table 3.3 *M. bovis* substitution rates (per genome per year) obtained from previous studies for bTB genomic epidemiology compared to current study results

Research study	Substitution rate	HPD 95%
Trewby et al. (2016)	0.20	[0.1-0.3]
Crispell et al. (2017)	0.53	[0.22-0.94]
Salvador et al. (2019)	0.37	[0.24–0.51]
Crispell et al. (2019)	0.28	[0.21-0.37]
Current study	0.36	[0.28-0.45]

The posterior distribution of the estimated evolutionary rate sampled in Beast2 under the Bayesian skyline demographic model for 302 *M. bovis* isolates is shown in Figure 3.7. The mean of the times of the most recent common ancestors (TMRCA) differ between the two phylogenetic analysis: using the constant population model, the TMRCA is 1968 while for the Skyline population model the TMRCA is 1984. However, the HPDs 95% interval of the Skyline population model overlaps totally with the HPD 95% interval of the constant population model: [1982, 1986] versus [1951, 1991], respectively. The Skyline population model was able to estimate the TMRCA in a much shorter interval of the posterior distribution than the constant population model (Table 3.4). Figure 3.8 illustrates the Bayesian skyline demographic plot, showing the effective population expansion in years 1990, with a second change in 2011-2012. Based on this analysis, the dataset was split into two main datasets, which represent the *M. bovis* population before and after the expansion occurred. Subset 1 included 124 *M. bovis* isolates collected from years 1986-2011, whereas subset 2 included 178 isolates collected from years 2012-2017.

Table 3.4 Substitution rates and TMRCA of two coalescent-based phylogenetic analyses (Tracer results)

Data	Demographic model	MRCA, year	Molecular Clock rate [HPD 95%], substitutions per genome per year
Full data 1986-2017	Constant	1968 [1951, 1991]	0.37 [0.27; 0.47]
Full data 1986-2017	Skyline	1984 [1982, 1988]	0.36 [0.28; 0.45]

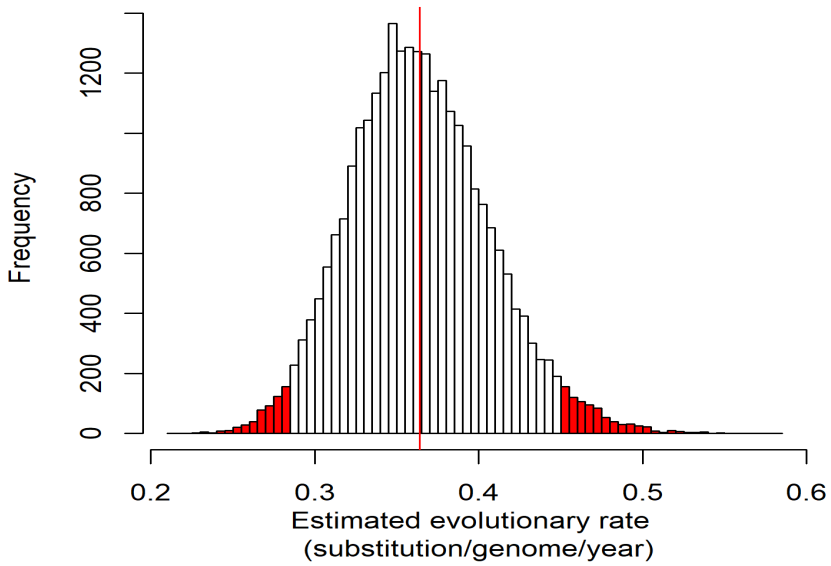


Figure 3.7 Posterior distribution of the estimated evolutionary rate of *M. bovis*

GTR substitution model, relaxed log normal molecular clock and the Skyline population model. Mean rate 0.36 substitutions per genome per year.

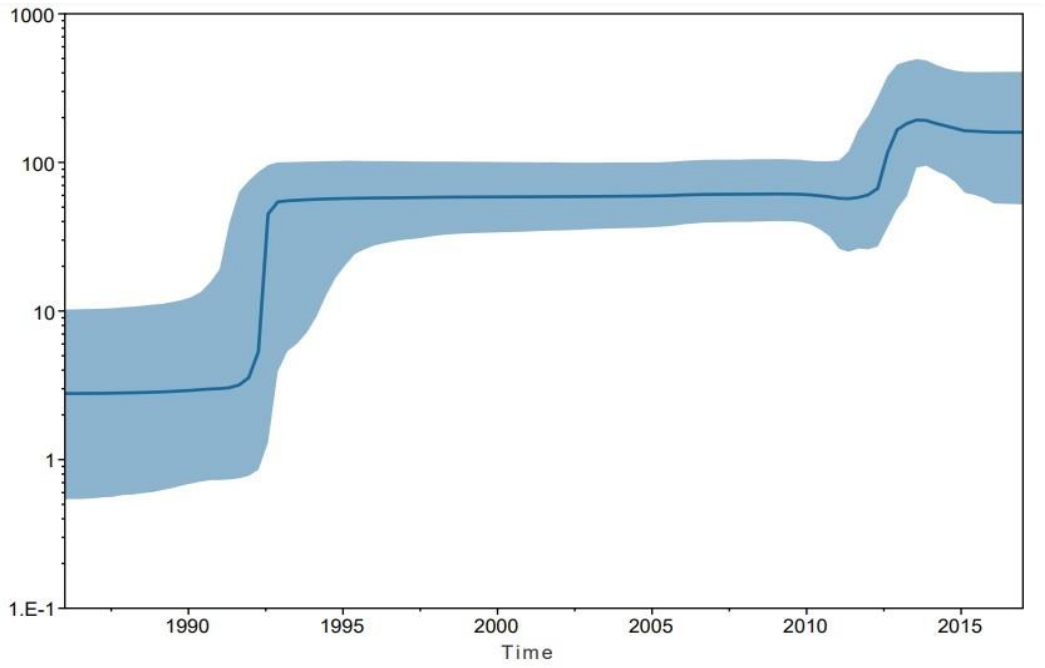


Figure 3.8 Bayesian skyline plot estimated from reconstructed phylogeny of 302 *M. bovis* isolates

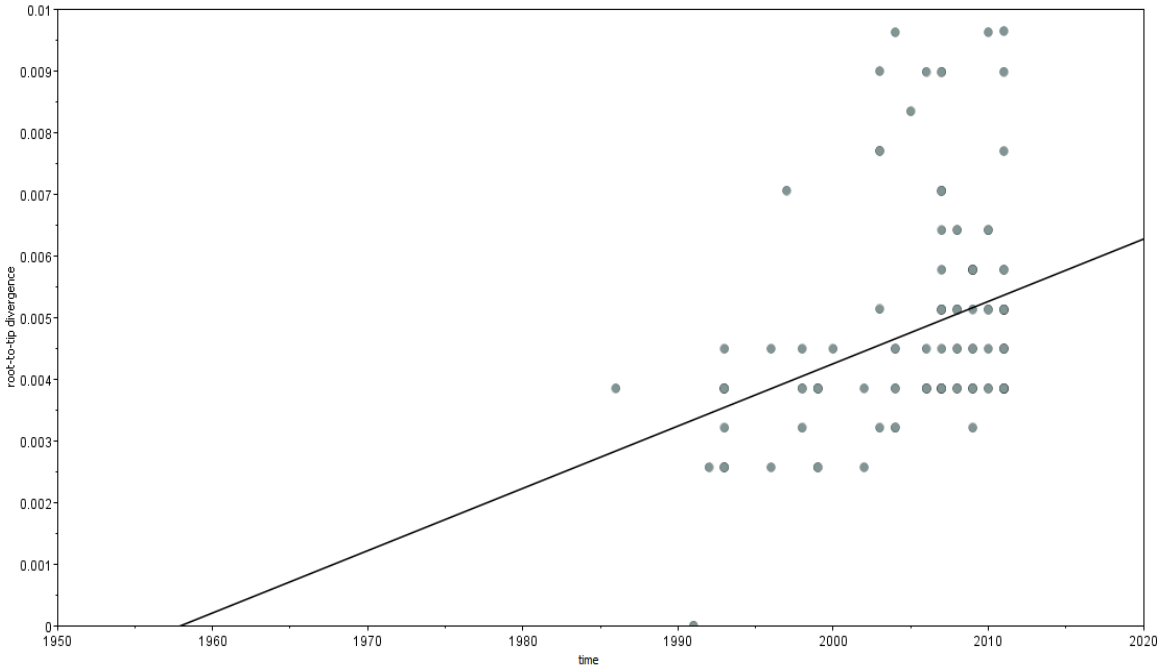
Isolates collected from cattle and badgers in 1986-2017 years, under GTR substitution model, relaxed log normal molecular clock model.

TempEst results for the *M. bovis* subsets 1 and 2 are shown on Figures 3.9 (a, b) indicating significant temporal signal in the two datasets. The statistical variables for the temporal analysis for the subsets are shown in Table 3.5.

Table 3.5 TempEst results for *M. bovis* endemic clade subset 1 and subset 2 data

	Subset 1 (1986-2011)	Subset 2 (2012-2017)
Date range	25	5
Slope (rate)	1.0105E-4	6.8597E-4
X-Intercept (TMRCA)	1957.9072	2001.6623
Correlation Coefficient	0.3519	0.3286
R squared	0.1239	0.108
Residual Mean Squared	2.5808E-6	5.8721E-6

a.



b.

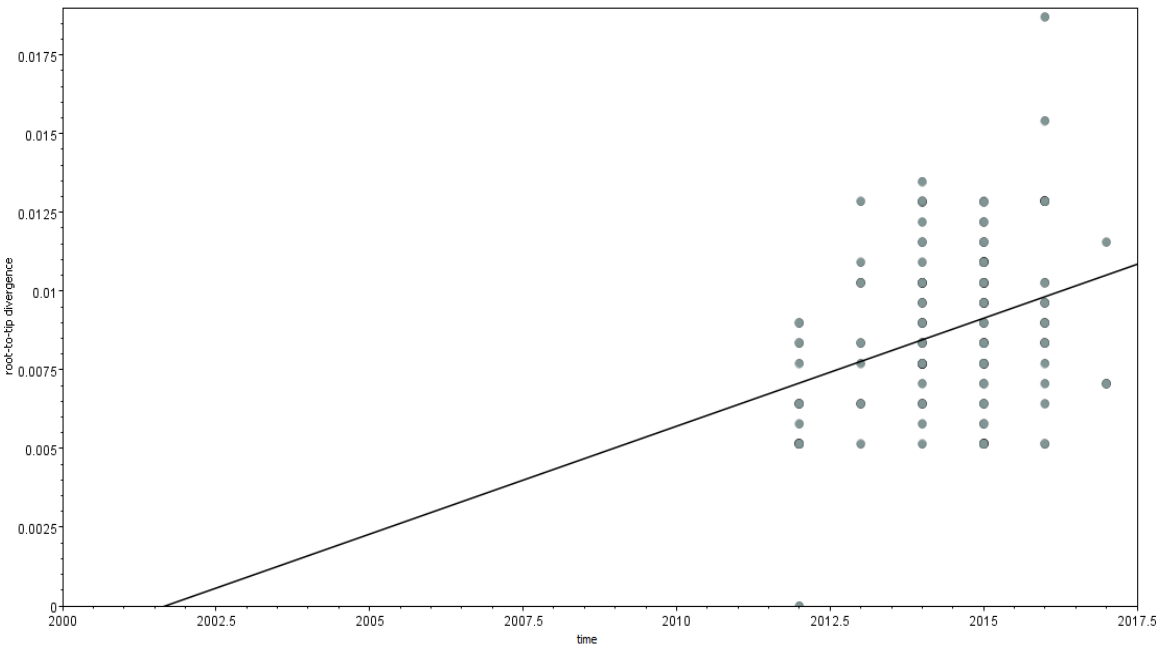


Figure 3.9 (a, b) TempEst plot root-to-tip divergence over time for ML trees

a. *M. bovis* subset 1 data, from time period between 1986 and 2011. b. *M. bovis* subset 2, 2012-2017

3.3.2 Investigation of inter-species transmission

The reconstruction of the time-calibrated phylogenetic trees with asymmetric ancestral trait reconstruction was performed on the three datasets (main, subset 1 and subset 2 datasets), using the models specified in Table 3.6.

Table 3.6 Ancestral trait reconstruction of *M. bovis* isolates collected from cattle and badgers.

Data	Number of isolates	Substitution model	Clock model	Population model
Full data 1986-2017	302	GTR	Relaxed Log Normal	• Constant population • Bayesian Skyline
Subset 1 1986-2011	124	GTR	Relaxed Log Normal	• Constant population • Bayesian Skyline
Subset 2 2012-2017	178	GTR	Relaxed Log Normal	• Constant population • Bayesian Skyline

The Tracer statistics after the results were combined for the independent runs for each model showed that chains converged and mixed adequately, after the initial burn-in (10%) was discarded. The Effective Sample Sizes (ESSs) as the correlation between sampled entries in the chain for all the parameters were higher than 200 and indicated statistically efficient sampling from a probability distribution. Table 3.7 demonstrates the mean values of the root heights (estimated the oldest point in the tree, years) and respective 95% HPD interval for the subset datasets.

Table 3.7 Estimated root height (Years) and clock rates for the subset datasets

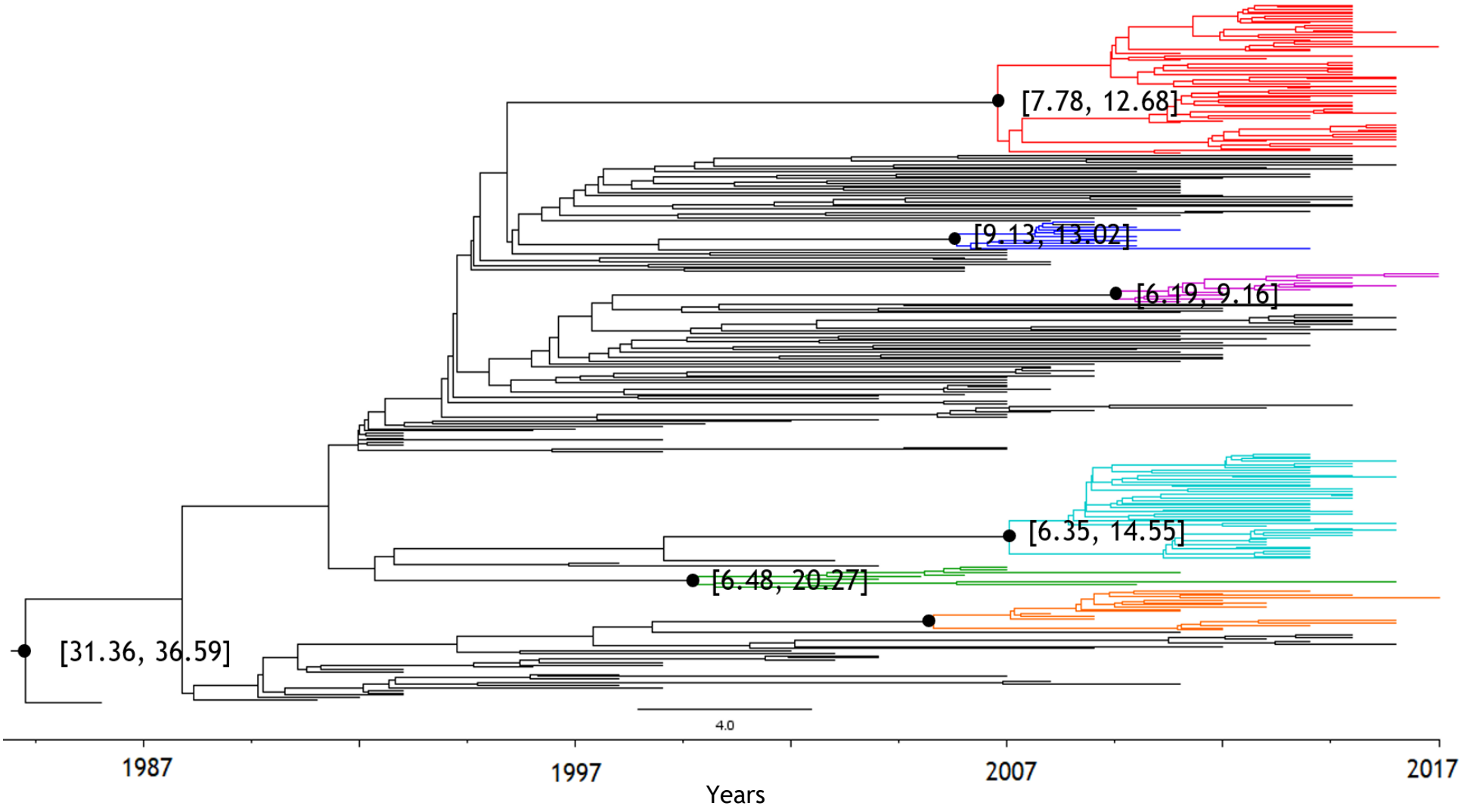
Data	Demographic model	Root Height (Years) [HPD 95%]
Subset 1 (1986-2011)	Constant	47.35 [28.24, 71.09]
	Skyline	30.26 [25.00, 46.79]
Subset 2 (2012-2017)	Constant	47.66 [15.21, 99.37]
	Skyline	16.23 [7.82, 30.23]

These analyses demonstrated similar results (when using the same models) between the different subset data and the full *M. bovis* dataset of the endemic

clade. The TMRCA dated between 1970s (for constant population model) and 1980s (for Bayesian skyline population model) for the full dataset, which included *M. bovis* isolates from years 1986-2017.

The time-calibrated maximum clade credibility tree under the GTR model, relaxed log normal and the Bayesian skyline population model with posterior probability (PP) support for major nodes is shown (Figure 3.10a). Host species were modelled as a discrete trait over the full *M. bovis* dataset genealogy by ancestral state reconstruction using Discrete Ancestral Trait Mapping (DATM) approach in Beast2.

a.


[7.78, 12.68]
[9.13, 13.02]
[6.19, 9.16]
[6.35, 14.55]
[6.48, 20.27]
[31.36, 36.59]
4.0
1987
1997
2007
2017
Years

b.

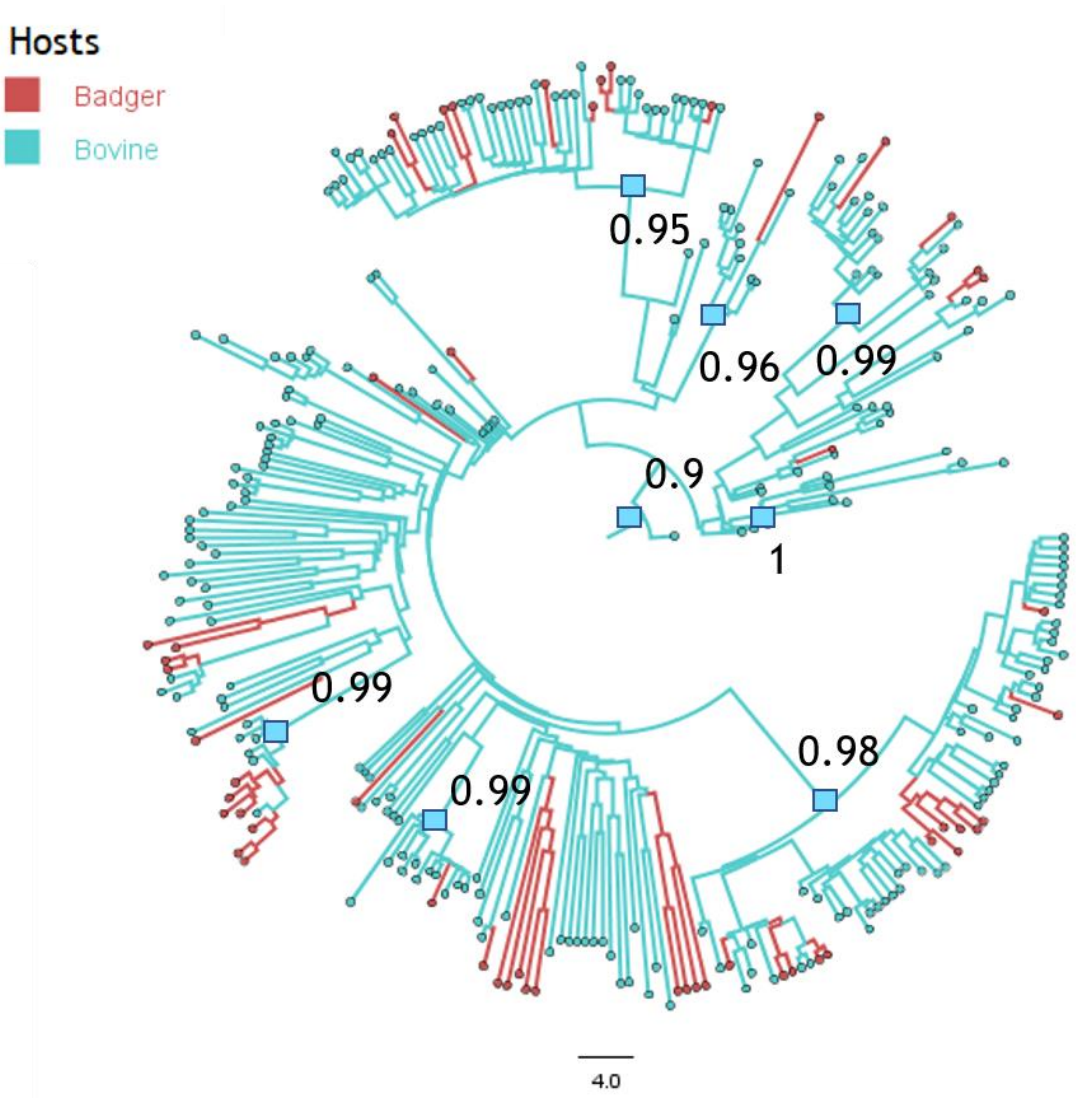


Figure 3.10 (a, b) Maximum clade credibility tree for 302 *M. bovis* isolates (*M. bovis* cattle isolates shown in blue, badger isolates in red) Bayesian skyline population size.

a. Ancestral nodes with higher posterior support (tree) (PP>0.95) shown as black circles, 95% HPD interval for TMRCA estimates for each clade shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badgers in red) states for the associated main clades shown in Figure 3.10a. Cattle was identified as the ancestral hosts for the MRCA of all the major clades and for the MRCA of all the isolates.

The reconstructed *M. bovis* phylogeny showed six subclades (coloured branches) with high posterior probability supports >0.95, and in each clade there were both cattle and badger associated host-states >0.9, suggesting the presence of possible cross-species transmission. Ancestral state reconstruction demonstrated host traits of MRCA supported with posterior probabilities >0.9 for major clades and 0.53 for the oldest ancestor (Figure 3.10b). No clade was found to be badger or

cattle specific. Clade 1 (shown in red) (Figure 3.10a) included the highest number of badger *M. bovis* isolates, which are very genetically close to those in cattle.

The estimated PP support for transitions between cattle and badgers and *vice-versa* are shown in Table 3.8. The calculated Bayes factors (BF) for each direction (asymmetric transition between hosts species) were almost identical in the constant population model, however it is important to note that calculation of the BF for cattle-badger transitions was giving an infinite value, therefore it was approximated to 0.999.

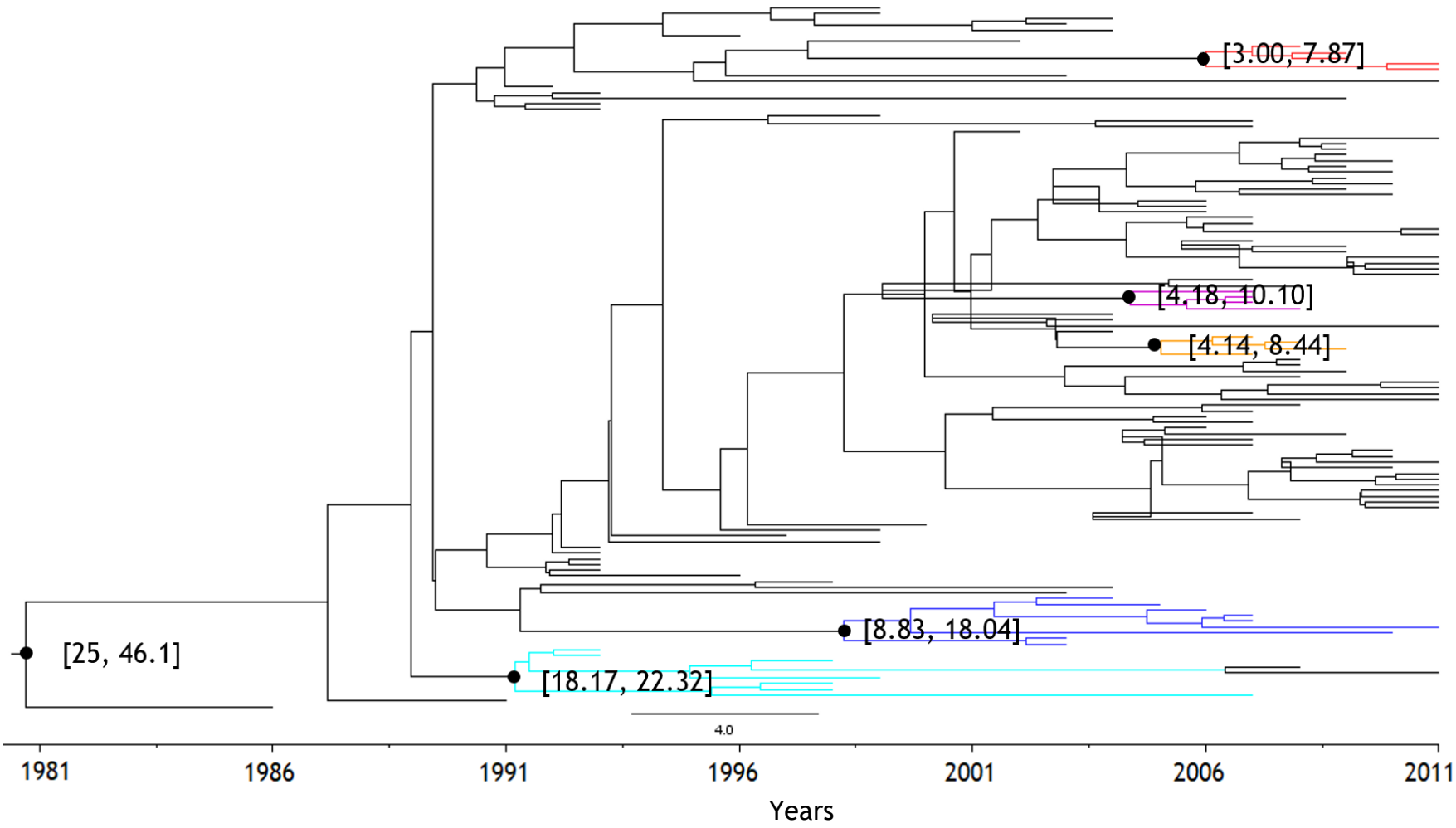
Table 3.8 Comparison of the posterior probabilities of *M. bovis* transitions between host species

Demographic model	Direction of transition	Estimated PP of transition between host species (asymmetric)	Bayes factor (BF)
Constant	Badger-Cattle	0.999	4092.33
	Cattle-Badger	1	4092.33
Skyline	Badger-Cattle	0.934	57.97
	Cattle-Badger	1	4092.33

There is high posterior probability support for *M. bovis* transitions for both cattle-badger and badger-cattle directions using the two population models (Table 3.8). These results suggest that there was *M. bovis* inter-species transmission between the two hosts present in the TVR area during 1986 and 2017.

M. bovis data subsets 1 and 2 were also used for ancestral hosts reconstruction using the two demographic models and the phylogenies demonstrated in Figures 3.11 (a, b) and 3.12 (a, b).

a.


[3.00, 7.87]
[4.18, 10.10]
[4.14, 8.44]
[8.83, 18.04]
[25, 46.1]
[18.17, 22.32]
4.0
1981
1986
1991
1996
2001
2006
2011
Years

b.

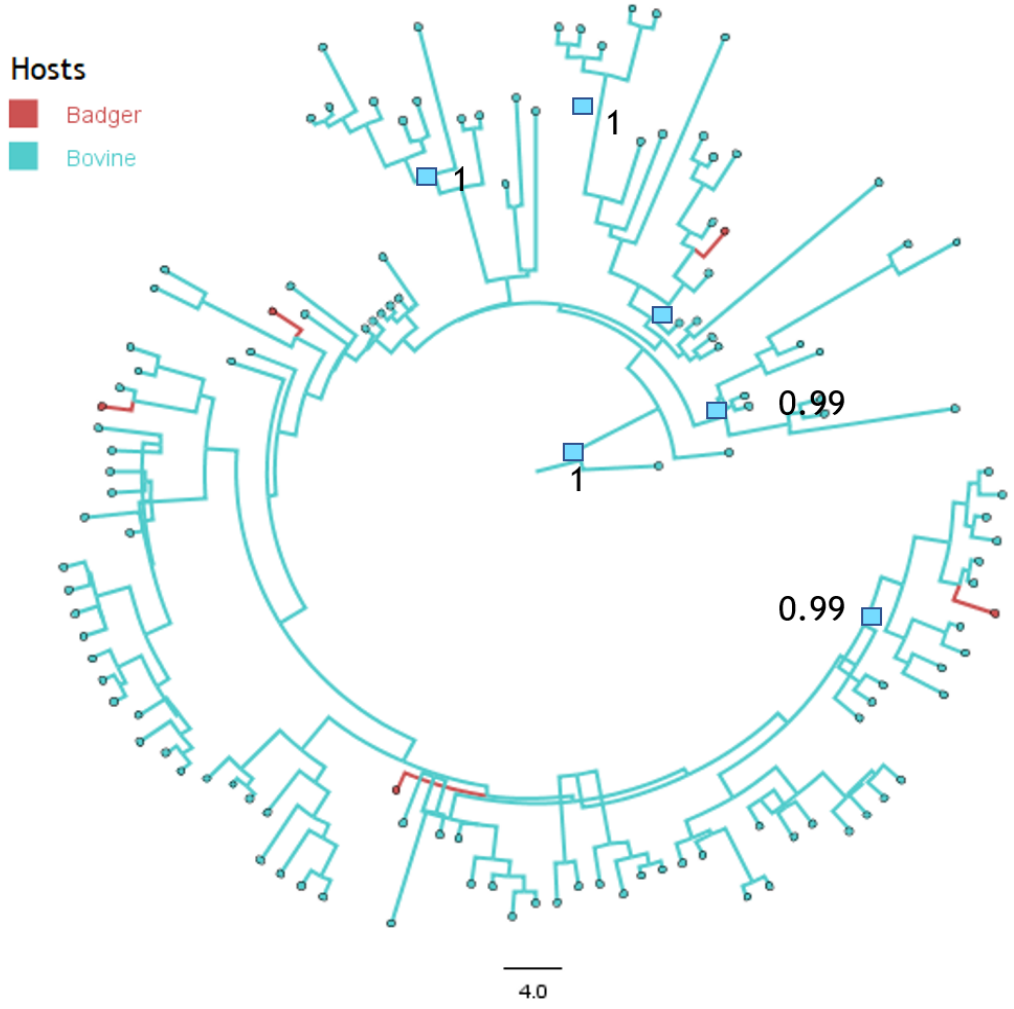
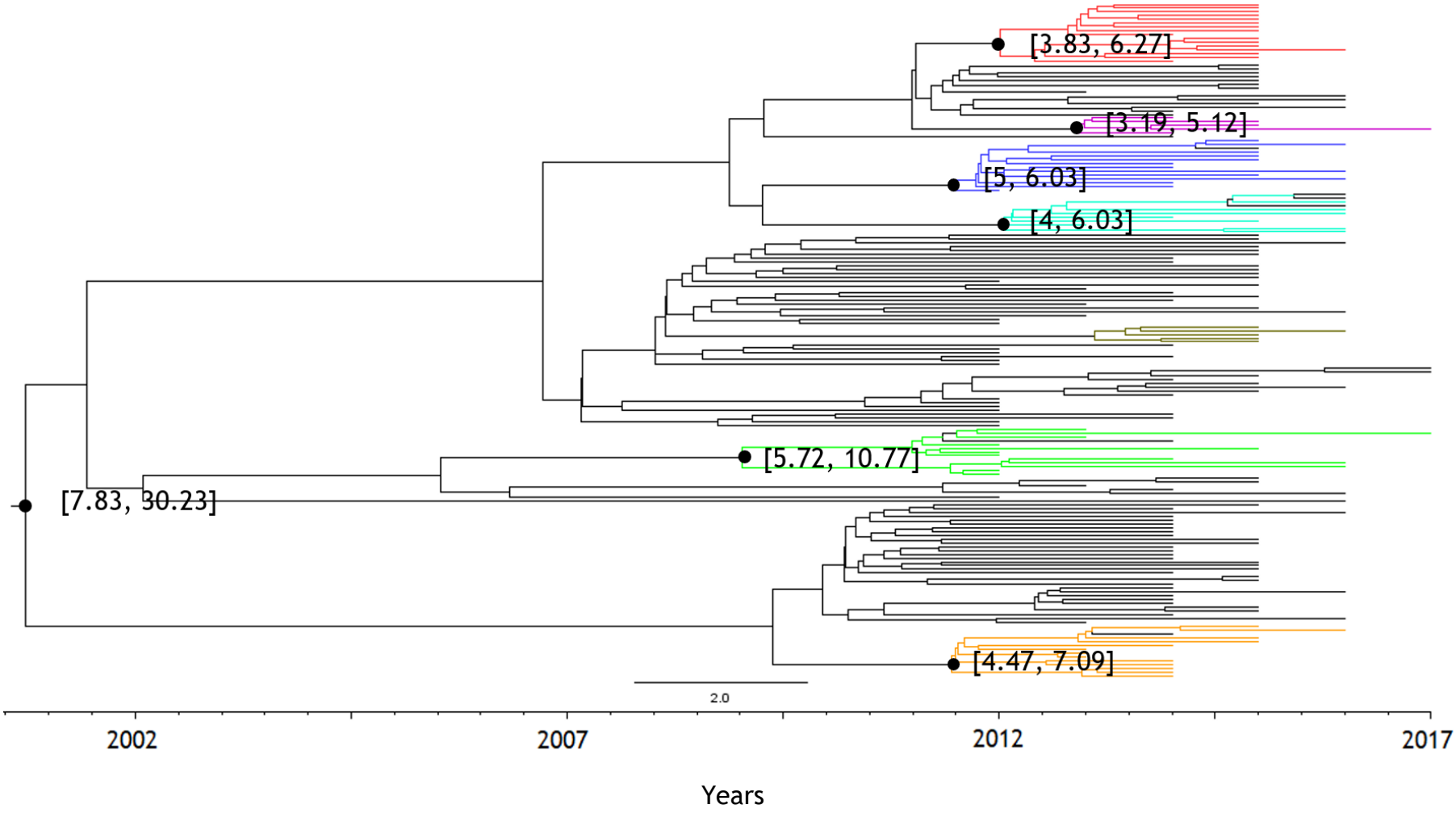


Figure 3.11 (a, b) Phylogenetic MCC trees for the subset 1

Bayesian skyline pop model. *M. bovis* data sampled from years 1986-2011. a. Ancestral nodes with higher posterior support (tree) (PP>0.95) are shown as black circles, 95% HPD interval for node TMRCA estimates of supported clades are shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badger in red) states for associated main clades shown in Figure 3.11a. Cattle was identified as the ancestral hosts for the oldest MRCA of all isolates (host associated PP>0.50).

