SPATIO-TEMPORAL DISTRIBUTION AND PERSISTENCE OF *MYCOBACTERIUM BOVIS* IN A BADGER POPULATION

Submitted by:

Clare Helen Benton

To the University of Exeter as a thesis for the Degree of Doctor of Philosophy in Biological Sciences

May 2017

(Signature) ……

ABSTRACT

Studying the dynamics of pathogen transmission within wildlife populations presents an array of challenges. Where populations are socially structured, this can influence parasite transmission, impacting on the effectiveness of disease management strategies. In this thesis, I focus on a well-studied social mammal, the European badger *(Meles meles)* which is a key wildlife reservoir of a disease of economic importance; bovine TB (caused by infection with *Mycobacterium bovis*). The social structuring, characteristic of high density badger populations, is of well-established importance in the transmission of bovine TB and has resulted in unexpected management outcomes. However, little is known about the role of kin structure or host genotype on transmission dynamics. In this thesis, I combine traditional spatial epidemiology and ecological analysis of a well-studied badger population with more novel genetic and genomic approaches. Firstly, I investigate the role of kin structure within badger social groups in determining early life infection risk (Chapter 3). Using host genotype data, I demonstrate that cubs who are related to infected adults experience enhanced infection risks. I then explore the role of badger genotype on outcomes of *M. bovis* exposure and demonstrate that inbred badgers are more likely to show evidence of progressive infection (Chapter 4). Where the social structure of badgers is stable and unmanaged, this is predicted to result in a stable spatial distribution of *M. bovis* infection. Motivated by an observation of change in the spatial distribution of *M. bovis* infection in the study population, in the absence of management, I characterise the attrition of a spatially stable infection distribution (Chapter 5). To explore the drivers of this, I detect changes in the genetic population structure (Chapter 6) and present evidence that the population has experienced a period of demographic flux. Finally, I use a novel dataset generated by whole genome sequencing of *M. bovis* isolates and present evidence of spatial spread of *M. bovis* infection across the study population (Chapter 7). To conclude, I discuss how my findings demonstrate how genetic and genomic approaches can complement traditional wildlife epidemiology approaches, how they contribute to our understanding of heterogeneity in transmission dynamics and discuss their implications for wildlife disease management.

ACKNOWLEDGMENTS

I'd like to start by thanking my supervisors, Dave Hodgson and Dez Delahay for all the time, advice and expertise they've provided me with over the last 6 years. Dave first taught me statistics when I was an undergrad (terrifyingly, back in 2003), through my Masters, when I was convinced I would never ever be able to work R, and it has been great to be in the position to learn from him again. Dez has provided so much, firstly encouraging my application for part time PhD study and offering so much expertise in wildlife disease management. Both Dave and Dez have been fantastically supportive, particularly following my return from maternity leave and I couldn't have asked for a better supervision experience. I'd also like to thank Robbie McDonald for forging the link between Woodchester Park and Exeter University which made my studies possible. Thanks go to our collaborators at Glasgow University; Rowland Kao and Roman Biek, Joseph Crispell and particularly I'd like to thank Hannah Trewby who has given me so much advice and has been an exceptionally patient guide to the world of whole genome sequencing. Thanks go to Al Wilson and Xav Harrison who both provided invaluable advice on all things genetics and Lucy Steward for useful discussions on badger social networks. Thanks also to Jenni McDonald and Chris Beirne for encouragement along the way.

Thanks to the whole team, present and past, at Woodchester Park for all you do, both in providing an amazing dataset to use and more importantly for making it a great place to work. Particular thanks go to my office mates; Andy Robertson for helping me out of many a loop related crisis and to both him and Freya Smith for being great sounding boards for ideas. I'd also like to thank my array of line managers over the 6 years who have all been extremely supportive of my PhD studies; Iain Trewby, Fiona Rogers, Gavin Wilson and Steve Carter. Thanks also go to my employer, APHA, for allowing me to undertake a PhD alongside my day job and to Defra, as the funder of the Woodchester Park project.

On a personal note, I'd like to thank my Dad for encouraging me into science and for teaching me the value of hard work and perseverance and my Mum for constant love and support. Many people have helped look after my son Ezra in these final writing up stages to give me time to work, my parents, my father-in-

law David and Great-Uncle Roger, Anna Firkins and others and I'm truly grateful. Thanks to friends at Trinity Cheltenham for support in so many ways. And finally, thanks to my son Ezra for sleeping well and generally being a happy little soul and to my wonderful husband Tim, for being my biggest champion and for encouraging me throughout.

ETHICAL STATEMENT

All work in this thesis involving the capture and sampling of live badgers was carried out by the Animal & Plant Health Agency under Natural England and UK Home Office licences, in accordance with the Animals (Scientific Procedures) Act 1986 and was subject to an internal ethical review process.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

Appendix

AUTHOR'S DECLARATIONS

This PhD studentship was funded by the Animal and Plant Health Agency (APHA) and The University of Exeter. All the chapters presented in this thesis were written by Clare Benton, with comments and edits by Dave Hodgson, and Dez Delahay. Data from the Woodchester Park study is used in chapters 3 – 7 of this thesis, and is funded by Defra.

Chapter 2: A version of this literature review is published in *European Journal of Wildlife Research*, Jan 2015. CB, DD and DH are grateful for the comments of the editor of the journal, Christian Gortazaar.

Chapter 3: A version of this chapter is published in *Proceedings of the Royal Society B: Biological Sciences,* July 2016. CB, DD and DH are grateful for the comments of 3 anonymous reviewers. Robbie McDonald, Alastair Wilson (University of Exeter), Graham Smith and Glyn Hewinson (APHA) provided comments and edits on this chapter. Andy Robertson provided assistance with data analysis.

Chapter 4: R code for HHC correlations and inbreeding vs multi-locus heterozygosity provided by Xavier Harrison (ZSL). Freya Smith (APHA) provided advice on interpretation of TB diagnostic tests.

Chapter 7: Sequence alignment, phylogenetic tree construction, lineage assignment, and bioinformatics for the *M. bovis* whole genome sequences were performed by Hannah Trewby (University of Glasgow). Transmission clusters between *M. bovis* isolates from badgers and cattle associated with the Woodchester Park study site were constructed by Joseph Crispell (University of Glasgow). Hannah Trewby provided edits and comments for this chapter.

CHAPTER 1: General Introduction

1.1 Introduction

The five data chapters within this thesis have been written as stand-alone pieces of work with the ultimate aim of being published as scientific papers. The purpose of this introductory chapter is to introduce the central themes which draw together each of these distinct pieces of work and to illustrate how each of them contribute to our understanding of disease epidemiology, both within the particular context of the disease system from which they are drawn, but also their more general implications. I will also suggest why each work area covered by this thesis may be of practical value in terms of informing management interventions within the specific context from which the data are collected. As four of the five data chapters rely on the use of molecular data, either at the level of the host or the pathogen, Chapter 2 consists of a stand-alone literature review exploring the applications of molecular epidemiological techniques to wildlife disease research.

1.2 Spatial Epidemiology

How does a pathogen spread across an area? What affects the rate of this spread? Why do clusters of disease appear in certain places? Spatial epidemiology aims to describe and analyse geographic variations in disease in relation to demographic, environmental, behavioural, genetic and infectious risk factors (Elliott and Wartenberg 2004). It is focused on understanding the causes and consequences of spatial heterogeneity in infectious diseases (Ostfeld, Glass et al. 2005). Often, in considering where disease is present and where it is not, we are trying to uncover underlying causes, predictors or risk factors (Doll 1980). Turning this on its head, we can use the observed distribution of disease as a signal for social contact, using the spatial arrangement of infection to tell us something about the social structure and demography of a population, although this is dependent on having an understanding of the transmission mechanisms relevant to the pathogen in question. Often, infected hosts are spatially clustered; this could be driven by the pathogen itself having limited dispersal, by vectors or reservoirs of infection

being spatially restricted or by susceptible hosts being spatially clumped (Ostfeld, Glass et al. 2005). Uncovering the mechanisms which govern the spatial pattern and rate of spread of pathogens (Ostfeld, Glass et al. 2005) across landscapes is crucial in informing disease modelling (Gaudart, Rebaudet et al. 2013) and aids prediction of rates of spread and disease emergence in new areas (Plowright, Eby et al. 2015). From the most simplistic approach of plotting cases of disease on a map to carrying out complex analyses such as Bayesian models of spatial diffusion, a wide range of approaches can be used to describe and understand the processes by which disease distributions change.

1.3 Social Structure and Pathogen Transmission

As well as spatial clusters of hosts resulting in clusters of infection, social structure within host populations can lead to further aggregation. Even though two individuals are spatially proximate to each other, if social boundaries exist which deter contacts between them, transmission opportunities may be limited (Loehle 1995). R_0, the 'basic reproductive number' is a term widely used by epidemiologists. It refers to the expected number of secondary cases caused by a typical infectious individual in a completely susceptible population (Macdonald 1952). If R_0 is less than 1, then the population of infected individuals cannot grow and the pathogen is expected to fade to extinction. However, if R_0 is greater than 1, then the pathogen can spread through the population (Anderson, May et al. 1992). In a randomly mixing population uniform transmission risk is assumed between all members of the population. However in the case of socially structured populations, contact rates between individuals will vary, depending on social membership. One suggested driver for the evolution of social boundaries is reduction in the risk of disease transmission and such mechanisms may be of comparable importance to physical and immunological barriers (Loehle 1995). However, as it has been frequently noted that disease and parasites are more common in larger social groups (Alexander 1974, Côté and Poulinb 1995), higher parasitism risk may represent a cost of social living. Avoidance of disease transmission is predicted to favour rigid membership of social groups, however this conflicts with inbreeding avoidance and the drive to move group in order to improve food

availability, mating opportunities or social position (Loehle 1995). An individual may face a trade-off between these different drivers.

Whether a disease persists or dies out in a socially structured population is likely to depend on whether it has an acute or chronic progression. When compared with an acute disease of the same R_0, chronic diseases with longer infectious periods allow more time for host mixing to occur and therefore are in contact with an effectively larger population. This means that a chronic disease is more likely to invade a structured population than an acute disease of the same R_0. (Cross, Lloyd-Smith et al. 2005). Studies of human pathogens have demonstrated that social structure, or 'community' structure within contact networks can have a major impact on disease dynamics (Salathé and Jones 2010, Cauchemez, Bhattarai et al. 2011). Within animal populations, the importance of social structure in transmission of infectious diseases is being increasingly recognised (see review in Tompkins, Dunn et al. 2011). An individual's position or rank within a socially structured population may influence the likelihood of it becoming infected (Böhm, Hutchings et al. 2009) as demonstrated in social animals such as meerkats (Drewe 2010). Certain "super-spreader" individuals within a population may contribute to a disproportionate number of secondary infections (Lloyd-Smith, Schreiber et al. 2005), due to a particular behavioural or biological trait or their position within a social network. For example, individuals who overcome social group boundaries are likely to be important in allowing pathogens to be transmitted between groups (Salathé and Jones 2010, Weber, Carter et al. 2013).

As well as being spatially clustered, individuals can be genetically clustered. Often the two are linked, with spatial clusters of individuals being made up of clusters of relatives, for example if young do not disperse (natal philopatry). Considering genetic distance (or relatedness) between individuals offers an alternative way of considering structure within a population. As well as the different contact rates within and between social groups, a relatedness structure also exists. This reflects the relatedness both of individuals within the same social group and between individuals in the group and those who have dispersed elsewhere. Where kin-biased association exists (i.e. relatives spend more time together or engage in closer contacts than non-relatives) this adds further heterogeneity to inter-individual contact rates. Kin structure, defined as

the spatial aggregation of related individuals (Hatchwell 2010), is proposed to increase individual disease transmission risk in directly transmitted pathogens (Dharmarajan, Beasley et al. 2012), because transmission rates are expected to be higher between related individuals than between non-related individuals.

Bovine tuberculosis (bTB), a chronic disease caused by infection with the pathogen *Mycobacterium bovis*, remains a critical issue in livestock farming in several parts of the world, including the UK. The European badger (*Meles meles*) is a known reservoir of the pathogen. Since the 1970s when badgers were first implicated as a potential source of the disease in cattle (Gallagher, Muirhead et al. 1976), there has been much debate over the most effective management strategy. UK governments of the day have commissioned several comprehensive reviews of the scientific evidence and a broad swathe of research (Krebs, Anderson et al. 1997, Bourne, Donnelly et al. 2007). There is no doubt that this is one of the most intensively studied disease systems in the world. Transmission of *M. bovis* between badgers and cattle is thought to be possible both directly, as a result of badgers and cattle coming into 'nose to nose' contact with each other (Little, Naylor et al. 1982), and indirectly, via contact with a shared infected environment, which is thought to be the predominant transmission route (Drewe, O'Connor et al. 2013, Woodroffe, Donnelly et al. 2016). In terms of transmission amongst badgers, the pathology of *M. bovis* infection observed in badgers *post mortem* suggests that aerosol based transmission is the most common route (Clifton-Hadley, Wilesmith et al. 1993, Gallagher and Clifton-Hadley 2000), although bite-wounding is also likely to be an important transmission route (Clifton-Hadley, Wilesmith et al. 1993, Gallagher and Clifton-Hadley 2000) and has been associated with a more acute presentation of disease (Gallagher and Nelson 1979, Clifton-Hadley, Wilesmith et al. 1993, Gallagher and Clifton-Hadley 2000).

It is well documented that the social structure typical of moderate to high density, managed and unmanaged badger populations can have a marked impact on the persistence and transmission of bTB (bovine TB) (Delahay, Langton et al. 2000, Carter, Delahay et al. 2007). At moderate to high densities, badgers live in social groups in defended territories, limiting population mixing, such that members of different social groups are less likely to come into close contact than members of the same social group (Weber, Carter

et al. 2013). This heterogeneity in contact behaviour is thought to drive the clustered distribution of *M. bovis* infection in badger populations (Delahay, Langton et al. 2000, Woodroffe, Donnelly et al. 2005). Additionally, bovine tuberculosis in badgers is a chronic disease, caused by a slow growing mycobacteria, with infected badgers persisting and reproducing successfully despite infection (Tomlinson, Chambers et al. 2013). The presence of stable social structure within a population favours prolonged close contact between individuals, which in studies of human tuberculosis, has been suggested to generally be required in order for transmission to occur (Beggs, Noakes et al. 2003) . Kin structure within social groups can add further complexity to transmission dynamics, as has been demonstrated in a small number of studies (Grear, Samuel et al. 2010, Dharmarajan, Beasley et al. 2012, Vander Wal, Edye et al. 2013). However, to date, the role of kin structure on infection transmission within badger social groups has not been explored. This is likely to be significant, as many badgers remain in their natal group throughout their lives (Roper 2010), resulting in marked kin structure within badger social groups. In my first data chapter, I explore how social structure and kin structure within the social group boundaries predict the likelihood of infection with *M. bovis* in early life. This has implications for disease management, as *M. bovis* infection in cubs is thought to be associated with more progressed disease, compared to that in adults (Gallagher and Clifton-Hadley 2000) and therefore a higher risk of onward transmission. This has been attributed to the underdeveloped immune system of cubs (Tomlinson, Chambers et al. 2013), such that individuals who are exposed to *M. bovis* in early life are less able to contain the infection and are more likely to progress to disseminated infection, resulting in excretion of *M. bovis* bacilli through a range of routes (Clifton-Hadley, Wilesmith et al. 1993, Gallagher and Clifton-Hadley 2000). An understanding of the risks of acquiring infection in early life also has key implications for the success of vaccination campaigns, which only have access to cubs once they are above the ground and less dependent.

1.4 Between and Within Individual Variation

As well as the role of social structure in creating heterogeneity in transmission rates within a population, further complexity is found when we look closer at variation amongst individuals, both in terms of their behaviour and their

immuno-genetic profiles. It has been stated that 'the contribution that host and pathogen genotypes make to disease outcomes has, until recently, been overlooked by policy makers, yet it is biologically untenable that genetic variation of both organisms does not play a role.' (Allen, Minozzi et al. 2010). Individual variation in immune response can be strongly influenced by variation in an individual's genetic background (Ardia, Parmentier et al. 2011). Mating between related individuals, known as inbreeding, results in a reduction in genetic variation in the resultant offspring. Inbreeding may be associated with a depressed immune response, if the loss of genetic variability is within loci involved in parasite defence (O'Brien and Evermann 1988). An individual's genetic background is an important determinant of variation in antibody responses and immune system activation (Ardia, Parmentier et al. 2011).

Alongside genetic sources of inter-individual variation in immune response, life history theory predicts that immune responses should be flexible and responsive to changing environmental contexts; if there is no immunological challenge present in the form of a pathogen which limits life history success then there is no selection pressure for an individual to invest in immune response (Ardia, Parmentier et al. 2011). Individuals who invest more in certain facets of immune response face trade-offs, as they have less resource to invest in other life-history functions, or even other aspects of the immune response (Ardia, Parmentier et al. 2011). It has been proposed that, although individuals can invest differently in immune responses, based on the selection pressures present in a given environment, underlying genetic and physiological factors can constrain this flexibility (Ardia, Parmentier et al. 2011) i.e. some individuals have a greater genetic bank from which to select an appropriate immune response than others. It can be difficult to discern when an individual is mounting a sub-optimal immune response because of a lack of underlying allelic diversity or due to a resource trade-off with another function (Ardia, Parmentier et al. 2011). Environmental conditions, such as food availability (Lochmiller, Vestey et al. 1993, Chandra 1996) and weather (Sevi, Annicchiarico et al. 2001) can also influence host condition, with sub-optimal conditions resulting in reduced immune responses. Immune responses can vary over an individual's life history; juveniles and adults can mount divergent responses to the same pathogen challenge and older individuals are predicted

to have less flexibility in their immune responses. Chronic stress has been shown to depress the immune system (Padgett and Glaser 2003), so individuals may have compromised immune responses at certain life stages associated with 'stressful' events such as dispersal and resource competition (Gallagher and Clifton-Hadley 2000). These phenomena create within individual variation in immune responses.

There is substantial evidence of a genetic component to inter-individual variation in susceptibility to infection with bacteria of the *M. tuberculosis* complex, of which *M. bovis* is a member (Lyons, Frodsham et al. 2009, Barreiro, Tailleux et al. 2012). However, less is known about the influence of genetic variation on *M. bovis* infection progression, with limited data from wild populations (Acevedo-Whitehouse, Vicente et al. 2005). Inbreeding depression has been linked to increased susceptibility to bovine TB infection in free-living (Dorman, Hatem et al. 2004, Trinkel, Cooper et al. 2011) and captive (Briles 2012) wildlife and livestock (Allen, Minozzi et al. 2010, Brotherstone, White et al. 2010, Vordermeier, Ameni et al. 2012). In wild boar, genetic heterozygosity was an important predictor of both risk of infection with bovine TB and progression of disease (Acevedo-Whitehouse, Vicente et al. 2005). A study investigating the immune responses of red deer experimentally challenged with *M. bovis* suggested that nearly half of the observed variation in response was attributable to host genetic variation (Griffin and Mackintosh 2000). However, no published studies to date investigate inbreeding depression and bovine TB progression in badgers. In Chapter 4 I use microsatellite marker data to explore whether individual genetic heterozygosity predicts the progression of *M. bovis* infection in exposed badgers. In terms of practical applications, uncovering whether there is any evidence of a genetic component to bTB progression in badgers may have important implications for modelling infection spread in badger populations, particularly at the moving edge of the epidemic front where populations may not have previously been exposed to *M. bovis* and may therefore be more naïve in terms of their immune responses. Population management strategies, such as culling, may also influence the extent of inbreeding within a population; either increasing it by reducing the number of available mates or decreasing it by enhancing population movement.

Additionally, the efficacy of vaccination may vary between individuals based on their genotype.

1.5 Changing spatial and temporal patterns

From exploring the role of social structure and individual variation on disease dynamics, in Chapter 5 I take a step back to consider the study population as a whole, capitalizing on its long temporal scale to look for spatial and temporal trends in *M. bovis* infection distribution. As discussed above, pathogens are often spatially clustered; driven by limited pathogen dispersal, spatially restricted reservoirs or hosts being spatially clumped (Ostfeld, Glass et al. 2005), with social structure and potentially kin structure generating further substructure. These drivers of aggregation can be altered by external factors. For example, climate change may cause a vector to move outside of its current spatial range (Kovats, Campbell-Lendrum et al. 2001), or population management can alter host social structure (Carter, Delahay et al. 2007) or ranging behaviour (Ramsey, Spencer et al. 2002, Riordan, Delahay et al. 2011, Blackwood, Streicker et al. 2013), resulting in changes to clustering patterns (Jenkins, Woodroffe et al. 2007).

Significant changes in the spatial distribution of *M. bovis* infections in badgers have been noted in response to population management (Jenkins, Woodroffe et al. 2007). Since they were first implicated as a potential source of TB transmission to UK cattle, badgers have been the subject of various culling strategies (Krebs, Anderson et al. 1997). In 1998 the Randomised Badger Culling Trial (RBCT) was initiated. This large-scale field experiment was designed to provide a scientifically robust evaluation of the effect of culling badgers on cattle TB breakdowns (Krebs, Anderson et al. 1997). Thirty 100 square kilometre study areas were matched into treatment triplets; 'reactive' areas where badgers were removed on and surrounding farms that had experienced disease in cattle (known as a herd breakdown), 'proactive' areas where as many badgers as possible were removed and 'control' areas, where no action was taken (Donnelly, Woodroffe et al. 2005). The 'reactive' treatment, which represented the most similar approach to the pre- RBCT badger culling strategies (Donnelly, Woodroffe et al. 2003) was linked to a 27% increase in the incidence of cattle herd breakdowns and was abandoned early for this reason.

In the proactive areas, cattle TB incidence was 23% lower, with the benefit increasing over successive culls, although in the land surrounding these areas, it increased by 25% (Donnelly, Wei et al. 2007). This increase was attributed to culling-induced changes in badger behaviour, a phenomenon which was termed 'social perturbation'. Culling was linked to disruption of the social organisation of the badger population, resulting in enhanced movements amongst animals surviving the cull, and potentially exacerbating disease spread (Carter, Delahay et al. 2007). Badger group territories became consistently larger in culled areas, demonstrating that culling had had marked impacts on the spatial organisation of badger populations (Woodroffe, Donnelly et al. 2006). Radio-tracking studies of the behaviour of badgers during the RBCT also identified disruption in territoriality in culled areas, and showed that surviving individuals moved between groups more often and had larger home ranges than animals in unculled areas (Riordan, Delahay et al. 2011). Spatial clustering of *M. bovis* infections in badgers taken during sequential culls reduced, which was attributed to surviving individuals ranging more widely (Jenkins, Woodroffe et al. 2007).

To assess when major changes have occurred to a system as a result of management interventions, it is valuable to have an understanding of the spatio-temporal distribution of disease under unmanaged conditions. Prompted by an untested observation that the spatial arrangement of *M. bovis* infections in the Woodchester badger population had destabilised from spatial persistent foci (Delahay, Langton et al. 2000), in my third data chapter I describe the spatio-temporal arrangement of *M. bovis* infections within this study population over a two decade period. I use a range of methods, including those employed during the RBCT, to assess whether the distribution of *M. bovis* infection has substantially changed in this unmanaged population.

1.6 Genetic Population Structure and Pathogen Transmission

Understanding gene flow through a host population can provide insights into contact events between spatially distant individuals which may go undetected by observational methods and are not logistically possible to capture using techniques such as radio-collaring over a long period (Altizer, Bartel et al. 2011). Population genetic approaches can characterize historic host dispersal

patterns and potentially help predict pathogen spread (Streicker, Winternitz et al. 2016). Differences in dispersal patterns between the sexes can mean that one sex may be more important in pathogen spread (Streicker, Winternitz et al. 2016). Investigating the spatial scale of genetic correlation among hosts can provide insights into the distribution and potential spread of disease (Blanchong, Robinson et al. 2016). Having explored the fluctuations in the spatio-temporal distribution of *M. bovis* in the study population, in Chapter 6 I use genetic approaches to determine whether these fluctuations may result from temporal change in movement behaviour in the Woodchester badger population. Previous studies on this population have linked inter-social group movement to TB incidence (Rogers, Delahay et al. 1998, Vicente, Delahay et al. 2007), with years of high movement followed by years of increased incidence risk (Rogers, Delahay et al. 1998). However, these studies have relied entirely on the capture-mark-recapture data. It has been previously noted that extra-group mating can be commonplace in badger populations (Carpenter, Pope et al. 2005, Annavi, Newman et al. 2014). Visualising the genetic population structure allows us to gain a different perspective on possible opportunities for *M. bovis* transmission amongst badgers, rather than relying solely on the capture dataset. Forays into other social group territories to seek mating opportunities may be ephemeral and may not be detected by trapping at main setts or by delineating social territory boundaries. Here, I use the longitudinal capture records in combination with host genotype data in order to gain a more complete picture of movement and demography within the Woodchester badger population. This will contribute to a currently limited body of work in which host genetic structure has been used to understand pathogen spread and host movement at a restricted spatial scale (Cullingham, Merrill et al. 2011, Mazé-Guilmo, Blanchet et al. 2016) rather than as a larger scale landscape genetics approach (Biek and Real 2010) looking at host population structure across regions (Blanchong, Samuel et al. 2008, Lee, Ruell et al. 2012), countries (Lang and Blanchong 2012, Vander Wal, Edye et al. 2013) or continents (Streicker, Winternitz et al. 2016). Also, as management interventions such as culling have been noted to result in changes in genetic population structure (Pope, Butlin et al. 2007), related to movement patterns, it will be valuable to quantify the changes that can occur within a unmanaged system as a result of natural demographic processes.

An alternative or ideally additional approach to considering host genetic population structure is to consider the genetic population structure of the pathogen itself (Blanchong, Robinson et al. 2016). Comparing pathogen genotypes is based on a fairly straightforward rationale. Mutations in conserved nucleotide sequences are identified that can be used to determine how samples are related to each other; closely related samples will share most of the same mutations (Blanchong, Robinson et al. 2016). As I explore in Chapter 2, the development of technological advances such as whole genome sequencing has allowed differentiation between even very closely related samples. However, whole genome sequencing has rarely been used to examine the transmission dynamics of a bacterial pathogen in wildlife (Kamath, Foster et al. 2016). For a pathogen such as *M. bovis* which is highly clonal (Smith, Gordon et al. 2006), with an extremely low mutation rate (Biek, O'Hare et al. 2012), there are few informative differences between isolates. More traditional genotyping approaches based on typing only small genomic regions are therefore unlikely to be able to distinguish between *M. bovis* isolates at a fine spatial scale. My final data chapter (Chapter 7) capitalizes on the recent availability of whole genome sequence data for *M. bovis* isolates from the Woodchester badger population. *M. bovis* sequence data from badgers has only recently become available, hence very little work has been published in this area (Biek, O'Hare et al. 2012). This final approach uses host genotype data as an explanatory factor for pathogen genetic structure. Pathogen genetic structure is expected to reflect host genetic structure, as dispersal of hosts drives the dispersal of pathogens (Mazé-Guilmo, Blanchet et al. 2016). A high and significant correlation between pairwise distances between genotypes of a host and those of the pathogen they are harboring would indicate that host and parasite dispersal rates are strongly related to each other (Nieberding, Durette-Desset et al. 2008). In this final chapter, I use the available sequence data in order to examine strain diversity within the badger population over a decade long period, to consider whether there is genetic evidence for spatial spread of infection across the study site and to further investigate the impact of kin structure of transmission. Finally, in Chapter 8 I discuss the overall findings from the thesis, place them within the context of the wider field of research, discuss their management implications and outline potential directions for future research.

CHAPTER 2: Literature Review

What has molecular epidemiology ever done for wildlife disease research?

2.1 Abstract

The increasing availability of novel molecular techniques has transformed the study of human health and disease epidemiology. However, uptake of such approaches has been more conservative in the field of wildlife disease epidemiology. I consider the reasons for this and discuss current and potential applications of molecular techniques in a variety of relevant areas within the field of wildlife disease research. These include conducting wildlife disease surveillance, identifying sources of pathogen emergence, uncovering host-pathogen dynamics and managing current outbreaks, including the development and monitoring of wildlife vaccines. I highlight key examples of applications of molecular epidemiological approaches to wildlife disease scenarios and draw parallels from human disease research to suggest potential future directions. The potential value of next generation sequencing technologies to the field of wildlife disease research is discussed and initial applications are highlighted, balanced against consideration of the challenges involved. Using a wide range of examples drawn from research into human, livestock and wildlife diseases I demonstrate the value of using molecular epidemiological approaches at all scales of wildlife disease research, from pathogen strains circulating at a global scale to intra-individual host pathogen dynamics. The potential future contribution of these technologies to the field of wildlife disease epidemiology is substantial. In particular they are likely to play an increasingly important role in helping us to address a principal challenge in the management of wildlife diseases which is how to tease apart the transmission dynamics of complex multi-host systems in order to develop effective and sustainable interventions.

2.2 Introduction

Since its emergence in the 1970's, the term 'molecular epidemiology' has appeared in a vast number of publications in a wide range of scientific disciplines. Initially the term was used primarily in the study of human cancer to describe the process of identifying biomarkers within populations which improved identification of subgroups at greater risk of developing disease (Vineis and Perera 2007). However, the term is now widely used in the field of infectious disease biology, where it has been defined as involving 'the various techniques derived from immunology, biochemistry and genetics for typing and sub-typing pathogens' (Tibayrenc 1998). A broader definition that goes some way to capturing the breadth of the subject is 'a science which utilises molecular biology to define the distribution of disease within a population and relies heavily on integration of traditional epidemiological approaches to identify the etiological determinants of this distribution' (Snow 2011). The influence of molecular epidemiology in the field of human health research has been extensive. Ongoing surveillance of the spatio-temporal distribution of disease strains has helped uncover drivers of disease transmission (Liu, Graber et al. 2008), infer the geographic origin of pathogens (Hemelaar, Gouws et al. 2011) and provided a baseline against which changes can be detected (Koopmans, Vinjé et al. 2000). In addition, transmission routes of zoonotic pathogens have been identified (Feng and Xiao 2011, Salyer, Gillespie et al. 2012), and the evolutionary provenance of pathogenic strains (Byrnes, Li et al. 2010) and antibiotic resistance mechanisms (Hoffmann, Minkah et al. 2007, Kumarasamy, Toleman et al. 2010) have been pinpointed. Molecular technologies have also been widely used in the development of vaccines against human pathogens and in responsive investigations of disease outbreaks (Gardy, Johnston et al. 2011, Rasko, Worsham et al. 2011, Grad, Lipsitch et al. 2012). In this review, I will explore similar applications of molecular technologies in the field of wildlife diseases and suggest directions for future applications.

2.3 Bibliometrics

A comprehensive literature search for journal articles published since 1980 using the term 'molecular epidemiology' revealed over 111,000 results. In

order to consider the growth of molecular approaches in epidemiology, the proportion of 'molecular epidemiology' articles within 'epidemiology' articles for the fields of clinical, livestock, zoonoses and wildlife research were calculated. Time series analysis and forecasting of publication trends illustrate the proportion of 'epidemiology' papers incorporating molecular epidemiological approaches for each of the four categories (Fig 2.1). When the overall numbers of 'molecular epidemiology' publications within each field are considered, unsurprisingly the vast majority fall within the field of human clinical research (Fig 2.2). The application of molecular epidemiological techniques in livestock, zoonoses and wildlife disease research has steadily grown from the mid 2000's onwards, but such studies still represent only a small proportion of the total observed. The origin of the term in human clinical research is indicated by its earlier appearance in this field, filtering into the other fields shortly afterwards (Fig 2.1). Time series forecasting indicates that the proportion of 'molecular epidemiology' publications is predicted to increase relatively steeply in the fields of livestock and zoonotic research, but the trend is less certain in the field of wildlife disease research as indicated by the wider confidence intervals (Fig 2.1).