Only four major monophyletic subclades (coloured) were found to have support (with an estimated posterior probability for the nodes >0.95) (Figures 3.11a and 3.12a). Circular phylogenetic trees (3.11b, 3.12b) represent the MCC trees under a model of asymmetric host species transitions, with branch colours associated with host species and host-state PP >0.50. From the 4 subclades, two are cattle specific.

a.


[3.83, 6.27]
[3.19, 5.12]
[5, 6.03]
[4, 6.03]
[5.72, 10.77]
[7.83, 30.23]
[4.47, 7.09]
2.0
2002
2007
2012
2017
Years

b.

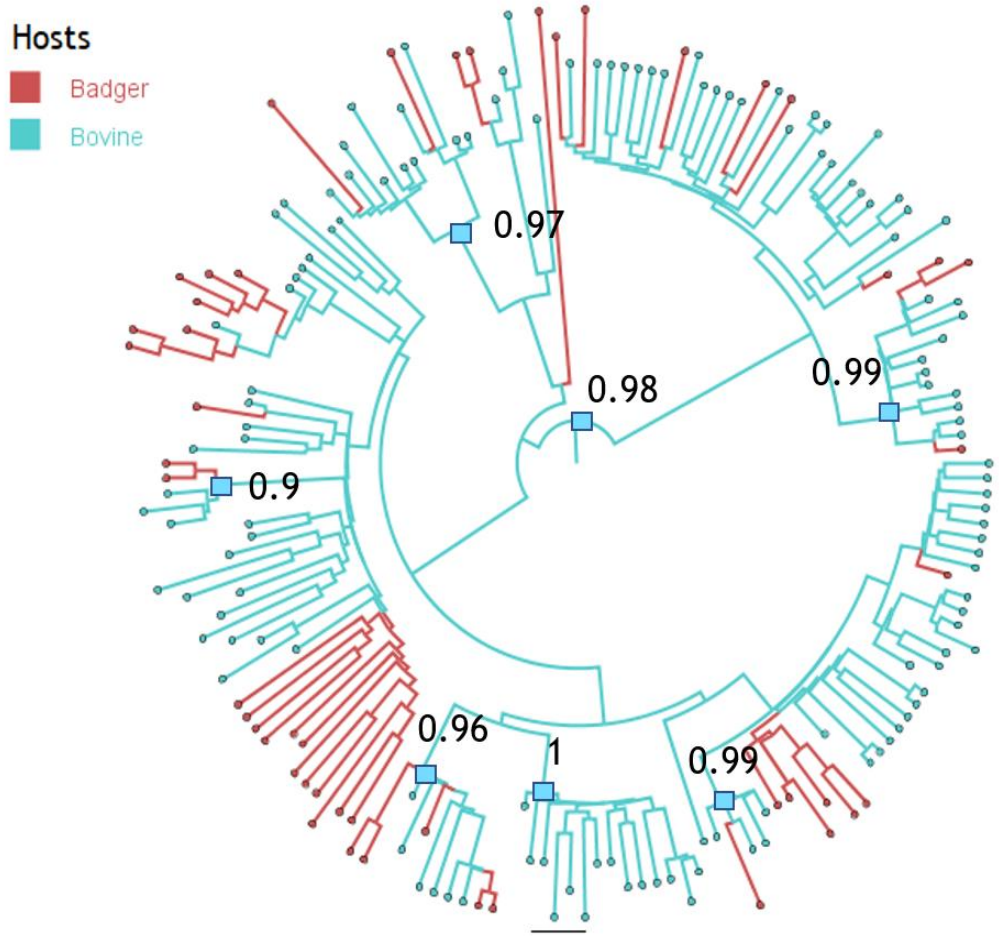


Figure 3.12 (a, b) Phylogenetic MCC trees of the subset 2

Bayesian skyline pop model. *M. bovis* data sampled from years 1986-2011. a. Ancestral nodes with higher posterior support (tree) (PP>0.95) shown as black circles, 95% HPD interval for node TMRCA estimates for supported clades are shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badger in red) states for associated main clades are shown in Figure 3.12a. Cattle was identified as the ancestral hosts for the oldest MRCA of all isolates (host associated PP>0.50).

Phylogenetic trees for discrete traits analysis using Constant population size model are described in Appendix 1 for Chapter 3.

The estimated posterior probability of *M. bovis* transitions between host species (asymmetric) for two data subsets with the estimated BF support are shown in Table 3.9 and 3.10.

Table 3.9 Comparison of the estimated posterior probability support and BFs for *M. bovis* transition between host-species in *M. bovis* subset 1 (1986-2011) under the asymmetric discrete trait model

Demographic model	Direction of transition	Estimated PP of transition between host species (asymmetric)	Bayes factor (BF)
Constant	Badger-Cattle	0.988	337.272
	Cattle-Badger	1	4092.332
Skyline	Badger-Cattle	0.965	112.944
	Cattle-Badger	1	4092.332

Table 3.10 Comparison of the estimated posterior probability support and correspondent BFs for *M. bovis* transition in subset 2 (2012-2017) under the asymmetric discrete trait model

Demographic model	Direction of *M. bovis* transition	Estimated PP of transition between host species (asymmetric)	Bayes factor (BF)
Constant	Badger-Cattle	0.998	2044.118
	Cattle-Badger	1	4092.332
Skyline	Badger-Cattle	0.619	6.655
	Cattle-Badger	1	4092.332

In subset1 and subset2 there are high posterior probability (with very good BF support) of *M. bovis* transitions from cattle to badgers (Table 3.9 and 3.10) and from badgers to cattle, with the exception of *M. bovis* transition from badger to cattle in subset 2 using the Skyline demographic model that only has BF moderate support (Tables 3.9 and 3.10). These results suggest that there are *M. bovis* exchange between the two hosts before and after population expansion.

3.3.3 SeqTrack and Outbreaker2 inter- and within-species events

Both transmission tree reconstruction models SeqTrack (Jombart et al., 2011) and *outbreaker2* (Campbell et al., 2018) analysis inferred the ancestry of the 302 *M. bovis* isolates in the endemic lineage by distinguishing (based on dates of

collection and genetic distances), ancestor isolates from descendant ones. SeqTrack as a graph approach identified 68 possible ancestors within 302 *M. bovis* isolates sampled from both cattle and badgers in years 1986-2017. From these 68 possible ancestors, 12 belonged to badgers' *M. bovis* and 56 from cattle. Regarding the estimated connections between isolates hosts, in total, I identified 72% of interactions within the cattle population, 14% from cattle to badgers, 10% from badgers to cattle and 4% within badgers.

The posterior consensus tree identified for *outbreaker2* found 115 unique ancestors, with one of them being from some common (unsampled) case that was estimated for the dataset (earlier in time than the oldest sampled *M. bovis)*. The number of inferred mutations between cases varied from 0 to 11, with the estimated mutation rate of 0.28×10^{-8} per site per generation of transmission. From the total of 115 identified *M. bovis* ancestors, 103 were isolated from cattle, 11 from badgers, and one common unsampled ancestor (host not identified). The proportion of transmission events between ancestor and descendant hosts was determined from the total number of interactions identified. This assumption is consistent with that previously identified by Walker et al. (2014), which showed that likely transmissions between and within hosts for *M. tuberculosis* are 0-12 SNPs apart. The majority of interactions, 79% were within cattle, 16% from cattle to badgers, 3% from badgers to cattle, and 2% within badgers (Figure 3.13). From the total of these interactions, 86% were only one generation apart from each other, whereas for the rest (14%) of cases there were up to five generations separating their ancestors and descendants, thus showing the proportion of cases that were missed (never sampled). The posterior support of the ancestors in the consensus tree was higher than 50% in more than one third of all 302 ancestor-descendant links. The consensus transmission tree reconstructed is shown as a transmission network plot (Figure 3.14) where nodes were coloured by host species (green for cattle, blue for badgers) and arrows represented at least 50% of support between connected cases.

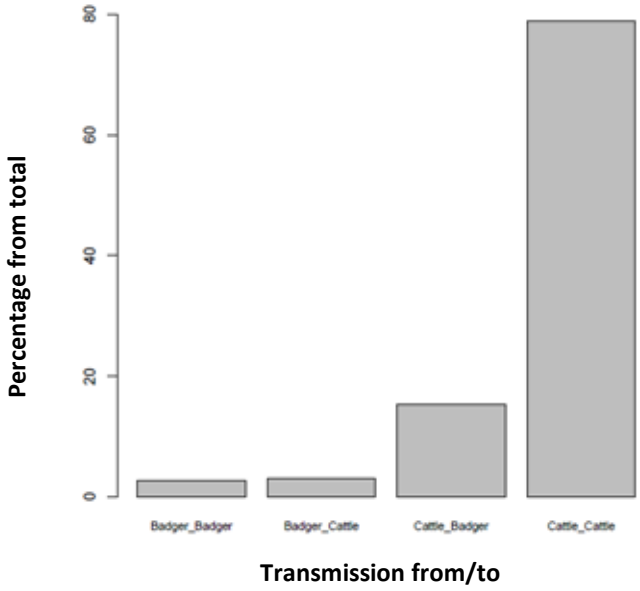


Figure 3.13 Transmission of *M. bovis* 302 isolates between two hosts species (cattle, badgers) obtained for the *outbreaker2* transmission tree reconstruction

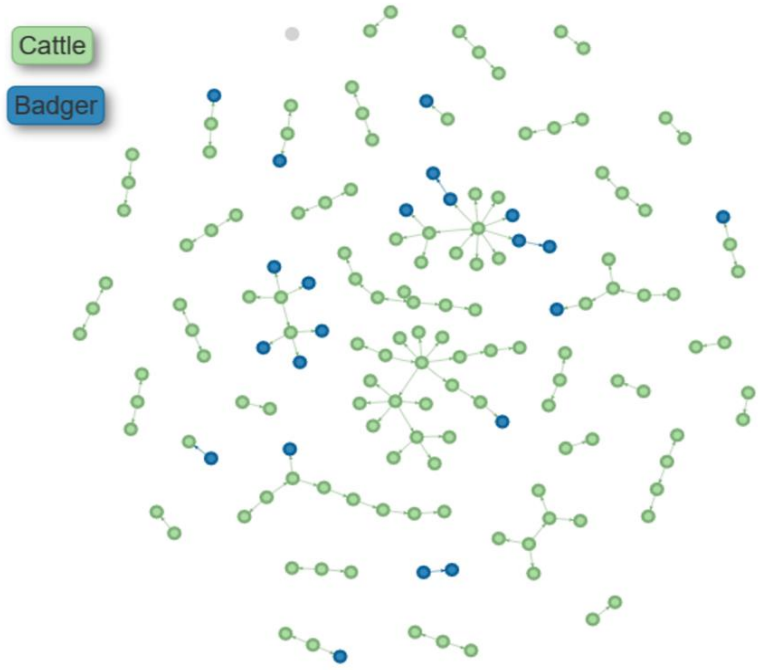


Figure 3.14 Graphic representation of the transmission tree reconstructed from *outbreaker2*

302 *M. bovis* isolates (endemic lineage), cattle hosts shown in green, badgers in blue. The arrows for connected nodes represent >50% posterior support.

3.4 Discussion

The TVR study was implemented to assess the *M. bovis* population in badgers and the prevalence of the infection in cattle within the area. Historical *M. bovis* isolates collected from sympatric cattle and RTA badgers were also sequenced to fulfil the temporal range required for the phylogenetic analyses. Thus, this is a first time *M. bovis* isolates were such densely sampled in Northern Ireland, and all isolates were whole genome sequenced from both cattle and badgers within this specific area. These data provide great opportunity to understand the main questions about bovine tuberculosis genomic epidemiology and transmission.

The distribution of *M. bovis* molecular types were represented in current dataset by 6 unique spoligotypes and 22 MLVA types. In previous studies by (Skuce et al., 2010) it was suggested that the diversity of *M. bovis* isolates can be grouped as eight major "strain families" circulating in NI. They represent founder MLVA type and spoligotypes 1.140, 2.142, 3.140, 4.140, 5.140, 6.263, 19.140 and 20.131. With most of the isolates in current dataset being found in 6.263 endemic lineage with little SNP diversity within the group, which was historically endemic for the TVR area, comparing to other lineages (Skuce et al., 2010, Skuce et al., 2020). Molecular typing techniques as MLVA and spoligotyping used have enough power to discriminate between isolates, but not very useful for the strains with little genetic diversity. These methods are still useful as low-priced and reliable laboratory typing methods. However, with the WGS data providing much higher resolution and the ability to capture rare variants it might be more efficient to use it for routine typing of *M. bovis*. The MLVA type 006 and spoligotype 263 were previously described as the most prevalent in the TVR area and County Down (Skuce et al., 2020). The presence of other *M. bovis* strain families in the area was most likely introduced from different regions, which was also confirmed by large enough SNP distances between lineages, greater than 0-20 SNPs (mean 7.5 SNPs) identified within the endemic lineage (Figure 3.4).

Smith et al. (2003) first introduced the term clonal expansion that explained the distribution of the *M. bovis* molecular types across Great Britain in multiple hosts. It was suggested that, for example, spoligotype 263 has originated in one area and later undergone clonal expansion in neighbouring territories, which led to rapid distribution of that spoligotype in a new area (Smith et al., 2003, Smith et al.,

2006). Therefore, analysing genomics of endemically circulating strains in the TVR area can help to understand the spread of the bTB infection in different Division Veterinary Office (DVO) regions; and study other lineages common in specific areas in the future. But different factors such as limitations in sampling procedures that cannot fully represent all possible isolates across the different time periods and complex epidemiology of *M. bovis* (slow generation time and the time period (months to years) for clinical signs to appear in animals) should be taken into consideration for any phylogenetic and/or transmission tree analyses.

M bovis is a slowly-evolving pathogen typically characterised by limited genetic diversity between phylogenetic clades with very few SNPs distances between isolates. Several studies for *Mycobacterium tuberculosis complex* (MTBC) epidemiology have suggested the minimum number of SNPs that are suggestive of possible transmission links between isolates. In a paper by Walker et al. (2013) the recommended threshold for the transmission of *M. tuberculosis* between two epidemiologically linked patients was suggested as from five to twelve SNPs. Similarly, isolates with ten SNPs difference were estimated as epidemiologically linked in *M. tuberculosis* and *M. bovis* studies (Bryant et al., 2013, Roetzer et al., 2013, Crispell et al., 2019). In the current study, a minimum pairwise genetic distances between *M. bovis* isolates within the endemic lineage 6.263 were estimated within 0-20 SNPs, with mean of 7.5 SNPs within endemic lineage, and the 0-11 SNPs threshold was used to identify the direct transmission links between two isolates for interpretation of the outbreaker2 results.

Before starting the phylogenetic analysis in the current study, the presence of molecular clock signal was assessed. This was done by the examination of the linear relationship between the evolutionary rate and sampling dates. In the current dataset, I examined the "temporal signal" of three datasets, including the full data and two subsets divided based on the time where *M. bovis* population expansion has occurred, and obtained temporal signal (R^2 =0.27) in the 302 *M. bovis* isolates within the 31 years period, but lower for the subsets (R^2 =0.12 and 0.10 for subset 1 and 2 respectively, Table 3.4 and Figures 3.7 (a, b)). These differences might be explained by the difference in sampling times covered in subset 1, 25 years and only 5 years in subset 2. The challenges in estimating the molecular signal for mycobacteria, which has large genome sizes but slow

mutation rates, have also been described in other studies (Menardo et al., 2019, van Tonder et al., 2021).

Phylogenetic analysis of the *M. bovis* dataset (n=302) of the endemic lineage indicated that bTB was introduced to the area in years 1960-1980 (independently of the demographic model used). The mean molecular clock rate was estimated to be between 0.36-0.37 substitutions per genome per year when estimated using the full endemic lineage data. This finding was similar to others in previous studies (Crispell et al., 2019, Salvador et al., 2019). Beast2 phylogenetic analyses estimated the most recent common ancestor that was introduced in 1970-1980s, which showed that estimated MRCAs overlap with the *M. bovis* historical expansions that happened in the 1980s (Robinson, 2015).

It is known that individual badgers do not tend to travel long distances compared to cattle that can be bought-in to the area from neighbouring regions (Woodroffe et al., 2006). Therefore, cattle movements may explain the occurrence of non-endemic *M. bovis* lineages from other home ranges in the TVR area. Some strain families are present only in cattle isolates (Figure 3.2b) which might be explained that they did not yet affect the wildlife species or that those were missed while sampling *M. bovis* from badgers. Therefore, focusing on the transmission dynamics of the endemic *M. bovis* lineage was crucial for the TVR intervention study.

3.4.1 Non-endemic M. bovis lineages

Four *M. bovis* badger isolates collected from two badgers were found in lineage 20.131 within the TVR, and 28 isolates from cattle from the neighbouring region of the home range of spoligotype 131 (Figure 3.3). A more thorough genomic and epidemiological investigation is required to understand the underlying processes of lineage 20.131 spread. MLVA type 004 found in multiple historical cattle and cattle and badger isolates from the TVR area were observed genetically and spatially close, which is also an indicative of within and between species transmission.

3.4.2 Phylogenetic reconstruction of M. bovis ancestral hosts

The inferred phylogenies reconstructed for DATM analysis of the 302 *M. bovis* isolates suggested that although badger and cattle host species were evenly distributed among different clades, the ancestors of the major supported clades and the MRCA of all data was estimated to be *M. bovis* from cattle hosts. This was also confirmed by the posterior probability (and associated BF) of transitions between cattle and badgers and the results of both SeqTrack and *outbreaker2*. These showed that cattle-cattle and cattle-badger transmissions play an important role in disease transmission dynamics.

Some of the limitations of this study are related to the differences in *M. bovis* sampling cases from cattle and badgers, and to the relative short period of the TVR intervention study. For slowly evolving pathogens such as *M. bovis*, it is desirable to have large periods of study to perform evolutionary dynamics analysis to allow for the pathogen to accumulate enough mutations over time and to allow for them to be tracked. This is a crucial point in order to allow for enough temporal signal in the data. Analyses such as these can only be performed if there are enough mutations accumulated over time. In this study, *M. bovis* isolates from badgers collected during the TVR study were densely sampled, which is a very different sampling from occasional infected badgers that were killed in road traffic accidents (Courcier et al., 2018). The number of isolates used in the studies like this should be proportion to the prevalence of the disease in each host (as best as possible), therefore, this sampling bias can cause inadequate estimates of prevalence of the disease in badgers over time, as well as influence the results related to the role of host species in the transmission process (since it is very likely that several ancestral isolates have been missed when the phylogenetic and transmission trees were reconstructed).

Based on the phylogenetic analyses findings, I identified the presence of temporal signal in the data and calculated the evolutionary rate of *M. bovis* within the 31 years of sampling. The Bayesian skyline demographic model demonstrated the possible *M. bovis* population expansion in the area in early 1990s and second in 2011, which was consistent with the estimation of MRCAs from 1970s-1980s and the data in subset 2 representing more densely sampled dataset, containing the

majority of badger isolates. The differences in substitution rates and times of MRCA showed that subset 1 data, which included historically collected *M. bovis* isolates in 1986-2011 covered longer time period, however lesser genomic data might not describe the *M. bovis* epidemiology very accurately. This also indicates the challenges in selecting the most suitable model parameters. The choice of the priors, substitution and clock models must be as accurate as possible and used with caution (Drummond and Bouckaert, 2015) when considering any phylogenetic studies. Even with the best fitted model parameters for specific data the analysis results might not always be correctly interpreted.

The results obtained from both phylogenetic and transmission tree methods suggest that there is an evidence of *M. bovis* inter-species transmission within the TVR area in years 1986-2017. The SeqTrack and *outbreaker2* results however, indicated that proportions of within cattle transmissions were much higher than cattle-badger and badger-cattle ones. The direct comparison of these methods was difficult because the methods focused on slightly different aspects. Phylogenies describe the ancestral relationships between sampled isolates and the MRCA sampled sometime in the past, whereas graph methods SeqTrack and *outbreaker2* identify transmission links not only across species but also within species (Jombart et al., 2011, Ypma et al., 2013, Campbell et al., 2018). The differences between used SeqTrack and *outreaker2,* that first is the graph method and assume that all ancestors and descendants should be present in the sampled data, and all ancestors were strictly sampled earlier in time, depending on the input dates of sampling and transmission links based on genetic distance between isolates (Jombart et al., 2011). *Outbreaker2* is more flexible and reconstruct transmission links based on genetic data, considering incubation time of the infection and infectious period, which is beneficial for *M. bovis* as slowly evolving pathogen (Campbell et al., 2018). However, the selection of input values for the parameters and priors used in outbreaker2 can influence the results. Using these models to study chronic and slow evolving diseases is complicated due to the long generation times and incubation periods associated to these diseases. For these reasons, it is suggested to use more recently developed techniques to directly combine the phylogenies and transmission trees in one model, such as TransPhylo (Didelot et al., 2021) which can reconstruct the transmission model using phylogenetic tree as a base and incorporate the host data. Similar analysis was

done for *M. bovis* data from the Randomised Badger Culling Trial and used to identify the *M. bovis* transmission clusters by incorporating genomic and animal movement data to estimate bTB transmission dynamics and directionality for cross-species transmission (van Tonder et al., 2021).

The distribution of *M. bovis* MLVA types throughout NI was demonstrated in several studies and indicated high levels of spatial clustering of the infection in cattle herds in specific regions, the so-called home-ranges (Skuce et al., 2020). This might be due to the association of *M. bovis* collected from cattle to the farm locations and not the actual infection in cattle. Whereas the badgers' bTB locations were exactly the spots where animals were trapped and captured. It is important for future analyses to improve the spatial data used for *M. bovis* studies and incorporate the locations that animals visited/lived and not only where they died. It is also crucial to improve our understanding of the *M. bovis* spatial distribution between DVOs and the home ranges of specific (most prevalent) MLVA types (Skuce et al., 2020). The use of WGS of MLVA types *M. bovis* is important for tracking transmission, but at a national scale the use of spatial distribution of MLVA types is still useful for bTB epidemiology. Therefore, a more thorough sampling of *M. bovis* within different molecular type groups and WGS of the isolates will be beneficial for further studies. Another possible investigation will be whether the *M. bovis* in different clades on reconstructed phylogenetic tree will be spatially more associated within each other, than between clades. Similar approach as I used in Chapter 2 can be applied to the phylogenetic clades data and the spatial locations of the isolates within the major clades.

The TVR study was the first implemented wildlife intervention study in NI, which provided an exceptional dataset from bTB infected badgers and their spatial locations. Overall, these findings showed that within a relatively small study area with an endemic history of *M. bovis* circulation, WGS data can provide good resolution for improving our understanding of bTB genomic epidemiology (Kao et al., 2016). In this study, transmission tree analysis using *outbreaker2* conducted for *M. bovis* data collected from cattle and badgers from an intensely sampled TVR area (Menzies et al., 2021) suggested that, while much of the transmission inferred is within cattle, *M. bovis* transmission is happening in both directions,

while both cattle and badgers are playing important roles in disease persistence and spread.

4 Chapter

Using molecular bacterial load assay to quantify *M. bovis* in cattle tissue samples

4 Chapter

4.1 Introduction

The Mycobacterium tuberculosis complex (MTBC) consists of genetically closely related species that can cause tuberculosis in humans and multiple animal species. There are non-tuberculous mycobacteria which do not cause tuberculosis and can be isolated from environmental samples, such water systems and soils (Falkinham, 1996, Griffith et al., 2007) (Falkinham, 1996; Griffith et al., 2007). Other tubercle bacilli of MTBC, such as *Mycobacterium tuberculosis*, *M. bovis*, *M. africanum*, *M. microti and M. canetti* are known to cause tuberculosis in humans (Forrellad et al., 2013). *M. bovis* Calmette-Guérin (BCG) is an attenuated vaccine strain that was introduced in 1921 by Albert Calmette and Camille Guérin and is the only vaccine available against pulmonary tuberculosis in humans (Calmette, 1922).

Tuberculosis caused by *Mycobacterium bovis* is an infectious zoonotic disease of various domestic and wildlife animals (Clifton-Hadley et al., 1993; Palmer & Waters, 2011). Bovine tuberculosis (bTB) is a chronic disease of animals that can lead to the development of granulomas in animals' lungs and lymph nodes and can affect other organs (Pollock et al., 2001; Philips & Ernst, 2012). There are several potential routes of *M. bovis* transmission in animals and humans, such as direct and indirect infection transmission. The formation and distribution of lesions in infected animals suggests that respiratory route, inhalation of the aerosols during direct contact between animals, is the most common in cattle (Neill et al., 1994, Neill et al., 2005). Whereas, excretion of *M. bovis* bacilli in animal urine and faeces is considered less important for the cattle-cattle *M. bovis* transmission, as alimentary lesions in infected cattle were less common. Animals can also be exposed indirectly through a potentially contaminated environment shared between animals, such as animal handling facilities and farm lands (Neill et al., 2001, Cassidy, 2006, Allen et al., 2021). In recent years there has been an increasing interest BTB is a slowly progressing disease and usually difficult to confirm early and solely by clinical signs in animals (Pollock and Neill, 2002).

Bovine tuberculosis still plays an important role and can have significant negative effects for farming economies in some countries, including where the disease is associated with local wildlife (Fitzgerald & Kaneene, 2013). In the United Kingdom

(UK) the main wildlife host of *M. bovis* infection is the European badger (*Meles meles*) (Delahay et al., 2001). Despite preventive measures in cattle, the presence of bTB infection in badgers and the close proximity of their natural habitats to cattle farms makes the situation complicated and sensitive. In the past 12 months (2021-2022) confirmed bTB prevalence in Northern Ireland (NI) cattle herds was around 9.95% and control efforts cost the government about £44 million (Department of Agriculture Environment and Rural Affairs, 2022). Many research studies found epidemiological associations between bTB infection in cattle and badgers (Woodroffe et al., 2005; Jenkins et al., 2007), and the relative importance of each hosts studied (Crispell et al., 2019). To improve the understanding of *M. bovis* infection in wildlife in NI, Test and Vaccinate or Remove (TVR) intervention study was undertaken in 2014-2018.

One of the major difficulties in *M. bovis* surveillance is implementation of the laboratory diagnostic methods to detect the bacterium (Allen et al., 2018). Current diagnosis of *M. bovis* in the United Kingdom includes compulsory tuberculosis skin testing and supplementary interferon gamma testing (IFN-γ) of blood samples (de la Rua-Domenech et al., 2006; Gormley et al., 2006), with the sensitivity (*Se*) of the skin test evaluated between 75.0% and 95.5% and specificity *(Sp*) 78.8% and 100% (de la Rua-Domenech et al., 2006); and for IFN-γ test the calculated *Se* was between 73.0% and 100%, and *Sp* between 85.0–99.6% (de la Rua-Domenech et al., 2006). Misidentification, as well as false positive or false negative identification of bTB infected animals will affect the disease persistence and have negative impacts on farm management, such as the need for additional testing and control measures and further transmission of the disease in case of false negative identification (Lahuerta-Marin et al., 2016). If cattle are identified as bTB positive, using the tuberculin test, the animals are culled and sent for post-mortem examination (de la Rua-Domenech et al., 2006; Godfray et al., 2018). A definitive confirmation of *M. bovis* is achieved by the isolation and culture-confirmation of the bacterium from animal specimens.

Bacteriological identification and culture of *M. bovis* is one of the core diagnostic techniques used in laboratory practices for bTB management. This method remains a ”gold standard” for *M. bovis* confirmation from the post-mortem specimens or clinical samples and aims to detect any viable mycobacterial cells that can potentially cause the disease (Allen, 1998). Mycobacterial culture is

costly, time-consuming and requires additional biosafety level 3 facilities and reagents compared to direct microscopy/histology. The latter is the fastest and cheapest way to identify the acid-fast bacilli in samples but is not as sensitive as culture and cannot differentiate between live and dead bacilli. The bacteriological mainstay of histological classification, gram staining, does not work for mycobacteria on account of their thick, waxy coat which prevents dye uptake. By contrast, mycobacteria are acid-fast bacteria, which means they can retain carbolfuchsin in their mycolic acid rich cell wall which forms the basis of the Ziehl-Neelsen staining method (Lahiri and Chatterjee, 1994) used to diagnose the presence of mycobacteria histologically.

Bacterial isolation of *M. bovis* from clinical samples requires specific pre-processing of the specimens, such as homogenization and decontamination. Culture media cannot inhibit all contaminating agents and therefore, decontamination step is essential for mycobacterial cells recovery (Corner et al., 2012a). Various studies have shown that it requires more than 3 weeks for primary isolation of *M. bovis* and even up to 12 weeks in some cases (Corner et al., 2012a, Gormley et al., 2014). This is caused by its slow replication rate (16-20 h) (Beste et al., 2009). Various factors, such as contamination of bacterial culture from specimens can affect the incubation period and the growth of *M. bovis* (Miller et al., 2002, Corner et al., 2012b). There are different types of enriched media which can be used for isolation of *M. bovis*, egg based, such as Stonebrink's medium and Löwenstein–Jensen (LJ) medium, agar based as Middlebrook 7H11, and tuberculosis blood agar (B83). And requires decontamination of specimens' step for the effective isolation of *M. bovis*. Ideally correct decontaminant should have no effect for *M. bovis* growth but be toxic for other agents contained in samples. Most commonly used reagents are hexadecylpyridinium chloride (HPC) (0,75%/0.075%), sodium hydroxide (NaOH) (2%) and oxalic acid (5%), N-acetyl-L-cysteine-sodium hydroxide (NALC-NaOH) (Allen, 1998, Verma and Kashyap, 2021).

Working with highly contagious materials require specific safety controls and operation in specialized laboratories (World Health Organization, 2012, Advisory Committee on Dangerous Pathogens, 2021). For safe handling of samples and possibility for work in laboratories with lower biosafety levels there is a need for methods that will inactivate the bacteria but will not affect sensitivity of the diagnostic techniques. It is mostly common to use chemical disinfectants, such as

aldehydes, peroxides and guanidium thiocyanate (Sambrook et al., 1989, Sagripanti et al., 2011) or physical (heat, radiation) or biological inactivation (bacteriophages). Implementation of *M. bovis* inactivation methods that will not affect the bacterial quantities and allow technicians to work with infected samples in BSL2 labs will provide more opportunities for research groups to conduct molecular tests for tuberculosis. Most mycobacteria are not easily inactivated and are resistant to many chemical disinfectants. Solutions containing chlorine, phenols and ethyl alcohol will work more efficiently (Rutala et al., 1991). Sabiiti et al. (2019) also demonstrated the efficiency of high temperature heating to inactivate *M. tuberculosis*.

Recently proposed molecular diagnostic technique, based on polymerase chain reaction (PCR) was accepted for *M. bovis* detection from animal tissue samples in Great Britain (Animal and Plant Health Agency, 2022). PCR techniques have made a significant contribution to the detection and characterisation of *M. bovis* since 1995. Wards et al. (1995) and Liébana et al. (1995) successfully used mycobacterial DNA as a target for rapid detection of *M. bovis* in bovine tissues. Insertion sequences, such as IS*6110* and IS*1081*, present in MTBC species have been widely used in many research studies for the detection and differentiation of the strains (Eisenach et al., 1990, Bhattacharya et al., 2003, Thacker et al., 2011, Nghiem et al., 2015). However, detection of *M. bovis* DNA in animal samples is not indicative of the viability of the bacterium, where these samples may be culture-negative (Hellyer et al., 1999). Working with fresh or frozen animal tissues is complex due to the presence of host DNA and nucleases destroying the bacterial nucleic acids and effective samples preparation procedures must be optimized. Furthermore, preservation and extraction of bacterial nucleic acids from these samples might also be complicated by other factors, such as temperature and storage regimes, and how the samples were shipped.

Currently used molecular techniques are mostly based on detection of *M. bovis* DNA and highly dependent on the efficiency of the extraction and nucleic acids purity. Amaro et al. (2008) compared three different techniques for the extraction of *M. bovis* DNA in animal tissue samples and indicated that mechanical disruption in combination with enzymatic lysis had better efficiency comparing to protocols using lysozyme and proteinase K; and cetyltrimethylammonium bromide (CTAB) extraction. Protocol based on mechanical cell disruption by bead beating

and DNA purification using magnetic separation was shown to yield a higher concentration of DNA (Mićić et al., 2016) and is successfully used for extraction of DNA for next generation sequencing (NGS) which requires removal of different contamination agents and isolation of fragments bigger than 100 bp. Similar study by Caldarelli-Stefano et al. (1999) used magnetic beads for *M. tuberculosis* DNA extraction and amplification of *IS6110* sequence by PCR from frozen and fixed tissues and implicated the high efficiency of the technique. Another advantage of the method was shorter time and that it does not require working with dangerous chemicals (Caldarelli-Stefano et al., 1999).

In contrast to DNA based detection, there was an increasing interest in finding the molecular target for PCR, that is present in multiple copies in mycobacterial cells and can indicate the viability of the bacteria. Bacterial RNA has been studied as having a shorter half-life and playing an important role in different processes and therefore vary in copy numbers and stability (Hellyer et al., 1999). The potential of use of RNA to identify live bacteria from clinical samples was investigated by van der Vliet et al. (1994). Recent studies showed the use of ribosomal RNA (rRNA) as a marker of "viability" of bacteria and used to measure the bacillary loads in sputum in response to antimicrobial treatment (Honeyborne et al., 2011). Some studies showed the use of several antigen genes as RNA targets to detect *Mycobacterium tuberculosis* complex by PCR, such as *mpb*70 and *mpb*64 (Young et al., 2005). The mycobacterial load assay (MBLA) is a culture-free method to detect and quantify viable *Mycobacterium tuberculosis* complex bacilli. It is a reverse transcriptase quantitative PCR (RT-qPCR) technique based on identification of *M. tuberculosis* complex 16S rRNA from human sputum samples (Honeyborne et al., 2011, Gillespie et al., 2017). This molecular approach was used to monitor the response to antibiotic treatment in tuberculosis patients with decline in 16S rRNA loads (Sabiiti et al., 2020a, Sabiiti et al., 2020b). Detection and quantitation of *M. bovis* with even very few copy numbers of bacilli, which can be obtained using MLBA technique potential indicator of ability to infect other animals, will be very valuable for the disease monitoring and management in cattle. Although there is no treatment being used against bovine tuberculosis in animals, the detection of viable mycobacteria in live animals is important because *M. bovis* can be shed and persist into the environment from infected animals (Young et al., 2005, Barbier et al., 2017). Moreover, quantifying the levels of

mycobacterial loads in cattle samples obtained using the MBLA method can potentially help to investigate how infectious the animals are. *M. bovis* is difficult to isolate using bacterial culture and might be missed during confirmation, therefore, complementing this method with extra data about bacterial viability, could be an in important step for bTB control management.

The aims of the current study were:

1. Investigate the use of MBLA assay as a diagnostic technique to detect *M. bovis* in animal tissue specimens
2. Quantify the bacterial loads of *M. bovis* 16S rRNA in infected animal tissues and compare with bacteriological culture
3. Investigate the correlation of mycobacterial loads with various parameters, such as animal tissue weights used for rRNA extraction, identification of lesions in sampled bovine tissues and different *M. bovis* genotypes.

To achieve these aims within the current research study, a programme of laboratory work to assess and validate the MBLA was performed using the following approaches:

1. Optimization of molecular bacterial load assay (MBLA) protocol for detection and quantitation of 16S ribosomal RNA directly from M. bovis BCG-spiked bovine tissue samples (non-infected).
2. Application of the MBL assay on 214 frozen bovine tissue samples collected from bovine tuberculosis (bTB) infected cattle in NI.
3. Use of heat inactivation of *M. bovis* for the MBLA use in BSL2 laboratories.
4. Investigation of *M. bovis* rRNA stability at room temperature in during different time periods, in order to examine the use of MBLA method within various conditions of storage, transportation and sampling of *M. bovis* samples.

4.2 Materials and methods

4.2.1 Materials and consumables

4.2.1.1 Equipment

1. Fridge and freezers
2. Laboratory scales
3. Class II Biosafety Cabinet
4. Homogenizer, Precellys 24 (Bertin Instruments, France)

5. Centrifuge for 1.5-2 ml tubes, Eppendorf model 5415R (Germany)
6. Thermoblock, Starlab N2400-4002 (UK)
7. Vortex, Starlab N2400-6110 (UK)
8. Spectrophotometer, Qubit 3 Fluorometer (Invitrogen, USA)
9. Real-time PCR thermocycler RotorGene Q with 72-well rotor (Qiagen, USA)

4.2.1.2 Laboratory consumables and ware

1. Laboratory pipettes
2. Sterile filtered tips (DNase/RNase-free) for different volumes
3. Disposable laboratory gloves and required PPE
4. Falcon tubes and racks, 15 and 50 mL
5. Measuring cups and cylinders, plastic and glass
6. Homogenization tubes 2 ml with screw caps Lysing Matrix Z and B (MP Biomedicals, USA)
7. RNase-free Microfuge Tubes (1.5 mL)
8. Single 0.2 ml PCR optical thin wall flat cap microtubes

4.2.1.3 Buffers and solutions

1. Phosphate-buffered saline
2. Lysozyme from chicken egg white
3. 2-Mercaptoethanol
4. Absolute ethanol (99-100%)
5. Molecular grade water, DNase and RNase free
6. Guanidine thiocyanate (GTC)
7. 1M Tris-HCl pH 7.5 (Sigma, USA)
8. N-Lauroylsarcosine sodium salt (Sigma, USA)
9. Sodium citrate tribasic dihydrate (Sigma, USA)

4.2.1.4 Culture media

1. Middlebrook 7H9 broth
2. Middlebrook 7H11 agar
3. ADC/OADC enrichments (BD Difco, UK)

4.2.1.5 Culture strains

1. *M. bovis* BCG NCTC 5692 (NCTC, UK)
2. *M. marinum* NCTC2275 (NCTC, UK)

4.2.1.6 RNA/DNA extraction

1. PureLink™ RNA Mini Kit (Invitrogen, USA)
2. TURBO DNA-free™ Kit (Invitrogen, USA)

3. JetSeq beads (Bioline, UK)
4. Wizard® Genomic DNA Purification Kit (Promega, USA)
5. Zymoclean Gel DNA Recovery Kits (Zymo Research, USA)

4.2.1.7 PCR reagents

1. Rotor-Gene Multiplex PCR NoROX Kit (400) (Qiagen, USA)
2. Primers and probes for 16S rRNA *Mycobacterium bovis* BCG (Eurofins, UK)
3. Primers and probes for 16S rRNA *Mycobacterium marinum* (Eurofins, UK)

4.2.1.8 Disinfectants

1. TristelFuse (Tristel Solutions Direct, UK)
2. RNase away (Thermo Fisher Scientific, UK)

This section covers the laboratory experiments conducted to detect and quantify *M. bovis* 16S rRNA in infected cattle tissues by MBLA.

4.2.2 Internal control

Mycobacterium marinum is non-tuberculous bacterium that is infectious for fish and humans, mostly through skin (Chen et al., 2017). This gram-positive and acid-fast bacterium is highly genetically similar with *M. tuberculosis*, however growing faster at 32^{0}C with generation time about 4 hours (Akram and Aboobacker, 2022). Therefore, the use of *M. marinum* as internal control in our study allowed us working with bacterium in BSL2 laboratory, and because of the high genetic similarities of 16S rRNA in *M. marinum* and *M. tuberculosis*, use it for the MBL assay was more accurate. The obtained NCTC cultures were grown in Middlebrook 7H9 broth (with ADC supplement) and Middlebrook 7H11 agar (with OADC supplement) for a week in 32^{0}C reaching the Log (exponential) phase (Figure 4.1). The use of IC is crucial for any PCR analysis to avoid the false-negative results of the assay and RNA extraction errors. When running the multiplex RT-qPCR assay we obtain the results in two compatible fluorescent channels for 16S rRNA of *M. bovis* and *M. marinum*. In case of a negative result in the target-sequence (*M. bovis* 16S rRNA), IC should always show a positive signal. The required concentration of IC was optimized as 10^6 CFU/ml at the optimal cycle threshold (C_t) of 25 in RT-qPCR. This was required to investigate if any factors during RNA extraction from animal tissues affected the signal. A separate standard curve for *M. marinum* was calculated in order to identify the optimal concentration of the

IC used for MBLA. The dynamic range for IC RNA quantitation was from C_t 15.89 to C_t 35.43 for *M. marinum* concentrations (10^9 to 10^3 CFU/ml respectively).

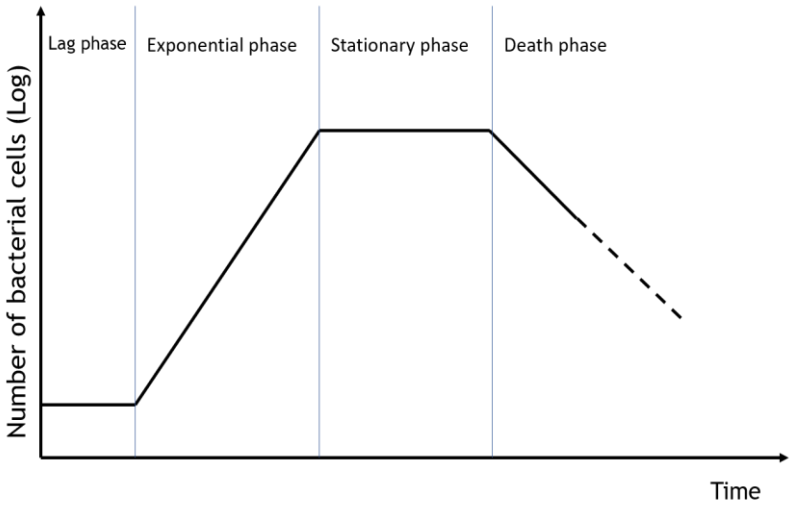


Figure 4.1 Bacterial growth curve in required conditions (i.e. single batch of the medium, temperature, pH, oxygen) (Paulton, 1991)

4.2.3 Sample collection

All samples (n=214) were collected in 2018 from the statutory testing of cattle for bTB in Northern Ireland (NI) and included animals from within and near (maximum 2 km) the TVR intervention area. This area was chosen because of the highest bTB prevalence in NI in 2011-2012 and was home to medium density badger population (5.6 badgers per km^2; Menzies et al. (2021)) which was larger than that found in other areas of NI (DAERA, 2018a, Menzies et al., 2021). No animals were culled specifically for the purposes of this study. However, cattle suspected positive for bTB during routine testing, either by tuberculin testing or interferon gamma testing (de la Rua-Domenech et al., 2006) were slaughtered at the abattoirs and inspected for the presence/absence of visible lesions, some animals that were identified having lesions during routine slaughter were so-called lesioned at routine slaughter (LRS) which by definition were skin test negative. If granulomas were found at post-mortem examination, they were scored using internal system of scoring bovine tuberculosis lesions.

4.2.4 Bacterial strains and culture

Laboratory analyses for the MBLA protocols, using *M. bovis* BCG and *M. marinum* strains and *M. bovis* 16S rRNA were conducted in a BSL2 laboratory, the One Health

Research Into Bacterial Infectious Diseases (OHRBID) laboratory, the University of Glasgow. Relevant COSHH forms were in place for the work with these pathogens and samples. Potential contamination and cross-contamination were prevented by using aseptic culture techniques. Disinfection and cleaning procedures were performed with bactericidal compounds effective in inactivation of *M. bovis,* such as Tristel-Fuse (Tristel Solutions Direct, UK), 1% bleach and 70% ethanol.

The *M. bovis* BCG strain (NCTC 5962) and *M. marinum* (NCTC 2275) were grown in two types of culture media used for the growth of mycobacteria, Middlebrook 7H9 broth (Sigma-Aldrich, USA) with ADC enrichment (bovine albumin, dextrose, and catalase) (BD Difco, UK); and Middlebrook 7H11 agar (Sigma-Aldrich, USA) with required OADC supplement (oleic acid, bovine albumin, sodium chloride, dextrose and catalase) (BD Difco, UK). Both liquid and solid media were resuspended according to the manufacturer's instructions.

Both reference strains were inoculated in 10 mL Middlebrook 7H9 broth. *M. marinum* was incubated aerobically at 30ºC, statically for up to one week, to reach late exponential growth phase. *M. bovis* BCG was incubated aerobically at 37ºC for up to 3-6 weeks. To check for isolate purity, the cultures were then inoculated onto 7H11 agar plates and incubated at their respective temperatures. Work with strain *M. bovis* BCG and IC can be performed in BSL2 laboratory (World Health Organization, 2012). Decimal dilutions *M. bovis* BCG cultures (10^9 to 10^0 cells/ml) in 7H9 Middlebrook broth with ADC enrichment in exponential phase of growth (after incubation for approximately 4 weeks) were prepared in 900 µL of PBS. For this, 1 mL of culture was homogenised by bead beating in Lysing matrix Z tubes (yttria-stabilized zirconium oxide beads) (MP Biomedicals, USA) to avoid bacteria clumping at 6000 rpm for 40 seconds, and 100 µL used for serial dilutions for M. bovis spiking of non-infected tissues.

4.2.5 Optimization of MBLA on M. bovis BCG-spiked bovine tissues

Fresh tissues (lungs and lymph nodes), without any evidence of *M. bovis* infection were obtained from the post-mortem room of the School of Veterinary Medicine, the University of Glasgow. Spiking of bovine tissues with known concentrations of *M. bovis* BCG was performed to evaluate, refine and validate the performance

characteristics (limits of detection) of the MBLA when applied to bovine tissue samples.

The tissues were cut into smaller pieces using sterile scalpels, to obtain ~ 100 mg each. Tissue samples were spiked with 1 mL of ten-fold dilutions of *M. bovis* BCG (described in previous section), within the concentration range of 10^9 – 10^2 CFU/mL. Two positive controls (PC) with high concentration of *M. bovis* BCG (10^9 CFU/ml) and two negative controls (NC) with non-spiked tissues (spiked with 1 mL PBS) were added to the analyses. All spiked tissues where then used of the extraction of *M. bovis* 16S rRNA extraction.

4.2.6 M. bovis 16S rRNA extraction from animal tissue samples

M. bovis BCG-spiked tissue samples were homogenised by bead beating in Lysing matrix Z tubes (2 mL) (MP Biomedicals, USA) for tissue grinding in Precellys 24 (Bertin Instruments, France) at 6000 rpm for 40 seconds for further RNA extraction as suggested in manufacturers' instructions.

RNA extraction was performed using PureLink RNA Mini Kit (Ambion, USA), purifying RNA from animal tissues. This is a column-based method for RNA purification from tissue samples, with two steps of washing and elution in 100-300 mL nuclease-free water. Extracted RNA samples were treated with DNase I (TURBO DNA-free kit; Ambion, USA) to remove any remaining DNA, according to manufacturers' protocols. All plastic laboratory consumables used for RNA extraction were RNase-free; filter pipette tips were used for molecular biology needs. To avoid contamination, work with RNAs was performed in a separate biological safety cabinet II. Lysis buffer, Wash buffer I and Wash buffer II were prepared in necessary volumes before extraction procedures. Solutions (not included in the kit) were prepared separately: 1% β-mercaptoethanol, 70% ethanol (in RNase-free grade water), absolute ethanol. RNase away solutions were used for decontamination purposes.

Final volume of RNAs extracted from tissues was 300 µL (100 µL (50 µLx2) for the tissue spiking experiment) due to the different tissue volumes used for MBL assay. Samples stored at -80ºC before transportation (in dry ice) for further RT-qPCR to the University of Glasgow, OHRBID laboratory.

4.2.7 DNase treatment

RNA extraction and purity when isolated from animal samples can be complicated by the presence of DNA fragments. To obtain a good quality RNA we used DNase treatment procedure after the extraction step with the DNA-*free*™ kit (Ambion, USA). The protocol included adding to the sample RNA, DNase digestion reagents and incubating at 37ºC; next, inactivation of the DNases and centrifugation at high speeds to pellet the reagents; after this processing, the supernatant contained extracted RNA.

4.2.8 Quantification and creation of standard curve from spiked samples.

Standard curves for quantitation of mycobacterial load based on 16S rRNA *M. bovis* BCG were created based on eight ten-fold serial dilutions (in three replications) of cultures reaching the end of exponential growth phase when spiked into bovine tissues. Standard curves were created on RotorGeneQ (Qiagen, USA) software by plotting average Ct values (cycle threshold, when the fluorescence signal of a PCR product was detected) against the respective mycobacterial concentrations (CFU/ml) in the culture media. Standard curves were used to estimate the mycobacterial load in bovine tissue samples and separately in *M. bovis* BCG and IC cultures to assess the performance of the RT-qPCRs. Detection and quantitation limits of the new MBL assay used for *M. bovis* isolates were evaluated. Standard curves were created for each RT-qPCR run and incorporated in data analysis for tested isolates (Gillespie et al., 2017).

4.2.9 RT-qPCR

M. bovis BCG NCTC 5692 (attenuated *M. bovis*) was used as positive control (PC), i.e. a surrogate of virulent *M. bovis* strains, and non-tuberculous fast-growing *Mycobacterium,* strain *M. marinum* NCTC 2275 as internal process control (IC) to control for the entire procedure including RNA extraction and RT-qPCR.

4.2.9.1 Preparation of RT-qPCR master mix

Work with RT-qPCR reagents was performed in a separate “clean” room to avoid contamination. Rotor-Gene Multiplex RT-qPCR kit (no ROX; Qiagen) was used with

species-specific primers and probes for *M. bovis* 16S rRNA and *M. marinum* 16S rRNA for MBLA analysis.

4.2.9.2 Preparation of primers and probes

The use of both MM and RT enzyme allows to run the reverse transcription and PCR reactions in one tube. Multiplexing allows different sets of primers and probes to be combined. Species-specific primers (Forward + Reverse) and TaqMan probes targeting the 16S rRNA of *M. marinum* (IC) and 16S rRNA of *M. bovis* were used. Primers and probes were synthesized by Eurofins (https://www.eurofins.com/). *P*rimers and probes were supplied in lyophilized tubes and were dissolved in 1/10 concentrations in RNase-free molecular grade water and supplied buffer (for probes) and stored in stock aliquots (100 µL) in -20^{0}C freezer. Detailed information on primers and probes sequences is shown in Table 4.1.

Table 4.1 Primers and probe sequences used for MBLA (from Gillespie et al. (2017)).

Name	Sequence (5'-3')
Mtb 16s Forward	GTGATCTGCCCTGCACTTC
Mtb 16s Reverse	ATCCCACACCGCTAAAGCG
IC MMtmRNA F	CGTCATCCTGGCTAGTTC
IC MMtmRNA R	CTACGGCATTCCCTCAAG
Mtb 16s probe	FAM-AGGACCACGGGATGCATGTCTTGT-BHQ1
IC MMtmRNA probe	HEX-AGT CCG CTA TGT CTC TGC TCG-BHQ1

Master mix (MM) is the solution of all reagents required for PCR and optimized for the Rotor-Gene cyclers. All components of the kit were already adapted for the RT-qPCR; the concentrations were calculated for the required number of samples (including two extra samples to avoid pipetting errors) final volume in each reaction was 20 µL (Table 4.2).

Table 4.2 All required components, RNA isolates and ICs calculated for each reaction included in the final master mix solution

Master mix	**Per 1 reaction (RT+)**
Rotor-Gene Multiplex RT-qPCR Master Mix	10 µL
Mtb 16s Forward	0.4 µL
Mtb 16s Reverse	0.4 µL

Mtb 16s probe	0.4 µL
IC MMtmRNA F	0.4 µL
IC MMtmRNA R	0.4 µL
IC MMtmRNA probe	0.4 µL
RT enzyme	0.2 µL
RNase free water	3.6 µL
Sample RNA	2 µL
IC RNA	2 µL
Total volume	20 µL

4.2.10 Setting up the Rotor-Gene Q instrument

Multiplex real-time PCR assays allow users to detect different fluorescent dyes. Green (FAM) and Yellow (HEX) reporter dyes were compatible with each other and detecting the signal at different wavelengths installed in Rotor Gene Q (Qiagen, USA) machine. PC and NC controls were also included for each MBLA run.

RT-qPCR protocol:

1. Prepare all required stock solutions Rotor-Gene Multiplex RT-qPCR master mix, target-specific primers and probes, RT mix (used from -20ºC immediately before use and returned immediately after use), RNase-free water.
2. Add required reagents (Table 1) and mix thoroughly, pipette in separate PCR tubes (RNase-free).
3. In a separate room add IC to each tube, add template RNA (in duplicate).
4. Set up the PCR instrument according to the program below.

Programming of the instrument was done according to the user manual:

a. Hold at 50°C, 15 min (reverse transcription)
b. Hold at 95°C, 15 min (Taq DNA polymerase activation)
c. Cycling, 40-45 cycles of 95°C, 15 s and 60°C, 15 s acquiring fluorescence at Green and Yellow channels.

4.2.11 Bovine tissues samples processing

The next section describes the real samples obtained from AFBI for detection and quantitation of *M. bovis* 16S rRNA using MBLA assay. The samples processing, RNA

extraction and samples inactivation was performed in the BSL3 laboratory in AFBI and further sent to the University of Glasgow for MBLA analysis.

Tuberculosis affected tissue samples (mostly lungs and lymph nodes) collected from bTB positive cattle during annual testing, were sent for further confirmation to the bacteriological laboratory and molecular typing at AFBI, NI. Obtained 214 *M. bovis* infected frozen (at -20 ºC) bovine tissues were stored at the laboratory. Each tissue was first defrosted and weighed, sample weights ranged between 400 mg to 7 g.

Tissues were homogenised with 5 M guanidium thiocyanate (GTC) solution (Thermo Fisher Scientific, USA) from Honeyborne et al. (2011) and Gillespie et al. (2017) including 17 mM N-Lauroylsarcosine sodium salt (Sigma-Aldrich, USA), 25 mM trisodium citrate (Sigma-Aldrich, USA) (pH 7.0 with 1 M HCl), 1% Tween 80, and 0.7% β-mercaptoethanol. Homogenisation was performed in two steps, first, using Stomacher 80 Biomaster; second, using bead beating in Lysing Matrix B 2 ml tubes (MP Biomedical, USA) with Precellys 24 (Bertin Instruments, France).

1. Samples were weighed and cut into smaller tissues using sterile scalpels.
2. Tissues were placed into plastic Stomacher bags for homogenisation.
3. Added 3 ml 5 M GTC and homogenised in Stomacher 80 for 2 minutes.
4. Supernatant (~1 ml) was transferred into homogenization tubes with lysing matrix B containing 0.1mm silica spheres for the second round of homogenisation with 500 µl Phosphate-Buffered Saline (PBS) buffer in Precellys 24, at 6000 rpm for 40 seconds.
5. Tubes were centrifuged to sediment the silica beads and supernatant transferred into new RNase-free 1.5 mL tubes for further *M. bovis* RNA extraction (described in section for optimization of MBLA).
6. Final volumes taken for RNA purification varied from 200 to 500 mL (depending on tissue volumes after homogenisations).
7. Extracted RNAs were DNase treated using the above described methods

All RNA samples undergone heat inactivation for 30 mins in 80ºC before sent to the University of Glasgow. MBLA analysis was performed for all 214 bTB infected tissues, according to the sections above and the results described in the Results section of this Chapter.

4.2.12 DNA extraction from bacteriological culture, comparison of techniques

DNA was extracted from *M. bovis* BCG cultures were grown for 4-6 weeks in 7H9 Middlebrook broth (with ADC supplement). Comparison of various techniques for DNA isolation from MTB complex strains were studied in previous research experiments (Wards et al., 1995, Amita et al., 2002, Amaro et al., 2008). Three commercial kits for mycobacterial DNA extraction for PC and IC isolates in combination with various homogenisation steps were compared. Wizard® Genomic DNA Purification Kit (Promega), Zymoclean Gel DNA Recovery Kit (Zymo Research, USA), bead based JetSeq Clean (Bioline) were tested. DNA extractions were performed using manufacturers' manuals and adapted for mycobacterial strains. The results were evaluated using PCR (targeting the gene that encodes 16S rRNA of *M. bovis*) for each isolate to estimate the best quality of DNA extracted.

4.2.13 The impact of heat treatment on 16S rRNA suitability as a marker of M. bovis loads

Work with viable mycobacteria of the tuberculosis complex requires BSL-3 facilities, which are not available in our Glasgow laboratory. Therefore, a heat inactivation study of *M. bovis* BCG (PC) culture was carried out to understand the stability of nucleic acids after heating at high temperatures. It was suggested that heat killing mycobacteria species could be useful to avoid working with live bacteria in laboratories that do not have access to BSL-3 (Sabiiti et al., 2019). Heating at 85ºC for 20 minutes in dry heat block was found suitable and rendered no viable bacilli in current study.

Cells from 1 ml of *M. bovis* BCG exponential phase cultures were resuspended in 0.1 ml 1XTE buffer and heat-inactivated at 85ºC for 20 minutes. Two sample aliquots were used for RNA extraction and two for DNA extraction with homogenisation with silica beads in TE buffer prior to lysis. RNA extraction was performed using PureLink Mini Kit (Ambion, USA), using protocol described previously. DNA was purified from the lysate using paramagnetic beads according to manufacturer's protocol JetSeq Clean (Bioline, UK). Two control samples of *M.*

bovis BCG were also used without any heat inactivation, and both RNA and DNA were extracted, respectively. Results compared the RNA yields using RT-qPCR and DNA levels using qPCR for heated/non heated samples.

4.2.14 RNA stability at room temperature (RT)

Reverse transcription quantitative PCR (RT-qPCR) reactions were conducted to investigate degradation of *M. bovis* BCG 16S rRNA stored at room temperature at different time points. RNA was purified from 3-6 weeks fresh *M. bovis* BCG culture in 7H9 Middlebrook broth (with ADC enrichment) using PureLink Mini Kit (Ambion, USA) with bead beating homogenisation in PBS buffer before lysis. Eight RNA extracts in total, from both *M. bovis* BCG and *M. marinum* were preserved at room temperature. *M. bovis* 16S rRNA loads were quantified at day 1, day 8, two weeks and one month after preservation at RT by comparing RT-qPCR Ct values.

4.2.15 Statistical analyses

A linear regression model was used to investigate any association between the results of the MBLA analysis and the tissue samples weights, as well as the presence/absence of visible lesions.

The test to compare if different molecular types of *M. bovis* were more likely to have relatively higher bacterial loads using the non-parametric Kruskal-Wallis (Kruskal and Wallis, 1952) was performed, and also estimated if there were any statistically significant differences in distribution of the MBLA results found in these *M. bovis* types pairwise using Kruskal-Wallis test (Kruskal & Wallis, 1952).

4.3 Results

4.3.1 Detection and quantitation limits of the MBL assay

4.3.1.1 Internal control

Multiplex RT-qPCR assay was performed in two compatible fluorescent channels for 16S rRNA of *M. bovis* and *M. marinum*, in case of a negative *M. bovis* test result in the target-sequence (*M. bovis* 16S rRNA), IC should always show a positive

signal. Figure 4.2 (a, b) shows the standard curve created for IC RNA, diluted in seven ten-fold dilutions (which correspond to 10^9-10^3 CFU/ml concentrations).

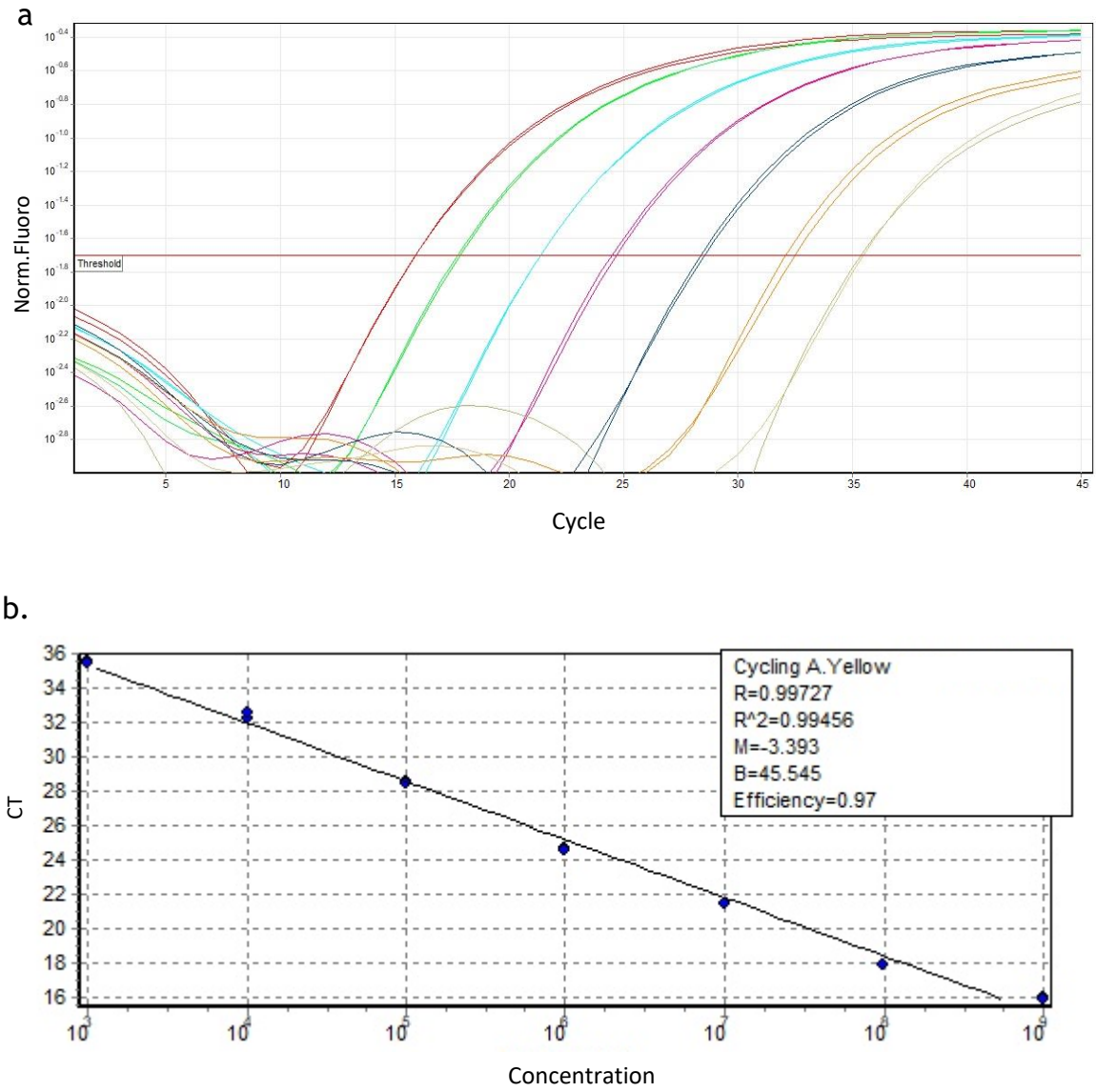


Figure 4.2 (a, b) Internal control RNA standard curve

a. Amplification curves (Log Scale) and cycle thresholds (Ct) for each dilution (10^9 -10^3 CFU/ml) (each colour represent amplification curves at different concentrations, from left to the right: red - 10^9 CFU/ml; green - 10^8 CFU/ml; light blue – 10^7 CFU/ml; purple – 10^6 CFU/ml; blue - 10^5 CFU/ml; orange -10^4 CFU/ml;); b. Calculated standard curve (R^2=0.994, slope=-3.393; efficiency=0.97)

According to the standard concentrations, IC in 10^6 CFU/ml was added in each tested sample for MBLA analysis. The comparison of the *M. bovis* RNA (FAM channel) and IC RNA (HEX channel) in two fluorescent channels shown in Table 4.3.

Table 4.3 Comparison table of RT-qPCR results shown for two fluorescent channels (FAM/HEX) in cycle threshold values for the IC performance

Sample	Extracted from	MBLA C_t value (FAM)	IC C_t value (HEX)
1	IC (Standard)	nd	24.67
2	Tissue sample 6	33.05	30.45
3	Tissue sample 10	29.69	31.1
4	Tissue sample 42	nd	29.81
5	Tissue sample 159	27.22	30.10

*nd-not detected

In the absence of *M. bovis* BCG the C_t value for the IC (10^6 CFU/ml) in HEX channel was 24.67, in the presence of the tissue samples fluctuated between C_t 29-31. Detection of the signal for IC was detected regardless of the sample result.

Four of the total 214 samples were characterized with no visible lesions present but MBLA assay detected *M. bovis* in three of them. These results suggest that if these infected animals were missed while undergoing post-mortem examinations (with no clinical signs), the MBL assay was able to detect viable mycobacteria. Most of the tested samples (98%) had visible lesions and were previously culture confirmed for *M. bovis*. From 210 NVL samples, 189 were MBLA positive and 21 tested negatives. In some cases, lesions might have been encapsulated and calcified (Cardona, 2015). Older calcified lesions are also more difficult to grind mechanically.

4.3.2 MBLA quantification range

Figure 4.3a shows the amplification curves and the cycle thresholds (C_t) of the standard curve performed for the MBLA analysis. Using mean C_t values for each concentration and each dilution, with the maximum of 1.59E+09 bacilli in ml (CFU/ml) detected at C_t 11.83 and lower quantitation limit at 1.68E+02 CFU/ml detected at C_t 35.54 (Figure 4.3 a, b).

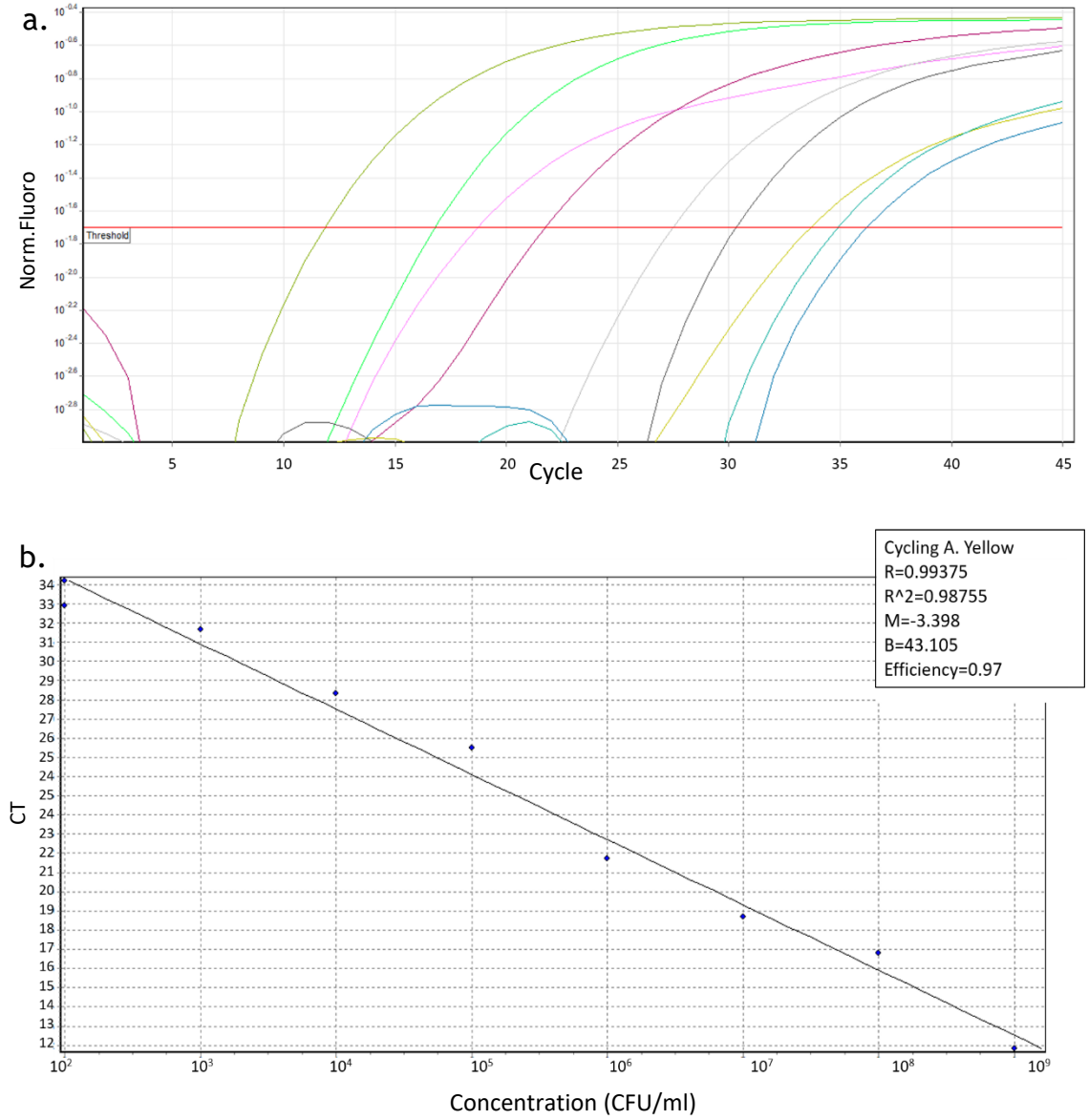


Figure 4.3 (a, b) Standard curve for *M. bovis* BCG

a. Amplification curves for eight serial dilutions of standard sample and C_t values (x-axis), fluorescence (y-axis) (each colour represent amplification curves at different concentrations, from left to the right: olive - 10^9 CFU/ml; green - 10^8 CFU/ml; pink – 10^7 CFU/ml; grey – 10^6 CFU/ml; dark blue - 10^5 CFU/ml; yellow -10^4 CFU/ml; aquamarine -10^3 CFU/ml; blue -10^2 CFU/ml); . b. Standard curve for *M. bovis* BCG with 97% efficiency (slope=-3.398, R2=0.99), concentrations (x-axis), Ct values (y-axis).