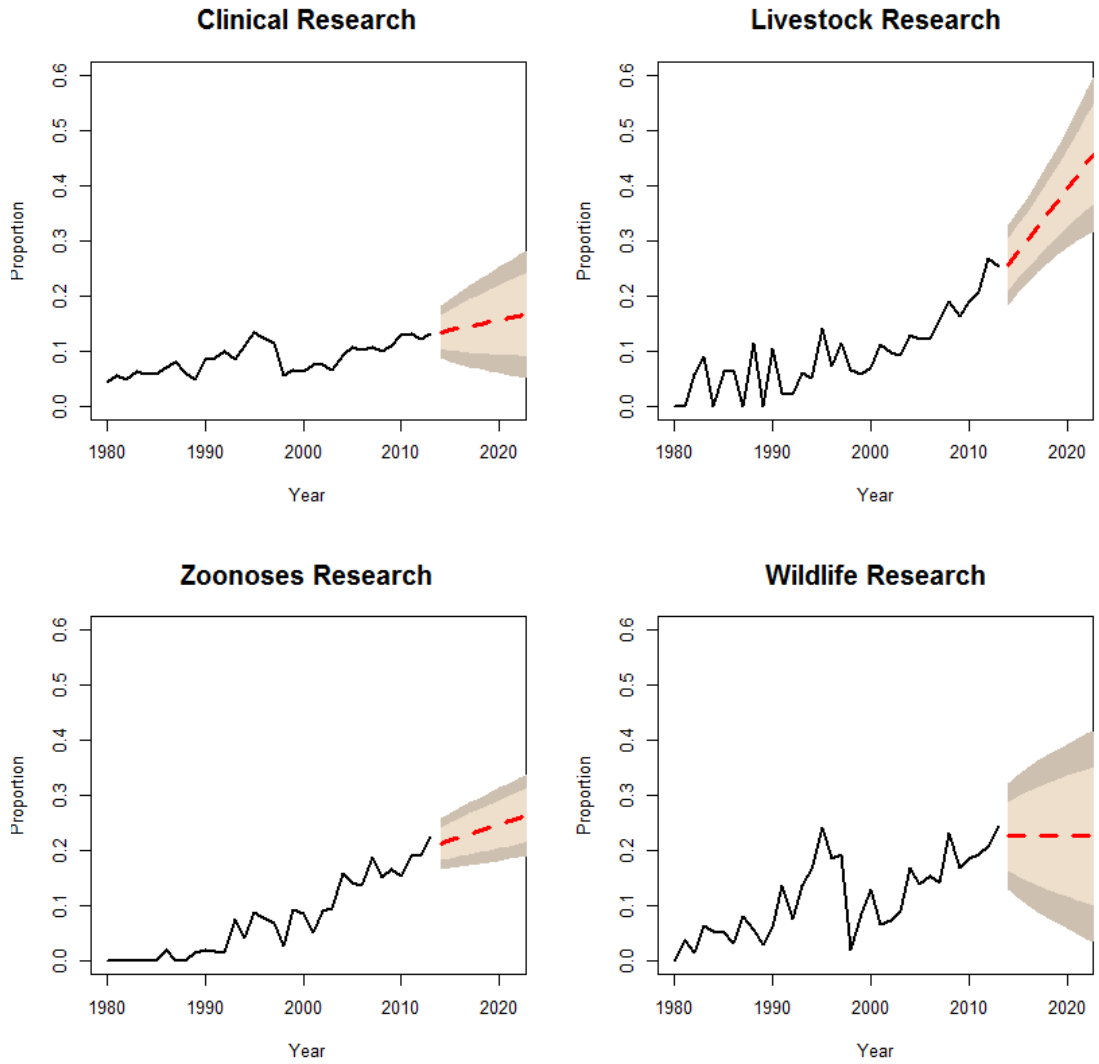


Fig 2.1 Time series analysis and forecasting of the proportion of 'epidemiology' journal articles that are 'molecular epidemiology' in the fields of clinical research, livestock research, zoonoses research and wildlife research. Based on ISI Web of Science search within topic field conducted in April 2014. Time series forecasting carried out using the 'forecast' package in R (Hyndman and Khandakar 2007).

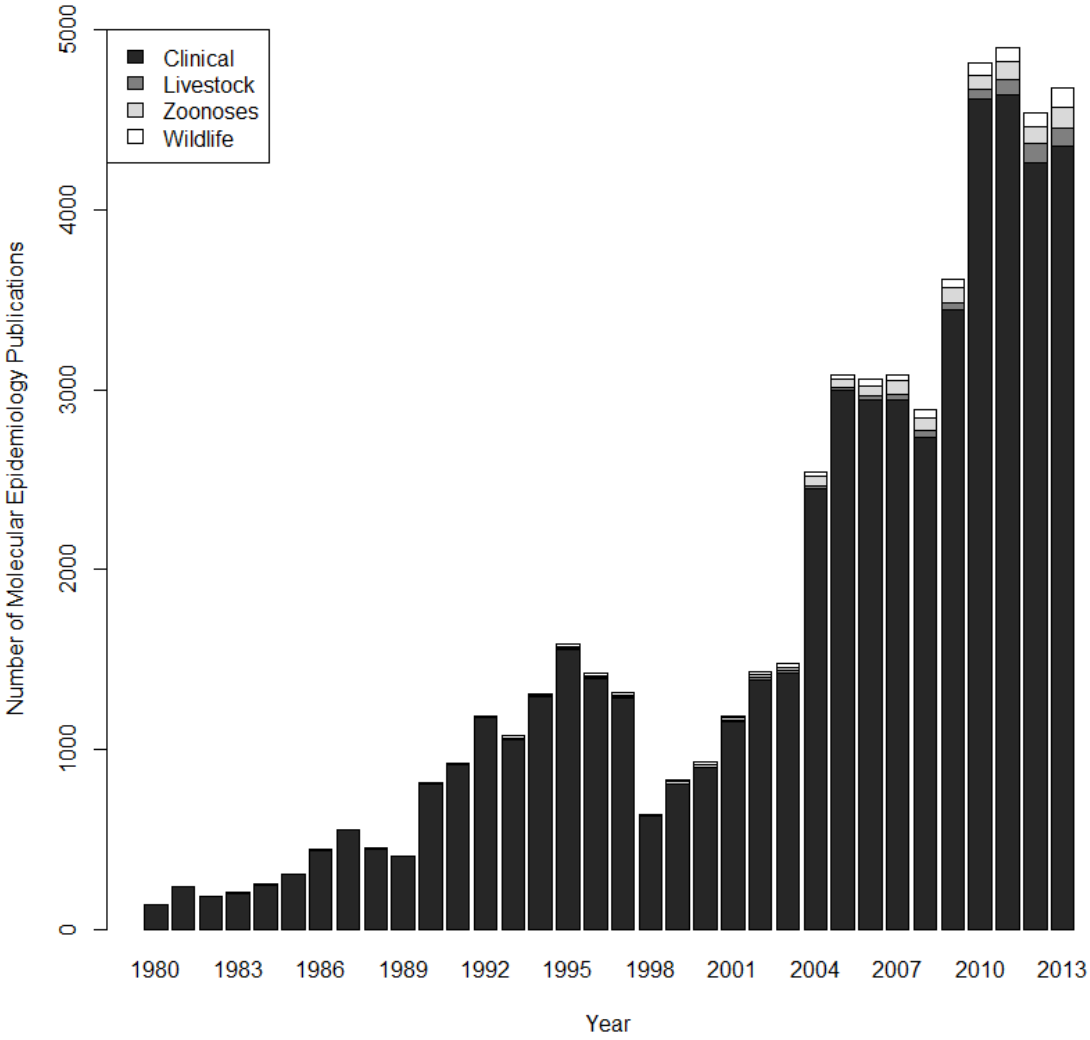


Fig 2.2 Number of 'molecular epidemiology' journal articles published in the fields of clinical research, livestock research, zoonoses research and wildlife research. Search conducted using ISI Web of Science in April 2014

The bias towards human disease is perhaps not surprising or unexpected given that molecular approaches were first developed in this field and human diseases naturally attract a greater level of attention and funding than diseases of animals. It may also reflect the greater focus on non-infectious diseases related to genetic or environmental factors in human systems, for example the large number of studies which employ molecular epidemiological approaches to identify biomarkers associated with cancer. In general, diseases of livestock or

wildlife have only been considered important when agriculture or human health is potentially threatened (Daszak, Cunningham et al. 2000), which is consistent with the smaller number of publications and conservative uptake of molecular epidemiological techniques in the field of non-zoonotic wildlife disease. There are also a range of practical challenges involved in any study of disease epidemiology in wild populations, including access to individuals, absence of validated diagnostic tests, logistics and costs of sampling, poor baseline surveillance and inherent uncertainty surrounding species ecology and behaviour (Delahay, Smith et al. 2008). Nevertheless, as we draw parallels with studies using molecular techniques in other fields, we will see that there are many potentially valuable applications of molecular methods to the epidemiology of disease in wildlife populations. In this review, I consider how molecular epidemiological approaches can help wildlife managers address key questions about disease dynamics and I suggest directions and opportunities for their wider application in this field.

2.4 Molecular methods

Although DNA based techniques have been in use for less than fifty years (Medini, Serruto et al. 2008) during recent decades a wide range of molecular techniques have emerged for the study of pathogens. One example is typing based on 16S ribosomal RNA (rRNA), in which the percentage of sequence similarity of rRNA molecules between samples is used to classify species (99% sequence identity is used as the cut-off between separate species) (Medini, Serruto et al. 2008). In alternative techniques, other genomic elements are sequenced and used as the basis of classification such as housekeeping gene fragments in MLST (multi-locus sequence typing) and enzyme profiles in MLEE (multi-locus enzyme electrophoresis) in bacteria. However, these systems can struggle to distinguish amongst very similar strains (Achtman 2001), such as members of the *Mycobacterium tuberculosis* complex (Frothingham 1995, Köser, Ellington et al. 2012) or strains of *Bacillus anthracis* (Keim, Smith et al. 2001), the causative agent of anthrax.

The choice of an appropriate molecular typing technique requires an understanding of the genomic structure of the pathogen in question. Bacterial genomes consist of a core genome, common to all strains, dispensable genes

which are not present in all strains and genomic islands; clusters of contiguous genes with a specialised function (e.g. virulence) (Relman 2011). Classical methods of classifying bacterial pathogens are based on phenotypic characteristics such as cellular structure (Medini, Serruto et al. 2008), colony morphology (Baron, Hummel et al. 1996) and antibiotic susceptibility (Harwood, Whitlock et al. 2000). In contrast, viral pathogens contain small genomes, are highly diverse and some can evolve very rapidly (Fierer, Breitbart et al. 2007). For example, the ability of the influenza virus to rapidly change its antigenic profile requires continual development of vaccines (Relman 2011). Genetic diversity in fungal pathogens depends on a) the mode of reproduction of the species; sexual or asexual and b) the presence of 'transposable elements'; mobile genetic elements which can insert themselves within genes, changing their structure and function (Daboussi 1997). Protozoans such as the human pathogens *Trypanosoma brucei* and *Leishmania major* have genomes with species specific surface antigens and variable strategies of invading hosts and evading immune responses. Conservation of gene order between species is high, indicating the presence of a strong selection pressure to conserve certain gene clusters and their associated function (El-Sayed, Myler et al. 2005).

The underlying genomic structure and diversity within a particular pathogen will influence the choice of typing method. All the typing techniques described above are fundamentally limited as they only examine a small section of the host genome. Depending on the biology of the pathogen, the section under examination will vary in the degree to which it is representative of the whole genome. For example, the typing methods traditionally used to categorize strains of *Mycobacterium bovis* (the causative agent of bovine TB) are spoligotyping (spacer-oligonucleotide typing) and VNTR (Variable Number Tandem Repeat) typing, both of which are based on small, hyper-variable genomic regions that are generally evolving at a higher rate than the rest of the genome (Joshi, Harris et al. 2012). Such methods are therefore potentially more useful for differentiating between species than detecting finer scale intra-specific variation, although this will depend largely on the genetic diversity within the particular pathogen complex under examination. Rapidly evolving viruses may generate sufficient diversity for intra-specific strains to be differentiated on the basis of more restricted areas of the genome than bacterial

pathogens, but only by examining the complete genome of a species can finer genetic structuring be uncovered (Medini, Serruto et al. 2008). The development of 'next generation' sequencing approaches, which base classification on the identification of SNPs (Single Nucleotide Polymorphisms - single base substitutions; insertions or deletions) that vary between individual genomes, has facilitated the rapid sequencing of whole genomes, opening the door for studies which were previously impossible. In order to construct a phylogenetic tree from a group of sequenced isolates, phylogenetically informative SNPs (i.e. those shared by two or more isolates) are identified through the examination of the maximum sequenceable genome of each isolate. SNPs occur in both coding and non-coding regions, but those in the latter are less likely to exert a phenotypic effect and therefore are less likely to be affected by selective pressures. Hence, whole genome sequencing infers phylogenetic relationships from the maximum amount of genetic information available.

The current next generation sequencing (NGS) technologies are based on breaking the original genome into fragments, which are then extensively sequenced to yield short read sequences. These reads can then either be mapped to a reference genome (where present) or assembled against each other (*"de novo"* assembly) in order to identify potential SNPs. SNPs which successfully pass the required quality checks can then be used to produce phylogenetic trees and inform transmission models. The cost and required infrastructure for these technologies have so far limited their widespread uptake (Metzker 2010). However, costs are rapidly falling (Köser, Ellington et al. 2012), uptake is increasing and the incorporation of NGS into routine human disease surveillance (Roetzer, Diel et al. 2013) and clinical diagnostics (Boyd 2013) appears to be imminent. Full sequence data has all the potential applications of strain typing but at a far higher resolution and gives the opportunity to determine the extent of the differences amongst strains rather than to simply distinguish them from one another. However, there are challenges in these approaches in relation to the storage, quality control and manipulation of the enormous amounts of data that they generate (Pop and Salzberg 2008). Also, lack of standardisation in bioinformatics protocols limits the extent to which sequences can be compared across laboratories (Köser,

Ellington et al. 2012) and technologies (Metzker 2010). Nevertheless, NGS has enormous potential to uncover fine-scale disease transmission dynamics, which may otherwise remain hidden to epidemiologists.

2.5 Disease surveillance

The importance of surveillance for wildlife diseases is well established (Artois, Bengis et al. 2009). Ongoing surveillance can act as an early warning system for outbreaks of new or emerging diseases, allowing pre-emptive management interventions and potentially helping to inform assessment of risk related to conservation interventions such as translocations of endangered populations (Artois, Bengis et al. 2009).

It is important to assess the extent of genetic diversity within a pathogen population as this has implications for how refined molecular tools need to be in order to investigate disease transmission events. For example, with a low diversity pathogen such as *Bacillus anthracis*, the most divergent strains are thought to be 99.99% similar in terms of nucleotide sequencing (Rasko, Worsham et al. 2011) and therefore most isolates will appear homogenous, no matter how rigorous the sequencing method. In contrast, within the HIV-1 virus there is a wealth of genetic diversity which is organised into 'subtypes' within which genetic variation can range between 8 and 17% and can be as much as 35% amongst sub-types (Hemelaar, Gouws et al. 2011). Genetically diverse populations of the virus, termed 'quasi-species', can be harboured by a single host individual (Domingo and Holland 1997). Only through ongoing surveillance of circulating strains can the intrinsic genetic diversity of a pathogen be captured.