Bacterial loads of *M. bovis* in 214 tissue samples were quantified using the above-described standard curve. Mycobacterial loads were calculated in correlation with sampled specimen weights (400 mg – 7 g) used in current analyses as CFU/g of the tested tissue. Most (n=192) RNA samples amplified within C_t 25-40, therefore recognized 90% of samples as positive for *M. bovis* rRNA. Quantitation of mycobacterial loads was possible in 155 samples (79% from MBLA positive) associated between 1.09E+06 CFU/g and 3.11E+02 CFU/g of *M. bovis* bacilli in tissues. This also indicated that bacterial load does not seem to be linked to tissue

weight, although variation in extraction protocols might affect the results. Current results demonstrated that MBLA rapid and sensitive method which, subject to extensive and successful validation and cost-benefit analyses and can be considered used as diagnostic technique for detection and quantitation of *M. bovis* in bovine tissue samples. These also showed that the optimized RNA extraction protocol can be used for sufficient RNA yields in molecular biology studies of bTB. The results of the MBLA analysis of sampled bovine tissues are shown in summary Table 4.4 (full MBLA results table shown in Appendix 2 for Chapter 4).

Table 4.4 Summary of MBLA results

	Tissue weight, g	**Volume used for extraction (tissue + 3 ml GTC), ml**	**CFU/g tissue (bacterial load)**
Minimum value	0.480	3.480	3.11E+02
1st Quartile	3.422	6.423	1.81E+03
Median	4.450	7.450	4.55E+03
Mean	4.399	7.399	2.13+04
3rd Quartile	5.475	8.475	1.19E+05
Maximum value	7.300	10.300	1.09E+06

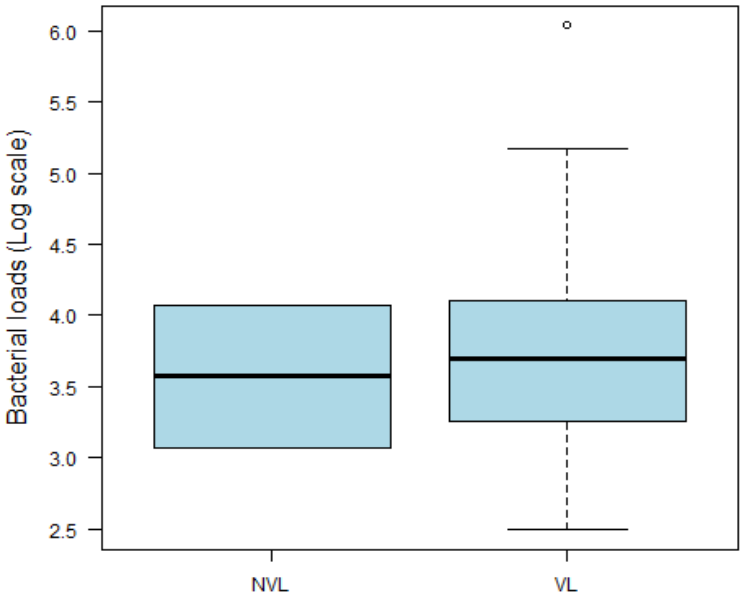
4.3.3 Impact of sample weight on measured bacterial load

In the 214 clinical, bTB positive samples from NI tested by MBLA, correlation between higher weights of tissues used for RNA extraction and mycobacterial loads were tested, tissue weights varied between 400 mg and 7 g (Table 4.4). Samples that were identified negative using MBLA and not within the quantitation limits, with missing tissue weight values, missing MLVA type information were excluded from the analysis. Results indicated that there was no correlation between the two values (correlation=-0.001), therefore demonstrating that bacterial loads of *M. bovis* in sampled tissues are not dependent on the tissue volumes (p-value=0.2484).

Visible lesions were identified in 210 samples during the post-mortem examination, and 4 tissues without visible lesions; where VL tissues were scored

from high (+++) to low (+) levels of lesion depending on histology score. These data were available for 62 of total number of samples, where 5 (+), 28 (++), 29 (+++) samples correspond to the scoring system. Examination whether appearance of visible lesions or higher scores indicate higher quantities of *M. bovis* bacilli in bovine tissues for 140 samples within MBLA quantitation range was evaluated. But no correlation between these characteristics and the MBL results (p-value=0.8235) were identified. It is important to note that due to the big difference in number of samples identified with NVL(2)/VL(138), statistical test might be underpowered to correctly answer the research question. Current results could be also affected by the number of samples in each VL scoring group, where from five positive for MBLA (scored "+") only one sample was identified within MBLA quantitation range. Results shown in Figure 4.4 (a, b).

a.



b.

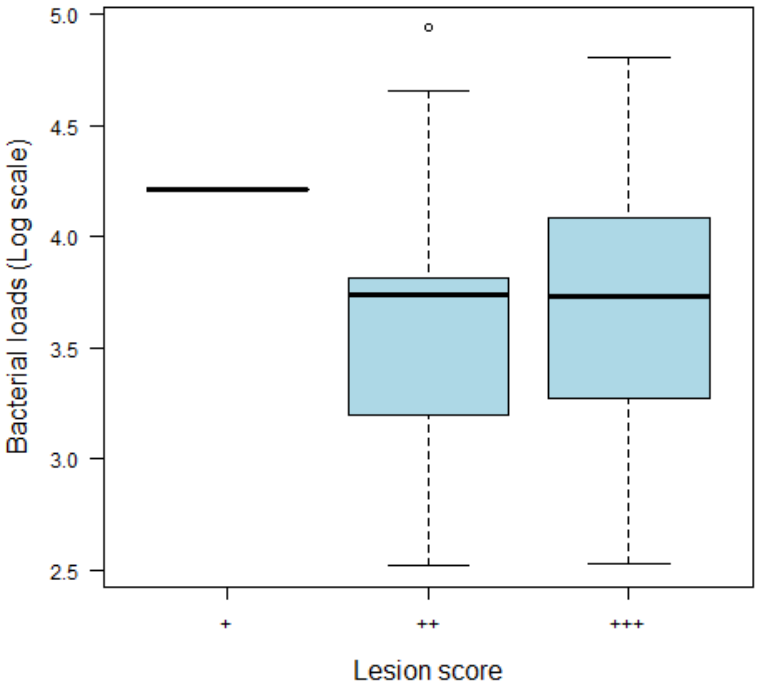


Figure 4.4 (a, b) Boxplot summary of the data for *M. bovis*, the distribution of bacterial loads

Bacterial loads (Log10 scale) in a. visible (n=138)/non-visible (n=2) lesions in tested tissue samples; b. lesion scores (data obtained from AFBI, NI)(+ (n=5), ++ (n=28), +++ (n=29).

The distribution of *M. bovis* MLVA types among tested samples were not equal. Therefore, we excluded groups with less than five isolates per type, MLVA types 001, 004, 006, 122 and 297 were retained. The results are shown in Figure 4.5. These non-parametric tests showed that there was no significant association

between mycobacterial loads and MLVA types and that our null hypothesis was not rejected (Kruskal-Wallis chi-squared = 5.1963, df = 4, p-value = 0.2677).

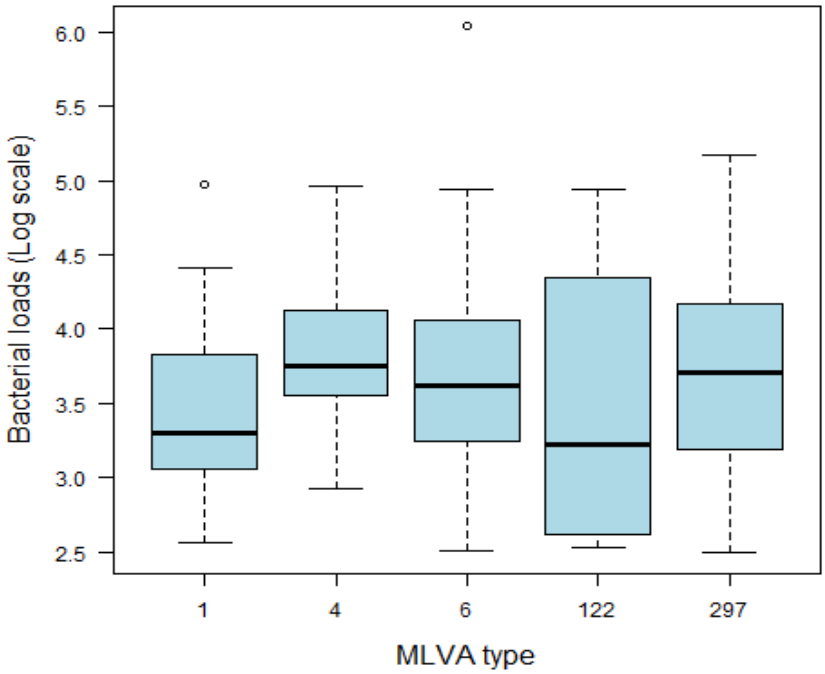


Figure 4.5 Distribution of the MBLA results in five *M. bovis* MLVA types used in statistical analysis

Boxplots demonstrate the distribution of the MBLA results in five *M. bovis* MLVA types used in statistical analysis (more than five samples per type).

4.3.4 Identification of M. bovis DNA in infected tissue samples

The use of enzymatic lysis, chemical lysis or physical lysis by bead beating linked with column or magnetic bead purification for *M. bovis* BCG DNA extraction, i.e. Zymoclean Gel DNA Recovery Kit (Zymo Research, USA), Wizard® Genomic DNA Purification Kit (Promega, USA), and JetSeq Clean (Bioline, UK) were investigated. Comparison of the DNA levels after extraction using real time PCR detecting *M. bovis* BCG DNA was done. And also, optimization of the homogenisation of the samples in TE buffer and GTC solutions using lysing matrix B (MP Biomedicals, USA) tubes in Precellys 24 (Bertin Instruments, France) was performed. Extraction protocols included in the manufacturers' kit and adapting different incubation times (at 37^{0}C) with proteinase K solution were used and DNA yields were measured using NanoDrop and Qubit spectrophotometry.

The highest *M. bovis* BCG PCR cycle thresholds were obtained using the magnetic beads extraction method, specifically JetSeq beads at C_t 14.36 and 14.48 for both samples. In comparison with Wizard (Promega, USA) kit with proteinase K solution, extracted DNA identified at C_t values 28.33 and 28.63, and when incubated overnight C_t 25.35 and 25.14, which showed that overnight incubation increases the DNA extraction efficiency. When testing homogenisation in different solutions the results showed that GTC might reduce the DNA extraction yields, most likely that it was not compatible with extraction enzymes and buffers. The DNA concentrations were extracted at average similar concentration and purity using all three kits. The results of the comparison tests were shown in Table 4.5.

Table 4.5 Comparison of *M. bovis* DNA extracted from infected tissue samples using RT-qPCR

Sample	Homogenised in TE buffer, C_t	Homogenised in GTC, C_t
Promega1	28.33	34.76
Promega2	28.63	37.35
Promega (incubated overnight) 1	25.35	32.76
Promega (incubated overnight) 2	25.14	32.55
JetSeq 1	14.36	21.42
Jet Seq 2	14.48	17.1
ZymoClean 1	11.24	24.27
ZymoClean 2	10.47	18.36

The comparison of extracted DNA and 16S rRNA of *M. bovis* in infected tissue samples were demonstrated in Table 4.6. These results suggested low extraction efficiencies of target DNA and higher extraction loads of RNA.

Table 4.6 Comparison of (RT+) qPCR results of *M. bovis* 16S rRNA and DNA extracted from infected tissue samples

Sample	*M. bovis* 16S rRNA (RT+ qPCR, C_t)	*M. bovis* DNA (RT- qPCR, C_t)
1	31.81	37.8
2	30.7	36.8

3	27.4	34

4.3.5 M. bovis rRNA stability after heat inactivation

This laboratory experiment was conducted to investigate whether *M. bovis* 16S rRNA can survive heat inactivation and still be detected in four tested isolates, also without the need of working in a high containment laboratory within current study. No viability of heat inactivated in four *M. bovis* BCG cultures was confirmed when no growth was detected after their subculture in 7H9 Middlebrook broth and incubation at 37ºC for 6 weeks. The viability of control *M. bovis* cultures, not undergoing any heat treatment, was confirmed by detecting visible growth within three weeks. Realtime qPCR results were compared and Ct values for heat treated and not treated isolates were obtained. *M. bovis* heat inactivation caused only a small decline in 16S rRNA and DNA loads. Not heated DNA isolates were amplified at C_t 13.56-15.51 and *M. bovis* 16S rRNA at C_t 10.6-13.96. After heating for 20 minutes at 85ºC DNA detected at C_t 16.79-17.3; RNA at C_t 16.43-14.63 Results are shown in Table 4.7.

Table 4.7 Results of *M. bovis* DNA/rRNA stability after heat inactivation

Sample	Not heat inactivated qPCR, Ct	Heat inactivated qPCR, Ct	Not heat inactivated RT+ qPCR, Ct	Heat inactivated RT+ qPCR, Ct
1 (DNA)	13.56	16.79		
2 (DNA)	15.51	17.3		
3 (RNA)			10.6	16.43
4 (RNA)			13.96	14.63

4.3.6 M. bovis rRNA stability at room temperature

We tested 16S rRNA stability at room temperature in two *M. bovis* isolates extracted from fresh culture to investigate how different storage conditions might affect the stability of RNA. The latter would be useful to estimate potential losses during transportation or sampling of *M. bovis* specimens. Mean C_t values were compared in this experiment for four duplicated samples of 16S rRNA extracted from BCG and *M. marinum*, four of which were homogenised in lysing matrix B

tubes (MP Biomedicals, USA) in GTC solution and four without GTC. At day 0 of storage at room temperature (used as control) mean values (for each technical duplicate) ranged between C_t 11.1 and 12.9 for BCG, and 14.1 to 14.5 for IC with no significant difference between GTC/no GTC preservation of cultures. Another qPCR was performed at day 8 of the experiment, with C_t values varying between 11.78 to 13.7 for BCG with very little degradation and 14.68 to 15.97 for IC. The results showed that after two weeks of storage, amounts of RNA were very similar to values after week one. Even after one month, the decrease in C_t showed at 16.33 to 18.9, the bacterial loads decreased only one Log in *M. bovis* BCG; and C_t 16.48 to 24.3 for IC. These results indicated that RNA extracted from *M. bovis* and *M. marinum* cultures are relatively stable at room temperature for at least four weeks and can be used for MBLA analysis. Results are shown in Table 4.8.
However, these results might be different for *M. bovis* RNA in clinical specimens.

Table 4.8 Results for RNA stability experiment

Mean Ct values shown for duplicated *M. marinum* and *M. bovis* BCG RNA stored at room temperature at different time points.

Sample	Day 0, C_t	Day 8, C_t	Week 2, C_t	Month 1, C_t
M. bovis BCG 1	12.95	13.7	14.25	16.33
M. bovis BCG 2	11.06	11.78	11.75	18.9
M. marinum 1	14.16	15.97	15.63	16.48
M. marinum 2	14.55	14.68	15.32	24.3

4.4 Discussion

The use of molecular diagnostic techniques for the efficient detection of *M. bovis* has been studied by different research groups (de la Rua-Domenech et al., 2006, Pfyffer, 2015, Ramos et al., 2015). One of the main goals of effective control/ eradication schemes for bTB in cattle is to ensure diagnostic methods, as statutory testing of all animals in NI, but supplementary testing in herds with known bTB problem, MBLA can be used in ways that balances the impact of both false positive and false negative results. No currently available tests provide 100% sensitivity and specificity (de la Rua-Domenech et al., 2006). Adding to this difficulty is the fact the epidemiology of bovine tuberculosis is very complex, involving largely

untested wildlife populations which can infect cattle and time consuming pipeline of tests to confirm the disease's presence (Morris et al., 1994, Neill et al., 2001). With the implementation of molecular biology methods such as those based on the detection of *M. bovis* nucleic acids (DNA, RNA), there are high-throughput laboratory techniques to confirm the causative agent of the infection in a shorter time (Animal and Plant Health Agency, 2022). PCR-based studies by Wards et al. (1995), Amaro et al. (2008) and Young et al. (2005) have demonstrated that the MTBC species can be directly identified in different specimens. These and other studies use amplification of *M. bovis* specific DNA, which is typically present in single gene copies. Therefore, we investigated the amplification of multicopy 16S rRNA in mycobacterial load assay (MBLA) for the detection and quantitation of *M. bovis* in bovine tissue samples, with the "prospect" of replacing the time-consuming culture-based diagnostic techniques.

4.4.1 MBLA performance

The MBLA has been implemented effectively as a method to monitor anti-TB therapy for human tuberculosis and quantifies *M. tuberculosis* bacterial load decline in response to antibiotic treatment, with the minimum detection limit of 10^2 CFU/ml of sputum (Honeyborne et al., 2014, Gillespie et al., 2017, Sabiiti et al., 2020a, Hai et al., 2021).

It was proved that MBLA can be used for *M. bovis* quantification within eight orders of magnitude. This is comparable to DNA-based quantitative methods for *M. bovis*, such as Taylor et al. (2007), Costa et al. (2013) and Zarden et al. (2013). On the other hand, mRNA-based RT-qPCR tends to have more narrow quantitative range.

In order to compare an alternative method with culture, the gold standard for mycobacterial identification, it is desirable for the alternative method to be able to detect single bacterial cells. The sensitivity of MBLA for the detection of *M. bovis* was 90% (identified in current study), comparing to DNA based PCR techniques (sensitivity 66.7-100%) (Thacker et al., 2011, Costa et al., 2013, Zarden et al., 2013) and mRNA based PCR methods (sensitivity 76.8-100%) (Therese et al., 2012, Montenegro et al., 2014, Atahan et al., 2020). Comparing to DNA and rRNA, mRNA has much shorter half-life (few minutes) (Belasco et al., 1986, Beggs et al., 1995) and was demonstrated as a good bacterial "viability" marker and drug

susceptibility of *M. tuberculosis* (Hellyer et al., 1999). The use of RNA based techniques to detect *M. bovis* directly from sampled specimens is not yet studied well. Several research groups studied the use of *M. bovis* 16S rRNA as RT-PCR target to differentiate between bTB and non-TB strains (Miller et al., 2002, Quan et al., 2017). Although non-infected samples were not used for the current analysis to calculate the statistical specificity of the technique, the bovine tissues (non-infectious) spiking with *M. bovis* BCG to simulate the infection in cattle and optimize the MBLA protocol was performed.

4.4.2 Internal control performance

M. marinum was implemented as internal control in order to monitor the RT-PCR reaction performance and to avoid any false negative results. The 16S rRNA sequence was sufficiently different from that of *M. bovis*, and the identification was enabled by detecting distinct fluorescence signal (from HEX fluorophore, detected in the yellow PCR instrument channel). The use of IC is crucial for any PCR analysis to avoid the false-negative results of the assay and RNA extraction errors. And the performance of the *M. marinum* used for MBLA analyses were shown in Table 4.3, where the IC was detected for both positive and negative samples.

Sample type can affect the performance of extraction methods and the subsequent sensitivity of a molecular assay. The optimized sample preparation and RNA extraction protocol for the use of bovine tissue samples and efficient extraction of RNA was based on previous studies (Honeyborne et al., 2014, Mićić et al., 2016). The use of optimal tissue preservation and preparation procedures were also demonstrated, in combination of several homogenisation steps and efficient lysis of the bacterial cell wall. The optimised protocol includes tissue incision mechanical cell wall disruption by high-speed bead-beating, RNA extraction and purification, and enzymatic DNA degradation. The estimation of detection and quantification limits of MBLA on bovine lung and lymph node tissues from uninfected animals was performed.

The use of MBLA for animal tissues showed high sensitivity of the method and was able to detect *M. bovis* in 90% of tested samples that were previously confirmed as culture positive for *M. bovis*. The 10% of samples that were detected as MBLA

negative might be explained by samples preparation errors or that the lesions sampled had lower mycobacterial loads than those parts used for bacteriological culture. Moreover, four out of 214 MBLA positive sampled tissues (lungs and lymph nodes) were not confirmed with visible lesions during post-mortem examination at the abattoir. This indicated that current technique might be effective in detecting positive samples that might be missed during other tests, although these were small numbers to make such assumptions. A simulation study by Ezeoke (2020) described the possible experimental design of the MBL assay for the detection of *M. bovis* 16S rRNA in animal tissues, indicating the number of required technical replicates (six) of the RNA serial dilutions for efficient detection and quantification of *M. bovis*. The generated standard curve to calculate *M. bovis* concentrations in tissue samples based on triplicates of each serial dilution used for tissue spiking. Total RNA was extracted from each spiked tissue (three replicates) and amplified in duplicates in RT-qPCR (total 6 replicates); which is consistent with the above-mentioned suggestions. The trial was useful to demonstrate the dynamic range of the MBL assay and was able to quantify mycobacterial loads in 79% of those that were MBLA positive. The 21% that were not quantified but positive for MBLA were also culture confirmed.

Firstly, these findings indicated no positive correlation between the tissue volumes and extracted mycobacterial loads. One potential explanation is that sampling of tissues for the analyses cannot predict the amounts of bacilli that were present (Table 4.4). It is important to mention that we recalculated the CFU/ml bacilli concentration in the amounts of tissue samples in CFU/g. RNA has the standardized protocol for the different sizes of the samples. And therefore, different sizes of tissues can contain higher or lower *M. bovis* loads. The second assumption was that the presence of granulomatous lesions can be indicative of the mycobacterial loads, however, we could not identify the association between the MBLA assay results and lesions found in samples (Figure 4.4a).

The quantitative MBLA was used to investigate whether different *M. bovis* MLVA genotypes correlated with bacterial loads in infected tissues. And used MLVA types with more than five samples and subset the data using only MLVA types 001, 004, 006, 122 and 297, no correlation between bacillary load in infected tissue and the genotype was identified (Figure 4.5).

Another advantage of MBLA assay is the use of *M. bovis* 16S rRNA as target for PCR amplification directly from bovine tissues, which was not fully studied yet. Some researchers demonstrated different mycobacterium RNA extraction protocols (Patel et al., 1991, Mangan et al., 1997, Butcher et al., 1998, Rustad et al., 2009) and its potential in using for PCR-based detection. Some studies were focused mostly on distinguishing between closely related MTBC species, or tuberculosis versus non-tuberculosis species. In study by Miller et al. (2002) the focus was to detect and discriminate the closely related *M. tuberculosis*, *M. bovis* and *M. avium* in formalin-fixed bovine tissues that were culture-negative. In the latter studies specific primers for *IS6110* to detect MTBC species, and 16S rRNA and *IS900* specific-primers to detect *M. avium* were used, and the results showed that in 58% of tissues mycobacteria could still be identified even though they were negative at culture (Miller et al., 2002). Another study by Quan et al. (2017) developed PCR method for direct differentiation of MTBC species (*M. tuberculosis*, *M. bovis* and *M. bovis* BCG) and non-tuberculous mycobacteria using 16S rRNA and Rv3873 (PPE protein) sequences with high specificity (100%) and detecting from 15 pg of genomic DNA. But currently known studies for quantitation of *M. bovis* directly from animal tissue samples are limited.

Research study by Honeyborne et al. (2011) initially implemented MBLA to detect *M. tuberculosis* 16S rRNA in human sputum samples. To my knowledge, MBLA was not yet used for *M. bovis* identification in sputum or animal tissue samples. There have been several studies examining the detection of MTBC species from tissue specimens, mostly using DNA targets for PCR amplification. *M. bovis* DNA extraction and PCR detection from different sample types were widely used in bTB diagnosis, specifically animal tissues, such as lungs, lymph nodes and other organs (Thacker et al., 2011, Costa et al., 2013). Costa et al. (2013) used a semi-nested PCR targeting the regions of insertion sequence *IS6110* to identify *M. bovis* isolated from fresh tissues collected from bovines, deer, foxes and wild boars. The results demonstrated high levels of sensitivity (98.2%) and specificity (88.7%) in detecting *M. bovis* DNA using this assay (Costa et al., 2013) . Another study by Thacker et al. (2011) used detection of *IS6110*_T insertion element of *M. bovis* using PCR in homogenised tissue samples, with the overall good performance of the molecular test. The assay detected *M. bovis* DNA in 20 of 30 infected cattle and not detected *M. bovis* in 18 control negative samples. Several studies assessed

the use of DNA extracted from milk samples as PCR template for *M. bovis* detection using PCR (Zumarraga et al., 2012, Franco et al., 2013, Zarden et al., 2013). Ingestion of raw milk can be a potential risk for zoonotic bTB transmission to humans in some developing countries, hence identification of *M. bovis* was important. However, the sensitivity of the tests was 0.03-50% (Franco et al., 2013, Zarden et al., 2013).

Previous studies demonstrated that *M. bovis* can be shed by animals (depending on species) in their urine, faeces and sputum into environment and can be transmitted to animals through indirect route, such as inhalation or ingestion of mycobacteria (Palmer et al., 2004, Corner et al., 2012b). The potential of detecting *M. bovis* DNA and RNA from environmental samples was shown by Young et al. (2005), that provided evidence of long-term persistence (after more than 4 months after infection) of *M. bovis* in soil microcosms from farms with known bTB history in Ireland. This showed that *M. bovis* can survive in environmental samples outside the hosts, with the optimal temperature 37^{0}C for mycobacterial survival (Young et al., 2005). Considering the above-mentioned findings and our results, the MBLA assay may find application for detection of *M. bovis* in different types of samples potentially. Moreover, to study the potential epidemiological importance of *M. bovis* surviving in faeces, soil, water that provides some confidence in actual viability of the bacteria (Allen et al., 2021).

Work with highly infectious bacteria such as *M. bovis* and suitable inactivation of the bacilli would allow researchers to work in lower BSL laboratories, while maintaining the nucleic acids intact for downstream experiments. Previous studies showed the utility of heat inactivation of MTBC bacteria for 20 minutes at 80, 85 and 95^{0}C (Doig et al., 2002, Sabiiti et al., 2019). Establishing experiments with heat inactivation of *M. bovis* BCG and comparing RNA and DNA amounts by qPCR showed that with very minimum losses we can extract nucleic acids (DNA/rRNA) suitable for amplification but not infectious which indicated the possibility of further work being carried out in a BSL2 laboratory. This also implies that clinical samples from animals obtained in the field, or tissue samples after post-mortem examination can be used for RNA-based techniques such as MBLA after heat killing of bacilli without impact on mycobacterial loads.

The tested stability of *M. bovis* BCG 16S rRNA at room temperature during extended time and our findings confirmed that if sampled bTB isolates were stored and transported in room temperature within up to one month, it should not affect significantly the identification of mycobacterial load.

Most diagnostic techniques for laboratory confirmation of bTB in cattle require well equipped laboratory facilities and levels of biological safety (World Health Organization, 2012). These include BSL3 laboratory facilities, biosafety cabinets, separate conditions for work with samples, bacteriology cultures and molecular biology tests. Moreover, bTB control policies include expenses for testing and removal of cattle, confirmation at post-mortem and bacteriology, and further risk management measures (Northern Ireland Audit Office, 2009, Skuce et al., 2010, Godfray et al., 2018). All these steps are time- and labour-intensive and add to the economic burden of the disease. Therefore, proposed MBLA analysis as specific, accurate, rapid and highly cost-effective technique that does not require specific BSL laboratory facilities, subject to extensive validation and cost-benefit analyses. Moreover, proper epidemiological study of sensitivity and specificity of the technique versus gold standard (or other approaches) could be useful for future analysis.

The estimated the economic cost of MBLA assay reagents as a diagnostic molecular biology laboratory test and calculated the approximate cost of the reagents for one *M. bovis* sample (bovine tissue/or potentially other types of samples). The price included the sample preparation, RNA extraction kit, DNase treatment, and RT-qPCR kit including the price of primers and probes to detect *M. bovis* 16S rRNA, which in total for one sample will cost around £15 (calculations not shown). Whereas, the bacteriological culture costs per one specimen would cost about 28-32$ (~23-26£) as estimated for LJ and MGIT cultures calculated in Mueller et al. (2008). It is important to note, that MBLA prices were calculated for the reagents ordered in 2018-2019. There was an increase in prices in years 2020-2022 for some consumables and changes in supplier companies.

Despite thorough RNA extraction and MBLA optimization before the use of MBL assay for *M. bovis* detection in infected tissues, some limitations were recognized. The amounts of bacilli extracted from infected tissues and granulomas (lesions) was not always in high abundance, which is seen from quantified bacterial loads.

Lesions and tissues with low numbers of mycobacteria is a well-recognised clinical picture and may affect the positivity of MBLA results (Domingo et al., 2014).

4.5 Conclusions

In conclusion, the use of the MBLA might be valuable supplementary test for routine laboratory confirmation of bTB and implemented as rapid and sensitive technique – however the method needs wider validation and characterisation to determine its utility. Optimised sample preparation and RNA/DNA extraction protocols showed high yields of *M. bovis* bacilli from bovine tissues can be easily used in laboratories. In particular, findings of this thesis Chapter regarding the use of 16S rRNA target as a "marker of bacterial viability" can be important in understanding the disease progress.

5 Chapter

Overall discussion

5 Chapter

5.1 Discussion

Bovine tuberculosis (bTB) is a chronic disease in cattle that has significant impact on the farming industry and the government economy in Northern Ireland (NI). Despite thorough disease control measures in the country, *Mycobacterium bovis* persists in both livestock and wildlife populations. There are several factors that make the situation with bTB complex, such as the presence of a wildlife host, the low sensitivity of diagnostic techniques and the time-consuming confirmation tests, and the difficulties with vaccination. Therefore, in order to effectively control this disease, it is crucial to improve our understanding of *M. bovis* spatial distribution, genomic epidemiology, transmission, and diagnosis.

One such direction is the use of both spatial distribution and genomic epidemiology of *M. bovis* to understand the disease persistence and spread in different host species. Currently used molecular typing techniques such as multilocus variable-number tandem-repeat analysis (MLVA) typing and whole genome sequencing (WGS) provide valuable information regarding the *M. bovis* population structure across NI. I used both techniques to study *M. bovis* spatial structure in the TVR area (Chapter 2), as well as to improve our understanding of *M. bovis* transmission dynamics between cattle and badgers in this region (Chapter 3). Another direction is the development of rapid and reliable molecular diagnostic techniques for *M. bovis* detection to improve our ability to detect viable mycobacteria without the need for long laboratory confirmation. In the current study, I introduced the novel molecular approach to identify *M. bovis* 16S ribosomal RNA directly from bovine tissues.

5.2 Chapters' summary

Chapter 1 introduced the history of bTB infection in the United Kingdom, as well as an overview of the available diagnostic tests for *M. bovis* detection and molecular typing. Data used in the current thesis (Chapters 2 and 3) included two types of *M. bovis* molecular (multiple locus variable number of tandem repeats analysis (MLVA)) and spatial data collected from cattle and badgers from the test and vaccinate or remove (TVR) intervention study area in Northern Ireland that was implemented in years 2014-2018, additionally included historically sampled

data from 1986-2013 surveillance cattle data and badgers killed in road traffic accidents. Laboratory analysis presented in Chapter 4 used *M. bovis* infected tissue samples collected in 2018 from the same area.

In Chapter 2, I investigated the presence of the spatial clustering of *M. bovis* in cattle and badgers, determined the MLVA type diversity present in cattle and badgers, and studied the relationship between genetic and spatial diversity of MLVA types between host-species. The presence of the same *M. bovis* MLVA types in both cattle and badger species as well as the close links between the spatial distribution of the bTB infections suggested the existence of inter-species transmission between the two host-species. Furthermore, I focused on the shared *M. bovis* MLVA types between cattle and badger species and demonstrated the high prevalence of MLVA type 006 among cattle and badgers (>51%), which was previously identified as historically endemic MLVA type in the TVR area, and identified as a founder MLVA type using goeBURST algorithm (Francisco et al., 2009). This demonstrated that current MLVA type highly influence the spatial distribution of other neighbouring MLVA types found in sympatric cattle and badgers. The predictions made to estimate the association of MLVA types present in cattle and badgers within study area was only possible in the absence of MLVA type 006. I discuss the beneficial use of WGS techniques for specific MLVA type 006 in this dataset to better understand the spread of *M. bovis* within the area and better discriminate isolates within such closely related bTB in cattle and badgers.

In Chapter 3, I showed that the integration of whole-genome sequencing data with phylogenetic tools allowed to obtain an unprecedent level of resolution for the investigation of *M. bovis* evolution and transmission dynamics across host species in the TVR area in NI. The use of host species as discrete traits for ancestral host state reconstruction has shown that *M. bovis* in the TVR area is transmitted in both directions (from cattle to badgers and from badgers to cattle) during the entire period 1986-2017 and during the time subperiods of 1986-2011 and 2012-2017. In this chapter, I also discuss that the results of these types of analyses should always be interpreted with caution due to slow evolutionary rate of *M. bovis* and possible sampling biases caused by a non-systematic data sampling over time. The estimated mean evolutionary rates 0.36-0.37 were similar to other research studies (Crispell et al., 2017, Crispell et al., 2019, Salvador et al., 2019).

Moreover, the estimated tMRCA (1970-1980s) of these phylogenies were consistent with the *M. bovis* historical population expansions descried previously (Robinson, 2015). I also used graph approach SeqTrack and *outbreaker2* tools to reconstruct the transmission trees of 302 *M. bovis* isolates (endemic lineage).These methods showed that even with differences in sample sizes collected from cattle and badger *M. bovis* there was some evidence of inter-species transmission, with most bTB transmissions coming from cattle ancestors.