Ongoing surveillance of strain diversity can inform the development of diagnostic tests which may be employed for the identification and potentially selective removal of infected individuals in livestock and wildlife populations. The removal of test positive individuals could potentially exert a selection pressure on pathogen populations, with selection favouring strains that produce a weak or negative diagnostic test response. Hence, diagnostic test development is ideally an ongoing process, which seeks to keep one step ahead of such selection pressure. An example from human health is that of *Neisseria meningitides*, the bacterial cause of Meningitis A. Molecular

approaches indicate that within this bacterial complex, there are a number of clonal groups, some of which cause disease and others which live commensally within human hosts (Achtman, van der Ende et al. 2001). Horizontal gene transfer between group members can generate genetic diversity. The identity of the most frequent genotype in a population can vary as changing forces of selection favour different strains (Achtman, van der Ende et al. 2001) with clear implications for the design of effective vaccines.

Comparison of pathogen genotypes in regions where infection is endemic versus those where infections only occur sporadically, may uncover genetic differences associated with the two scenarios, related for example to pathogenicity factors. Molecular typing may also have important applications in detecting the emergence of new pathogenic strains in populations. In the case of bacteria, the jump from benign to pathogenic could potentially occur relatively rapidly, through the acquisition of a genomic island which codes for a pathogenicity factor (Hacker and Carniel 2001) and molecular typing may aid the detection of such events.

2.6 Phylogeography

Molecular techniques are widely used to describe the spatio-temporal distribution of variant pathogen strains. For example, the characteristic home ranges of *M. bovis* genotypes in cattle have been mapped across the affected areas of the UK (Smith, Dale et al. 2003). Routine mapping of this kind may identify the appearance of atypical strains in an area; this may indicate that a 'novel' transmission event has occurred (e.g. the translocation of an infected host animal from another region). Geographic differences in virulence between pathogen strains may also occur, as has been identified for the fungal pathogen *Cryptococcus gatti* (Byrnes, Li et al. 2010). Spatial mapping exercises can also tell us something about the evolution of pathogen strains as geographically dispersed genotypes may be considered more likely to be ancestral strains than those with a restricted home range (Smith, Dale et al. 2003). Examining the prevalence of disease in a region can also be used to infer risk factors which could inform management strategies. Incorporating molecular information into these investigations can provide greater insight into possible causes than

simply comparing populations with and without disease (Cowled, Ward et al. 2012).

By examining the strains that are appearing at the moving edge of an epidemic front it may be possible to gain insights into the factors that are driving disease spread. For example, molecular epidemiology may be a useful tool in determining the proximate causes of new cases of bovine tuberculosis infection in UK cattle at the fringes of the endemic areas, helping to distinguish whether infection is seeded from livestock movements or the presence of infected wildlife. A very different example, focused on conservation of a highly threatened species is provided by devil facial tumour disease (DFTD) in Tasmanian devils *(Sarcophilus harrisii)*, where identifying the location of the disease front has informed management options. Geographic differences have been noted in the epidemiology and population effects of DFTD on devils. Genotyping techniques have recently been applied suggesting that distinct lineages of the pathogen may have different epidemiological outcomes (Hamede, Pearse et al. 2015), hence pathogen genetics may contribute to observed inter-population differences.

When phylogenetic trees of a particular pathogen are overlaid with epidemiological data (such as geographic location of outbreaks) they can be used to map spatial disease spread. This can help epidemiologists infer where transmission events have occurred and therefore potentially to predict and manage future disease risks. For example, examining the geographic localisation of strains of *Mycobacterium leprae*, the bacterial cause of leprosy in humans, indicated that global disease spread was most likely linked with historic human migration patterns and trade routes (Monot, Honore et al. 2009). Epidemiological linkages between particular populations or geographic locations can be identified if shared genotypes are recorded more often than would be expected by chance (Archie, Luikart et al. 2009). Host geography has also been found to play a role in rates of pathogen evolution. In the case of *Lyssavirus* (rabies) in bat populations, rates of viral evolution by nucleotide substitution vary depending on whether the host species is in a temperate or tropical environment, which may be related to differences in the seasonality of bat activity and the influence of climate on rates of virus transmission (Streicker, Lemey et al. 2012). Examining pathogen phylogenies can provide an

understanding of rates of new strain emergence, helping epidemiologists to predict and prepare for new disease outbreaks. Also, where transmission rates vary between strains of the same pathogen, either due to differences in infectivity amongst strains or the availability of susceptible hosts, this could be identified through considering rates of spread. Phylogeographic investigations have been conducted on a wide range of human pathogens, including the zoonotic bacteria *Yersinia pestis* (Vogler, Chan et al. 2011), Dengue Virus (Nunes, Faria et al. 2012) and influenza (Chang, Ding et al. 2007). In the case of vector-borne diseases, the same approach can be used to investigate distributions, as carried out in a study of *Triatoma infestans*, the primary insect vector of Chagas disease (Perez de Rosas, Segura et al. 2011). Phylogeographic approaches have also been used, albeit to a lesser extent, in wildlife and livestock diseases, for example to consider the ecological drivers behind foot and mouth disease (FMD) in cattle (de Carvalho, Santos et al. 2013), rates of viral evolution driving infectious bursal disease virus in farmed poultry (Cortey, Bertran et al. 2012) and the role of the global expansion of fish farming in the spread of salmonid proliferative kidney disease (Henderson and Okamura 2004).

2.7 Roots of emergence

The construction of pathogen phylogenetic trees has made an enormous contribution to the study of human disease, leading to the emergence of the field of evolutionary medicine (Bull 1994). Virulence is known to differ amongst pathogen strains and this variation is the result of evolutionary processes. Genetic signatures in pathogen phylogenies allow us to look back at the underlying ecological selection pressures which have previously been exerted on a pathogen, and shaped its evolution (Biek and Real 2010). Correct inference of ancestry (i.e. determining which strains of a particular pathogen are ancestral and which are descendant) is key to building a clear picture of pathogen population structures (Medini, Serruto et al. 2008). For example, the population structure of *Mycobacterium bovis* genotypes in the UK suggests a 'clonal expansion' of genotype evolution from a common ancestor, through a combination of selection and 'ecological opportunity' as invasion into new geographic areas occurred (Smith, Dale et al. 2003). Inferring ancestry is also extremely valuable for dating disease transmission events and tracing cross-

species transmission in multi-host disease complexes, such as SIV /HIV and Hepatitis B in humans and non-human primates (Starkman, MacDonald et al. 2003, Neel 2010). If a pathogen has been transmitted from one species to another, the phylogeny within the recipient species should be nested within that of the source species (Archie, Luikart et al. 2009). Disease introduction through migration or translocation events can be suggested where there is a genetic mismatch with resident strains, as was recently been inferred for some species of blood parasites in wild birds in Japan where strains of *Leucocytozoon* from migratory and resident birds were phylogenetically separated (Yoshimura, Koketsu et al. 2014). Hence, phylogenetic investigation can be used to identify risk factors for future disease outbreaks.

A substantial body of work exists where whole genome sequencing has been applied to the study of human viral pathogens, such as influenza and HIV (Henn, Boutwell et al. 2012) and in recent years this approach has also been applied to investigations of viral pathogens in wildlife, including the detection of Highlands J Virus in a critically endangered species of crane (Ip, Wiley et al. 2014), the development of a genome database of orbiviruses (Maan, Belaganahalli et al. 2013) and an investigation into encephalitis cases in captive polar bears (Szentiks, Tsangaras et al. 2014). Recently, the phylogeny of the pathogenic fungus *Geomyces destructans*, the causative agent of white-nose disease in bats (Blehert 2011) was produced, indicating that the introduction into North America originated in Europe (Leopardi, Blake et al. 2015). A number of pathogens of veterinary importance have had at least one isolate sequenced, including African swine fever virus (Chapman, Darby et al. 2011), *Mycoplasma haemofelis* (the causative agent of Feline Infectious Anaemia) (Barker, Darby et al.) and *Streptococcus equi* (Paillot, Darby et al. 2010), although these studies have focused primarily on describing pathogenicity factors, rather than epidemiological outcomes.

The potential for next-generation sequencing to infer the origin and population structure of veterinary and wildlife pathogens is substantial. We may expect uptake to be initially greater in the fields of zoonotic and livestock diseases, where the potential human 'cost' is perceived to be higher. As illustrated above, phylogenetic approaches offer so much more than an opportunity to delve backwards into the evolutionary history of a pathogen. They can also help us to

understand the drivers of the current distribution of pathogens and help us predict their likely distribution in the future.

2.8 Routes of transmission

When investigating the dynamics of infection in a given host population we reasonably assume that transmission is more likely to have occurred between individuals infected with the same strain of a pathogen than amongst those infected with different strains (Wylie, Cabral et al. 2005). Pathogen genotyping can therefore help to rule out or implicate particular transmission pathways, which may be valuable in tracing the initial source of infection and preventing further disease spread. The availability of next generation sequencing technologies has allowed contact networks and transmission pathways to be inferred with greater confidence and accuracy (Gardy, Johnston et al. 2011). Given the relatively recent availability of these technologies, and their decreasing cost, their full potential in the field of human health has yet to be realised (Walker, Ip et al. 2013) and to date their use in relation to livestock and wildlife diseases has been limited. However, there are notable examples such as studies of TB in cattle and badgers in the UK (Biek, O'Hare et al. 2012), TB in cattle and livestock in New Zealand (Crispell, Zadoks et al. 2017), brucellosis in livestock and wildlife (Foster, Beckstrom-Sternberg et al. 2009, Kamath, Foster et al. 2016) and MRSA in livestock (Price, Stegger et al. 2012). In studies of human pathogens such as *M. tuberculosis* (Cook, Sun et al. 2007, Gardy, Johnston et al. 2011, Walker, Ip et al. 2013) MRSA (Harris, Cartwright et al. 2013), *Clostridium difficile* (Eyre, Cule et al. 2013) and *Chylamidia trachomatis* (Wylie, Cabral et al. 2005), genotyping of pathogenic isolates has informed contact tracing, suggested the existence of undetected carriers and helped to both construct and verify the conclusions of social network analysis of disease outbreaks. Clinical disease outbreaks in human populations are often treated on a ‘case by case’ basis, on the understanding that no two events are epidemiologically identical. On the other hand, wildlife disease managers are often called upon to use simple management strategies to tackle disease in multiple socially-structured populations, without information on the particular transmission dynamics in each situation. The overlaying of data on pathogen strain diversity onto ecological information could be used in wildlife populations to assess transmission rates in relation to population structure (e.g. social

groups, herds etc.). In the case of the European badger the prevailing social structure in high density populations has been associated with the clustering of infection within social groups (Delahay, Langton et al. 2000). Disruption of this social structure, as observed following culling, leads to a reduction in this clustering, as surviving individuals range more widely (Jenkins, Woodroffe et al. 2007). Further understanding of the role of social behaviour in the spread of infection may be achievable by investigating the genetic diversity of *M. bovis* strains in badger populations. If social structure acts as a barrier to disease spread then we would expect the genetic distance between *M. bovis* strains within badger social groups to be shorter than that observed between individuals in different social groups. Wherever wildlife is implicated as a reservoir of zoonotic and/or livestock disease, such approaches may be valuable in identifying chains of disease transmission between species and could potentially indicate the direction of disease transmission.

In order to make meaningful inferences about transmission dynamics, a pathogen must be acquiring mutations within an epidemiologically meaningful timeframe and the genotyping method applied must have the ability to detect this variation (Grad, Lipsitch et al. 2012). Epidemiologists studying pathogens with very little variation between strains will require a typing method that is able to detect small differences between isolates. Where discrimination between isolates is not possible using conventional methods, whole genome sequencing (WGS) may be the only tool suitable for looking at fine-scale transmission dynamics. The exceptionally high level of genetic resolution achievable using WGS means that even sequencing a restricted number of isolates can reveal a wealth of epidemiologically valuable information. Where access to long term wildlife studies is possible, a 'phylodynamic' approach (Grenfell, Pybus et al. 2004) of overlaying pathogen phylogenies onto well documented epidemiological systems is potentially very powerful. Novel molecular approaches are not a replacement for traditional epidemiological investigations, but are complimentary, allowing a finer scale approach. For this reason, long term, well-studied epidemiological systems are the ideal scenarios in which to explore the contribution of cutting-edge sequencing to uncovering the drivers of disease transmission. Examples of such well-studied systems include TB infection in wild badgers (Delahay, Langton et al. 2000) and meerkats (Drewe

2010), chronic wasting disease in white-tailed deer (Williams, Miller et al. 2002) and DFTD in Tasmanian devils (Hamede, Bashford et al. 2009). It is important to note however that, even with the added resolution provided by WGS, there are considerable challenges to identifying pathogen transmission chains. The point at which mutations are acquired in a given transmission sequence is unknown and when mutation rates are slow compared to pathogen generation time, closely related isolates may appear genetically identical as they lack informative mutations (Kao, Haydon et al. 2014).

2.9 Host-pathogen dynamics

Pathogens can have widely differing effects in different host species, as is the case for the Squirrel Parapox Virus which causes severe disease in the European Red Squirrel (*Sciurus vulgaris)* but has no observed effects on the North American Grey Squirrel (*Sciurus carolinensis)* (Sainsbury, Nettleton et al. 2000). Variability in the observed costs of pathogen infection has also been observed amongst individuals of the same species. Heterogeneities in susceptibility to infection among individuals can affect the estimation of the transmission parameter R_0 (the basic reproductive number) (Hudson, Rizzoli et al. 2002). In such instances molecular techniques may allow us to distinguish between differences in pathogenicity which arise from strain variation and those which reflect heterogeneity in host immune responses. Scaling up these effects can impact on host population dynamics as regulation by a pathogen requires its *per capita* impact to outweigh the intrinsic population growth rate (Hudson, Rizzoli et al. 2002). If the *per capita* impact on host fitness is widely variable amongst individuals, then inferring population regulation is more complicated. Variation in how a pathogen physiologically affects individuals within a population has implications for onward transmission and persistence of disease (Cross, Lloyd-Smith et al. 2005). For example, inter-individual variation in the amount or concentration of pathogenic material excreted and the duration over which this occurs, is likely to affect the number of secondary cases observed. Ignoring this individual heterogeneity and assuming that each infected individual contributes to the same number of secondary infections can lead to highly inaccurate estimations of R_0. Molecular techniques can also allow us to examine individual variation in susceptibility and resistance within a host population by assessing the genetic basis of the immune response; an

approach known as immunogenetics. Recently, individuals who have apparently recovered from infection have been identified in Tanzanian devils (to DFTD) (Wright, Willet et al. 2017) and in European red squirrels (Chantrey, Dale et al. 2014)(to infection with parapox virus). In both of these cases, recovery from infection had never previously been documented. Individuals with greater allelic diversity within immune genes such as the major histocompatibility complex (MHC) are able to mount an appropriate immune response against a greater variety of pathogens (Castro-Prieto, Wachter et al. 2012). Accounting for heterogeneity in individual susceptibility and for differential strain pathogenicity is likely to allow R_0 to be estimated with a greater degree of accuracy.