Chapter 4 described the optimization and use of the molecular bacterial load assay (MBLA) for the detection and quantitation of *M. bovis*. Little research was known to directly extract *M. bovis* 16S rRNA from animal tissue samples and quantification of their mycobacterial loads. Molecular bacterial load assay is a culture-free method for detection of *M. tuberculosis* in human sputum samples (Honeyborne et al., 2011). As a potential to replace time-consuming culture-based diagnostic techniques by the use of 16S rRNA as a target for RT-qPCR, which has a higher number of copies in a cell and a shorter half-life than DNA. This method is more sensitive to changes in bacterial load and can detect viable bacilli in sputum specimens. This method was not used previously for the detection and quantitation of *M. bovis*. I optimized the use of MBLA for *M. bovis* identification directly from bovine tissue samples and analysed 214 samples using this method and identified *M. bovis* 16S *rRNA* in 90% of tissue samples. Quantification of mycobacterial loads was possible in 79% of MBLA positive samples, while 21% negatives were previously culture confirmed. I demonstrated the use of *M. marinum* non-tuberculosis mycobacteria as an internal control for the RT-qPCR. I performed series of additional laboratory analyses to evaluate the use of different *M. bovis* DNA extraction methods, inactivation of mycobacterial RNA by heating and estimated the RNA stability at room temperature for transportation and storage purposes.

5.3 Spatial distribution of *M. bovis* MLVA types in cattle and badgers

Several studies have demonstrated the spatial distribution of *M. bovis* molecular types in cattle and badgers across Great Britain and Northern Ireland (Woodroffe et al., 2005, Jenkins et al., 2007, Milne et al., 2020). In Chapter 2, using MLVA *M. bovis* types, I showed that there are similarities in *M. bovis* molecular types circulating in both cattle and badger populations, and that there is a strong spatial

association between them. In the Skuce et al. (2020) bTB surveillance study in cattle herds it was demonstrated that molecular types are strongly associated with the Division Veterinary Office (DVO) regions across NI, which is consistent with the findings of the current study. Authors suggested that the local transmission in cattle is playing a key role in bTB transmission. In current Chapter, geospatial data was restricted to cattle farm locations for *M. bovis* found in livestock and to locations of trapped badgers within the intervention study and of badgers killed in road traffic accidents (RTA). Therefore, these locations are related to the time of death of the animal and they are not representative of their home range.

A total of 36 MLVA types were found from the 1248 *M. bovis* isolates, and described in Chapter 2, from which only 6 were found in badger hosts. This is not surprising since cattle populations in NI are much higher than the badger ones and, in general, there is fewer data collected from badger. Some *M. bovis* MLVA types were specific for badgers, indicating the importance of badger-badger transmissions. With the limitations in *M. bovis* sampling from badgers in previous years, before the intervention study, little was known about how much badgers influenced bTB prevalence in cattle in NI. Therefore, it was important to investigate the role of badger species in *M. bovis* transmission to cattle.

Estimation of the kernel densities of *M. bovis* MLVA types that were shared between the two hosts demonstrated a high spatial overlap between cattle and badgers MLVA types 004, 006 and 297 (Figure 2.6). The analysis also predicted the spatial distribution of MLVA types in cattle using only badger data and vice-versa, which indicated the existence of cross-species transmission in the area. In Chapter 3, I created a maximum likelihood tree of the *M bovis* MLVA types (Figure 3.3) and showed that types 006, 122 and 297 were the part of the same clade in the spoligotype 263 group, while the MLVA type 004 isolates were part of a distinct clade. It is most likely that MLVA type 004 was brought-in to the TVR area via cattle movements from different regions. The spatial distribution of *M. bovis* in cattle and badgers showed that overall, distances between infected badgers were shorter than within both cattle and badgers (Figure 2.3). The presence of wider population of *M. bovis* MLVA types in cattle, in some cases in single isolates, indicated that the introduction of those molecular types is more likely via cattle movements from long distance regions (Green et al., 2008, Milne et al., 2019), and not from badger species, which are likely to move only for shorter distances.

5.4 Use of whole genome sequencing for *M. bovis* epidemiology

The phylogenetic analyses in Chapter 3 described the *M. bovis* genetic populations circulating in the TVR area. I showed that even with such a slowly evolving pathogen as *M. bovis,* the presence of temporal signal in all datasets was still possible. I also estimated the substitution rate which was similar to what was shown in previous studies by Trewby et al. (2016), Crispell et al. (2019).

Even with thorough *M. bovis* sampling in cattle and wildlife within the TVR intervention study in 2014-2018, it is likely that some infected cases in cattle and in badgers were never detected, leading to an underestimation of the disease prevalence. This might indicate that conclusions regarding the phylogenetic inferences and intra- and inter-species transmission of bTB as well as the directionality of such transmissions must be interpreted with care and are very specific to the dataset analysed. Data included in this study was systematic and covered a long time-period of thirty-one years. The observed bTB sampling in livestock and wildlife populations during the TVR project was very close to the expected sample sizes planned to cover at the beginning of the study (DAERA, 2018a, Menzies et al., 2021). However, the prevalent number of *M. bovis* samples was collected from cattle species than from badgers, and the percentage of cattle-cattle transmissions was demonstrated when reconstructed the transmission trees using non-phylogenetic methods, such as *Seqtrack* and *outbreaker2*. The use of phylogenetic and transmission trees reconstruction tools is widely studied for different pathogens. However, complex epidemiology and evolution of *M. bovis* must be considered when using such techniques and applied with caution.

Very little genetic diversity was identified for inter-SNP distances within 6.263 endemic clade studied (mean ~7.5 SNPs) in this case, the use of WGS is important where other molecular typing methods, such as MLVA and spoligotyping do not have enough resolution. However, one possible direction for future development of diagnostic technique that can specifically target this short but informative regions specific for the endemic strain family is using SNP-genotyping methods. The use of this technique was described in studies for bacterial genotyping and differentiation between isolates (Moorhead et al., 2003, Amoako et al., 2017) and allows researchers to target specific pathogen-associated SNPs using PCR based

methods. This will be less labour consuming than WGS, but still useful for bTB epidemiology as useful indicator of particular lineage.

M. bovis is a member of *M. tuberculosis* complex that causes bovine tuberculosis in a wide host range, primarily cattle but also in various wildlife species, and in humans. The ability of the bacterium to adapt in different host environments is known to be affected by evolutionary factors, such as changes in virulence and host associated factors (Allen, 2017). The use of WGS and bioinformatics methods provide unprecedent opportunities in determining the basis of host adaptation. For further research direction, it is important to use comparative genomics and genome assembly tools to discover the mechanisms that are responsible for the adaptation of *M. bovis* to different hosts, in particular to European badger, as a maintenance host of *M. bovis* in NI. Furthermore, it is also important to determine the underlying gene signatures (if they exist) associated to each host-species. For that, we can reconstruct the *M. bovis* pan-genome (composed by core and accessory genes) to identify and quantify the set of genes that compose *M. bovis* extracted from each species. Dissimilar host-specific genetic signatures would indicate that the bacteria went through evolutionary events such as gene loss/gain and adapted to a specific host. Specifically, pan and core genomes of *M. tuberculosis* were studied in human adapted strains and specific virulence genes were identified that could be important for vaccine development (Yang et al., 2018, Zakham et al., 2021).

5.5 Molecular diagnostic methods for detection of *M. bovis*

Together with traditional methods of bTB confirmation it is important to use highly sensitive diagnostic techniques that can detect mycobacteria directly from collected samples. In Chapter 1, I describe different diagnostic tools that are currently used for *M. bovis* detection in laboratory. Complex immune response of different animal host to *M. bovis* infection has the potential for research studies in a long term, which can be used for further development of bTB diagnostic methods in various hosts. Complex *M. bovis* pathogenicity and host-pathogen interactions create additional challenges for the disease diagnosis. Another challenge is to culture *M. bovis* where some proportion of collected samples remain unconfirmed.

In Chapter 4, I introduced the MBLA assay molecular technique that can be used as part of *M. bovis* laboratory confirmation and can also be used for different types of samples (such as environmental). I performed experimental analyses and collected samples from bovine lungs and lymph nodes that were *M. bovis* culture-confirmed. One of the limitations of current study was obtaining tissues after *post-mortem* examinations of carcasses, where some granulomatous lesions could be missed or obtained with lower levels of mycobacteria (it is difficult to predict which samples contain higher loads of *M. bovis)*. To resolve this potential limitation, I optimized the protocol with spiking of non-infectious tissue samples with known concentrations of *M. bovis* BCG. The detection range of MBLA showed that the method was able to detect very low levels of bacilli in samples.

Sensitivity of bacteriological culture methods are on average lower than molecular detection techniques (Courcoul et al., 2014), and are used as the gold standard technique because it detects only "viable" bacteria. Most currently used PCR based *M. bovis* detection methods are based on detection of mycobacterial DNA and mRNA. In the current study, I used 16S ribosomal RNA as a marker of bacterial viability. This might be helpful in identification of live mycobacteria in animals and potentially quantify the infectivity doses of *M. bovis* in cattle and levels of disease severity for further research purposes. Even though I could not identify the correlation between different phenotypic characteristics of the animal samples used for MBLA. I showed that there was no significant association between the tissue weights used for *M. bovis* 16S rRNA extraction, lesioned and not lesioned tissues and the bacterial loads quantified in the tested samples. This might be explained by missing data for some samples, where most likely statistical tests were underpowered. Obtaining more detailed information about the infected tissues might resolve the problem.

Although MBLA assay demonstrated high sensitivity (90%) and quantified bacterial loads of *M. bovis* in >73% of positive bovine tissue samples, a more thorough sampling of negative tissues for MBLA specificity testing would be necessary to complement the current analysis. MBLA has a potential to be used as an accurate *M. bovis* diagnostic technique in various host species samples and can be potentially be optimized for use on environmental samples, such as soils. The

importance of *M. bovis* in environmental and soil samples has been reported by Young et al. (2005) and reviewed by Allen et al. (2021).

5.6 Conclusions

Overall, this PhD thesis aimed to answer main epidemiological questions related to bovine tuberculosis in cattle and badgers in NI, based on three different approaches: *M. bovis* spatial distribution, *M. bovis* molecular epidemiology and transmission, and *M. bovis* laboratory diagnosis. Using the experience of previous research studies from countries with known history of bovine tuberculosis, it was possible to infer important knowledge about the infection persistence and spread of *M. bovis* in Northern Ireland. The opportunity to obtain novel data from the first wildlife intervention study (TVR) implemented in NI provided strong evidence of the bi-directional transmission of M. bovis in this complex multi-host system composed by cattle and badgers.

The association of *M. bovis* molecular types found in both cattle and badgers and their close spatial distribution demonstrated that there might be some local epidemiological and environmental factors that influence the persistence of the infection in the area. Moreover, the presence of the endemic lineage that is circulating in the South-East of NI confirms these assumptions. This study provided an in-depth analysis of the evolutionary and transmission patterns of *M. bovis* isolates found in cattle and badgers within the TVR area showing that cattle to badgers transmissions played a major role in *M. bovis* transmission dynamics. The experimental analysis for MBLA used for *M. bovis* isolates provides support to use PCR-based techniques as a diagnostic test. The detection and quantification of *M. bovis* 16S rRNA as a marker of 'viability' of the bacteria showed that MBLA can be used for different types of samples, such as contaminated environmental samples.

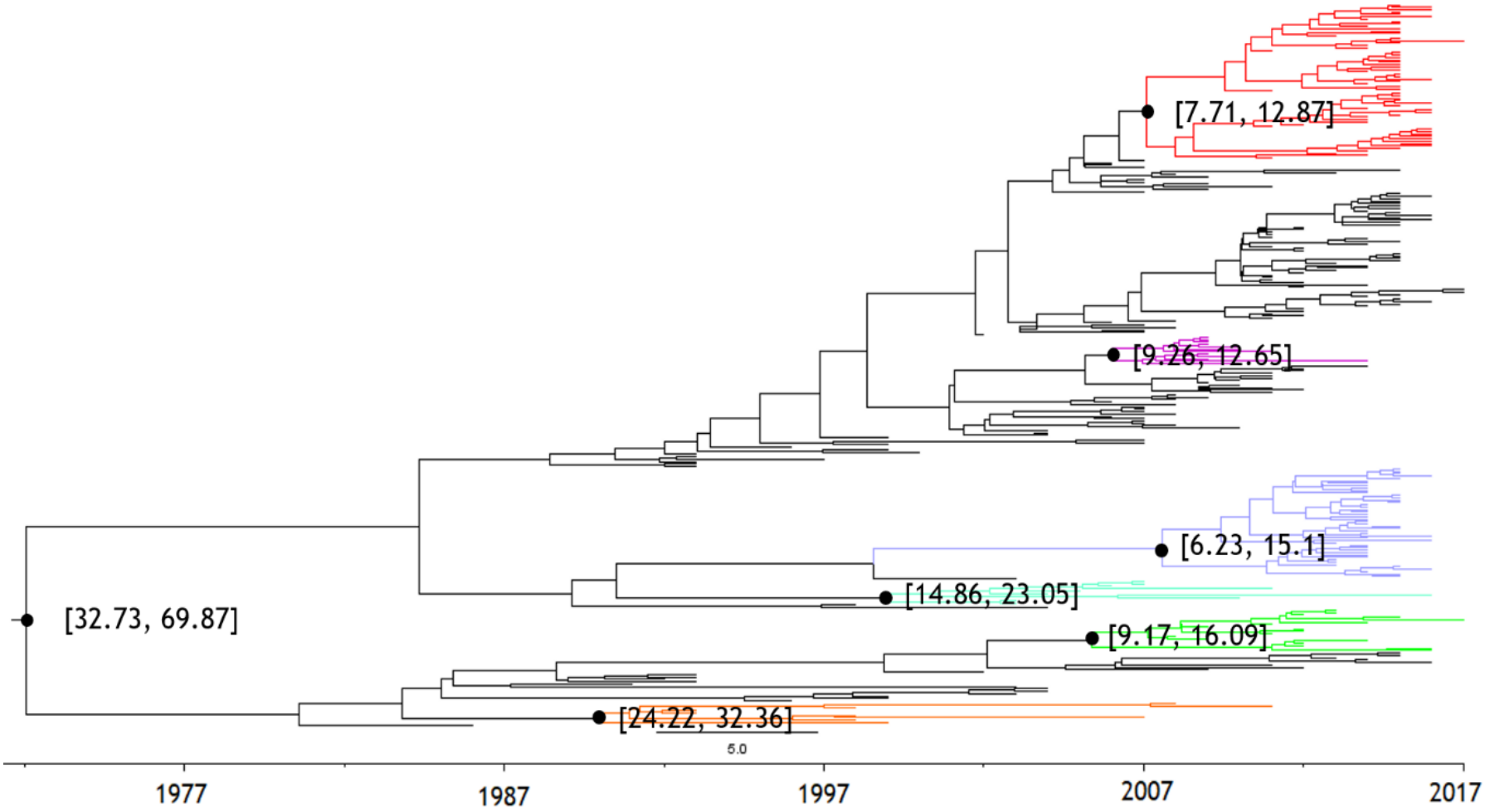
Appendix 1

Chapter 3

Time-calibrated maximum clade credibility tree under the GTR model, relaxed log normal and the coalescent constant population model with posterior probability (PP) support for major nodes is shown (Figure 1a). Host species were modelled as a discrete trait over the full *M. bovis* dataset genealogy by ancestral state reconstruction using Discrete Ancestral Trait Mapping (DATM) approach in Beast2 (Figure 1b).

M. bovis data subsets 1 and 2 were also used for ancestral hosts reconstruction using the Constant population size demonstrated in Figures 2 (a, b) and 3 (a, b).

a.


[7.71, 12.87]
[9.26, 12.65]
[6.23, 15.1]
[14.86, 23.05]
[32.73, 69.87]
[9.17, 16.09]
[24.22, 32.36]
5.0
1977
1987
1997
2007
2017

Years

b.

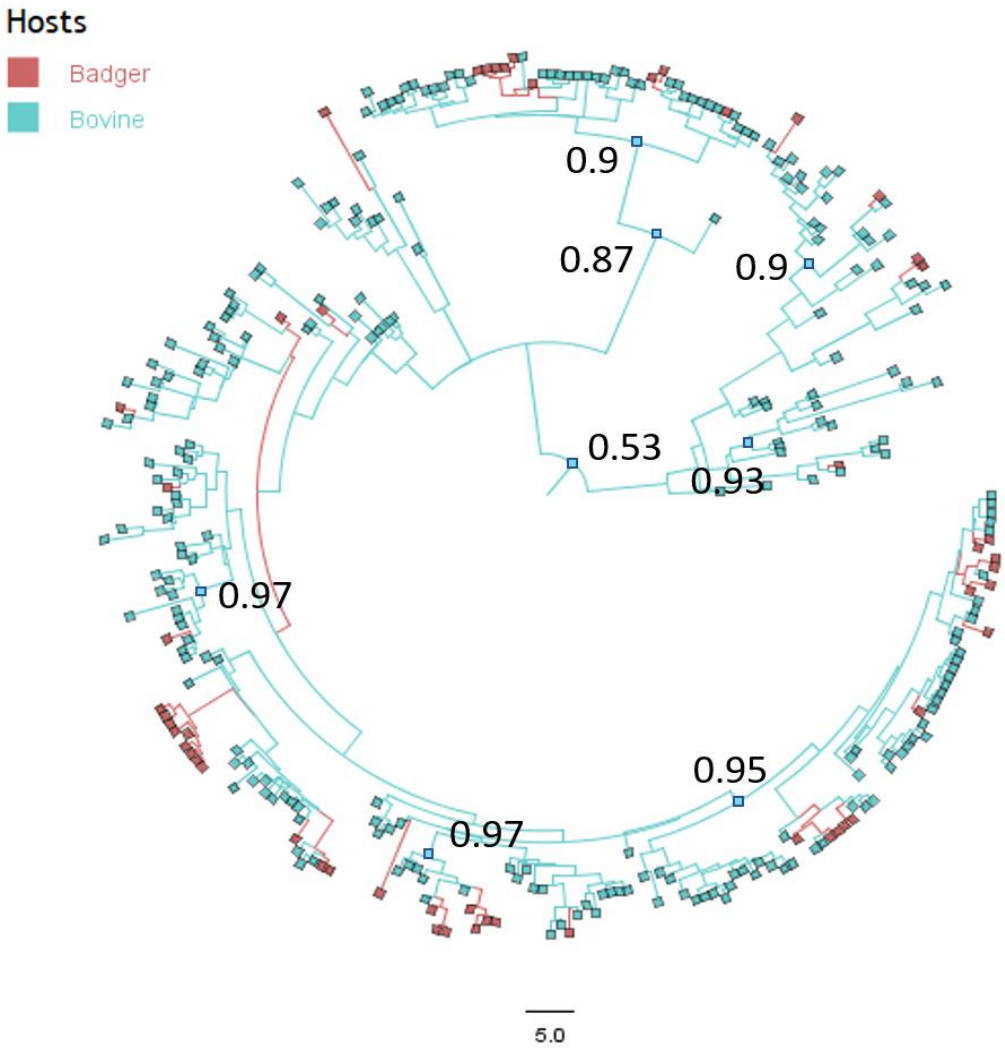
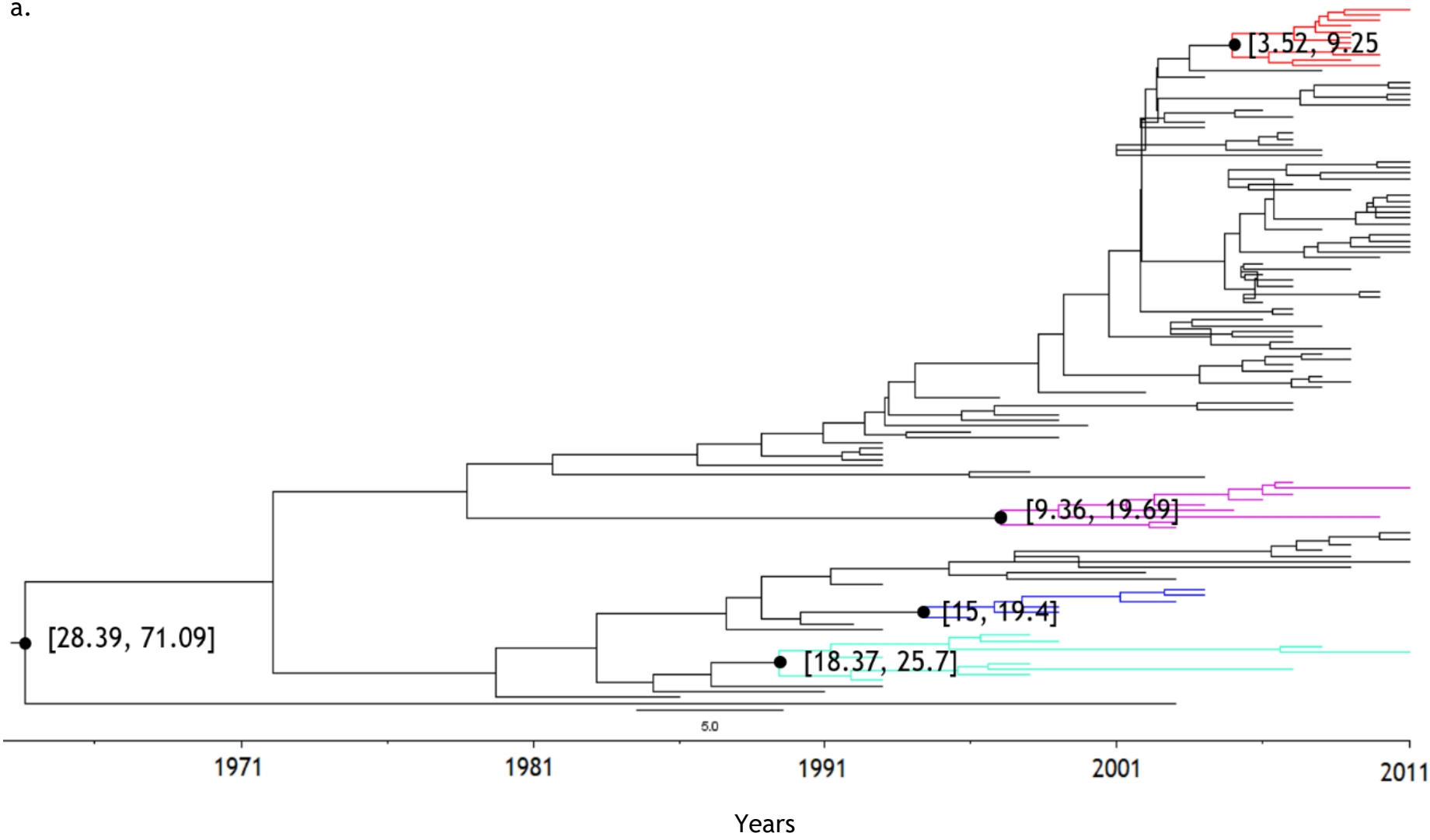


Figure 1 (a, b) Maximum clade credibility tree for 302 *M. bovis* isolates

Constant population size a. Ancestral nodes with higher posterior support (tree) (PP>0.95) shown as black circles, 95% HPD interval for TMRCA estimates for each clade shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badger in red) states for the associated main clades shown in Figure 1a. Cattle was identified as the ancestral hosts for the MRCA of all the major clades and for the MRCA of all the isolates.

a.


[3.52, 9.25
[9.36, 19.69]
[15, 19.4]
[18.37, 25.7]
[28.39, 71.09]
5.0
1971
1981
1991
2001
2011
Years

b.

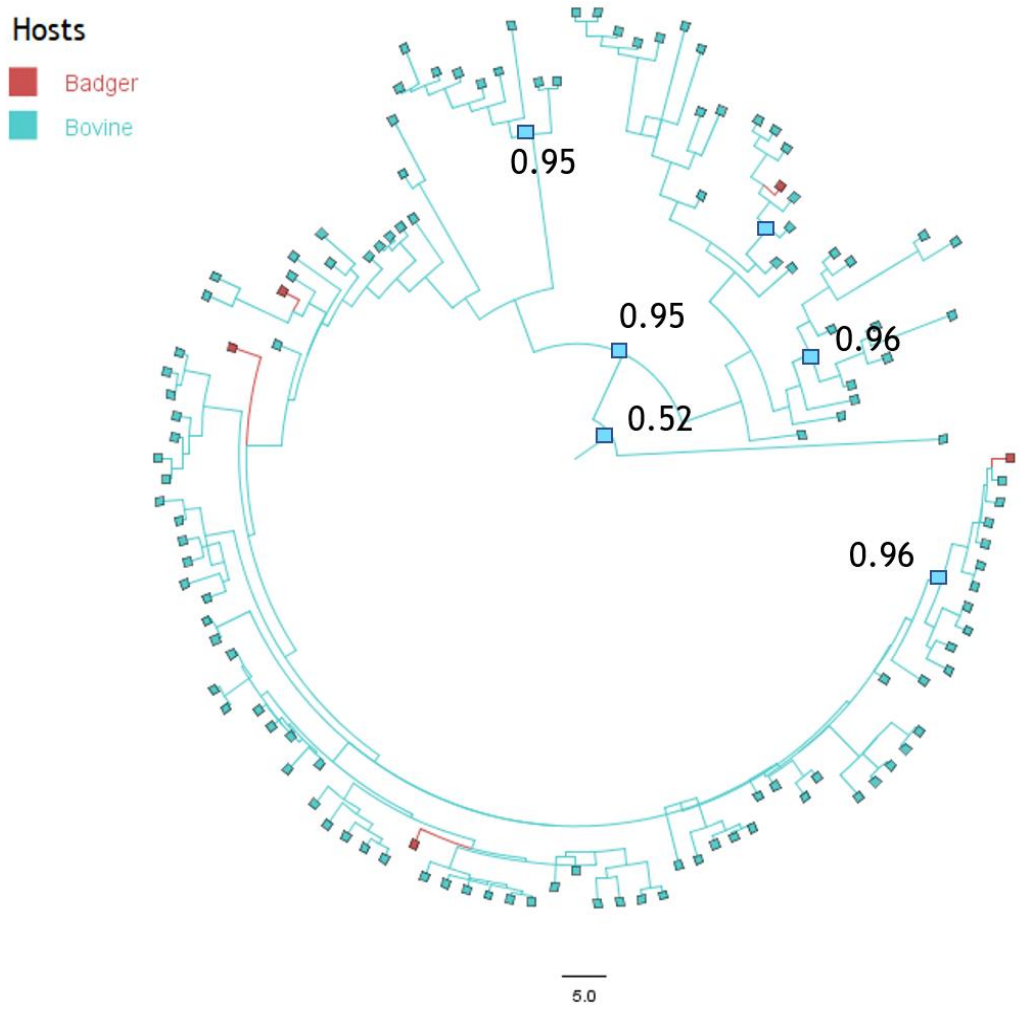


Figure 2 (a, b) Phylogenetic MCC trees for the subset 1 Constant Population size

M. bovis data sampled from years 1986-2011. a. Ancestral nodes with higher posterior support (tree) (PP>0.95) are shown as black circles, 95% HPD interval for node TMRCA estimates of supported clades are shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badger in red) states for associated main clades shown in Figure 2a. Cattle was identified as the ancestral hosts for the oldest MRCA of all isolates (host associated PP>0.50).

Only four major monophyletic subclades (coloured) were found to have support (with an estimated posterior probability for the nodes >0.95) (Figures 3.11a and 3.12a). Circular phylogenetic trees (3.11b, 3.12b) represent the MCC trees under a model of asymmetric host species transitions, with branch colours associated with host species and host-state PP >0.50. From the 4 subclades, two are cattle specific.

a.


[6.48, 20.27]
[5.58, 13.1]
[6.5, 20.71]
[4.22, 12.67]
[15.21, 99.73]
[4.6, 11.35]
5.0
1977
1987
1997
2007
2017
Years

b.

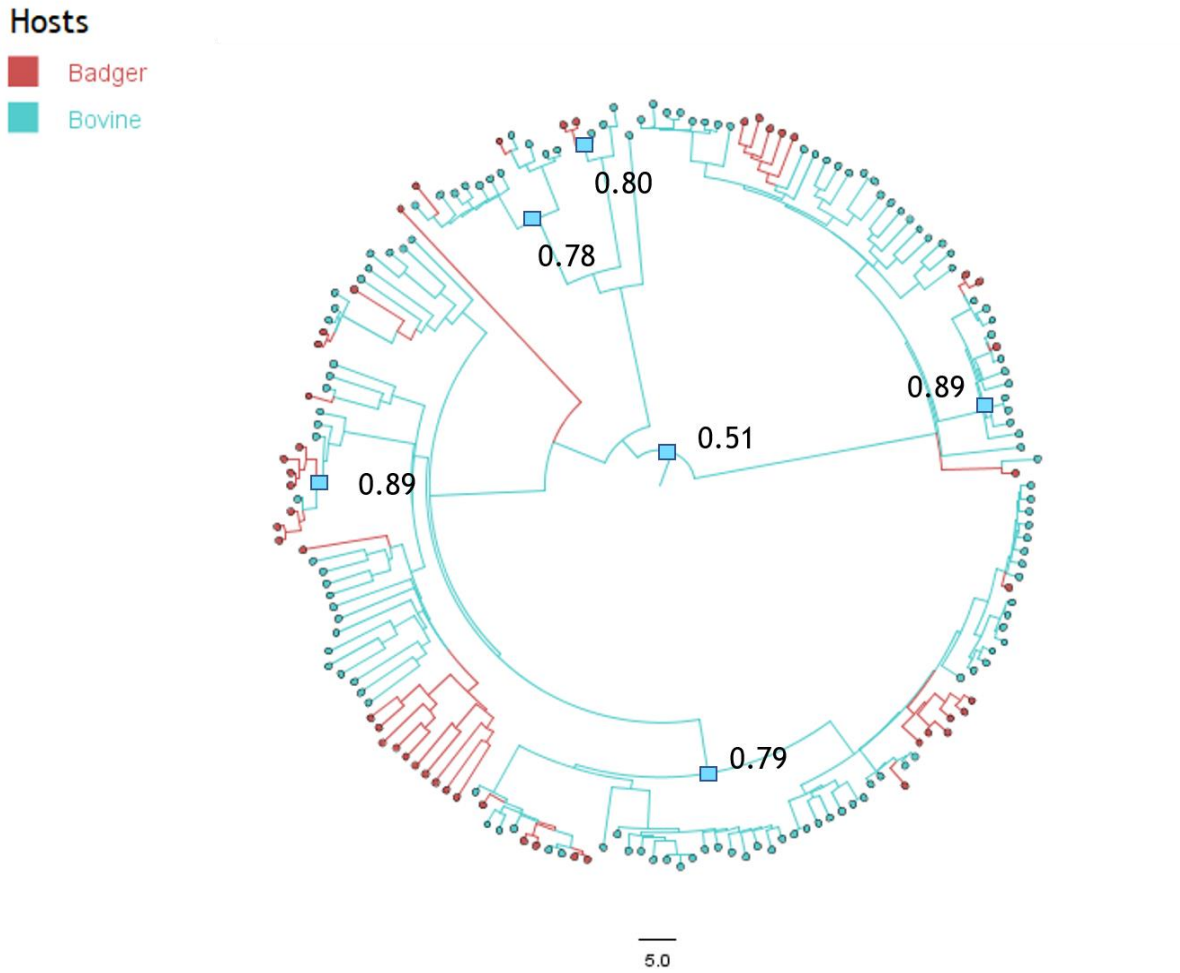


Figure 3 (a, b) Phylogenetic MCC trees of the subset 2 Constant population demographic model

M. bovis data sampled from years 1986-2011. a. Ancestral nodes with higher posterior support (tree) (PP>0.95) shown as black circles, 95% HPD interval for node TMRCA estimates for supported clades are shown in brackets. b. Discrete trait model (asymmetric) with branches and nodes (squares) annotated with their most probable (PP>0.5) host species (cattle in blue, badger in red) states for associated main clades are shown in Figure 3.12a. Cattle was identified as the ancestral hosts for the oldest MRCA of all isolates (host associated PP>0.50).

Appendix 2

Chapter 4

MBLA results for direct detection and quantitation of *Mycobacterium bovis* 16S rRNA from 214 animal tissue samples collected from bTB positive cattle in Northern Ireland.

Table 1. MBLA results table (including data about tissue samples)

Sample	Tissue, g	GTC, ml	Total volume for extraction	For extraction, mkl	RNA elution, mkl	RT+qPCR, CT	MBLA result	PCR (CFU/ml)	CFU/ extraction	Tissue dilution factor	CFU/g tissue	Lesions, NVL/VL	Lesion score	LJ /Mgit acid fast	MLVA type	MLVA type, spoligotype
1	5	3	8	200	300	31.81	positive	2.10E+03	4.20E+02	0.125	3.36E+03	VL	NA	NA	6	6.263
2	4.3	3	7.3	500	300	30.7	positive	4.48E+03	2.24E+03	0.295	7.61E+03	VL	+++	+/+	1	1.14
3	5.8	3	8.8	400	300	27.4	positive	4.18E+04	1.67E+04	0.264	6.34E+04	VL	+++	+/+	3	3.14
4	6	3	9	100	300	31.29	positive	2.99E+03	2.99E+02	0.067	4.49E+03	VL	NA	NA	4	4.14
5	3.3	3	6.3	300	300	32.47	positive	1.34E+03	4.02E+02	0.157	2.56E+03	VL	NA	NA	117	117.14
6	3.2	3	6.2	500	300	33.05	positive	9.07E+02	4.54E+02	0.258	1.76E+03	VL	NA	NA	6	6.263
7	4.4	3	7.4	500	300	31.69	positive	2.28E+03	1.14E+03	0.297	3.83E+03	VL	NA	NA	4	4.14
8	5.2	3	8.2	500	300	31.83	positive	2.08E+03	1.04E+03	0.317	3.28E+03	VL	NA	NA	4	4.14
9	2.6	3	5.6	200	300	33.03	positive	9.21E+02	1.84E+02	0.093	1.98E+03	VL	NA	NA	6	6.263
10	2.67	3	5.67	500	300	29.69	positive	8.85E+03	4.43E+03	0.235	1.88E+04	VL	NA	NA	6	6.263
11	5.3	3	8.3	500	300	32.62	positive	1.22E+03	6.10E+02	0.319	1.91E+03	VL	+++	+/+	1	1.14
12	4.4	3	7.4	500	300	32.38	positive	1.43E+03	7.15E+02	0.297	2.41E+03	VL	NA	NA	6	6.263
13	5.1	3	8.1	300	300	32.74	positive	1.13E+03	3.39E+02	0.189	1.79E+03	VL	NA	NA	4	4.14
14	5.7	3	8.7	400	300	32.35	positive	1.46E+03	5.84E+02	0.262	2.23E+03	VL	NA	NA	297	297.263
15	6.5	3	9.5	400	300	32.75	positive	1.11E+03	4.44E+02	0.274	1.62E+03	VL	NA	NA	6	6.263
16	6.4	3	9.4	400	300	32.84	positive	1.05E+03	4.20E+02	0.272	1.54E+03	VL	NA	NA	119	119.14
17	4.3	3	7.3	500	300	33.68	positive	5.95E+02	2.98E+02	0.295	1.01E+03	VL	NA	NA	6	6.263
18	5.4	3	8.4	300	300	29.94	positive	7.46E+03	2.24E+03	0.193	1.16E+04	VL	NA	NA	6	6.263
19	5	3	8	500	300	30.5	positive	5.12E+03	2.56E+03	0.313	8.19E+03	VL	NA	NA	6	6.263
20	7	3	10	400	300	27.82	positive	3.14E+04	1.26E+04	0.280	4.49E+04	VL	++	+/+	6	6.263
21	4	3	7	400	300	33.56	positive	6.42E+02	2.57E+02	0.229	1.12E+03	VL	++	+/+	1	1.14
22	3.2	3	6.2	500	300	30.68	positive	4.52E+03	2.26E+03	0.258	8.76E+03	VL	NA	NA	6	6.263
23	4.7	3	7.7	500	300	31.86	positive	2.04E+03	1.02E+03	0.305	3.34E+03	VL	++	+/+	1	1.14

24	3.8	3	6.8	500	300	30.14	positive	6.54E+03	3.27E+03	0.279	1.17E+04	VL	NA	NA	4	4.14
25	5.3	3	8.3	300	300		negative	NA	NA	NA	NA	VL	NA	NA	297	297.263
26	4.4	3	7.4	400	300	31.95	positive	1.92E+03	7.68E+02	0.238	3.23E+03	VL	NA	NA	6	6.263
27	4.3	3	7.3	300	300	35.81	positive	NA	NA			VL	NA	NA	6	6.263
28	4.9	3	7.9	500	300	27.89	positive	2.99E+04	1.50E+04	0.310	4.82E+04	VL	NA	NA	6	6.263
29	5.3	3	8.3	500	300	30.72	positive	4.40E+03	2.20E+03	0.319	6.89E+03	VL	NA	NA	6	6.263
30	4.2	3	7.2	200	300	32.94	positive	9.82E+02	1.96E+02	0.117	1.68E+03	VL	NA	NA	6	6.263
31	4.17	3	7.17	400	300	32.38	positive	1.43E+03	5.72E+02	0.233	2.46E+03	VL	NA	NA	6	6.263
32	5.05	3	8.05	300	300	30.9	positive	3.89E+03	1.17E+03	0.188	6.20E+03	VL	NA	NA	1	1.14
33	5.8	3	8.8	200	300	32.42	positive	1.39E+03	2.78E+02	0.132	2.11E+03	VL	++	+/+	1	1.14
34	2.76	3	5.76	200	300	36.55	positive	NA	NA	NA	NA	VL	+++	NA	1	1.14
35	3.9	3	6.9	100	300	29.24	positive	1.20E+04	1.20E+03	0.057	2.12E+04	VL	+++	NA	297	297.263
36	3.2	3	6.2	200	300	33.15	positive	8.49E+02	1.70E+02	0.103	1.64E+03	VL	+++	NA	2	2.142
37	2.7	3	5.7	300	300	33.13	positive	8.64E+02	2.59E+02	0.142	1.82E+03	VL	+++	NA	1	1.14
38	2.6	3	5.6	100	300	31.6	positive	2.43E+03	2.43E+02	0.046	5.23E+03	VL	++	NA	4	4.14
39	1.9	3	4.9	400	300	35.35	positive	1.91E+02	7.64E+01	0.155	4.93E+02	VL	++	NA	122	122.263
40	3.49	3	6.49	500	300	38.17	positive	NA	NA	NA	NA	VL	NA	NA	4	4.14
41	3	3	6	300	300	26.55	positive	7.41E+04	2.22E+04	0.150	1.48E+05	VL	NA	NA	297	297.263
42	4.4	3	7.4	500	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	297	297.263
43	2.4	3	5.4	200	300	39.69	positive	NA	NA	NA	NA	VL	+	NA	297	297.263
44	5	3	8	500	300	33.77	positive	5.59E+02	2.80E+02	0.313	8.94E+02	VL	++	NA	4	4.14
45	3.5	3	6.5	200	300	30.15	positive	6.50E+03	1.30E+03	0.108	1.21E+04	VL	NA	NA	11	11.145
46	3.8	3	6.8	500	300	33.17	positive	8.37E+02	4.19E+02	0.279	1.50E+03	VL	+++	NA	1	1.14
47	2.7	3	5.7	300	300	23.69	positive	5.15E+05	1.55E+05	0.142	1.09E+06	VL	NA	NA	6	6.263
48	4.7	3	7.7	200	300	32.63	positive	1.21E+03	2.42E+02	0.122	1.98E+03	VL	NA	NA	1	1.14
49	3.1	3	6.1	200	300	30.1	positive	6.72E+03	1.34E+03	0.102	1.32E+04	VL	NA	NA	1	1.14
50	3	3	6	300	300	33.71	positive	5.82E+02	1.75E+02	0.150	1.16E+03	NVL	NA	+/+	4	4.14
51	2.4	3	5.4	100	300	30	positive	7.19E+03	7.19E+02	0.044	1.62E+04	VL	+	NA	6	6.263
52	2.8	3	5.8	300	300	31.38	positive	2.81E+03	8.43E+02	0.145	5.82E+03	VL	++	NA	6	6.263

53	5.2	3	8.2	300	300	34.95	positive	2.51E+02	7.53E+01	0.190	3.96E+02	VL	NA	NA	6	6.263
54	1.82	3	4.82	400	300	35.66	positive	NA	NA	NA	NA	VL	++	NA	6	6.263
55	4.3	3	7.3	300	300	30.93	positive	3.83E+03	1.15E+03	0.177	6.50E+03	VL	++	NA	6	6.263
56	4.4	3	7.4	300	300	29.28	positive	1.17E+04	3.51E+03	0.178	1.97E+04	VL	+++	+/+	4	4.14
57	3.5	3	6.5	200	300	36.87	positive	NA	NA	NA	NA	VL	+++	+/+	4	4.14
58	5.2	3	8.2	100	300	29.32	positive	1.14E+04	1.14E+03	0.063	1.80E+04	VL	+++	+/+	3	3.14
59	3.8	3	6.8	300	300	34.43	positive	3.56E+02	1.07E+02	0.168	6.37E+02	VL	+++	-/+	6	6.263
60	2.8	3	5.8	100	300	nd	negative	NA	NA	NA	NA	NVL	NA	-/+	122	122.263
61	3.5	3	6.5	300	300	30.16	positive	6.43E+03	1.93E+03	0.162	1.19E+04	NVL	NA	+/-	297	297.263
62	4.2	3	7.2	200	300	31.91	positive	1.97E+03	3.94E+02	NA	NA	VL	++	+/+	1	1.14
63	4.3	3	7.3	300	300	30.15	positive	6.47E+03	1.94E+03	NA	NA	VL	NA	NA	3	3.14
64	3.3	3	6.3	200	300	33.24	positive	8.02E+02	1.60E+02	NA	NA	VL	NA	NA	9	9.273
65	5.5	3	8.5	200	300	31.29	positive	3.00E+03	6.00E+02	NA	NA	VL	NA	NA	4	4.14
66	4.4	3	7.4	300	300	34.89	positive	2.62E+02	7.86E+01	NA	NA	VL	NA	NA	6	6.263
67	4.6	3	7.6	100	300	34.11	positive	4.43E+02	4.43E+01	NA	NA	VL	NA	NA	6	6.263
68	3.3	3	6.3	100	300	26.23	positive	9.21E+04	9.21E+03	NA	NA	VL	+++	+/+	10	10.14
69	4.1	3	7.1	100	300	31.94	positive	1.93E+03	1.93E+02	NA	NA	VL	+++	+/+	6	6.263
70	5.3	3	8.3	200	300	31.69	positive	2.28E+03	4.56E+02	NA	NA	VL	NA	NA	5	5.14
71	4.8	3	7.8	500	300	31.49	positive	2.61E+03	1.31E+03	NA	NA	VL	NA	NA	4	4.14
72	5.5	3	8.5	500	300	27.57	positive	3.72E+04	1.86E+04	NA	NA	VL	NA	NA	1	1.14
73	4.9	3	7.9	100	300	33.56	positive	6.45E+02	6.45E+01	NA	NA	VL	NA	NA	5	5.14
74	5.5	3	8.5	100	300	29.45	positive	1.04E+04	1.04E+03	NA	NA	VL	NA	NA	6	6.263
75	6.1	3	9.1	300	300	34.37	positive	3.73E+02	1.12E+02	NA	NA	VL	NA	NA	6	6.263
76	5.1	3	8.1	200	300	32.87	positive	1.03E+03	2.06E+02	NA	NA	VL	NA	NA	6	6.263
77	7.2	3	10.2	500	300	39.96	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
78	4.7	3	7.7	300	300	38.28	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
79	3.8	3	6.8	200	300	38.56	positive	NA	NA	NA	NA	VL	NA	NA	4	4.14
80	5.6	3	8.6	200	300	38.66	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
81	2.5	3	5.5	500	300	39.85	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263

82	5	3	8	500	300	39.18	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
83	5.9	3	8.9	100	300	40.49	negative	NA	NA	NA	NA	VL	++	+/A	6	6.263
84	6.7	3	9.7	200	300	nd	negative	NA	NA	NA	NA	VL	+	+/+	122	122.263
85	7.1	3	10.1	500	300	37.4	positive	NA	NA	NA	NA	VL	NA	NA	4	4.14
86	7.1	3	10.1	400	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	4	4.14
87	6.9	3	9.9	500	300	38.09	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
88	5.6	3	8.6	400	300	38.94	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
89	4.7	3	7.7	100	300	41.3	negative	NA	NA	NA	NA	VL	++	+/A	6	6.263
90	6.6	3	9.6	500	300	40.23	negative	NA	NA	NA	NA	VL	++	+/+	6	6.263
91	5.4	3	8.4	200	300	40.03	negative	NA	NA	NA	NA	VL	++	+/A	6	6.263
92	7.1	3	10.1	300	300	40.05	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
93	2.7	3	5.7	200	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	5	5.14
94	6.4	3	9.4	200	300	39.73	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
95	5.3	3	8.3	100	300	39.66	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
96	2.2	3	5.2	300	300	42.28	negative	NA	NA	NA	NA	VL	NA	NA	297	297.263
97	2.9	3	5.9	200	300	41.05	negative	NA	NA	NA	NA	VL	NA	NA	122	122.263
98	3.9	3	6.9	300	300	43.44	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
99	1.7	3	4.7	500	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
100	6.1	3	9.1	300	300	38.72	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
101	2.5	3	5.5	300	300	39.61	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
102	2.9	3	5.9	200	300	40.52	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
103	4.2	3	7.2	500	300	39.43	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
104	5.6	3	8.6	500	300	38.42	positive	NA	NA	NA	NA	VL	NA	NA	117	117.14
105	2.9	3	5.9	500	300	36.78	positive	NA	NA	NA	NA	VL	++	+/+	4	4.14
106	6.4	3	9.4	500	300	39.62	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
107	3.1	3	6.1	200	300	41.59	negative	NA	NA	NA	NA	VL	++	+/+	297	297.263
108	0.79	3	3.79	500	300	38.04	positive	NA	NA	NA	NA	NVL	NA	+/+	6	6.263
109	4.6	3	7.6	100	300	36.9	positive	NA	NA	NA	NA	VL	+	+/+	4	4.14
110	6.09	3	9.09	100	300	35.99	positive	NA	NA	NA	NA	VL	+	+/+	4	4.14

111	3.5	3	6.5	500	300	36.95	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
112	3.5	3	6.5	300	300	33.62	positive	6.19E+02	1.86E+02	0.162	1.15E+03	VL	NA	NA	1	1.14
113	7.3	3	10.3	500	300	33.99	positive	4.80E+02	2.40E+02	0.354	6.77E+02	VL	NA	NA	1	1.14
114	4.7	3	7.7	500	300	35.15	positive	2.20E+02	1.10E+02	0.305	3.60E+02	VL	NA	NA	1	1.14
115	2.5	3	5.5	500	300	34.04	positive	4.65E+02	2.33E+02	0.227	1.02E+03	VL	NA	NA	1	1.14
116	4.2	3	7.2	200	300	36.36	positive	NA	NA	NA	NA	VL	NA	NA	1	1.14
117	3.9	3	6.9	200	300	36.81	positive	NA	NA	NA	NA	VL	NA	NA	6	6.263
118	5.4	3	8.4	200	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
119	6.6	3	9.6	200	300	35.14	positive	2.21E+02	4.42E+01	0.138	3.21E+02	VL	NA	NA	6	6.263
120	5.8	3	8.8	200	300	41.13	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
121	3.7	3	6.7	100	300	29.83	positive	8.05E+03	8.05E+02	0.055	1.46E+04	VL	NA	NA	6	6.263
122	2.2	3	5.2	300	300	32.13	positive	1.70E+03	5.10E+02	0.127	4.02E+03	VL	NA	NA	6	6.263
123	3.4	3	6.4	100	300	29.37	positive	1.10E+04	1.10E+03	0.053	2.07E+04	VL	NA	NA	6	6.263
124	1.1	3	4.1	200	300	33.67	positive	5.98E+02	1.20E+02	0.054	2.23E+03	VL	NA	NA	6	6.263
125	4.02	3	7.02	200	300	32.75	positive	1.11E+03	2.22E+02	0.115	1.94E+03	VL	NA	NA	6	6.263
126	2.3	3	5.3	300	300	33.62	positive	6.20E+02	1.86E+02	0.130	1.43E+03	VL	NA	NA	6	6.263
127	5.7	3	8.7	500	300	43.71	negative	NA	NA	NA	NA	VL	NA	NA	6	6.263
128	6.5	3	9.5	500	300	34.23	positive	4.09E+02	2.05E+02	0.342	5.98E+02	VL	NA	NA	6	6.263
129	5.9	3	8.9	200	300	39.35	positive	NA	NA	NA	NA	VL	++	NA	297	297.263
130	6.2	3	9.2	100	300	nd	negative	NA	NA	NA	NA	VL	++	NA	297	297.263
131	4.9	3	7.9	200	300	35.34	positive	1.93E+02	3.86E+01	0.124	3.11E+02	VL	NA	NA	297	297.263
132	4.3	3	7.3	100	300	31.3	positive	2.97E+03	2.97E+02	0.059	5.04E+03	VL	NA	NA	297	297.263
133	4.9	3	7.9	100	300	33.27	positive	7.84E+02	7.84E+01	0.062	1.26E+03	VL	NA	NA	297	297.263
134	1.5	3	4.5	300	300	34.8	positive	2.78E+02	8.34E+01	0.100	8.34E+02	VL	NA	NA	297	297.263
135	4.9	3	7.9	200	300	33.34	positive	7.48E+02	1.50E+02	0.124	1.21E+03	VL	NA	NA	297	297.263
136	5.3	3	8.3	100	300	33.75	positive	5.67E+02	5.67E+01	0.064	8.88E+02	VL	NA	NA	297	297.263
137	4.4	3	7.4	100	300	33.95	positive	4.95E+02	4.95E+01	0.059	8.33E+02	VL	NA	NA	4	4.14
138	5	3	8	200	300	nd	negative	NA	NA	NA	NA	VL	NA	NA	297	297.263
139	4.3	3	7.3	500	300	32.5	positive	1.32E+03	6.60E+02	0.295	2.24E+03	VL	NA	NA	297	297.263

140	7.2	3	10.2	200	300	33.43	positive	7.05E+02	1.41E+02	0.141	9.99E+02	VL	NA	NA	297	297.263
141	6.5	3	9.5	200	300	36.25	positive	NA	NA	NA	NA	VL	+++	NA	122	122.263
142	5.6	3	8.6	500	300	35.15	positive	2.20E+02	1.10E+02	0.326	3.38E+02	VL	+++	NA	122	122.263
143	4.5	3	7.5	100	300	35.29	positive	1.99E+02	1.99E+01	0.060	3.32E+02	VL	++	NA	6	6.263
144	5.4	3	8.4	300	300	37.57	positive	NA	NA	NA	NA	VL	+++	NA	6	6.263
145	4	3	7	500	300	34.8	positive	2.79E+02	1.40E+02	0.286	4.88E+02	VL	NA	NA	297	297.263
146	6.5	3	9.5	200	300	33.04	positive	9.15E+02	1.83E+02	0.137	1.34E+03	VL	NA	NA	6	6.263
147	3.7	3	6.7	200	300	38.47	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
148	1.9	3	4.9	300	300	33.62	positive	6.16E+02	1.85E+02	0.116	1.59E+03	VL	++	+/+	27	27.14
149	6.3	3	9.3	200	300	29.11	positive	1.32E+04	2.64E+03	0.135	1.95E+04	VL	++	+/+	4	4.14
150	3.5	3	6.5	300	300	31.73	positive	2.22E+03	6.66E+02	0.162	4.12E+03	VL	NA	NA	6	6.263
151	3.8	3	6.8	300	300	27.32	positive	4.42E+04	1.33E+04	0.168	7.91E+04	VL	NA	NA	297	297.263
152	1.53	3	4.53	500	300	31.72	positive	2.25E+03	1.13E+03	0.169	6.66E+03	VL	+++	NA	1	1.14
153	5.6	3	8.6	500	300	31.05	positive	3.54E+03	1.77E+03	0.326	5.44E+03	VL	++	-/+	561	561.14
154	6	3	9	500	300	31.82	positive	2.09E+03	1.05E+03	0.333	3.14E+03	VL	+++	+/+	6	6.263
155	5.9	3	8.9	200	300	30.85	positive	4.05E+03	8.10E+02	0.133	6.11E+03	VL	++	+/+	6	6.263
156	4.6	3	7.6	300	300	30.7	positive	4.48E+03	1.34E+03	0.182	7.40E+03	VL	NA	NA	297	297.263
157	2.2	3	5.2	500	300	31.08	positive	3.46E+03	1.73E+03	0.212	8.18E+03	VL	NA	NA	6	6.263
158	2.2	3	5.2	200	300	27.93	positive	2.92E+04	5.84E+03	0.085	6.90E+04	VL	NA	NA	6	6.263
159	3.5	3	6.5	200	300	27.22	positive	4.71E+04	9.42E+03	0.108	8.75E+04	VL	NA	NA	6	6.263
160	4.8	3	7.8	200	300	27.23	positive	4.71E+04	9.42E+03	0.123	7.65E+04	VL	NA	NA	6	6.263
161	4.3	3	7.3	500	300	28.95	positive	1.47E+04	7.35E+03	0.295	2.50E+04	VL	NA	NA	297	297.263
162	5.04	3	8.04	300	300	30.51	positive	5.10E+03	1.53E+03	0.188	8.14E+03	VL	NA	NA	4	4.14
163	4.2	3	7.2	300	300	30.65	positive	4.62E+03	1.39E+03	0.175	7.92E+03	VL	NA	NA	4	4.14
164	5.9	3	8.9	300	300	31.98	positive	1.87E+03	5.61E+02	0.199	2.82E+03	VL	NA	NA	4	4.14
165	3.7	3	6.7	500	300	31.58	positive	2.46E+03	1.23E+03	0.276	4.45E+03	VL	NA	NA	4	4.14
166	5.3	3	8.3	500	300	29.47	positive	1.03E+04	5.15E+03	0.319	1.61E+04	VL	NA	NA	4	4.14
167	4.4	3	7.4	300	300	30.56	positive	4.92E+03	1.48E+03	0.178	8.27E+03	VL	NA	NA	5	5.14
168	5.3	3	8.3	100	300	27.44	positive	4.07E+04	4.07E+03	0.064	6.37E+04	VL	NA	NA	6	6.263

169	3.1	3	6.1	500	300	31.08	positive	3.46E+03	1.73E+03	0.254	6.81E+03	VL	+++	NA	4	4.14
170	2.6	3	5.6	300	300	32.56	positive	1.27E+03	3.81E+02	0.139	2.74E+03	VL	+++	NA	4	4.14
171	4.5	3	7.5	300	300	31.17	positive	3.25E+03	9.75E+02	0.180	5.42E+03	VL	NA	NA	4	4.14
172	4.9	3	7.9	500	300	26.93	positive	5.73E+04	2.87E+04	0.310	9.24E+04	VL	NA	NA	4	4.14
173	4.9	3	7.9	200	300	29.54	positive	9.84E+03	1.97E+03	0.124	1.59E+04	VL	+++	NA	297	297.263
174	4.9	3	7.9	200	300	36.02	positive	NA	NA	NA	NA	VL	+++	NA	297	297.263
175	3.5	3	6.5	200	300	31.22	positive	3.14E+03	6.28E+02	0.108	5.83E+03	VL	+++	NA	4	4.14
176	2.5	3	5.5	100	300	31.73	positive	2.23E+03	2.23E+02	0.045	4.91E+03	VL	+++	NA	4	4.14
177	2.5	3	5.5	500	300	31.53	positive	2.54E+03	1.27E+03	0.227	5.59E+03	VL	NA	NA	4	4.14
178	1.9	3	4.9	300	300	33.34	positive	7.46E+02	2.24E+02	0.116	1.92E+03	VL	NA	NA	297	297.263
179	4.5	3	7.5	200	300	31.65	positive	2.35E+03	4.70E+02	0.120	3.92E+03	VL	NA	NA	297	297.263
180	2.2	3	5.2	500	300	30.65	positive	4.61E+03	2.31E+03	0.212	1.09E+04	VL	NA	NA	297	297.263
181	5.8	3	8.8	500	300	31.98	positive	1.88E+03	9.40E+02	0.330	2.85E+03	VL	NA	NA	297	297.263
182	5.3	3	8.3	200	300	28.38	positive	2.15E+04	4.30E+03	0.128	3.37E+04	VL	NA	NA	297	297.263
183	5.5	3	8.5	200	300	37.91	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
184	2.9	3	5.9	100	300	31.05	positive	3.54E+03	3.54E+02	0.049	7.20E+03	VL	NA	NA	297	297.263
185	2.8	3	5.8	500	300	31.82	positive	2.09E+03	1.05E+03	0.241	4.33E+03	VL	NA	NA	6	6.263
186	1.3	3	4.3	500	300	30.85	positive	4.05E+03	2.03E+03	0.151	1.34E+04	VL	NA	NA	297	297.263
187	5.5	3	8.5	200	300	30.7	positive	4.48E+03	8.96E+02	0.129	6.92E+03	VL	NA	NA	4	4.14
188	5.4	3	8.4	300	300	31.08	positive	3.46E+03	1.04E+03	0.193	5.38E+03	VL	NA	NA	4	4.14
189	3.7	3	6.7	300	300	27.93	positive	2.92E+04	8.76E+03	0.166	5.29E+04	VL	NA	NA	4	4.14
190	4.1	3	7.1	300	300	27.22	positive	4.71E+04	1.41E+04	0.173	8.16E+04	VL	NA	NA	4	4.14
191	3.6	3	6.6	300	300	27.23	positive	4.71E+04	1.41E+04	0.164	8.64E+04	VL	++	NA	122	122.263
192	4	3	7	300	300	28.95	positive	1.47E+04	4.41E+03	0.171	2.57E+04	VL	++	NA	1	1.14
193	4	3	7	500	300	32.12	positive	1.71E+03	8.55E+02	0.286	2.99E+03	VL	+++	+/+	297	297.263
194	5.3	3	8.3	300	300	30.65	positive	4.62E+03	1.39E+03	0.192	7.24E+03	VL	NA	NA	6	6.263
195	5.7	3	8.7	100	300	31.98	positive	1.87E+03	1.87E+02	0.066	2.85E+03	VL	NA	NA	4	4.14
196	5.6	3	8.6	300	300	31.58	positive	2.46E+03	7.38E+02	0.195	3.78E+03	VL	NA	NA	6	6.263
197	5.9	3	8.9	200	300	29.47	positive	1.03E+04	2.06E+03	0.133	1.55E+04	VL	NA	NA	4	4.14

198	5.5	3	8.5	500	300	30.56	positive	4.92E+03	2.46E+03	0.324	7.60E+03	VL	NA	NA	4	4.14
199	5.4	3	8.4	500	300	27.44	positive	4.07E+04	2.04E+04	0.321	6.33E+04	VL	NA	NA	4	4.14
200	4.6	3	7.6	300	300	31.08	positive	3.46E+03	1.04E+03	0.182	5.72E+03	VL	++	+/+	122	122.263
201	0.48	3	3.48	500	300	32.56	positive	1.27E+03	6.35E+02	0.069	9.21E+03	VL	+++	NA	297	297.263
202	5.8	3	8.8	500	300	34.04	positive	4.66E+02	2.33E+02	0.330	7.07E+02	VL	NA	NA	6	6.263
203	4.6	3	7.6	500	300	26.93	positive	5.73E+04	2.87E+04	0.303	9.47E+04	VL	NA	NA	1	1.14
204	3.09	3	6.09	500	300	29.54	positive	9.84E+03	4.92E+03	0.254	1.94E+04	VL	NA	NA	297	297.263
205	5.8	3	8.8	300	300	36.02	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263
206	5.6	3	8.6	500	300	31.22	positive	3.14E+03	1.57E+03	0.326	4.82E+03	VL	NA	NA	6	6.263
207	NA	3	NA	500	300	31.73	positive	2.23E+03	1.12E+03	NA	2.23E+03	VL	NA	NA	114	114.14
208	NA	3	NA	500	300	31.53	positive	2.54E+03	1.27E+03	NA	2.54E+03	VL	+++	NA	1	1.14
209	NA	3	NA	300	300	33.34	positive	7.46E+02	2.24E+02	NA	7.46E+02	VL	++	NA	1	1.14
210	NA	3	NA	500	300	31.65	positive	2.35E+03	1.18E+03	NA	2.35E+03	VL	+++	NA	6	6.263
211	NA	3	NA	500	300	30.65	positive	4.61E+03	2.31E+03	NA	4.61E+03	VL	NA	NA	1	1.14
212	NA	3	NA	500	300	31.98	positive	1.88E+03	9.40E+02	NA	1.88E+03	VL	NA	NA	3	3.14
213	NA	3	NA	300	300	28.38	positive	2.15E+04	6.45E+03	NA	2.15E+04	VL	NA	NA	122	122.263
214	NA	3	NA	300	300	37.91	positive	NA	NA	NA	NA	VL	NA	NA	297	297.263

*nd-not detected

*NA – not applicable

Bibliography

1992. Protection of Badgers Act 1992.

ABERNETHY, D. A., DENNY, G. O., MENZIES, F. D., MCGUCKIAN, P., HONHOLD, N. & ROBERTS, A. R. 2006. The Northern Ireland programme for the control and eradication of Mycobacterium bovis. *Vet Microbiol,* 112**,** 231-7.

ADVISORY COMMITTEE ON DANGEROUS PATHOGENS 2021. The Approved List of biological agents. Crown copyright.

AGRI-FOOD AND BIOSCIENCES INSTITUTE 2014. Field Survey for the Presence of Badger Setts in Co. Down, Northern Ireland. Report to TB & Brucellosis Policy Branch, Department of Agriculture and Rural Development.

AKHMETOVA, A., GUERRERO, J., MCADAM, P., SALVADOR, L. C. M., CRISPELL, J., LAVERY, J., PRESHO, E., KAO, R. R., BIEK, R., MENZIES, F., TRIMBLE, N., HARWOOD, R., PEPLER, P. T., ORAVCOVA, K., GRAHAM, J., SKUCE, R., DU PLESSIS, L., THOMPSON, S., WRIGHT, L., BYRNE, A. & ALLEN, A. R. 2021. Genomic epidemiology of Mycobacterium bovis infection in sympatric badger and cattle populations in Northern Ireland. *bioRxiv***,** 2021.03.12.435101.

AKRAM, S. M. & ABOOBACKER, S. 2022. Mycobacterium Marinum. *StatPearls*. Treasure Island (FL).

ALLEN, A. R. 2017. One bacillus to rule them all? - Investigating broad range host adaptation in Mycobacterium bovis. *Infect Genet Evol,* 53**,** 68-76.

ALLEN, A. R., DALE, J., MCCORMICK, C., MALLON, T. R., COSTELLO, E., GORDON, S. V., HEWINSON, R. G., SKUCE, R. A. & SMITH, N. H. 2013. The phylogeny and population structure of Mycobacterium bovis in the British Isles. *Infect Genet Evol,* 20**,** 8-15.

ALLEN, A. R., FORD, T. & SKUCE, R. A. 2021. Does Mycobacterium tuberculosis var. bovis Survival in the Environment Confound Bovine Tuberculosis Control and Eradication? A Literature Review. *Vet Med Int,* 2021**,** 8812898.

ALLEN, B. W. 1998. Mycobacteria. *In:* PARISH, T. & STOKER, N. G. (eds.) *Mycobacteria Protocols*. Totowa, NJ: Humana Press.

ALVAREZ, A. H., ESTRADA-CHAVEZ, C. & FLORES-VALDEZ, M. A. 2009. Molecular findings and approaches spotlighting Mycobacterium bovis persistence in cattle. *Vet Res,* 40**,** 22.

AMARO, A., DUARTE, E., AMADO, A., FERRONHA, H. & BOTELHO, A. 2008. Comparison of three DNA extraction methods for Mycobacterium bovis, Mycobacterium tuberculosis and Mycobacterium avium subsp avium. *Letters in Applied Microbiology,* 47**,** 8-11.

AMITA, J., VANDANA, T., GULERIA, R. & VERMA, R. 2002. Qualitative evaluation of mycobacterial DNA extraction protocols for polymerase chain reaction. *Mol Biol Today,* 3**,** 43-9.

AMOAKO, K. K., THOMAS, M. C., JANZEN, T. W. & GOJI, N. 2017. Rapid SNP Detection and Genotyping of Bacterial Pathogens by Pyrosequencing. *In:* WHITE, S. J. & CANTSILIERIS, S. (eds.) *Genotyping: Methods and Protocols*. New York, NY: Springer New York.

ANIMAL AND PLANT HEALTH AGENCY 2022. New post-mortem PCR test to identify TB infection rolled out

ARANAZ, A., LIEBANA, E., MATEOS, A., DOMINGUEZ, L., VIDAL, D., DOMINGO, M., GONZOLEZ, O., RODRIGUEZ-FERRI, E. F., BUNSCHOTEN, A. E., VAN EMBDEN, J. D. & COUSINS, D. 1996. Spacer oligonucleotide typing of Mycobacterium bovis strains from cattle and other animals: a tool for studying epidemiology of tuberculosis. *J Clin Microbiol,* 34**,** 2734-40.

ARNOLD, M. E., COURCIER, E. A., STRINGER, L. A., MCCORMICK, C. M., PASCUAL-LINAZA, A. V., COLLINS, S. F., TRIMBLE, N. A., FORD, T., THOMPSON, S., CORBETT, D. & MENZIES, F. D. 2021. A Bayesian analysis of a Test and Vaccinate or Remove study to control bovine tuberculosis in badgers (Meles meles). *PLoS One,* 16**,** e0246141.

ASHFORD, R. T., ANDERSON, P., WARING, L., DAVE, D., SMITH, F., DELAHAY, R. J., GORMLEY, E., CHAMBERS, M. A., SAWYER, J. & LESELLIER, S. 2020. Evaluation of the Dual Path Platform (DPP) VetTB assay for the detection of Mycobacterium bovis infection in badgers. *Prev Vet Med,* 180**,** 105005.

ATAHAN, E., SARIBAS, S., DEMIRCI, M., BABALıK, A., AKKUS, S., BALıKCı, A., SATANA, D., ZIVER, T., DINC, H. O., KESKIN, M., OZBEY, D., KOCAK, B. T., GAREAYAGHI, N., KIRMUSAOGLU, S., TOKMAN, H. B. & KOCAZEYBEK, B. 2020. Evaluating the effectiveness of anti-tuberculosis treatment by detecting Mycobacterium tuberculosis 85B messenger RNA expression in sputum. *Journal of Infection and Public Health,* 13**,** 1490-1494.

BAKER, L., BROWN, T., MAIDEN, M. C. & DROBNIEWSKI, F. 2004. Silent nucleotide polymorphisms and a phylogeny for Mycobacterium tuberculosis. *Emerg Infect Dis,* 10**,** 1568-77.

BALSEIRO, A., GONZALEZ-QUIROS, P., RODRIGUEZ, O., COPANO, M. F., MEREDIZ, I., DE JUAN, L., CHAMBERS, M. A., DELAHAY, R. J., MARREROS, N., ROYO, L. J., BEZOS, J., PRIETO, J. M. & GORTAZAR, C. 2013. Spatial relationships between Eurasian badgers (Meles meles) and cattle infected with Mycobacterium bovis in Northern Spain. *Veterinary Journal,* 197**,** 739-745.

BARBIER, E., ROCHELET, M., GAL, L., BOSCHIROLI, M. L. & HARTMANN, A. 2017. Impact of temperature and soil type on Mycobacterium bovis survival in the environment. *PLoS One,* 12**,** e0176315.

BEGGS, M. L., CRAWFORD, J. T. & EISENACH, K. D. 1995. Isolation and sequencing of the replication region of Mycobacterium avium plasmid pLR7. *J Bacteriol,* 177**,** 4836-40.

BELASCO, J. G., NILSSON, G., VON GABAIN, A. & COHEN, S. N. 1986. The stability of E. coli gene transcripts is dependent on determinants localized to specific mRNA segments. *Cell,* 46**,** 245-51.

BENTLEY, S. D. & PARKHILL, J. 2015. Genomic perspectives on the evolution and spread of bacterial pathogens. *Proc Biol Sci,* 282**,** 20150488.

BESTE, D. J., ESPASA, M., BONDE, B., KIERZEK, A. M., STEWART, G. R. & MCFADDEN, J. 2009. The genetic requirements for fast and slow growth in mycobacteria. *PLoS One,* 4**,** e5349.

BHATTACHARYA, B., KARAK, K., GHOSAL, A. G., ROY, A., DAS, S., DANDAPAT, P., KHETAWAT, D., MONDAL, D. K., BHATTACHARYA, S. & CHAKRABARTI, S. 2003. Development of a new sensitive and efficient multiplex polymerase chain reaction (PCR) for identification and differentiation of different mycobacterial species. *Tropical Medicine & International Health,* 8**,** 150-157.

BIEK, R., O'HARE, A., WRIGHT, D., MALLON, T., MCCORMICK, C., ORTON, R. J., MCDOWELL, S., TREWBY, H., SKUCE, R. A. & KAO, R. R. 2012. Whole Genome Sequencing Reveals Local Transmission Patterns of Mycobacterium bovis in Sympatric Cattle and Badger Populations. *Plos Pathogens,* 8.

BIEK, R., PYBUS, O. G., LLOYD-SMITH, J. O. & DIDELOT, X. 2015. Measurably evolving pathogens in the genomic era. *Trends Ecol Evol,* 30**,** 306-13.