It has been suggested that immunogenetic data should be used more widely to complement wildlife management decisions, particularly where populations are restricted or fragmented with a limited gene pool, as is often the case for highly endangered species (Acevedo-Whitehouse and Cunningham 2006). A key example of this approach is investigation into the spread by biting of a contagious cancer amongst Tasmanian devils which is thought to have caused a 90% population decline (Siddle, Marzec et al. 2010). The absence of an immune response in infected devils is thought to be linked to the limited genetic diversity within their MHC complex (Siddle, Kreiss et al. 2007). Examination of MHC genetic diversity within devil populations in other areas of Tasmania, where the disease is absent or at low prevalence, have identified some unique profiles which may confer disease resilience. If this were the case, then selective breeding and translocation of resilient individuals has been suggested as a means of controlling disease spread (Hamede, Lachish et al. 2012). In contrast, as MHC profiles are likely to be adapted to local pathogenic selection pressures, a poor choice of origin or destination could leave a translocated individual unable to cope with a different pathogen community and so at a selective disadvantage (Castro-Prieto, Wachter et al. 2012). Hence, the application of sequence-based approaches to assess the immunogenetic characteristics of populations of endangered species may have the potential to increase the likelihood of successful translocation (Boyce, Weisenberger et al. 2011).

Molecular epidemiology also allows us to zoom in to an even finer scale than that of inter-individual variation, and to consider intra-individual host-pathogen

dynamics. Infectious pathogens persist in the context of a co-evolutionary arms race with the host (Acevedo-Whitehouse and Cunningham 2006) which can be considered as a habitat 'patch' occupied by a parasite 'community' (Hudson, Rizzoli et al. 2002). Where an individual host is infected with multiple strains of the same pathogen, strain competition can occur, with certain strains favoured owing to their faster growth rate or ability to grow in a certain tissue (Bull 1994). It is interesting to consider however, that the traits which allow a particular strain to dominate within the host environment, may not necessarily optimise onward transmission although they may increase pathogen virulence (Bull 1994). However, outcomes of multiple infection on evolution of virulence and subsequent effects on individual host fitness are variable (Rankin, Bargum et al. 2007). The application of suitable molecular techniques to detect multiple strains of a pathogen within a host and to detect within-host pathogen strain evolution or strain competition may have important implications, as for example the scale of competition between bacterial strains is thought to influence the evolution of virulence (Griffin, West et al. 2004). Comparative genomics studies, in which the aim is to link genetic sequence differences between strains with phenotypic differences in the host (e.g. differential pathogenicity), have acquired valuable additional resolution from the development of next generation sequencing technologies (Hu, Wang et al. 2012).

Understanding variations in the impact of pathogens both amongst and within individuals may be critical to achieving effective management at the population level. Furthermore, disregarding heterogeneity in host responses and failing to acknowledge within host-pathogen dynamics (such as the role of multiple infection) may result in unexpected, potentially adverse, outcomes of management interventions. Molecular approaches have much to offer at both scales of analysis.

2.10 Vaccine development and monitoring

Vaccination is currently being used or considered as a management option in several high profile wildlife disease scenarios, including the control of bovine TB in badgers in the UK (Chambers, Rogers et al. 2010, Carter, Chambers et al. 2012), chlamydia in koalas in Australia (Carey, Timms et al. 2010) and haemorrhagic disease and myxomatosis in rabbits in Europe (Spibey, McCabe

et al. 2012). Molecular epidemiology has a key role to play in the development of effective vaccines for wildlife and monitoring their impacts on disease epidemiology. Vaccine development against human pathogens has greatly benefited from technological advances in gene sequencing, and now the sequences of many pathogens are available. This has led to the emergence of the field of 'reverse vaccinology' which typically involves mining the pathogen sequence for antigens that may be suitable as vaccine targets (Serruto, Serino et al. 2009). In pan-genome reverse vaccinology, multiple isolates of a pathogen species are considered. This is based on the idea of the existence of a 'pan-genome', which acknowledges that any single isolate of a pathogen does not exhibit all the genetic diversity within that species, especially if they are capable of generating genetic diversity through recombination or horizontal gene transfer. Consequently, it is necessary to sequence multiple genomes in order to get a better measure of the entire genomic repertoire of a species (Tettelin, Riley et al. 2008). In comparative reverse vaccinology, sequences of pathogenic strain isolates are compared with those of non-pathogenic isolates of the same species, in order to identify antigens associated with pathogenicity. The first human pathogen for which a vaccine has been developed and recently licensed using this approach is Serogroup B Meningitis (*N.meningitidis*), responsible for 80% of meningitis cases in Europe (Santolaya, O'Ryan et al. 2012).

As well as informing the development of vaccines, genome sequencing technologies also have applications for monitoring the effectiveness of vaccine deployment. As an increasing proportion of a population is immunised the selection pressure favouring strain variants that are unaffected by vaccination will grow. The emergence of these strains, known as 'escape mutants' could be monitored by sequencing isolates before and after vaccination to look for new mutations related to immunity in the targeted proteins (Seib, Dougan et al. 2009). This sort of approach could be extremely valuable in monitoring the impacts of vaccination in wildlife populations. Comparing the genetic diversity of pathogen populations before, during and after vaccine deployment could provide valuable information on the potential emergence of resistant strains or differential vaccine performance against variant strains. Strain typing could also help monitor reversion to virulence of live vaccines, as vaccination has on

occasion been observed to result in clinical disease. This was recently reported in a red fox (*Vulpes vulpes*) in which strain typing was able to identify the live rabies vaccine as the aetiological agent (Hostnik, Picard-Meyer et al. 2014). Live vaccines, such as the oral rabies vaccine, may be derived from multiple strains. Genetic characterisation of these strains can uncover the genetic basis for their attenuation (Geue, Schares et al. 2008). Population coverage of vaccines which may be horizontally transmissible within a population (Angulo and Juan 2007) could also be monitored using molecular diagnostics.

Understanding the antigenic diversity of a pathogen is key in vaccine design, and is only possible through ongoing surveillance as the most frequent antigenic strain of a pathogen in a population may change in response to the selection pressure of immunisation (Achtman 2001), favouring new antigenic types which are able to evade the acquired immunity of the host (Bull 1994). Pathogens such as HIV, malaria and influenza have particularly high antigenic diversity (Buckee, Recker et al. 2011). A vaccine must either induce cross-immunity to all antigenic strains of a pathogen circulating within a population, or different vaccines may be required according to which antigenic strain is predominant in a given situation. Different strains of a pathogen may also vary in terms of the magnitude or type of immune response invoked (Wedlock, Denis et al. 2007). In the case of human seasonal influenza, it has been suggested that it is the changing immune response within the host population which creates the conditions for the emergence of the next dominant strain (Recker, Pybus et al. 2007). Molecular typing approaches offer powerful tools for furthering our understanding of antigenic diversity in wildlife populations and the role of vaccination in disease control.

2.11 Identifying reservoirs

Another key challenge for wildlife managers is identifying populations that may act as reservoirs of infection for livestock, humans or other wildlife of conservation or economic importance. Assessment of the risks of onward spread requires a clear understanding of transmission dynamics within and amongst the species concerned (Hudson, Rizzoli et al. 2002). Disease reservoirs can potentially be composed of one or more epidemiologically connected populations or environments where the pathogen can be

permanently maintained and genetic pathogen characterisation can be a powerful tool for identifying such reservoirs (Haydon, Cleaveland et al. 2002). Molecular techniques may be of value in inferring transmission routes amongst multiple host species, although confusion can arise if pathogens are capable of remaining infectious in the environment. Inter-specific transmission has been inferred through strain comparison of *Giardia* (Feng and Xiao 2011) and *Cryptosporidium* in human and animal hosts (Xiao and Ryan 2004). Transmission between wildlife and commercially important livestock can also be inferred through comparing pathogen genotypes, as demonstrated for bovine TB in cattle and badgers in the UK (Woodroffe, Donnelly et al. 2005, Biek, O'Hare et al. 2012, Goodchild, Watkins et al. 2012), bovine TB in cattle and wildlife in New Zealand (Crispell, Zadoks et al. 2017) and for *Babesia bovis* and *B. bigemia*, the bacteria responsible for Cattle Tick Fever in white-tailed deer in the USA (Holman, Carroll et al. 2011). Recently, whole genome sequencing was used to demonstrate multiple instances of cross species transmission of *Brucella abortus* infection between reservoir hosts (Kamath, Foster et al. 2016).

Molecular typing techniques provide valuable insights into multi-host systems when considering populations of conservation concern. Examples include hookworm and Feline Leukaemia Virus transmission from domestic cats *(Felis catus)* to the critically endangered Iberian lynx *(Lynx pardinus)* (Meli, Cattori et al. 2009, Millan and Blasco-Costa 2012) and transmission of canine parvovirus and rabies from domestic dogs *(Canis lupus familiaris)* to endangered African wild dogs *(Lycaon pictus)* (Woodroffe, Prager et al. 2012). Strain typing of pathogens can also indicate the presence of an undetected wildlife reservoir, or even multiple reservoirs, where strain diversity appears too high to have been generated by mutation alone. However, in order to make such assessments, a sufficient number of samples should ideally be available from all host species in the system under study, and any host-related variation in pathogen mutation rates should be known (Kao, Haydon et al. 2014).

2.12 Management strategies

One of the greatest challenges faced by wildlife disease managers is unpredictability in the outcome of interventions. This is in part due to the fundamental challenges of working with free-ranging wildlife, but is exacerbated

by a lack of reporting when unintended outcomes occur, which has limited the degree to which we can learn from past interventions (Lloyd-Smith, Cross et al. 2005). Coupling genetic information from hosts and pathogens, with ecological factors can help to predict patterns of disease emergence, spread and control (Biek and Real 2010). Employing molecular approaches can help managers to monitor the epidemiological impacts of interventions with a potentially high degree of resolution and hence allow a more informed approach to refining management actions. Where a novel or re-emerging pathogen appears and wildlife populations are implicated, either as the reservoir or target of disease, managers may be called on to advise on potential interventions. Rapid molecular typing can quickly reveal a wealth of information about a disease outbreak and help to identify true transmission events, trace individual contacts and identify the true source of a particular pathogen. In the field of public health, molecular strain typing has been used to trace the source for outbreaks of a wide range of pathogens including *E.coli* (Grad, Lipsitch et al. 2012), TB (Gardy, Johnston et al. 2011), pneumonia (Snitkin, Zelazny et al. 2012) and even deliberately introduced pathogens associated with bioterrorism (Rasko, Worsham et al. 2011). Molecular epidemiological investigations during an outbreak can also suggest the existence of undetected carriers through using pathogen phylogenies in association with social network analysis, as conducted in investigations of human TB outbreaks (Walker et al. 2013) and can help identify super-spreading individuals who make a disproportionately large contribution to secondary infections (Woolhouse et al. 1997). Through the comparison of multiple isolates from the same host individual over time, pathogen micro-evolution can be examined (Gardy et al. 2011). Understanding the rate at which a pathogen can acquire mutations has important implications for choosing appropriate diagnostic tests, predicting the emergence of new strains and informing potential intervention strategies, such as vaccination. Comparison of pathogen strains prior to and during an outbreak can indicate whether the epidemic is due to a genetic change in the pathogen or rather to some social or environmental trigger (Gardy et al. 2011). The genetic diversity amongst isolates associated with a particular disease outbreak can also provide information about the size of the initial infection; limited diversity among isolates may indicate a population bottleneck has occurred, suggesting the outbreak could have been caused by few initially infected individuals (Grad et al. 2012).

However, this requires pre-existing knowledge regarding what level of diversity is typical for that pathogen, which highlights the importance of ongoing disease surveillance.

Molecular epidemiological investigations have been carried out on a wide range of disease outbreaks in livestock, including Newcastle disease in poultry (Gould et al. 2001), FMD in cattle (Cottam et al. 2006) and bluetongue virus in sheep (Maan et al. 2004), (Barros et al. 2007). Of these, only the investigation of FMD employed complete genome sequencing. Examples of molecular epidemiological investigations in wildlife include studies on outbreaks of phocine distemper in seals on the Danish coast (Line Nielsen et al. 2009), salmonella in passerines (Hernandez et al. 2012), viruses of anthropogenic origin in protected ape populations (Köndgen et al. 2008) and the source of DFTD in Tasmanian devils (Murchison et al. 2012).

A major wildlife disease outbreak which represents a real threat to global biodiversity is the recent emergence of amphibian chytridiomycosis, caused by the fungus *Batrachochytrium dendrobatidis*. This pathogen has been isolated from all continents where its hosts are found (Fisher et al. 2009) and is thought to be the principal cause of decline in over 200 species of amphibian. There is substantial variation in observed host responses to infection with some species appearing to be resistant whilst others succumb quickly to its lethal effects, and virulence has been found to vary amongst strains (Blaustein et al. 2005). The full genomes of two geographically diverse chytrid isolates were sequenced and used to identify areas of variation within the genome. Low sequence diversity was observed between the two isolates, but genomic areas with some variation were targeted by multi-locus sequence typing of a global set of chytrid isolates which were used to create a phylogenetic tree, illustrating the geographic origin of each isolate and its host species (James et al. 2009). From examining the tree, it was apparent that all the isolates could feasibly have originated from a single clonal lineage as, in one host animal, the same chytrid sequence diversity existed as was found in the whole global sample (James et al. 2009). This molecular signature is consistent with the rapid spread of a novel pathogen, and hence movement of animals for trade purposes has been suggested as a potential explanation for its current global distribution (Fisher et al. 2009).

Combining molecular epidemiological approaches, in particular high resolution sequencing, with traditional epidemiological techniques may be a powerful approach in disease outbreak investigation. This is made even more powerful where data is also available from background pathogen surveillance.

2.13 Discussion

The value of molecular epidemiology in the study of human disease is well-established. We now have at least one complete genomic sequence for nearly all bacteria responsible for human disease. An extraordinary amount of genetic diversity has been uncovered, including variation from within apparently clonal cultures (Medini, Serruto et al. 2008). Phylogenetic tools can be applied to genetic sequence data within open source packages such as BEAST (Drummond, Suchard et al. 2012) providing powerful insights into pathogen spread within host populations. Also, pathogen sequence data can improve the performance of disease transmission models by reducing the number of candidate transmission trees (Kao, Haydon et al. 2014), as pathogen genetic data is integrated with epidemiological data to inform transmission model construction within a Bayesian framework (Jombart, Cori et al. 2014). Despite the particular challenges involved in applying molecular technologies in the field of wildlife disease, the emergence of a number of high profile zoonotic diseases and dramatic declines in the abundance of some wildlife populations in recent years have raised awareness of this area of study (Daszak, Tabor et al. 2004) and the application of cutting edge molecular tools is increasing.

The application of molecular technologies poses significant challenges even when used in tandem with traditional epidemiological approaches. Techniques such as whole genome sequencing produce huge amounts of data which can be expensive to store and computationally costly to handle. However the availability of online ‘cloud’ storage provides a potential solution (Baker 2010) and as uptake of these technologies increases we can expect further developments in data storage and handling capabilities. DNA amplification required for next generation sequencing can introduce sequencing errors and a lack of standardised quality control procedures between laboratories can add uncertainty to sequence data (Kao, Haydon et al. 2014). Particular challenges for the application of molecular approaches in wildlife populations include the

presence of multiple hosts and the possibility of environmental persistence of the pathogen. Such circumstances mean that even whole genome sequencing cannot pick out individual transmission pathways as there will often be multiple routes by which the same pattern of genotypes could have arisen.

Future developments in molecular technology could have exciting applications in the field of wildlife disease. Rapid, field sequencing of isolates from populations and their environment (for example using hand-held sequencers) could allow a 'forensic' approach to investigating disease outbreaks, in which localised management might be tailored to the particular source of infection. This could be useful for example in the case of bTB in UK cattle, where the source of infection is likely to vary widely between herds and geographic locations. In the field of human health, interest is growing in 'precision medicine' whereby the entire human genome of a patient is sequenced and a tailored health plan produced based on the patient's particular genetic composition. It is conceivable that human genome sequencing will become a routine procedure at birth, allowing the development of 'personalised programmes of lifelong health promotion' (Tonellato, Crawford et al. 2011). As we have seen by considering a variety of examples above, technological advances first developed in the field of human health are subsequently employed in livestock and wildlife. It is plausible therefore, that as sequencing costs fall, individual level, genome tailored approaches may become attractive for the management of disease in wildlife species of very high conservation value. The management of DFTD in individual Tasmanian devils might be a case in point (see Table 2.1).

Developments in the field of metagenomics, in which multiple microbe genomes could be sequenced directly from environmental samples (Doolittle and Zhaxybayeva 2010) may provide valuable tools for wildlife disease management and research. Such an approach could be used to screen for pathogens prior to translocation of threatened species or for clarifying transmission routes where a pathogen can persist in the environment. The latter would for example be useful in studies of bTB transmission amongst wildlife and livestock, where there is a potential role for environmental contamination with *M. bovis* (Duffield and Young 1985).

In summary molecular technologies allow us to consider pathogens at a wide range of spatial and temporal scales; from individual host-pathogen dynamics, to global patterns of strain diversity. Following their emergence in the field of human health they have begun to be adopted for the purposes of investigation and management of disease in wildlife. At the present time these tools have a range of applications in wildlife disease research from the local investigation of disease outbreaks to unearthing the evolutionary history and global spread of pathogens. The potential future contribution of these technologies to the field of wildlife disease epidemiology is substantial. In particular they are likely to play an increasingly important role in helping us to address a principal challenge in the management of wildlife diseases which is how to tease apart the transmission dynamics of complex multi-host systems in order to develop effective and sustainable interventions.

Table 2.1 Summary of applications of molecular epidemiology to wildlife disease research, including key examples

Application	Summary	Examples of use in wildlife
Disease Surveillance	Ongoing surveillance of circulating pathogen strains can help to capture the intrinsic genetic diversity of the pathogen, informing the use and development of appropriate diagnostics	Avian influenza viruses in wild birds (Hoye et al. 2010) Bovine tuberculosis in wildlife (Romero et al. 2008)
Phylogeography	Spatio-temporal mapping of variant pathogen strains can highlight the appearance of atypical strains in an area, identify geographic differences in pathogen virulence and help to infer risk factors	Devil facial tumour disease (Hamede et al. 2012) Rabies virus (Streicker et al. 2012) Foot & Mouth Disease (de Carvalho et al. 2013)
Roots of Emergence	The construction of phylogenetic trees from genotype data can indicate which pathogen strains are ancestral and which are descendant. This can be useful in dating disease transmission between populations and multiple host species.	*Leucocytozoon* blood parasites in wild birds (Yoshimura et al. 2014) White Nose Disease in bats (Chibucos et al. 2013) Orbiviruses (Maan et al. 2013)
Routes of Transmission	Pathogen genotyping can help to rule out or implicate particular transmission pathways, which may be valuable in tracing the initial source of infection and preventing further disease spread	Bovine TB in badgers and cattle (Biek et al. 2012) Brucellosis (Foster et al. 2009)
Host-Pathogen Dynamics	Molecular techniques can allow us to examine individual variation in susceptibility and resistance within a host population by assessing the genetic basis of the immune response	Devil facial tumour disease (Siddle et al. 2010) MHC differentiation in Namibian cheetahs (Castro-Prieto et al. 2012) Bovine TB resistance in wild boar (Acevedo-Whitehouse et al. 2005)
Vaccine Development and Monitoring	Molecular epidemiology has a key role to play in the development of effective vaccines for wildlife and monitoring their impacts on disease epidemiology	Reversion to virulence of rabies vaccine strain (Hostnik et al. 2014) Oral rabies vaccine strains in wildlife (Geue et al. 2008)
Identifying Reservoirs	Molecular techniques may be of value in inferring transmission routes amongst multiple host species. Strain typing of pathogens can also indicate the presence of an undetected wildlife reservoir, or even multiple reservoirs, where strain diversity appears too high to have been generated by mutation alone	Rabies in African wild dogs (Woodroffe et al. 2012) Feline Leukaemia Virus in Iberian lynx (Millan and Blasco-Costa 2012), (Meli et al. 2009) Cattle Tick Fever (Holman et al. 2011)
Management Strategies	Employing molecular approaches can help managers to identify the sources of disease outbreaks and to monitor the epidemiological impacts of interventions with a potentially high degree of resolution	Phocine Distemper in seals (Line Nielsen et al. 2009) Salmonella in passerines (Hernandez et al. 2012) Viruses in protected ape populations (Köndgen et al. 2008)

CHAPTER 3: Data Chapter

Blood thicker than water: Kinship, disease prevalence and group size drive divergent patterns of infection risk in a social mammal

3.1 Abstract

The importance of social- and kin-structuring of populations for the transmission of wildlife disease is widely assumed but poorly described. Social structure can help dilute risks of transmission for group members, and is relatively easy to measure, but kin-association represents a further level of population sub-structure that is harder to measure, particularly when association behaviours happen underground. Here, using epidemiological and molecular genetic data from a wild, high-density population of the European badger (*Meles meles*), I quantify the risks of infection with *Mycobacterium bovis* (the causative agent of tuberculosis; TB*)* in cubs. The risk declines with increasing size of its social group, but this net dilution effect conceals divergent patterns of infection risk. Cubs only enjoy reduced risk when social groups have a higher proportion of test negative individuals. Cubs suffer higher infection risk in social groups containing resident infectious adults, and these risks are exaggerated when cubs and infectious adults are closely related. I further identify key differences in infection risk associated with resident infectious males and females. I link my results to parent-offspring interactions and other kin-biased association, but also consider the possibility that susceptibility to infection is heritable. These patterns of infection risk help to explain the observation of a herd-immunity effect in badgers following low-intensity vaccination campaigns. They also reveal kinship and kin-association to be important, and often hidden, drivers of disease transmission in social mammals.

3.2 Introduction

Understanding disease transmission within wildlife populations has important applications in the fields of emerging zoonotic diseases (Begon, Hazel et al. 1999, Jones, Patel et al. 2008), biodiversity conservation (Daszak, Cunningham et al. 2000) and livestock health (Gortázar, Ferroglio et al. 2007). Increasingly, the importance of behavioural heterogeneity and social structure on disease transmission between individuals is being recognised, with these individual level differences scaling up to determine infection dynamics at the population scale (Tompkins, Dunn et al. 2011). However, in wild populations, capturing these individual behavioural differences and quantifying the resultant effects on disease transmission is challenging, particularly when behavioural associations happen out of sight, e.g. in underground setts.

Heterogeneity in individual transmission rates, defined as variability in the contribution of individual hosts to overall rates of pathogen spread (Paull, Song et al. 2011), is a key driver of disease dynamics. Several studies have demonstrated a relationship between individual contact dynamics and the transmission of infectious diseases (see review in Tompkins, Dunn et al. 2011)). For example, an individual's position within a socially structured population may influence the likelihood of it becoming infected (Böhm, Hutchings et al. 2009) as demonstrated in social animals such as meerkats (Drewe 2010). Certain "super-spreader" individuals within a population may contribute to a disproportionate number of secondary infections (Lloyd-Smith, Schreiber et al. 2005), due to a particular behavioural or biological trait or their position within a social network. Kin-biased social behaviours have been demonstrated in a variety of species (Möller, Beheregaray et al. 2006, Perry, Manson et al. 2008). These can include a wide range of behaviours, such as parental care of young, mutual grooming (Schino 2001), foraging (Brown and Brown 1996), and helping to raise young in the case of co-operative breeders (Russell and Hatchwell 2001). This enhanced contact between related individuals is likely to have important implications for disease transmission, as these kin-biased social behaviours afford potential opportunities for pathogen transfer. Generally kin structure, defined as the spatial aggregation of related individuals (Hatchwell 2010), is proposed to increase individual disease transmission risk in directly transmitted pathogens (Dharmarajan, Beasley et al. 2012), because

transmission rates are expected to be higher between related individuals than between non-related individuals (e.g. Canine Distemper Virus in raccoon populations (Dharmarajan, Beasley et al. 2012) and Chronic Wasting Disease in white-tailed deer (Grear, Samuel et al. 2010)). A greater understanding of heterogeneities in individual disease risk could help to inform management interventions and improve the estimation of parameters in epidemiological models to facilitate more ecologically realistic simulations and predictions (McDonald, Bailey et al. 2016).

Bovine tuberculosis remains a critical issue in livestock farming in several parts of the world, including the UK. The European badger *(Meles meles)* is the key wildlife reservoir of bovine TB *(*caused by *Mycobacterium bovis)* in the UK and, as such, has been subjected to a range of control interventions including culling and vaccination, with the aim of reducing disease transmission to cattle populations. However, it is well documented that the social structure typical of moderate to high density, managed and unmanaged badger populations can have a marked impact on the persistence and transmission of TB (Delahay, Langton et al. 2000, Carter, Delahay et al. 2007). As badgers live in social groups within defended territories, this can limit population mixing, such that members of different social groups are less likely to come into close contact than members of the same social group. This heterogeneity in contact behaviour is thought to drive the clustered distribution of *M. bovis* infection in badger populations (Delahay, Langton et al. 2000, Woodroffe, Donnelly et al. 2005). This relationship between population structure and TB dynamics has been implicated in the unexpected outcomes of management interventions to control TB in badgers and cattle, such that reductions in badger density do not result in proportional reductions in TB transmission (Woodroffe, Donnelly et al. 2006, Carter, Delahay et al. 2007, Jenkins, Woodroffe et al. 2007, Riordan, Delahay et al. 2011, Bielby, Donnelly et al. 2014). If social structure limits the spread of TB in badger populations, resulting in a naturally aggregated distribution of infection, then disruption of this social structure may carry with it the possibility of enhanced transmission (Carter, Delahay et al. 2007). Social network analysis has suggested that infected badgers occupy a social position within badger populations such that they facilitate transmission of infection between social groups (Weber, Carter et al. 2013).

Within a socially structured population, it is expected that mixing occurs at two scales: 'local' mixing, involving high levels of contact between members of the same social group; and 'global' mixing, involving occasional mixing with individuals outside the social group (Ball, Mollison et al. 1997, Böhm, Hutchings et al. 2009). In badgers, local mixing is likely to increase the risk of infection amongst cubs born into social groups harbouring infected adults. Furthermore, within the social group a kinship structure will exist, perhaps yielding heterogeneity in contact rates at a finer scale among group members. 'Pseudo-vertical transmission' whereby disease transmission occurs via lactation of offspring by infected dams or via the prolonged and repeated periods of close social contact both pre- and post-emergence from the underground sett environment, has been suggested to play an important role in the maintenance of *M. bovis* infection within badger social groups (Anderson and Trewhella 1985, Cheeseman, Wilesmith et al. 1988). The importance of the social group environment on early life infection risk in badger cubs has been supported by field trials using the now licenced injectable BadgerBCG vaccine (Chambers, Rogers et al. 2010): the risk of TB infection in unvaccinated badger cubs decreased significantly as the proportion of vaccinated individuals in their social group increased (Carter, Chambers et al. 2012). Other studies have shown that the presence of infectious females (i.e. those detected as excreting *Mycobacterium bovis*) within a social group helps to predict the incidence of infection in cubs (Delahay, Langton et al. 2000, Tomlinson, Chambers et al. 2013) consistent with pseudo-vertical transmission. However, no study to date has considered the impact of kin structure within badger social groups on individual infection risk to cubs.

Here I determine the impact of kinship and infection prevalence in social groups on the infection risk to young badgers present in the social group. I incorporate individual genotype data to account for kin structure within badger social groups, and TB diagnostic tests of adults and cubs to determine infection prevalence and transmission. I predict that cubs born into social groups where resident excretor badgers are present will be at higher risk of testing positive to TB in their first year than cubs born into social groups where excretor badgers are not present, but further that this effect will be greater when resident excretors are related to the cub.

3.3 Methods

All data used in these analyses were collected from the long-term capture-mark-recapture study at Woodchester Park in Gloucestershire. Badgers from this study population have been routinely trapped, up to four times a year, since 1976 (Delahay, Walker et al. 2013). Trapped badgers are brought back to a sampling facility, anaesthetised (for full details see Appendix) and a range of clinical samples taken (oesophageal and tracheal aspirates, faeces, urine, swabs of bite wounds or abscesses) for the detection of *M. bovis* by microbiological culture (Gallagher and Horwill 1977). Blood samples are collected and used for TB diagnostic testing. Diagnosis of infection is made at the individual level, with no reference to other social group members. The use of multiple diagnostic tests to determine disease status in this study helps to address the shortcomings in sensitivity of the tests when used in isolation (Drewe, Tomlinson et al. 2010). Between 1990 and 2005, the Brock ELISA antibody test (Goodger, Nolan et al. 1994) and the culture of clinical samples were the diagnostic tests used to assign TB status to individual badgers. From 2006 onwards, the Brock ELISA was replaced with the improved Stat-Pak antibody test (Chambers, Crawshaw et al. 2008) and the gamma interferon (IFN) test for T-cell responses to *M. bovis* was introduced (Dalley, Davé et al. 2008). The combination of diagnostic tests used provides a biologically meaningful picture of the progression of disease within an individual (Tomlinson 2013). It is thought that the cell-mediated response (as measured by the gamma interferon test) is the first line response to *M. bovis* exposure, whereas the serological response (as measured by the ELISA test and StatPak) takes time to develop as infection progresses (Tomlinson 2013). Some individuals then go on to become 'infectious', characterised by the excretion of *M. bovis* bacteria via various routes (Pritchard, Stuart et al. 1986, Graham, Smith et al. 2013). Due to these changes in diagnostic test use, study period was included as a co-variate in these analyses, with Study Period 1 identifying data from 1990-2005 and Study Period 2 identifying data from 2006-2011. Culture from clinical samples is the only diagnostic approach that has been used throughout the entire study period.

Selection of cubs:

In order to select a cohort of cubs for this analysis, first year data from the wider population study was selected for badgers first caught as cubs in the population between 1990-2011, yielding 1413 cubs for which genotype data were available. A cub which received a positive test result from any of the diagnostic tests used in their first year was assigned the status 'test-positive' whereas a cub with only negative test results in their first year was assigned the status 'test-negative'. Cubs were assigned to their assumed birth social group, based on the identity of the group in which they were first trapped. R software (v 3.0.2) (R Core Development Team 2013) was used to associate these cubs with data (disease status in that year, and sex) of adult badgers (≥1 year old) trapped in the same social group in the same year. Many individuals were trapped more than once during a calendar year, but each was assigned to just one social group following established assignment rules (Vicente, Delahay et al. 2007). For each cub, the following metrics were collated: number of resident female 'excretors' (females from whom at least one *M. bovis* positive culture had been isolated from a clinical sample from a prior trapping event, divided into 'relatives' and 'non-relatives' – defined below); number of resident male 'excretors' (males from whom at least one *M. bovis* positive culture had been isolated from a clinical sample from a prior trapping event, divided into 'relatives' and 'non-relatives'); number of resident 'blood test positive' females (females who had at least one positive result to a TB blood test (ELISA, StatPak or gamma interferon) from a prior trapping event, divided into 'relatives' and 'non-relatives'); and number of resident 'blood test positive' males (males who had at least one positive result to a TB blood test (ELISA, StatPak or gamma interferon) from a prior trapping event, divided into 'relatives' and 'non-relatives')

Genotyping:

On first capture of an individual, a hair sample was routinely taken and stored in 80% ethanol before being submitted for DNA extraction and genotyping (Carpenter, Pope et al. 2005). Genotype data were available for animals trapped from 1990 until 2011 inclusive. 22 microsatellite markers were used, each with 4-7 alleles, to derive genotypes for 1413 cubs and 470 adults resident in their social group of birth.