BOGAERTS, B., DELCOURT, T., SOETAERT, K., BOARBI, S., CEYSSENS, P. J., WINAND, R., VAN BRAEKEL, J., DE KEERSMAECKER, S. C. J., ROOSENS, N. H. C., MARCHAL, K., MATHYS, V. & VANNESTE, K. 2021. A Bioinformatics Whole-Genome Sequencing Workflow for Clinical Mycobacterium tuberculosis Complex Isolate Analysis, Validated Using a Reference Collection Extensively Characterized with Conventional Methods and In Silico Approaches. *J Clin Microbiol,* 59.

BOUCKAERT, R., VAUGHAN, T. G., BARIDO-SOTTANI, J., DUCHENE, S., FOURMENT, M., GAVRYUSHKINA, A., HELED, J., JONES, G., KUHNERT, D., DE MAIO, N., MATSCHINER, M., MENDES, F. K., MULLER, N. F., OGILVIE, H. A., DU PLESSIS, L., POPINGA, A., RAMBAUT, A., RASMUSSEN, D., SIVERONI, I., SUCHARD, M. A., WU, C. H., XIE, D., ZHANG, C., STADLER, T. & DRUMMOND, A. J. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *Plos Computational Biology,* 15.

BOUVEYRON, C., FAUVEL, M. & GIRARD, S. 2015. Kernel discriminant analysis and clustering with parsimonious Gaussian process models. *Statistics and Computing,* 25**,** 1143-1162.

BRITES, D. & GAGNEUX, S. 2017. The Nature and Evolution of Genomic Diversity in the Mycobacterium tuberculosis Complex. *Adv Exp Med Biol,* 1019**,** 1-26.

BROOKS-POLLOCK, E., ROBERTS, G. O. & KEELING, M. J. 2014. A dynamic model of bovine tuberculosis spread and control in Great Britain. *Nature,* 511**,** 228-31.

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**,** 3684-9.

BRYANT, J. M., SCHURCH, A. C., VAN DEUTEKOM, H., HARRIS, S. R., DE BEER, J. L., DE JAGER, V., KREMER, K., VAN HIJUM, S. A. F. T., SIEZEN, R. J., BORGDORFF, M., BENTLEY, S. D., PARKHILL, J. & VAN SOOLINGEN, D. 2013. Inferring patient to patient transmission of Mycobacterium tuberculosis from whole genome sequencing data. *Bmc Infectious Diseases,* 13.

BUDDLE, B. M., ALDWELL, F. E., PFEFFER, A. & DE LISLE, G. W. 1994. Experimental Mycobacterium bovis infection in the brushtail possum (Trichosurus

vulpecula): pathology, haematology and lymphocyte stimulation responses. *Vet Microbiol,* 38**,** 241-54.

BUDDLE, B. M., DE LISLE, G. W., GRIFFIN, J. F. T. & HUTCHINGS, S. A. 2015. Epidemiology, diagnostics, and management of tuberculosis in domestic cattle and deer in New Zealand in the face of a wildlife reservoir. *New Zealand Veterinary Journal,* 63**,** 19-27.

BUDDLE, B. M., LIVINGSTONE, P. G. & DE LISLE, G. W. 2009. Advances in ante-mortem diagnosis of tuberculosis in cattle. *New Zealand Veterinary Journal,* 57**,** 173-180.

BUTCHER, P. D., MANGAN, J. A. & MONAHAN, I. M. 1998. Intracellular Gene Expression. *In:* PARISH, T. & STOKER, N. G. (eds.) *Mycobacteria Protocols.* Totowa, NJ: Humana Press.

CALDARELLI-STEFANO, R., VAGO, L., BONETTO, S., NEBULONI, M. & COSTANZI, G. 1999. Use of magnetic beads for tissue DNA extraction and IS6110 Mycobacterium tuberculosis PCR. *Molecular pathology : MP,* 52**,** 158-160.

CALMETTE, A. 1922. *L'infection bacillaire et la tuberculose chez l'homme et chez les animaux: processus d'infection, et de defense, étude biologique et expérimentale*, Masson.

CAMPBELL, E. L., BYRNE, A. W., MENZIES, F. D., MCBRIDE, K. R., MCCORMICK, C. M., SCANTLEBURY, M. & REID, N. 2019. Interspecific visitation of cattle and badgers to fomites: A transmission risk for bovine tuberculosis? *Ecol Evol,* 9**,** 8479-8489.

CAMPBELL, F., DIDELOT, X., FITZJOHN, R., FERGUSON, N., CORI, A. & JOMBART, T. 2018. outbreaker2: a modular platform for outbreak reconstruction. *BMC Bioinformatics,* 19**,** 363.

CARDONA, P. J. 2015. The key role of exudative lesions and their encapsulation: lessons learned from the pathology of human pulmonary tuberculosis. *Front Microbiol,* 6**,** 612.

CASSIDY, J. P. 2006. The pathogenesis and pathology of bovine tuberculosis with insights from studies of tuberculosis in humans and laboratory animal models. *Veterinary Microbiology,* 112**,** 151-161.

CASSIDY, J. P., BRYSON, D. G., GUTIERREZ CANCELA, M. M., FORSTER, F., POLLOCK, J. M. & NEILL, S. D. 2001. Lymphocyte subtypes in experimentally induced early-stage bovine tuberculous lesions. *J Comp Pathol,* 124**,** 46-51.

CASSIDY, J. P., BRYSON, D. G., POLLOCK, J. M., EVANS, R. T., FORSTER, F. & NEILL, S. D. 1998. Early lesion formation in cattle experimentally infected with Mycobacterium bovis. *Journal of Comparative Pathology,* 119**,** 27-44.

CAUGANT, D. A., LEVIN, B. R. & SELANDER, R. K. 1981. Genetic diversity and temporal variation in the E. coli population of a human host. *Genetics,* 98**,** 467-90.

CERES, K. M., STANHOPE, M. J. & GRÖHN, Y. T. 2022. A critical evaluation of Mycobacterium bovis pangenomics, with reference to its utility in outbreak investigation. *Microbial Genomics,* 8.

CHARRAD, M., GHAZZALI, N., BOITEAU, V. & NIKNAFS, A. 2014. NbClust: An R Package for Determining the Relevant Number of Clusters in a Data Set. *Journal of Statistical Software,* 61**,** 1 - 36.

CHEESEMAN, C., MALLINSON, P. J., RYAN, J. & WILESMITH, J. W. 1993. Recolonisation by badgers in Gloucestershir. *In:* HAYDEN, T. J. (ed.) *The Badger*. Dublin: Royal Irish Academy.

CHEN, L., LIU, Z., SU, Y., WANG, D., YIN, B., SHU, B., ZHANG, J., ZHU, X. & JIA, C. 2017. Characterization of Mycobacterium marinum infections in zebrafish wounds and sinus tracts. *Wound Repair Regen,* 25**,** 536-540.

CHIEN, H. P., YU, M. C., WU, M. H., LIN, T. P. & LUH, K. T. 2000. Comparison of the BACTEC MGIT 960 with Lowenstein-Jensen medium for recovery of mycobacteria from clinical specimens. *Int J Tuberc Lung Dis,* 4**,** 866-70.

CLIFTON-HADLEY, R. S., WILESMITH, J. W. & STUART, F. A. 1993. Mycobacterium bovis in the European badger (Meles meles): epidemiological findings in tuberculous badgers from a naturally infected population. *Epidemiol Infect,* 111**,** 9-19.

COLEMAN, J. D. & COOKE, M. M. 2001. Mycobacterium bovis infection in wildlife in New Zealand. *Tuberculosis (Edinb),* 81**,** 191-202.

COLIJN, C. & GARDY, J. 2014. Phylogenetic tree shapes resolve disease transmission patterns. *Evol Med Public Health,* 2014**,** 96-108.

COOK, H. C. 1997. Origins of ... tinctorial methods in histology. *J Clin Pathol,* 50**,** 716-20.

CORNER, L., MELVILLE, L., MCCUBBIN, K., SMALL, K. J., MCCORMICK, B. S., WOOD, P. R. & ROTHEL, J. S. 1990. Efficiency of inspection procedures for the detection of tuberculous lesions in cattle. *Aust Vet J,* 67**,** 389-92.

CORNER, L. A., GORMLEY, E. & PFEIFFER, D. U. 2012a. Primary isolation of Mycobacterium bovis from bovine tissues: conditions for maximising the number of positive cultures. *Vet Microbiol,* 156**,** 162-71.

CORNER, L. A., O'MEARA, D., COSTELLO, E., LESELLIER, S. & GORMLEY, E. 2012b. The distribution of Mycobacterium bovis infection in naturally infected badgers. *Vet J,* 194**,** 166-72.

COSTA, P., FERREIRA, A. S., AMARO, A., ALBUQUERQUE, T., BOTELHO, A., COUTO, I., CUNHA, M. V., VIVEIROS, M. & INACIO, J. 2013. Enhanced detection of tuberculous mycobacteria in animal tissues using a semi-nested probe-based real-time PCR. *PLoS One,* 8**,** e81337.

COURCIER, E. A., MENZIES, F. D., STRAIN, S. A. J., SKUCE, R. A., ROBINSON, P. A., PATTERSON, I. A. P., MCBRIDE, K. R., MCCORMICK, C. M., WALTON, E., MCDOWELL, S. W. J. & ABERNETHY, D. A. 2018. Monitoring Mycobacterium bovis in Eurasian badgers (Meles meles) killed by vehicles in Northern Ireland between 1998 and 2011. *Vet Rec,* 182**,** 259.

COURCIER, E. A., PASCUAL-LINAZA, A. V., ARNOLD, M. E., MCCORMICK, C. M., CORBETT, D. M., O'HAGAN, M. J. H., COLLINS, S. F., TRIMBLE, N. A., MCGEOWN, C. F., MCHUGH, G. E., MCBRIDE, K. R., MCNAIR, J., THOMPSON, S., PATTERSON, I. A. P. & MENZIES, F. D. 2020. Evaluating the application of the dual path platform VetTB test for badgers (Meles meles) in the test and vaccinate or remove (TVR) wildlife research intervention project in Northern Ireland. *Res Vet Sci,* 130**,** 170-178.

COURCOUL, A., MOYEN, J.-L., BRUGÈRE, L., FAYE, S., HÉNAULT, S., GARES, H. & BOSCHIROLI, M.-L. 2014. Estimation of Sensitivity and Specificity of Bacteriology, Histopathology and PCR for the Confirmatory Diagnosis of Bovine Tuberculosis Using Latent Class Analysis. *PLOS ONE,* 9**,** e90334.

CRISPELL, J., BENTON, C. H., BALAZ, D., DE MAIO, N., AHKMETOVA, A., ALLEN, A., BIEK, R., PRESHO, E. L., DALE, J., HEWINSON, G., LYCETT, S. J., NUNEZ-GARCIA, J., SKUCE, R. A., TREWBY, H., WILSON, D. J., ZADOKS, R. N., DELAHAY, R. J. & KAO, R. R. 2019. Combining genomics and epidemiology to analyse bi-directional transmission of Mycobacterium bovis in a multi-host system. *Elife,* 8.

CRISPELL, J., ZADOKS, R. N., HARRIS, S. R., PATERSON, B., COLLINS, D. M., DE-LISLE, G. W., LIVINGSTONE, P., NEILL, M. A., BIEK, R., LYCETT, S. J., KAO, R. R. & PRICE-CARTER, M. 2017. Using whole genome sequencing to

investigate transmission in a multi-host system: bovine tuberculosis in New Zealand. *Bmc Genomics,* 18.

DAERA 2018a. The Test and Vaccinate or Remove (TVR) Wildlife Intervention Research Project. Department of Agriculture, Environment and Rural Affairs

DAERA 2018b. Tuberculosis disease statistics in Northern Ireland 2018. *In:* AFFAIRS, D. O. A. E. A. R. (ed.).

DAERA 2021. Tuberculosis disease statistics in Northern Ireland 2021.

DANECEK, P., BONFIELD, J. K., LIDDLE, J., MARSHALL, J., OHAN, V., POLLARD, M. O., WHITWHAM, A., KEANE, T., MCCARTHY, S. A., DAVIES, R. M. & LI, H. 2021. Twelve years of SAMtools and BCFtools. *Gigascience,* 10.

DE LA RUA-DOMENECH, R., GOODCHILD, A. T., VORDERMEIER, H. M., HEWINSON, R. G., CHRISTIANSEN, K. H. & CLIFTON-HADLEY, R. S. 2006. Ante mortem diagnosis of tuberculosis in cattle: a review of the tuberculin tests, gamma-interferon assay and other ancillary diagnostic techniques. *Res Vet Sci,* 81**,** 190-210.

DE SOUZA FIGUEIREDO, E. E., CARVALHO, R. C., SILVESTRE, F. G., LILENBAUM, W., FONSECA, L. S., SILVA, J. T. & PASCHOALIN, V. M. 2010. Detection of Mycobacterium bovis DNA in nasal swabs from tuberculous cattle by a multiplex PCR. *Braz J Microbiol,* 41**,** 386-90.

DE VOS, V., BENGIS, R. G., KRIEK, N. P., MICHEL, A., KEET, D. F., RAATH, J. P. & HUCHZERMEYER, H. F. 2001. The epidemiology of tuberculosis in free-ranging African buffalo (Syncerus caffer) in the Kruger National Park, South Africa. *Onderstepoort J Vet Res,* 68**,** 119-30.

DEAN, G. S., RHODES, S. G., COAD, M., WHELAN, A. O., COCKLE, P. J., CLIFFORD, D. J., HEWINSON, R. G. & VORDERMEIER, H. M. 2005. Minimum infective dose of Mycobacterium bovis in cattle. *Infect Immun,* 73**,** 6467-71.

DELAHAY, R. J., CHEESEMAN, C. L. & CLIFTON-HADLEY, R. S. 2001. Wildlife disease reservoirs: the epidemiology of Mycobacterium bovis infection in the European badger (Meles meles) and other British mammals. *Tuberculosis,* 81**,** 43-49.

DELOGU, G., BRENNAN, M. J. & MANGANELLI, R. 2017. PE and PPE Genes: A Tale of Conservation and Diversity. *Adv Exp Med Biol,* 1019**,** 191-207.

DIDELOT, X., FRASER, C., GARDY, J. & COLIJN, C. 2017. Genomic Infectious Disease Epidemiology in Partially Sampled and Ongoing Outbreaks. *Mol Biol Evol,* 34, 997-1007.

DIDELOT, X., KENDALL, M., XU, Y., WHITE, P. J. & MCCARTHY, N. 2021. Genomic Epidemiology Analysis of Infectious Disease Outbreaks Using TransPhylo. *Curr Protoc,* 1, e60.

DIDELOT, X. & MAIDEN, M. C. 2010. Impact of recombination on bacterial evolution. *Trends Microbiol,* 18, 315-22.

DOIG, C., SEAGAR, A. L., WATT, B. & FORBES, K. J. 2002. The efficacy of the heat killing of Mycobacterium tuberculosis. *J Clin Pathol,* 55, 778-9.

DOMINGO, M., VIDAL, E. & MARCO, A. 2014. Pathology of bovine tuberculosis. *Res Vet Sci,* 97 Suppl, S20-9.

DONNELLY, C. A. & NOUVELLET, P. 2013. The contribution of badgers to confirmed tuberculosis in cattle in high-incidence areas in England. *PLoS Curr,* 5.

DONNELLY, C. A., WEI, G., JOHNSTON, W. T., COX, D. R., WOODROFFE, R., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., GETTINBY, G., GILKS, P., JENKINS, H. E., LE FEVRE, A. M., MCINERNEY, J. P. & MORRISON, W. I. 2007. Impacts of widespread badger culling on cattle tuberculosis: concluding analyses from a large-scale field trial. *Int J Infect Dis,* 11, 300-8.

DONNELLY, C. A., WOODROFFE, R., COX, D. R., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., WEI, G., GETTINBY, G., GILKS, P., JENKINS, H., JOHNSTON, W. T., LE FEVRE, A. M., MCINERNEY, J. P. & MORRISON, W. I. 2006. Positive and negative effects of widespread badger culling on tuberculosis in cattle. *Nature,* 439, 843-6.

DONNELLY, C. A., WOODROFFE, R., COX, D. R., BOURNE, J., GETTINBY, G., LE FEVRE, A. M., MCINERNEY, J. P. & MORRISON, W. I. 2003. Impact of localized badger culling on tuberculosis incidence in British cattle. *Nature,* 426, 834-7.

DRUMMOND, A. J. & BOUCKAERT, R. R. 2015. *Bayesian Evolutionary Analysis with BEAST,* Cambridge, Cambridge University Press.

DRUMMOND, A. J., RAMBAUT, A., SHAPIRO, B. & PYBUS, O. G. 2005. Bayesian Coalescent Inference of Past Population Dynamics from Molecular Sequences. *Molecular Biology and Evolution,* 22, 1185-1192.

DUONG, T. 2007. ks: Kernel Density Estimation and Kernel Discriminant Analysis for Multivariate Data in R. *Journal of Statistical Software,* 21, 1 - 16.

EDWARDS, D. J., POPE, B. J. & HOLT, K. E. 2015. RedDog: comparative analysis pipeline for large numbers of bacterial isolates using high-throughput sequences. In prep. (working title only).

EISENACH, K. D., CAVE, M. D., BATES, J. H. & CRAWFORD, J. T. 1990. Polymerase chain reaction amplification of a repetitive DNA sequence specific for Mycobacterium tuberculosis. *J Infect Dis,* 161, 977-81.

ESTER, M. K., H-P.; SANDER, J.; XU, X. 1996. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. *Proc. 2nd Int. Conf. Knowledge Discovery and Data Mining (KDD'96),* 226-231.

EZEOKE, E. 2020. *Mycobacterial Load Assay for Rapid Identification of Mycobacterium Bovis in Bovine Tuberculosis Cases.* MSc thesis, university of Glasgow.

FAKSRI, K., TAN, J. H., CHAIPRASERT, A., TEO, Y.-Y. & ONG, R. T.-H. 2016. Bioinformatics tools and databases for whole genome sequence analysis of Mycobacterium tuberculosis. *Infection, Genetics and Evolution,* 45, 359-368.

FALKINHAM, J. O., 3RD 1996. Epidemiology of infection by nontuberculous mycobacteria. *Clin Microbiol Rev,* 9, 177-215.

FARRELL, D., CRISPELL, J. & GORDON, S. V. 2020. Updated functional annotation of the Mycobacterium bovis AF2122/97 reference genome. *Access Microbiol,* 2, acmi000129.

FEIL, E. J., LI, B. C., AANENSEN, D. M., HANAGE, W. P. & SPRATT, B. G. 2004. eBURST: Inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *Journal of Bacteriology,* 186, 1518-1530.

FITZGERALD, S. D. & KANEENE, J. B. 2013. Wildlife reservoirs of bovine tuberculosis worldwide: hosts, pathology, surveillance, and control. *Vet Pathol,* 50, 488-99.

FLYNN, J. L., CHAN, J. & LIN, P. L. 2011. Macrophages and control of granulomatous inflammation in tuberculosis. *Mucosal Immunology,* 4, 271-278.

FORRELLAD, M. A., KLEPP, L. I., GIOFFRÉ, A., SABIO Y GARCIA, J., MORBIDONI, H. R., SANTANGELO, M. D. L. P., CATALDI, A. A. & BIGI, F. 2013. Virulence factors of the Mycobacterium tuberculosis complex. *Virulence,* 4**,** 3-66.

FRANCISCO, A. P., BUGALHO, M., RAMIREZ, M. & CARRICO, J. A. 2009. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics,* 10**,** 152.

FRANCO, M. M., PAES, A. C., RIBEIRO, M. G., DE FIGUEIREDO PANTOJA, J. C., SANTOS, A. C., MIYATA, M., LEITE, C. Q., MOTTA, R. G. & LISTONI, F. J. 2013. Occurrence of mycobacteria in bovine milk samples from both individual and collective bulk tanks at farms and informal markets in the southeast region of Sao Paulo, Brazil. *BMC Vet Res,* 9**,** 85.

FROTHINGHAM, R. & MEEKER-O'CONNELI, W. A. 1998. Genetic diversity in the Mycobacterium tuberculosis complex based on variable numbers of tandem DNA repeats. *Microbiology,* 144**,** 1189-1196.

GALAGAN, J. E. 2014. Genomic insights into tuberculosis. *Nature Reviews Genetics,* 15**,** 307-320.

GARCIA, M. D. S., MELO, A. F., CARVALHO, G. F., POMIM, G. P., NEVES, P. M. D. S., SILVA, R. A. B., OLIVEIRA, R. O. D. & FRIAS, D. F. R. 2021. Epidemiology of bovine tuberculosis in South America. *Research, Society and Development,* 10**,** e8610917936.

GARDY, J. L., JOHNSTON, J. C., SUI, S. J. H., COOK, V. J., SHAH, L. N., BRODKIN, E., REMPEL, S., MOORE, R., ZHAO, Y. J., HOLT, R., VARHOL, R., BIROL, I., LEM, M., SHARMA, M. K., ELWOOD, K., JONES, S. J. M., BRINKMAN, F. S. L., BRUNHAM, R. C. & TANG, P. 2011. Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak. *New England Journal of Medicine,* 364**,** 730-739.

GARNIER, T., EIGLMEIER, K., CAMUS, J. C., MEDINA, N., MANSOOR, H., PRYOR, M., DUTHOY, S., GRONDIN, S., LACROIX, C., MONSEMPE, C., SIMON, S., HARRIS, B., ATKIN, R., DOGGETT, J., MAYES, R., KEATING, L., WHEELER, P. R., PARKHILL, J., BARRELL, B. G., COLE, S. T., GORDON, S. V. & HEWINSON, R. G. 2003. The complete genome sequence of Mycobacterium bovis. *Proc Natl Acad Sci U S A,* 100**,** 7877-82.

GILLESPIE, S. H., SABIITI, W. & ORAVCOVA, K. 2017. Mycobacterial Load Assay. *Methods Mol Biol,* 1616**,** 89-105.

GODFRAY, C., DONNELLY, C., HEWINSON, G., WINTER, M. & WOOD, J. 2018. Bovine TB Strategy Review.

GODFRAY, H. C., DONNELLY, C. A., KAO, R. R., MACDONALD, D. W., MCDONALD, R. A., PETROKOFSKY, G., WOOD, J. L., WOODROFFE, R., YOUNG, D. B. & MCLEAN, A. R. 2013. A restatement of the natural science evidence base relevant to the control of bovine tuberculosis in Great Britain. *Proc Biol Sci,* 280**,** 20131634.

GOODCHILD, A. V., WATKINS, G. H., SAYERS, A. R., JONES, J. R. & CLIFTON-HADLEY, R. S. 2012. Geographical association between the genotype of bovine tuberculosis in found dead badgers and in cattle herds. *Veterinary Record,* 170**,** 259-259.

GORMLEY, E. & CORNER, L. A. L. 2017. Pathogenesis of Mycobacterium bovis Infection: the Badger Model As a Paradigm for Understanding Tuberculosis in Animals. *Front Vet Sci,* 4**,** 247.

GORMLEY, E., CORNER, L. A. L., COSTELLO, E. & RODRIGUEZ-CAMPOS, S. 2014. Bacteriological diagnosis and molecular strain typing of Mycobacterium bovis and Mycobacterium caprae. *Research in Veterinary Science,* 97**,** S30-S43.

GORMLEY, E., DOYLE, M. B., FITZSIMONS, T., MCGILL, K. & COLLINS, J. D. 2006. Diagnosis of Mycobacterium bovis infection in cattle by use of the gamma-interferon (Bovigam (R)) assay. *Veterinary Microbiology,* 112**,** 171-179.

GORTAZAR, C., DELAHAY, R. J., MCDONALD, R. A., BOADELLA, M., WILSON, G. J., GAVIER-WIDEN, D. & ACEVEDO, P. 2012. The status of tuberculosis in European wild mammals. *Mammal Review,* 42**,** 193-206.

GRANGE, J. M. 2001. Mycobacterium bovis infection in human beings. *Tuberculosis (Edinb),* 81**,** 71-7.

GREEN, D. M., KISS, I. Z., MITCHELL, A. P. & KAO, R. R. 2008. Estimates for local and movement-based transmission of bovine tuberculosis in British cattle. *Proc Biol Sci,* 275**,** 1001-5.

GRIFFIN, J. M., WILLIAMS, D. H., KELLY, G. E., CLEGG, T. A., O'BOYLE, I., COLLINS, J. D. & MORE, S. J. 2005. The impact of badger removal on the control of tuberculosis in cattle herds in Ireland. *Prev Vet Med,* 67**,** 237-66.

GRIFFITH, D. E., AKSAMIT, T., BROWN-ELLIOTT, B. A., CATANZARO, A., DALEY, C., GORDIN, F., HOLLAND, S. M., HORSBURGH, R., HUITT, G., IADEMARCO, M. F., ISEMAN, M., OLIVIER, K., RUOSS, S., VON REYN, C. F., WALLACE, R. J.,

JR. & WINTHROP, K. 2007. An official ATS/IDSA statement: diagnosis, treatment, and prevention of nontuberculous mycobacterial diseases. *Am J Respir Crit Care Med,* 175**,** 367-416.

GROENEN, P. M., BUNSCHOTEN, A. E., VAN SOOLINGEN, D. & VAN EMBDEN, J. D. 1993. Nature of DNA polymorphism in the direct repeat cluster of Mycobacterium tuberculosis; application for strain differentiation by a novel typing method. *Molecular microbiology,* 10**,** 1057-1065.

GUIMARAES, A. M. S. & ZIMPEL, C. K. 2020. Mycobacterium bovis: From Genotyping to Genome Sequencing. *Microorganisms,* 8.

GUTIERREZ, C. M. M. & GARCIA, M. J. F. 1993. Comparison of Ziehl-Neelsen staining and immunohistochemistry for the detection of Mycobacterium bovis in bovine and caprine tuberculous lesions. *J Comp Pathol,* 109**,** 361-70.

GUTIERREZ, M. C., BRISSE, S., BROSCH, R., FABRE, M., OMAIS, B., MARMIESSE, M., SUPPLY, P. & VINCENT, V. 2005. Ancient origin and gene mosaicism of the progenitor of Mycobacterium tuberculosis. *Plos Pathogens,* 1**,** 55-61.

HAHSLER, M., PIEKENBROCK, M. & DORAN, D. 2019. dbscan: Fast Density-Based Clustering with R. *Journal of Statistical Software,* 91**,** 1–30.

HAI, H. T., SABIITI, W., THU, D. D. A., PHU, N. H., GILLESPIE, S. H., THWAITES, G. E. & THUONG, N. T. T. 2021. Evaluation of the molecular bacterial load assay for detecting viable Mycobacterium tuberculosis in cerebrospinal fluid before and during tuberculous meningitis treatment. *Tuberculosis,* 128**,** 102084.

HALL, M., WOOLHOUSE, M. & RAMBAUT, A. 2015. Epidemic reconstruction in a phylogenetics framework: transmission trees as partitions of the node set. *PLoS computational biology,* 11**,** e1004613.

HASEGAWA, M., KISHINO, H. & YANO, T.-A. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *Journal of Molecular Evolution,* 22**,** 160-174.

HATHERELL, H. A., COLIJN, C., STAGG, H. R., JACKSON, C., WINTER, J. R. & ABUBAKAR, I. 2016. Interpreting whole genome sequencing for investigating tuberculosis transmission: a systematic review. *BMC Med,* 14**,** 21.

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.

HELLYER, T. J., DESJARDIN, L. E., HEHMAN, G. L., CAVE, M. D. & EISENACH, K. D. 1999. Quantitative analysis of mRNA as a marker for viability of Mycobacterium tuberculosis. *Journal of Clinical Microbiology,* 37**,** 290-295.

HERMANS, P. W., VAN SOOLINGEN, D., BIK, E. M., DE HAAS, P. E., DALE, J. W. & VAN EMBDEN, J. D. 1991. Insertion element IS987 from Mycobacterium bovis BCG is located in a hot-spot integration region for insertion elements in Mycobacterium tuberculosis complex strains. *Infect Immun,* 59**,** 2695-705.

HILTY, M., DIGUIMBAYE, C., SCHELLING, E., BAGGI, F., TANNER, M. & ZINSSTAG, J. 2005. Evaluation of the discriminatory power of variable number tandem repeat (VNTR) typing of Mycobacterium bovis strains. *Vet Microbiol,* 109**,** 217-22.

HONEYBORNE, I., MCHUGH, T. D., PHILLIPS, P. P., BANNOO, S., BATESON, A., CARROLL, N., PERRIN, F. M., RONACHER, K., WRIGHT, L., VAN HELDEN, P. D., WALZL, G. & GILLESPIE, S. H. 2011. Molecular bacterial load assay, a culture-free biomarker for rapid and accurate quantification of sputum Mycobacterium tuberculosis bacillary load during treatment. *J Clin Microbiol,* 49**,** 3905-11.

HONEYBORNE, I., MTAFYA, B., PHILLIPS, P. P., HOELSCHER, M., NTINGINYA, E. N., KOHLENBERG, A., RACHOW, A., ROJAS-PONCE, G., MCHUGH, T. D., HEINRICH, N. & PAN AFRICAN CONSORTIUM FOR THE EVALUATION OF ANTI-TUBERCULOSIS, A. 2014. The molecular bacterial load assay replaces solid culture for measuring early bactericidal response to antituberculosis treatment. *J Clin Microbiol,* 52**,** 3064-7.

HUMBLET, M. F., BOSCHIROLI, M. L. & SAEGERMAN, C. 2009. Classification of worldwide bovine tuberculosis risk factors in cattle: a stratified approach. *Vet Res,* 40**,** 50.

HUMPHREY, H. M., ORLOSKI, K. A. & OLEA-POPELKA, F. J. 2014. Bovine tuberculosis slaughter surveillance in the United States 2001–2010: assessment of its traceback investigation function. *BMC Veterinary Research,* 10**,** 182.

ILLUMINA 2017. TruSight Tumor 170 Reference Guide.

JAIN, A. K., MURTY, M. N. & FLYNN, P. J. 1999. Data clustering: A review. *Acm Computing Surveys,* 31**,** 264-323.

JENKINS, H. E., WOODROFFE, R. & DONNELLY, C. A. 2010. The duration of the effects of repeated widespread badger culling on cattle tuberculosis following the cessation of culling. *PLoS One,* 5**,** e9090.

JENKINS, H. E., WOODROFFE, R., DONNELLY, C. A., COX, D. R., JOHNSTON, W. T., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., GETTINBY, G., GILKS, P., HEWINSON, R. G., MCINERNEY, J. P. & MORRISON, W. I. 2007. Effects of culling on spatial associations of Mycobacterium bovis infections in badgers and cattle. *Journal of Applied Ecology,* 44**,** 897-908.

JOHNSON, L., DEAN, G., RHODES, S., HEWINSON, G., VORDERMEIER, M. & WANGOO, A. 2007. Low-dose Mycobacterium bovis infection in cattle results in pathology indistinguishable from that of high-dose infection. *Tuberculosis,* 87**,** 71-76.

JOMBART, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics,* 24**,** 1403-1405.

JOMBART, T., CORI, A., DIDELOT, X., CAUCHEMEZ, S., FRASER, C. & FERGUSON, N. 2014. Bayesian Reconstruction of Disease Outbreaks by Combining Epidemiologic and Genomic Data. *Plos Computational Biology,* 10.

JOMBART, T., EGGO, R. M., DODD, P. J. & BALLOUX, F. 2011. Reconstructing disease outbreaks from genetic data: a graph approach. *Heredity (Edinb),* 106**,** 383-90.

KAMERBEEK, J., SCHOULS, L., KOLK, A., VAN AGTERVELD, M., VAN SOOLINGEN, D., KUIJPER, S., BUNSCHOTEN, A., MOLHUIZEN, H., SHAW, R., GOYAL, M. & VAN EMBDEN, J. 1997. Simultaneous detection and strain differentiation of Mycobacterium tuberculosis for diagnosis and epidemiology. *J Clin Microbiol,* 35**,** 907-14.

KAO, R. R., HAYDON, D. T., LYCETT, S. J. & MURCIA, P. R. 2014. Supersize me: how whole-genome sequencing and big data are transforming epidemiology. *Trends in Microbiology,* 22**,** 282-291.

KAO, R. R., PRICE-CARTER, M. & ROBBE-AUSTERMAN, S. 2016. Use of genomics to track bovine tuberculosis transmission. *Rev Sci Tech,* 35**,** 241-58.

KARLSON, A. G. & LESSEL, E. F. 1970. MYCOBACTERIUM BOVIS NOM. NOV. *International Journal of Systematic Bacteriology,* 20**,** 273-282.

KASSAMBARA, A. & MUNDT, F. 2020. Factoextra: Extract and Visualize the Results of Multivariate Data Analyses. R Package Version 1.0.7.

KLOOG, I., HAIM, A. & PORTNOV, B. A. 2009. Using kernel density function as an urban analysis tool: Investigating the association between nightlight exposure and the incidence of breast cancer in Haifa, Israel. *Computers Environment and Urban Systems,* 33, 55-63.

KOCH, R. 1890. An Address on Bacteriological Research. *Br Med J,* 2, 380-3.

KREMER, K., VAN SOOLINGEN, D., FROTHINGHAM, R., HAAS, W. H., HERMANS, P. W. M., MARTIN, C., PALITTAPONGARNPIM, P., PLIKAYTIS, B. B., RILEY, L. W., YAKRUS, M. A., MUSSER, J. M. & VAN EMBDEN, J. D. A. 1999. Comparison of methods based on different molecular epidemiological markers for typing of Mycobacterium tuberculosis complex strains: Interlaboratory study of discriminatory power and reproducibility. *Journal of Clinical Microbiology,* 37, 2607-2618.

KRIEGEL, H. P., KROGER, P., SANDER, J. & ZIMEK, A. 2011. Density-based clustering. *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery,* 1, 231-240.

KRUSKAL, W. H. & WALLIS, W. A. 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association,* 47, 583-621.

LAHIRI, K. K. & CHATTERJEE, S. K. 1994. A SIMPLE COLD STAINING METHOD FOR ACID FAST BACILLI. *Medical journal, Armed Forces India,* 50, 256-258.

LAHUERTA-MARIN, A., MCNAIR, J., SKUCE, R., MCBRIDE, S., ALLEN, M., STRAIN, S. A. J., MENZIES, F. D., MCDOWELL, S. J. W. & BYRNE, A. W. 2016. Risk factors for failure to detect bovine tuberculosis in cattle from infected herds across Northern Ireland (2004-2010). *Res Vet Sci,* 107, 233-239.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9, 357-9.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., DURBIN, R. & GENOME PROJECT DATA PROCESSING, S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25, 2078-9.

LIEBANA, E., ARANAZ, A., ALDWELL, F. E., MCNAIR, J., NEILL, S. D., SMYTH, A. J. & POLLOCK, J. M. 2000. Cellular interactions in bovine tuberculosis: release of active mycobacteria from infected macrophages by antigen-stimulated T cells. *Immunology,* 99, 23-9.

LIU, X., GUTACKER, M. M., MUSSER, J. M. & FU, Y. X. 2006. Evidence for recombination in Mycobacterium tuberculosis. *J Bacteriol,* 188, 8169-77.

LIVINGSTONE, P. G., HANCOX, N., NUGENT, G. & DE LISLE, G. W. 2015. Toward eradication: the effect of Mycobacterium bovis infection in wildlife on the evolution and future direction of bovine tuberculosis management in New Zealand. *N Z Vet J,* 63 Suppl 1**,** 4-18.

MACQUEEN, J. B. Some Methods for classification and Analysis of Multivariate Observations. Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability, 1967 Berkeley. University of California Press, 281-297.

MAECHLER, M., ROUSSEEUW, P., STRUYF, A., HUBERT, M. & HORNIK, K. 2021. cluster: Cluster Analysis Basics and Extensions. R package version 2.1.2

MALONE, K. M., FARRELL, D., STUBER, T. P., SCHUBERT, O. T., AEBERSOLD, R., ROBBE-AUSTERMAN, S. & GORDON, S. V. 2017. Updated Reference Genome Sequence and Annotation of Mycobacterium bovis AF2122/97. *Genome Announc,* 5.

MANGAN, J. A., SOLE, K. M., MITCHISON, D. A. & BUTCHER, P. D. 1997. An effective method of RNA extraction from bacteria refractory to disruption, including mycobacteria. *Nucleic Acids Res,* 25**,** 675-6.

MATHEMA, B., KUREPINA, N. E., BIFANI, P. J. & KREISWIRTH, B. N. 2006. Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev,* 19**,** 658-85.

MCCALLAN, L., BROOKS, C., BARRY, C., COUZENS, C., YOUNG, F. J., MCNAIR, J. & BYRNE, A. W. 2021. Serological test performance for bovine tuberculosis in cattle from herds with evidence of on-going infection in Northern Ireland. *PLoS One,* 16**,** e0245655.

MENARDO, F., DUCHENE, S., BRITES, D. & GAGNEUX, S. 2019. The molecular clock of Mycobacterium tuberculosis. *PLoS Pathog,* 15**,** e1008067.

MENZIES, F. D., MCCORMICK, C. M., O'HAGAN, M. J. H., COLLINS, S. F., MCEWAN, J., MCGEOWN, C. F., MCHUGH, G. E., HART, C. D., STRINGER, L. A., MOLLOY, C., BURNS, G., MCBRIDE, S. J., DOYLE, L. P., COURCIER, E. A., MCBRIDE, K. R., MCNAIR, J., THOMPSON, S., CORBETT, D. M., HARWOOD, R. G. & TRIMBLE, N. A. 2021. Test and vaccinate or remove: Methodology and preliminary results from a badger intervention research project. *Vet Rec,* 189**,** e248.

MENZIES, F. D. & NEILL, S. D. 2000. Cattle-to-cattle transmission of bovine tuberculosis. *Vet J,* 160**,** 92-106.

MIĆIĆ, M., WHYTE, J. D. & KARSTEN, V. 2016. High Performance Bead Beating Based Lysing, Homogenization and Grinding for DNA, RNA and Proteins Extraction with FastPrep® Systems. *In:* MICIC, M. (ed.) *Sample Preparation Techniques for Soil, Plant, and Animal Samples*. New York, NY: Springer New York.

MILLER, J. M., JENNY, A. L. & PAYEUR, J. B. 2002. Polymerase chain reaction detection of Mycobacterium tuberculosis complex and Mycobacterium avium organisms in formalin-fixed tissues from culture-negative ruminants. *Veterinary Microbiology,* 87**,** 15-23.

MILNE, G., ALLEN, A., GRAHAM, J., KIRKE, R., MCCORMICK, C., PRESHO, E., SKUCE, R. & BYRNE, A. W. 2020. Mycobacterium bovis Population Structure in Cattle and Local Badgers: Co-Localisation and Variation by Farm Type. *Pathogens,* 9.

MILNE, M. G., GRAHAM, J., ALLEN, A., MCCORMICK, C., PRESHO, E., SKUCE, R. & BYRNE, A. W. 2019. Variation in Mycobacterium bovis genetic richness suggests that inwards cattle movements are a more important source of infection in beef herds than in dairy herds. *BMC Microbiol,* 19**,** 154.

MONAGHAN, M. L., DOHERTY, M. L., COLLINS, J. D., KAZDA, J. F. & QUINN, P. J. 1994. The tuberculin test. *Vet Microbiol,* 40**,** 111-24.

MONTENEGRO, R. A., GUARINES, K. M., MONTENEGRO, L. M., LIRA, L. A., FALCÃO, J., MELO, F. L., SANTOS, F. C., NASCIMENTO, A. L., ZUZARTE, M. S., LEITE, R. C. & SCHINDLER, H. C. 2014. Assessment of messenger RNA (mRNA) of Mycobacterium tuberculosis as a marker of cure in patients with pulmonary tuberculosis. *J Appl Microbiol,* 117**,** 266-72.

MOORE, D. F. & JAMES, A. M. 1982. Growth studies on Mycobacterium BCG: oxygen preference. *Microbios,* 35**,** 151-9.

MOORHEAD, S. M., DYKES, G. A. & CURSONS, R. T. 2003. An SNP-based PCR assay to differentiate between Listeria monocytogenes lineages derived from phylogenetic analysis of the sigB gene. *J Microbiol Methods,* 55**,** 425-32.

MORRIS, R. S., PFEIFFER, D. U. & JACKSON, R. 1994. The epidemiology of Mycobacterium bovis infections. *Vet Microbiol,* 40**,** 153-77.

MUELLER, D. H., MWENGE, L., MUYOYETA, M., MUVWIMI, M. W., TEMBWE, R., MCNERNEY, R., GODFREY-FAUSSETT, P. & AYLES, H. M. 2008. Costs and cost-effectiveness of tuberculosis cultures using solid and liquid media in a developing country. *Int J Tuberc Lung Dis,* 12**,** 1196-202.

NAVIN, T. R., MCNABB, S. J. & CRAWFORD, J. T. 2002. The continued threat of tuberculosis. *Emerg Infect Dis,* 8**,** 1187.

NCUBE, P., BAGHERI, B., GOOSEN, W. J., MILLER, M. A. & SAMPSON, S. L. 2022. Evidence, Challenges, and Knowledge Gaps Regarding Latent Tuberculosis in Animals. *Microorganisms,* 10**,** 1845.

NEILL, S. D., BRYSON, D. G. & POLLOCK, J. M. 2001. Pathogenesis of tuberculosis in cattle. *Tuberculosis,* 81**,** 79-86.

NEILL, S. D., POLLOCK, J. M., BRYSON, D. B. & HANNA, J. 1994. Pathogenesis of Mycobacterium bovis infection in cattle. *Vet Microbiol,* 40**,** 41-52.

NEILL, S. D., SKUCE, R. A. & POLLOCK, J. M. 2005. Tuberculosis--new light from an old window. *J Appl Microbiol,* 98**,** 1261-9.

NGHIEM, M. N., NGUYEN, B. V., NGUYEN, S. T., VO, T. T. & NONG, H. V. 2015. A Simple, Single Triplex PCR of IS6110, IS1081, and 23S Ribosomal DNA Targets, Developed for Rapid Detection and Discrimination of Mycobacterium from Clinical Samples. *J Microbiol Biotechnol,* 25**,** 745-52.

NIGSCH, A., GLAWISCHNIG, W., BAGO, Z. & GREBER, N. 2018. Mycobacterium caprae Infection of Red Deer in Western Austria-Optimized Use of Pathology Data to Infer Infection Dynamics. *Front Vet Sci,* 5**,** 350.

NORTHERN IRELAND AUDIT OFFICE 2009. The Control of Bovine Tuberculosis in Northern Ireland. Belfast: Northern Ireland Audit Office.

NORTHERN IRELAND AUDIT OFFICE 2018. Eradicating Bovine TB in Northern Ireland.

O'HAGAN, M. J., MATTHEWS, D. I., LAIRD, C. & MCDOWELL, S. W. 2016. Herd-level risk factors for bovine tuberculosis and adoption of related biosecurity measures in Northern Ireland: A case-control study. *Vet J,* 213**,** 26-32.

O'REILLY, L. M. & DABORN, C. J. 1995. The epidemiology of Mycobacterium bovis infections in animals and man: a review. *Tuber Lung Dis,* 76 Suppl 1**,** 1-46.

O'MAHONY, D. 2015. Badger (Meles meles) contact metrics in a medium density population. *Mammalian Biology - Zeitschrift für Säugetierkunde,* 80.

OLEA-POPELKA, F. J., FLYNN, O., COSTELLO, E., MCGRATH, G., COLLINS, J. D., O'KEEFFE, J., KELTON, D. F., BERKE, O. & MARTIN, S. W. 2005. Spatial relationship between Mycobacterium bovis strains in cattle and badgers in four areas in Ireland. *Preventive Veterinary Medicine,* 71**,** 57-70.

OLEA-POPELKA, F. J., GRIFFIN, J. M., COLLINS, J. D., MCGRATH, G. & MARTIN, S. W. 2003. Bovine tuberculosis in badgers in four areas in Ireland: does tuberculosis cluster? *Prev Vet Med,* 59**,** 103-11.

PALMER, M. V. & WATERS, W. R. 2011. Bovine tuberculosis and the establishment of an eradication program in the United States: role of veterinarians. *Vet Med Int,* 2011**,** 816345.

PALMER, M. V., WATERS, W. R. & WHIPPLE, D. L. 2004. Investigation of the transmission of Mycobacterium bovis from deer to cattle through indirect contact. *Am J Vet Res,* 65**,** 1483-9.

PATANE, J. S., MARTINS, J., BEATRIZ CASTELAO, A., NISHIBE, C., MONTERA, L., BIGI, F., ZUMARRAGA, M. J., CATALDI, A. A., FONSECA JUNIOR, A., ROXO, E., LUIZA, A., OSORIO, A. R., JORGE UFMS, K. S., THACKER, T. C., ALMEIDA, N. F., ARAUJO, F. R. & SETUBAL, J. C. 2017. Patterns and processes of Mycobacterium bovis evolution revealed by phylogenomic analyses. *Genome Biol Evol*.

PATEL, B. K. R., BANERJEE, D. K. & BUTCHER, P. D. 1991. Extraction and characterization of mRNA from mycobacteria: implication for virulence gene identification. *Journal of Microbiological Methods,* 13**,** 99-111.

PAYEUR, J. B., CHURCH, S., MOSHER, L., ROBINSON-DUNN, B., SCHMITT, S. & WHIPPLE, D. 2002. Bovine tuberculosis in Michigan wildlife. *Ann N Y Acad Sci,* 969**,** 259-61.

PEREA, C., CERES, K., DURÁN, U., TORTOSA LA OSA, S., ALONSO, B., PERERA, A., GRÖHN, Y., BRAVO, D., STUBER, T., THACKER, T. & ROBBE-AUSTERMAN, S. 2023. Whole-genome sequencing to investigate Mycobacterium bovis strains circulating in the Dominican Republic. PREPRINT (Version 1) available at Research Square.

PFYFFER, G. E. 2015. Mycobacterium: General Characteristics, Laboratory Detection, and Staining Procedures. *Manual of Clinical Microbiology*.

PHILIPS, J. A. & ERNST, J. D. 2012. Tuberculosis pathogenesis and immunity. *Annu Rev Pathol,* 7**,** 353-84.

POLLOCK, J. M., MCNAIR, J., WELSH, M. D., GIRVIN, R. M., KENNEDY, H. E., MACKIE, D. P. & NEILL, S. D. 2001. Immune responses in bovine tuberculosis. *Tuberculosis (Edinb),* 81**,** 103-7.

POLLOCK, J. M. & NEILL, S. D. 2002. Mycobacterium bovis infection and tuberculosis in cattle. *Vet J,* 163**,** 115-27.

POLLOCK, J. M., POLLOCK, D. A., CAMPBELL, D. G., GIRVIN, R. M., CROCKARD, A. D., NEILL, S. D. & MACKIE, D. P. 1996. Dynamic changes in circulating and antigen-responsive T-cell subpopulations post-Mycobacterium bovis infection in cattle. *Immunology,* 87**,** 236-241.

PRITCHARD, D. G. 1988. A century of bovine tuberculosis 1888-1988: conquest and controversy. *J Comp Pathol,* 99**,** 357-99.

PYBUS, O. G., SUCHARD, M. A., LEMEY, P., BERNARDIN, F. J., RAMBAUT, A., CRAWFORD, F. W., GRAY, R. R., ARINAMINPATHY, N., STRAMER, S. L., BUSCH, M. P. & DELWART, E. L. 2012. Unifying the spatial epidemiology and molecular evolution of emerging epidemics. *Proc Natl Acad Sci U S A,* 109**,** 15066-71.

QUAN, Z., HAIMING, T., XIAOYAO, C., WEIFENG, Y., HONG, J. & HONGFEI, Z. 2017. Development of one-tube multiplex polymerase chain reaction (PCR) for detecting Mycobacterium bovis. *The Journal of veterinary medical science,* 78**,** 1873-1876.

R CORE TEAM. 2021. *R: A language and environment for statistical computing* [Online]. Vienna: R Foundation for Statistical Computing. Available: https://www.R-project.org/ [Accessed].

RAMBAUT, A. 2010. FigTree v1.3.1. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh.

RAMBAUT, A., DRUMMOND, A. J., XIE, D., BAELE, G. & SUCHARD, M. A. 2018. Posterior Summarization in Bayesian Phylogenetics Using Tracer 1.7. *Systematic Biology,* 67**,** 901-904.

RAMBAUT, A., LAM, T. T., CARVALHO, L. M. & PYBUS, O. G. 2016. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evolution,* 2.

RAMOS, D. F., SILVA, P. E. & DELLAGOSTIN, O. A. 2015. Diagnosis of bovine tuberculosis: review of main techniques. *Braz J Biol,* 75**,** 830-7.

REID, N., ETHERINGTON, T. R., WILSON, G. J., MONTGOMERY, W. I. & MCDONALD, R. A. 2012. Monitoring and population estimation of the European badger Meles meles in Northern Ireland. *Wildlife Biology,* 18**,** 46-57.

REIS, A. C. & CUNHA, M. V. 2021. Genome-wide estimation of recombination, mutation and positive selection enlightens diversification drivers of Mycobacterium bovis. *Sci Rep,* 11**,** 18789.

REYES, J. F., CHAN, C. H. S. & TANAKA, M. M. 2012. Impact of homoplasy on variable numbers of tandem repeats and spoligotypes in Mycobacterium tuberculosis. *Infection Genetics and Evolution,* 12**,** 811-818.

ROBINSON, P. A. 2015. A history of bovine tuberculosis eradication policy in Northern Ireland. *Epidemiology and Infection,* 143**,** 3182-3195.

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**,** 866-72.

ROETZER, A., DIEL, R., KOHL, T. A., RUCKERT, C., NUBEL, U., BLOM, J., WIRTH, T., JAENICKE, S., SCHUBACK, S., RUSCH-GERDES, S., SUPPLY, P., KALINOWSKI, J. & NIEMANN, S. 2013. Whole Genome Sequencing versus Traditional Genotyping for Investigation of a Mycobacterium tuberculosis Outbreak: A Longitudinal Molecular Epidemiological Study. *Plos Medicine,* 10.

ROMANOWSKI, G., LORENZ, M. G., SAYLER, G. & WACKERNAGEL, W. 1992. Persistence of free plasmid DNA in soil monitored by various methods, including a transformation assay. *Appl Environ Microbiol,* 58**,** 3012-9.

ROPER, T. J. & LÜPS, P. 1993. Disruption of territorial behaviour in badgers *Meles meles*. *Zeitschrift für Säugetierkunde,* 58**,** 252-255.

RORING, S., SCOTT, A., BRITTAIN, D., WALKER, I., HEWINSON, G., NEILL, S. & SKUCE, R. 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. *Journal of Clinical Microbiology,* 40**,** 2126-2133.

ROSSI, G., CRISPELL, J., BROUGH, T., LYCETT, S. J., WHITE, P. C. L., ALLEN, A., ELLIS, R. J., GORDON, S. V., HARWOOD, R., PALKOPOULOU, E., PRESHO, E. L., SKUCE, R., SMITH, G. C. & KAO, R. R. 2022. Phylodynamic analysis of an emergent Mycobacterium bovis outbreak in an area with no previously known wildlife infections. *Journal of Applied Ecology,* 59**,** 210-222.

RUSTAD, T. R., ROBERTS, D. M., LIAO, R. P. & SHERMAN, D. R. 2009. Isolation of Mycobacterial RNA. *In:* PARISH, T. & BROWN, A. C. (eds.) *Mycobacteria Protocols: Second Edition*. Totowa, NJ: Humana Press.

RUTALA, W. A., COLE, E. C., WANNAMAKER, N. S. & WEBER, D. J. 1991. Inactivation of Mycobacterium tuberculosis and Mycobacterium bovis by 14 hospital disinfectants. *Am J Med,* 91**,** 267S-271S.

SABIITI, W., AZAM, K., ESMERALDO, E., BHATT, N., RACHOW, A. & GILLESPIE, S. H. 2019. Heat Inactivation Renders Sputum Safe and Preserves Mycobacterium tuberculosis RNA for Downstream Molecular Tests. *Journal of Clinical Microbiology,* 57.

SABIITI, W., AZAM, K., FARMER, E. C. W., KUCHAKA, D., MTAFYA, B., BOWNESS, R., ORAVCOVA, K., HONEYBORNE, I., EVANGELOPOULOS, D., MCHUGH, T. D., KHOSA, C., RACHOW, A., HEINRICH, N., KAMPIRA, E., DAVIES, G., BHATT, N., NTINGINYA, E. N., VIEGAS, S., JANI, I., KAMDOLOZI, M., MDOLO, A., KHONGA, M., BOEREE, M. J., PHILLIPS, P. P. J., SLOAN, D., HOELSCHER, M., KIBIKI, G. & GILLESPIE, S. H. 2020a. Tuberculosis bacillary load, an early marker of disease severity: the utility of tuberculosis Molecular Bacterial Load Assay. *Thorax,* 75**,** 606-608.

SABIITI, W., MTAFYA, B., DE LIMA, D. A., DOMBAY, E., BARON, V. O., AZAM, K., ORAVCOVA, K., SLOAN, D. J. & GILLESPIE, S. H. 2020b. A Tuberculosis Molecular Bacterial Load Assay (TB-MBLA). *J Vis Exp*.

SAGRIPANTI, J. L., HÜLSEWEH, B., GROTE, G., VOSS, L., BÖHLING, K. & MARSCHALL, H. J. 2011. Microbial inactivation for safe and rapid diagnostics of infectious samples. *Appl Environ Microbiol,* 77**,** 7289-95.

SALVADOR, L. C. M., O'BRIEN, D. J., COSGROVE, M. K., STUBER, T. P., SCHOOLEY, A. M., CRISPELL, J., CHURCH, S. V., GROHN, Y. T., ROBBE-AUSTERMAN, S. & KAO, R. R. 2019. Disease management at the wildlife-livestock interface: Using whole-genome sequencing to study the role of elk in Mycobacterium bovis transmission in Michigan, USA. *Mol Ecol,* 28**,** 2192-2205.

SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. *Molecular cloning: a laboratory manual*, Cold spring harbor laboratory press.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A,* 74**,** 5463-7.

SCHLIEP, K. P. 2010. phangorn: phylogenetic analysis in R. *Bioinformatics,* 27**,** 592-593.

SHERIDAN, M. 2011. Progress in tuberculosis eradication in Ireland. *Vet Microbiol,* 151, 160-9.

SILVERMAN, B. W. 1986. *Density estimation for statistics and data analysis,* London ; New York, Chapman and Hall.

SKUCE, R., BREADON, E., ALLEN, A., MILNE, G., MCCORMICK, C., HUGHES, C., RUTHERFORD, D., SMITH, G., THOMPSON, S., GRAHAM, J., HARWOOD, R. & BYRNE, A. 2020. Longitudinal dynamics of herd-level Mycobacterium bovis MLVA type surveillance in cattle in Northern Ireland 2003-2016. *Infect Genet Evol,* 79, 104131.

SKUCE, R. A., MALLON, T. R., MCCORMICK, C. M., MCBRIDE, S. H., CLARKE, G., THOMPSON, A., COUZENS, C., GORDON, A. W. & MCDOWELL, S. W. J. 2010. Mycobacterium bovis genotypes in Northern Ireland: herd-level surveillance (2003 to 2008). *Veterinary Record,* 167, 684-689.

SKUCE, R. A., MCCORRY, T. P., MCCARROLL, J. F., RORING, S. M. M., SCOTT, A. N., BRITTAIN, D., HUGHES, S. L., HEWINSON, R. G. & NEILL, S. D. 2002. Discrimination of Mycobacterium tuberculosis complex bacteria using novel VNTR-PCR targets. *Microbiology-Sgm,* 148, 519-528.

SKUCE, R. A., MCDOWELL, S. W., MALLON, T. R., LUKE, B., BREADON, E. L., LAGAN, P. L., MCCORMICK, C. M., MCBRIDE, S. H. & POLLOCK, J. M. 2005. Discrimination of isolates of Mycobacterium bovis in Northern Ireland on the basis of variable numbers of tandem repeats (VNTRs). *Vet Rec,* 157, 501-4.

SKUCE, R. A. & NEILL, S. D. 2001. Molecular epidemiology of Mycobacterium bovis: exploiting molecular data. *Tuberculosis (Edinb),* 81, 169-75.

SMITH, N. H., DALE, J., INWALD, J., PALMER, S., GORDON, S. V., HEWINSON, R. G. & SMITH, J. M. 2003. The population structure of Mycobacterium bovis in Great Britain: clonal expansion. *Proc Natl Acad Sci U S A,* 100, 15271-5.

SMITH, N. H., GORDON, S. V., DE LA RUA-DOMENECH, R., CLIFTON-HADLEY, R. S. & HEWINSON, R. G. 2006. Bottlenecks and broomsticks: the molecular evolution of Mycobacterium bovis. *Nat Rev Microbiol,* 4, 670-81.

SMITH, N. H. & UPTON, P. 2012. Naming spoligotype patterns for the RD9-deleted lineage of the Mycobacterium tuberculosis complex; www.Mbovis.org. *Infection, Genetics and Evolution,* 12, 873-876.

SMITTIPAT, N. & PALITTAPONGARNPIM, P. 2000. Identification of possible loci of variable number of tandem repeats in Mycobacterium tuberculosis. *Tubercle and Lung Disease,* 80, 69-74.

STEINLEY, D. & BRUSCO, M. J. 2007. Initializing K-means Batch Clustering: A Critical Evaluation of Several Techniques. *Journal of Classification,* 24**,** 99-121.

STUCKI, D. & GAGNEUX, S. 2013. Single nucleotide polymorphisms in Mycobacterium tuberculosis and the need for a curated database. *Tuberculosis (Edinb),* 93**,** 30-9.

SUKUMAR, S., KRISHNAN, A. & BANERJEE, S. 2021. An Overview of Bioinformatics Resources for SNP Analysis. *In:* SINGH, V. & KUMAR, A. (eds.) *Advances in Bioinformatics*. Singapore: Springer Singapore.

SUPPLY, P., ALLIX, C., LESJEAN, S., CARDOSO-OELEMANN, M., RUSCH-GERDES, S., WILLERY, E., SAVINE, E., DE HAAS, P., VAN DEUTEKOM, H., RORING, S., BIFANI, P., KUREPINA, N., KREISWIRTH, B., SOLA, C., RASTOGI, N., VATIN, V., GUTIERREZ, M. C., FAUVILLE, M., NIEMANN, S., SKUCE, R., KREMER, K., LOCHT, C. & VAN SOOLINGEN, D. 2006. Proposal for standardization of optimized mycobacterial interspersed repetitive unit-variable-number tandem repeat typing of Mycobacterium tuberculosis. *J Clin Microbiol,* 44**,** 4498-510.

SUPPLY, P., MAZARS, E., LESJEAN, S., VINCENT, V., GICQUEL, B. & LOCHT, C. 2000. Variable human minisatellite-like regions in the Mycobacterium tuberculosis genome. *Molecular Microbiology,* 36**,** 762-771.

TAVARÉ, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. *Lectures on mathematics in the life sciences,* 17**,** 57-86.

TAYLOR, G. M., WORTH, D. R., PALMER, S., JAHANS, K. & HEWINSON, R. G. 2007. Rapid detection of Mycobacterium bovis DNA in cattle lymph nodes with visible lesions using PCR. *BMC Vet Res,* 3**,** 12.

TB STRATEGIC PARTNERSHIP GROUP 2016. Bovine Tuberculosis Eradication Strategy for Northern Ireland.

THACKER, T. C., HARRIS, B., PALMER, M. V. & WATERS, W. R. 2011. Improved specificity for detection of Mycobacterium bovis in fresh tissues using IS6110 real-time PCR. *BMC Veterinary Research,* 7**,** 50.

THERESE, K. L., GAYATHRI, R., DHANUREKHA, L., SRIDHAR, R., MEENAKSHI, N., MADHAVAN, H. N., MANOJ, S. E. & VINAYAGAM, A. K. 2012. Detection of Mycobacterium tuberculosis directly from sputum specimens & phenotypic drug resistance pattern of M. tuberculosis isolates from suspected tuberculosis patients in Chennai. *Indian J Med Res,* 135**,** 778-82.

THOEN, C., LOBUE, P. & DE KANTOR, I. 2006. The importance of Mycobacterium bovis as a zoonosis. *Veterinary Microbiology,* 112**,** 339-345.

TREWBY, H., WRIGHT, D., BREADON, E. L., LYCETT, S. J., MALLON, T. R., MCCORMICK, C., JOHNSON, P., ORTON, R. J., ALLEN, A. R., GALBRAITH, J., HERZYK, P., SKUCE, R. A., BIEK, R. & KAO, R. R. 2016. Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis. *Epidemics,* 14**,** 26-35.

TUYTTENS, F. A. M., MACDONALD, D. W., ROGERS, L. M., CHEESEMAN, C. L. & RODDAM, A. W. 2000. Comparative study on the consequences of culling badgers (Meles meles) on biometrics, population dynamics and movement. *J. Anim. Ecol.,* 69**,** 567–580.

VAN DER VLIET, G. M., SCHEPERS, P., SCHUKKINK, R. A., VAN GEMEN, B. & KLATSER, P. R. 1994. Assessment of mycobacterial viability by RNA amplification. *Antimicrob Agents Chemother,* 38**,** 1959-65.

VAN HELDEN, P. D., VICTOR, T. C., WARREN, R. M. & VAN HELDEN, E. G. 2001. Isolation of DNA from Mycobacterium tubercolosis. *Methods Mol Med,* 54**,** 19-30.

VAN TONDER, A. J., THORNTON, M. J., CONLAN, A. J. K., JOLLEY, K. A., GOOLDING, L., MITCHELL, A. P., DALE, J., PALKOPOULOU, E., HOGARTH, P. J., HEWINSON, R. G., WOOD, J. L. N. & PARKHILL, J. 2021. Inferring Mycobacterium bovis transmission between cattle and badgers using isolates from the Randomised Badger Culling Trial. *PLOS Pathogens,* 17**,** e1010075.

VERMA, G. & KASHYAP, P. 2021. Decontamination Method For Tuberculosis: A Review. *J Tuberc,* 4**,** 1025.

WALKER, T. M., LALOR, M. K., BRODA, A., ORTEGA, L. S., MORGAN, M., PARKER, L., CHURCHILL, S., BENNETT, K., GOLUBCHIK, T., GIESS, A. P., DEL OJO ELIAS, C., JEFFERY, K. J., BOWLER, I., LAURENSON, I. F., BARRETT, A., DROBNIEWSKI, F., MCCARTHY, N. D., ANDERSON, L. F., ABUBAKAR, I., THOMAS, H. L., MONK, P., SMITH, E. G., WALKER, A. S., CROOK, D. W., PETO, T. E. A. & CONLON, C. P. 2014. Assessment of Mycobacterium tuberculosis transmission in Oxfordshire, UK, 2007-12, with whole pathogen genome sequences: an observational study. *Lancet Respir Med,* 2**,** 285-292.

WALLIS, R. S., ELLNER, J. J. & SHIRATSUCHI, H. 1992. Macrophages, mycobacteria and HIV: the role of cytokines in determining mycobacterial virulence and regulating viral replication. *Research in Microbiology,* 143**,** 398-405.

WARDS, B. J., COLLINS, D. M. & DE LISLE, G. W. 1995. Detection of Mycobacterium bovis in tissues by polymerase chain reaction. *Vet Microbiol,* 43**,** 227-40.

WATERS, W. R., WHELAN, A. O., LYASHCHENKO, K. P., GREENWALD, R., PALMER, M. V., HARRIS, B. N., HEWINSON, R. G. & VORDERMEIER, H. M. 2010. Immune responses in cattle inoculated with Mycobacterium bovis, Mycobacterium tuberculosis, or Mycobacterium kansasii. *Clin Vaccine Immunol,* 17**,** 247-52.

WIRTH, T., HILDEBRAND, F., ALLIX-BEGUEC, C., WOLBELING, F., KUBICA, T., KREMER, K., VAN SOOLINGEN, D., RUSCH-GERDES, S., LOCHT, C., BRISSE, S., MEYER, A., SUPPLY, P. & NIEMANN, S. 2008. Origin, spread and demography of the Mycobacterium tuberculosis complex. *PLoS Pathog,* 4**,** e1000160.

WOOD, P. R. & JONES, S. L. 2001. BOVIGAM: an in vitro cellular diagnostic test for bovine tuberculosis. *Tuberculosis (Edinb),* 81**,** 147-55.

WOODROFFE, R., DONNELLY, C. A., COX, D. R., GILKS, P., JENKINS, H. E., JOHNSTON, W. T., LE FEVRE, A. M., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., GETTINBY, G., HEWINSON, R. G., MCINERNEY, J. P., MITCHELL, A. P., MORRISON, W. I. & WATKINS, G. H. 2009. Bovine tuberculosis in cattle and badgers in localized culling areas. *J Wildl Dis,* 45**,** 128-43.

WOODROFFE, R., DONNELLY, C. A., HAM, C., JACKSON, S. Y., MOYES, K., CHAPMAN, K., STRATTON, N. G. & CARTWRIGHT, S. J. 2016. Badgers prefer cattle pasture but avoid cattle: implications for bovine tuberculosis control. *Ecol Lett,* 19**,** 1201-8.

WOODROFFE, R., DONNELLY, C. A., JENKINS, H. E., JOHNSTON, W. T., COX, D. R., BOURNE, F. J., CHEESEMAN, C. L., DELAHAY, R. J., CLIFTON-HADLEY, R. S., GETTINBY, G., GILKS, P., HEWINSON, R. G., MCINERNEY, J. P. & MORRISON, W. I. 2006. Culling and cattle controls influence tuberculosis risk for badgers. *Proc Natl Acad Sci U S A,* 103**,** 14713-7.

WOODROFFE, R., DONNELLY, C. A., JOHNSTON, W. T., BOURNE, F. J., CHEESEMAN, C. L., CLIFTON-HADLEY, R. S., COX, D. R., GETTINBY, G., HEWINSON, R. G., LE FEVRE, A. M., MCINERNEY, J. P. & MORRISON, W. I.

2005. Spatial association of Mycobacterium bovis infection in cattle and badgers Meles meles. *Journal of Applied Ecology,* 42**,** 852-862.

WOODROFFE, R., MACDONALD, D. & DA SILVA, J. 1995. Dispersal and philopatry in the European badger, Meles meles. *Journal of Zoology,* 237**,** 227-239.

WORLD HEALTH ORGANIZATION 2012. *Tuberculosis Laboratory Biosafety Manual,* Geneva.

WORLD HEALTH ORGANIZATION (WHO) 2020. Global tuberculosis report.

WORLD HEALTH ORGANIZATION (WHO), WORLD ORGANISATION FOR ANIMAL HEALTH (OIE) & FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS (FAO) 2017. Roadmap for Zoonotic Tuberculosis.

YANG, T., ZHONG, J., ZHANG, J., LI, C., YU, X., XIAO, J., JIA, X., DING, N., MA, G., WANG, G., YUE, L., LIANG, Q., SHENG, Y., SUN, Y., HUANG, H. & CHEN, F. 2018. Pan-Genomic Study of Mycobacterium tuberculosis Reflecting the Primary/Secondary Genes, Generality/Individuality, and the Interconversion Through Copy Number Variations. *Front Microbiol,* 9**,** 1886.

YOUNG, J. S., GORMLEY, E. & WELLINGTON, E. M. 2005. Molecular detection of Mycobacterium bovis and Mycobacterium bovis BCG (Pasteur) in soil. *Appl Environ Microbiol,* 71**,** 1946-52.

YPMA, R. J., BATAILLE, A. M., STEGEMAN, A., KOCH, G., WALLINGA, J. & VAN BALLEGOOIJEN, W. M. 2012. Unravelling transmission trees of infectious diseases by combining genetic and epidemiological data. *Proc Biol Sci,* 279**,** 444-50.

YPMA, R. J. F., VAN BALLEGOOIJEN, W. M. & WALLINGA, J. 2013. Relating Phylogenetic Trees to Transmission Trees of Infectious Disease Outbreaks. *Genetics,* 195**,** 1055-+.

ZAKHAM, F., SIRONEN, T., VAPALAHTI, O. & KANT, R. 2021. Pan and Core Genome Analysis of 183 Mycobacterium tuberculosis Strains Revealed a High Inter-Species Diversity among the Human Adapted Strains. *Antibiotics (Basel),* 10.

ZARDEN, C. F., MARASSI, C. D., CARVALHO, A. C., FIGUEIREDO, E. E. & LILENBAUM, W. 2013. Bacteriological and molecular detection of Mycobacterium bovis in cattle with inconclusive results to intradermal tuberculin tests. *Epidemiol Infect,* 141**,** 1390-3.

ZIMPEL, C. K., PATANÉ, J. S. L., GUEDES, A. C. P., DE SOUZA, R. F., SILVA-PEREIRA, T. T., CAMARGO, N. C. S., DE SOUZA FILHO, A. F., IKUTA, C. Y., NETO, J.

S. F., SETUBAL, J. C., HEINEMANN, M. B. & GUIMARAES, A. M. S. 2020. Global Distribution and Evolution of Mycobacterium bovis Lineages. *Frontiers in Microbiology,* 11.

ZUMARRAGA, M. J., SOUTULLO, A., GARCIA, M. I., MARINI, R., ABDALA, A., TARABLA, H., ECHAIDE, S., LOPEZ, M., ZERVINI, E., CANAL, A. & CATALDI, A. A. 2012. Detection of Mycobacterium bovis-infected dairy herds using PCR in bulk tank milk samples. *Foodborne Pathog Dis,* 9**,** 132-7.