Relatedness:

The MicroDrop Programme (Wang and Rosenberg 2012) was used to impute missing data in the microsatellite data set. Deviations from Hardy-Weinberg equilibrium for each of the 22 microsatellite markers were tested on the MicroDrop-corrected dataset using the hwtest function in the 'adegenet' package (Jombart 2008); none was identified. The Bartlett test of homogeneity in the same package was also used to confirm homogeneity of variance among loci ($P = 0.78$). Data from all microsatellite markers were therefore used to calculate a relatedness matrix. Relatedness between cubs and resident adult members of their birth social group was estimated using the R package 'Demerelate' (v 0.8-1) (Kraemer, Gerlach et al. 2013). Bootstrap iterations were set to 100. Relatedness was calculated using the Queller and Goodnight r_{xy} relatedness estimator (Queller and Goodnight 1989). This provides an unbiased estimate of relatedness based on the population allele frequencies, and ranges from -1 to 1 with negative and positive values indicating lower- and greater-than-average relatedness, respectively (Grear, Samuel et al. 2010). A negative relatedness value indicates that a pair of individuals had a relatedness coefficient lower than the average pairwise relatedness coefficient calculated from the whole genotyped population. The mean relatedness estimate for the Woodchester Park population was close to 0, meeting the assumptions of the Queller and Goodnight estimator. Pairs of cubs and resident adults where the relatedness coefficient was ≥ 0.25 were assigned the status 'relatives' as 0.25 is the relatedness coefficient between half-sibs (Hedrick 2011). Where the relatedness coefficient was < 0.25, pairs were assigned the status 'non-relatives'. Potential misclassification rates were estimated based on previous simulations (Blouin, Parsons et al. 1996) which considered the number of loci used (22) and the average heterozygosity of these loci (0.68). In the current dataset it was estimated that 4% of pairs of unrelated individuals may be misclassified as full sibling pairs (full siblings should have an expected relatedness value of 0.5), and 17% of pairs of unrelated individuals may be misclassified as half-sibling pairs. The ability to distinguish between full siblings and unrelated individuals was therefore high (96%) and half siblings were correctly distinguished from unrelated individuals more than 80% of the time (Blouin, Parsons et al. 1996).

Modelling individual infection risk:

All statistical analyses in this chapter were conducted in R version 3.0.2 (R Core Development Team 2013). In order to investigate factors relating to cub infection risk (at a variety of complexities / scales) three distinct analyses were carried out, all consisting of generalised linear mixed models constructed via the R package 'lme4' (v1.0-5) (Bates 2010). In all cases, social group identity and year were included as random effects. Cub infection status was categorised as a binary response variable, with '1' indicating that at least one positive diagnostic test result had been recorded for that individual in year one and '0' indicating only negative test results being recorded. All analyses were performed with individual cub as the sampling unit. Cubs from social groups where genotype data from less than 3 adults in the group were available were excluded from the analysis, resulting in a dataset of 1362 cubs. Care was taken throughout the analysis and interpretation to avoid the term "infected" or "uninfected": issues of test sensitivity mean that some "test-negative" cubs are in fact infected. All the diagnostic tests employed have limitations in terms of their sensitivity and specificity, which have been explored in detail elsewhere (Drewe, Tomlinson et al. 2010, Buzdugan, Chambers et al. 2016). A recent multi-event model constructed for this population demonstrated that at a given sampling event, a truly uninfected individual has a 94% probability of testing negative to the gamma interferon and StatPak tests and on culture. Under these circumstances, the risk of wrongly classifying a truly uninfected badger as infected is low. We can therefore have some confidence that those individuals identified as 'test-positive' in the current study are not truly uninfected individuals. To our knowledge, the probability of false negative diagnosis is not influenced by phenotypic traits or social group structure, therefore it was assumed that false negativity affects all infected cubs equally.

In the first analysis, the effect of social group size on the risk of each cub testing positive to a diagnostic TB test in their first year was investigated, with social group size and study period included as fixed effects. Wald's chi squared tests were used to assess significance of fixed effects.

To investigate effects of social group composition on the risk of cubs testing positive to a diagnostic TB test in their first year, cub infection status was

regressed against the number of individuals test-positive to any of the diagnostic tests in the social group and the number of individuals test-negative to all of the diagnostic tests in the social group (as fixed effects), along with study period. Wald's chi squared tests were used to assess significance of fixed effects.

Finally, the effects of social group composition and relatedness structure on the risk of cubs testing positive to a diagnostic TB test in their first year were teased apart, using multi-model inference with model averaging. A global mixed effects model included the following fixed effects: the number of resident female excretors (divided into 'relatives' and 'non-relatives' of each cub); the number of resident male excretors (divided into 'relatives' and 'non-relatives'); the number of resident blood test positive females (divided into 'relatives' and 'non-relatives'); the number of resident blood test positive males (divided into 'relatives' and 'non-relatives'); and the number of test negative group members. Small sample sizes of excretor adults prevented us from using analyses that considered relatedness as a continuous variable (Grear, Samuel et al. 2010). Model averaging was carried out using the 'MuMIn' package (v 1.9.13) (Barton 2011) on the model set generated from the global model, applying a threshold delta-AICc value of 6 units, as this is the threshold suggested to be 95% sure that the most parsimonious model is included in the top model set (Richards 2005, Richards, Whittingham et al. 2011). Parameter estimates and their confidence intervals were extracted from the top model set identified through the model averaging procedure. Concerns regarding possible collinearity of the explanatory variables were addressed using correlation testing among all fixed effects in the global model; the mean correlation was 0.06 and the strongest correlation was only 0.36. The explanatory variables did not suffer Variance Inflation Factors greater than 10 and single term regression models produced parameter estimates that resembled the results of model averaging in terms of sign, size and significance (O'brien 2007).

To investigate alternative model structures two additional models were constructed for comparison with the global model described above. Firstly, to test whether test-negative badgers were differentially affecting cub infection risk a fully complex model in which test-negative badgers were disaggregated by relatedness and sex was constructed. Secondly, to test whether sex was

adding any information to the model, a model was constructed in which effects were collapsed across sexes throughout the model (i.e. grouping together related culture positive males and females etc.). Both of these models had higher AIC values than the global model described above (fully complex model, delta AIC=15, sex removed model, delta AIC =7), thus supporting the selection of the original model structure proposed for further study.

3.4 Results

Of the 1362 cubs included in this analysis, 184 yielded a positive test result within their first year (14%). Summary statistics of social group size and the number of adults in each relatedness – disease category are given in Table 3.1. In the first analysis, cub risk declined with increasing group size (Wald's $\chi^2_{(1)} = 6.0$, $p = 0.01$), indicating that cubs born into larger social groups were at a lower risk of yielding a positive test result in their first year (Fig 3.1a). Study period did not influence the risk of cubs testing positive (Wald's $\chi^2_{(1)} = 2.6$, $p = 0.11$). In the second analysis, where group size was elaborated into the number of 'test negative' and 'test-positive' individuals present in the cub's natal social group, cub risk increased with increasing numbers of test positive individuals (Wald's $\chi^2_{(1)} = 35.4$, $p < 0.01$, Fig 3.1b) but declined with increasing numbers of test negative individuals (Wald's $\chi^2_{(1)} = 38.0$, $p < 0.01$, Fig 3.1c).

Table 3.1 Summary statistics of social group sizes and numbers of adults in disease-relatedness categories.

Category	**Mean Value (SD)**	**Median Value**	**Range**
Social Group Size	11 (5.3)	10	3-31
No. of cubs	3 (2.2)	3	1-12
Proportion of test positive cubs	0.15 (0.3)	0	0-1
		Proportion of cubs with at least 1 positive in category	
Related Residents			
Culture positive ♀	0.04 (0.2)	0.04	0-2
Culture positive ♂	0.02 (0.1)	0.02	0-2
Blood test positive ♀	0.13 (0.4)	0.1	0-4
Blood test positive ♂	0.09 (0.3)	0.07	0-3
Test negative ♀	0.7 (1.1)	0.4	0-8
Test negative ♂	0.5 (0.9)	0.4	0-6
Unrelated Residents			
Culture positive ♀	0.31 (0.6)	0.2	0-4
Culture positive ♂	0.17 (0.4)	0.15	0-3
Blood test positive ♀	0.75 (1.2)	0.4	0-6
Blood test positive ♂	0.44 (0.7)	0.3	0-4
Test negative ♀	4.6 (3.4)	0.9	0-18
Test negative ♂	3.2 (2.3)	0.9	0-11

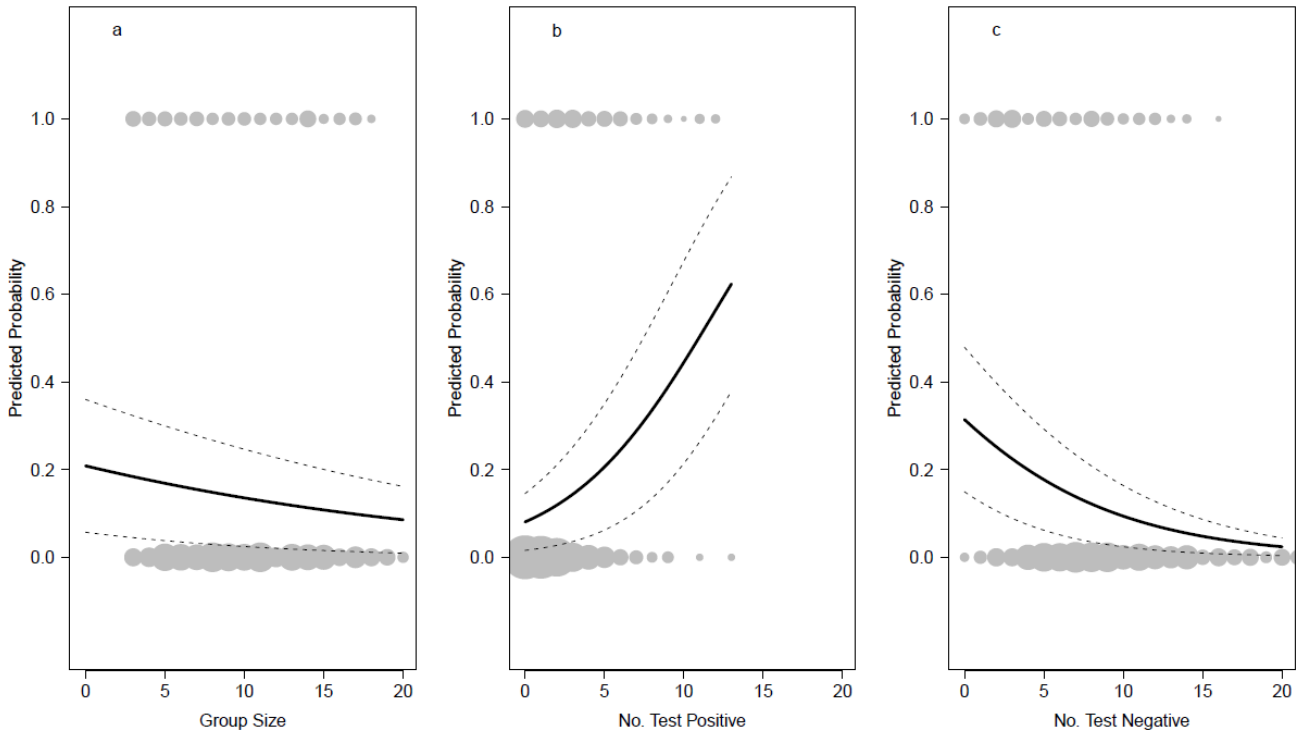


Fig 3.1 (a) Net dilution of the risk of badger cubs testing positive to tests for bovine tuberculosis with increasing social group size. (b) Increased risk of cubs testing positive within their first year with increasing number of test-positive individuals resident in their social group. (c) Reduced risk of cubs testing positive in their first year with increasing number of test-negative individuals resident in their social group. Bold lines indicate line of best fit, dashed lines indicate 95% confidence intervals. Circles summarise the raw data, with the size of symbol proportionally scaled to the number of individuals in that category (smallest point indicates 1 data point, largest point indicates 373 data points).

In the final analysis, where test positive badgers were broken down into the categories described above, model averaging indicated that several variables were important predictors of cub infection risk (Table 3.2). The risk of a cub becoming test-positive in its first year increased most markedly with changes in the number of related excretors of both sexes (Figs 3.2 & 3.3). The presence of one related male excretor in their birth social group increases the predicted probability of that cub testing positive within their first year by 26%, whereas the presence of a related female excretor increases the probability by 15%. Much lower risks are associated with unrelated male or female excretors (6% and 4% respectively; barely credibly different from zero (Fig 3.2 & 3.3)). The probability of test positivity in cubs increased in the presence of blood-test positive female relatives in the social group (Fig 3.4, presence of one blood test positive female increases risk by 4%), but was not influenced by blood-test positive male relatives, nor by blood-test positive unrelated individuals of either sex (Fig 3.2).

Table 3.2 Details of top 10 models with a ΔAICc<6 predicting the odds of cubs testing TB positive in their first year (1). Social group and Study Year were included as random effects. 'Blood test positive' indicates individuals who yielded a positive test result to either the ELISA, StatPak or gamma interferon diagnostic test. Each row in the table indicates a model, with a + indicating the inclusion of a given variable within each model. Degrees of freedom, ΔAICc, model weight and R^2 values are also included for each model. Marginal R^2 (R^2M) represents the variance explained by fixed factors and Conditional R^2 (R^2C) represents the variance explained by both fixed and random factors (Barton 2011).

Model	Related Excretor Males	Non-Related Excretor Males	Related Excretor Females	Non-Related Excretor Females	Related blood test positive Females	Non-Related blood test positive Females	Related blood test positive Males	Non-related blood test positive Males	Uninfected	Study Period	df	ΔAIC_C	ω_i	R^2C	R^2M
1	+	+	+	+	+	+		+	+	+	12	0.00	0.09	0.35	0.28
2	+	+	+	+	+	+		+	+		11	0.21	0.09	0.35	0.28
3	+	+	+	+	+			+	+		10	0.31	0.08	0.36	0.27
4	+	+	+	+	+			+	+	+	11	0.53	0.07	0.35	0.27
5	+	+	+	+	+	+			+		10	0.58	0.07	0.36	0.27
6	+	+	+	+	+	+			+	+	11	0.71	0.07	0.36	0.27
7	+	+	+	+	+	+	+	+	+	+	13	0.93	0.06	0.35	0.28
8	+	+	+	+	+	+	+	+	+		12	1.37	0.05	0.35	0.27
9	+	+	+	+	+	+	+		+	+	12	1.52	0.04	0.36	0.27
10	+	+	+	+	+	+	+		+		11	1.64	0.04	0.36	0.27

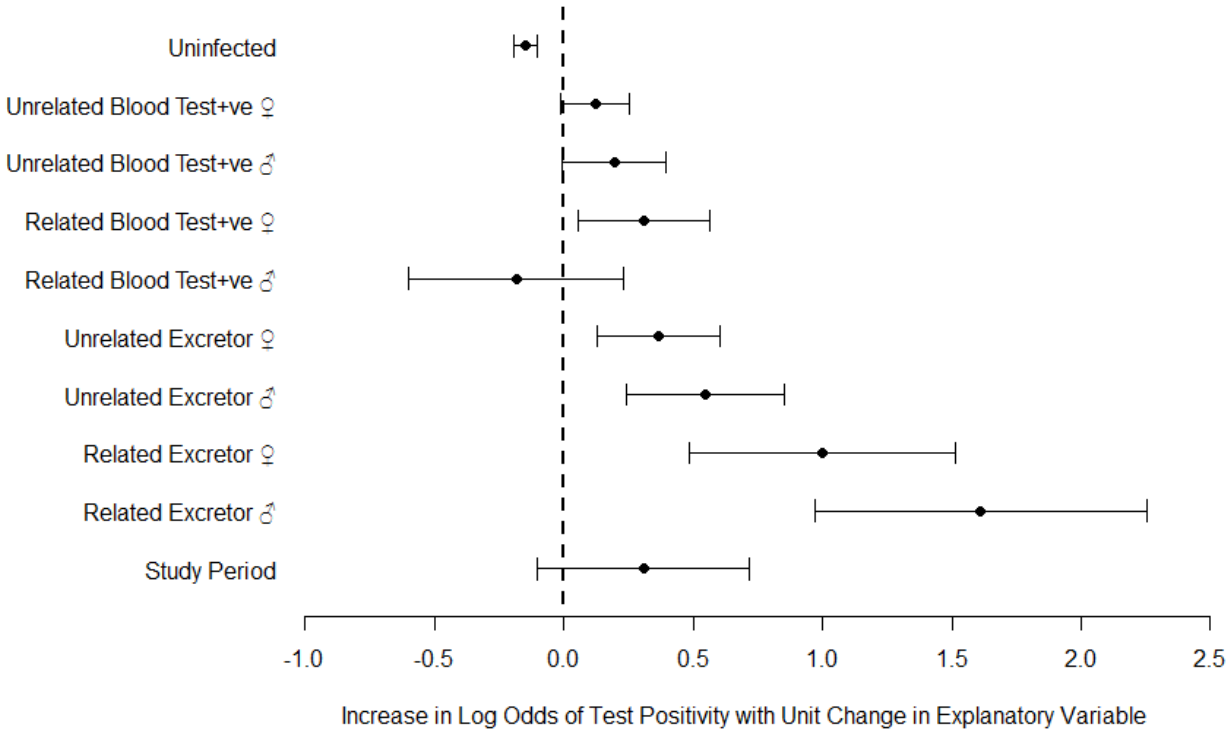


Fig 3.2 Factors affecting the risk of badger cubs testing positive for bovine tuberculosis in their first year (1990-2005). Average model coefficients (log odds) calculated for variables included in the top model set (Supplementary Table 1). Arrows indicate 95% confidence intervals. Model-averaged regression slopes are considered important if they are consistent and directional (i.e. their confidence intervals do not span zero).

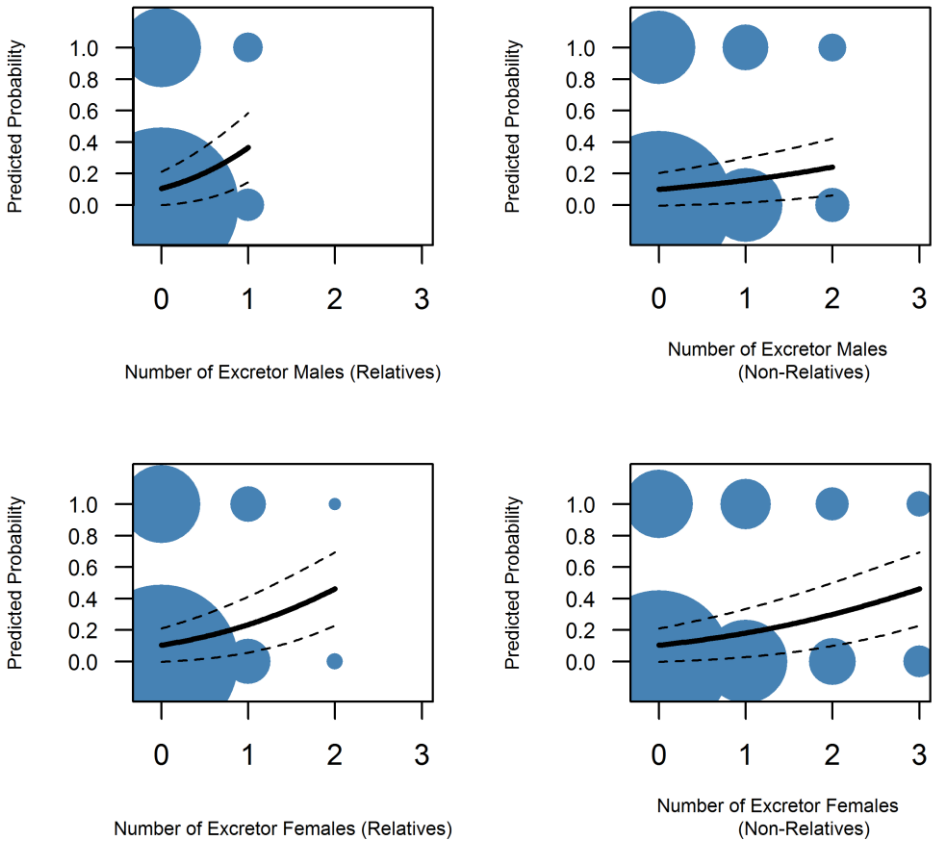


Fig 3.3 Predicted probability of a cub testing positive for TB in its first year with increasing numbers of excretor relatives and non-relatives resident in its social group. Bold lines indicate the line of best fit, dashed lines indicate 95% confidence intervals. Circles summarise raw data, with the size of symbol proportionally scaled to the number of individuals in the category (smallest point indicates 3 data points, largest point indicates 1336 data points).

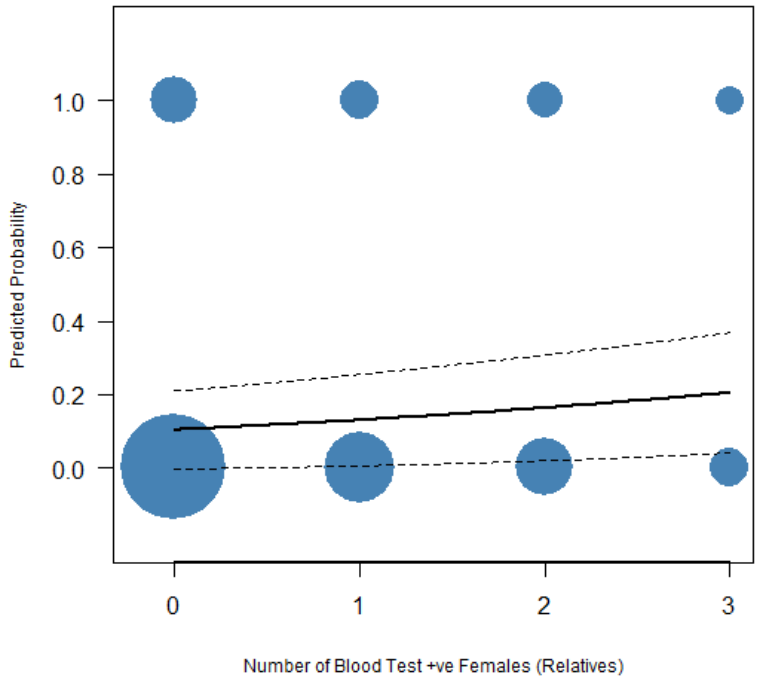


Fig 3.4 Predicted probability of a cub testing positive for TB in its first year with increasing numbers of blood test positive female relatives resident in its social group. Bold line indicates the line of best fit, dashed lines indicate 95% confidence intervals. Circles summarise raw data, with the size of symbol proportionally scaled to the number of individuals in the category (smallest point indicates 5 data points, largest point indicates 1277 data points).

3.5 Discussion

The findings of this study highlight the potential complexities of transmission dynamics within wild animal populations. When the influence of badger social group size on transmission risks was considered in isolation, I found that cubs born into larger groups were at a lower risk of yielding a positive test result in their first year, indicating net negative density dependence and therefore an important dilution effect on transmission. This is consistent with previous studies in which *M. bovis* prevalence was found to be consistently higher in small social groups (Woodroffe, Donnelly et al. 2009). When this simple measure of group size was decomposed according to the test history of resident badgers, the risk of a test-positive result in cubs was positively related to the number of test positive residents and was only diluted by test-negative residents. This is consistent with the herd immunity effect demonstrated in a vaccinated badger population, whereby the infection risk in unvaccinated badger cubs was reduced where more than a third of their birth social group was vaccinated (Carter, Chambers et al. 2012). The observation of divergent infection risks, associated with numbers of test positive versus test negative individuals, highlights the dangers of relying on population-level metrics (such as host density) to reveal transmission dynamics (Tompkins, Dunn et al. 2011), which in reality may be driven by processes operating at a finer scale.

Further complexity was revealed when social group composition was broken down into kin- and non-kin-structure. The number of related female badgers in a cub's natal social group that were excreting *M. bovis* bacteria was positively associated with the risk of that cub testing positive during its first year. This is consistent with infection risk driven by kin-biased association, i.e. closer, more prolonged or more regular contacts between cubs and female relatives than non-relative female group members. Previous studies within high density badger populations have indicated that females are more likely to be related to other individuals in their social group (Dugdale, Macdonald et al. 2008), perhaps because female badgers are less likely to permanently leave their natal group than males (Cheeseman, Cresswell et al. 1988, Rogers, Delahay et al. 1998). Therefore a cub may be born into a group where multiple female excretor

relatives are present, including their mother and sisters from previous years' litters. Cubs are born and suckled by their mothers during their first twelve weeks of life (Roper 2010) and may be particularly susceptible to infection in early life when their immune systems are under-developed, making them vulnerable to high infective doses of *M. bovis* from infectious excretor dams (Tomlinson, Chambers et al. 2013). Behavioural monitoring using radio collars shows that females, including younger and non-breeding females, use main setts more during this period than sub-adult and adult males (Weber, Bearhop et al. 2013). Cubs may therefore be exposed to infection, both from their mother and from other female badgers present in the main sett prior to emergence. Following emergence from the sett, which occurs at around 8 weeks of age, cubs only spend short periods of time above the ground (Roper 2010) and will remain closely associated with their mothers after emergence, until they are capable of independent foraging. Above ground, anecdotal evidence exists of non-breeding adult females babysitting (Woodroffe 1993) and allogrooming cubs, although these behaviours did not appear to be kin-biased (Dugdale, Ellwood et al. 2010). Overall the evidence for alloparental care in badgers is considered to be weak (Roper 2010). In addition to excretor females posing a risk to resident cubs, I also demonstrated that the number of female relatives in a social group who yielded a positive result to a serological or gamma interferon test was associated with a slight but significant increase in the risk of resident cubs testing positive in their first year. This was not the case for unrelated females or sero/ gamma interferon positive males. As expected, this risk was far lower than for cubs where related or non-related excretor females were resident, reflecting the particular epidemiological importance of infectious individuals in maintaining infection within the social group.

In contrast to previous work (Tomlinson, Chambers et al. 2013), the presence of excretor males in a cub's social group was a greater risk factor for cub infection risk than that of female excretors. This result is somewhat surprising, given our understanding of the greater intensity of cub-female behavioural interactions. Paternal care has not been documented in the European badger (Evans, Macdonald et al. 1989) and is not supported by observational studies (Dugdale, Ellwood et al. 2010). The primary route of bTB transmission between badgers is considered to be via the respiratory system, such that close and

prolonged contact between individuals in setts may facilitate transmission (Weber, Bearhop et al. 2013). Male badgers use more of the underground space than females (Kowalczyk, Zalewski et al. 2004), therefore excretor male badgers might be more responsible for contaminating the underground sett environment than female excretors. However this does not explain the difference in risk presented by related and non-related male badgers. Alternatively, opportunities for disease transmission might be due to above-ground contact as cubs become integrated into the social group following emergence. An observational study of cub social integration following emergence noted that as cubs matured they spent more time and engaged in play-fights more frequently with adult and sub-adult male group members (and less with female group members) (Fell, Buesching et al. 2006). We do not yet know whether these behaviours are kin-biased.

I have shown that the risk to cubs of acquiring infection depends on within-social-group structuring, particularly linked to kin and sex. The patterns observed are consistent with the 'herd immunity' effect in badger social groups, where the risk of TB infection in unvaccinated badger cubs decreased by nearly 80% when more than a third of the social group were vaccinated against TB (Carter, Chambers et al. 2012). Vaccinating a modest proportion of the adults in a badger social group may protect unvaccinated cubs indirectly by reducing their contact with infected adults. The results presented here suggest that kinship with vaccinated adults will provide cubs with even greater levels of protection.

There is a possible alternative explanation for the higher risk experienced by cubs that have a culture positive relative in their social group: susceptibility to bTB might be heritable. We know that cattle breeds show differential susceptibility to bTB infection (Vordermeier, Ameni et al. 2012), and that heritability of bTB resistance in farmed red deer (Mackintosh, Qureshi et al. 2000) and of bTB disease outcomes in cattle (le Roex, van Helden et al. 2013) can be high. No published work is currently available on genetic variation in bTB susceptibility in badgers and other wildlife hosts. As the full pedigree of the Woodchester Park badger population emerges in the near future, it will allow us to tease apart the influence of kin-biased behaviour and heritability in bTB transmission dynamics.

These findings have clear relevance for the understanding, modelling, prediction and management of disease in socially- and kin-structured host populations. Social structure can have major impacts on the success of strategies to manage or control disease prevalence and transmission (Woodroffe, Donnelly et al. 2006, Carter, Delahay et al. 2007, Riordan, Delahay et al. 2011) and the identification of kinship and disease prevalence as mediators of density-dependent transmission could provide important insights to disease management via targeted vaccination or removal campaigns (McCallum, Barlow et al. 2001). Kin structure is often hard to identify, and the behavioural interactions that drive direct transmission of disease are often hidden from observation, but their importance to patterns of disease transmission make the advent of molecular tools for wildlife disease all the more relevant (Benton, Delahay et al. 2014).

I have confirmed the epidemiological importance of infectious individuals in the maintenance and persistence of infection in groups of social mammals. I have demonstrated that kin structure causes within-group heterogeneities in infection risks for cubs, either through kin biased association favouring disease transmission, heritable susceptibility, or a combination of the two. Given that strategies for the management of disease in wild mammal populations can perturb social and kinship structures, these key drivers of disease transmission should be considered during the design and delivery of management strategies in wildlife reservoirs of disease. More generally, these findings highlight the potential for conflicting impacts of density, disease prevalence, and social- and kin- structure, on the transmission of disease. In badgers, blood is thicker than water because kinship with test positive individuals can counteract the dilution effect of reduced infection risk with increasing group size.

CHAPTER 4: Data Chapter

Badger genetic profiles predict progression of bovine TB *(Mycobacterium bovis)* infection

4.1 Abstract

Inbreeding and/or genome-wide homozygosity may be associated with compromised immune responses, whereby more homozygous individuals are less able to contain or tolerate infection once exposed. Little is known about the genetic basis of bovine TB progression in wild species. In this chapter I present data from a longitudinal population study of European badgers *(Meles meles)*; a key wildlife reservoir of bovine TB, caused by *Mycobacterium bovis*. I investigate associations between marker-based inbreeding estimates, describing the probability of an individual inheriting two identical alleles from a single ancestor, and individual diagnostic TB test results, capturing aspects of the cell-mediated and humoral immune responses to bovine TB exposure. Exposed badgers with higher inbreeding coefficients were more likely to test positive to an antibody test at a given capture event (indicative of progressed disease). I also found evidence of single locus effects predicting the likelihood of an exposed badger becoming culture positive (indicative of infectiousness). A significant interaction between age and inbreeding coefficient was noted in the antibody response model, with inbreeding costs more apparent among older animals. This finding is consistent with age-specific inbreeding depression in a wild mammal and therefore lends support to the mutation accumulation model of senescence.

4.2 Introduction

Host genetics play a critical role in determining the outcome of pathogen infection at an individual level. Matings between related individuals have been observed to result in offspring which are generally less viable than the population mean; this phenomenon is referred to as 'inbreeding depression' (Darwin 1876, Charlesworth and Charlesworth 1987, Slate, David et al. 2004). If such inbreeding effects are powerful, they may go undetected as the most inbred individuals do not survive and may therefore go undetected within a study population (Keller and Waller 2002). However inbreeding depression is commonly sub-lethal and may vary in fitness costs among life-history stages, sexes and environmental conditions (Keller, Reid et al. 2008). Within human populations, inbreeding has been linked to the onset of a range of diseases (Rudan, Rudan et al. 2003). Understanding the causes and consequences of inbreeding depression has important applications to conservation (Hedrick and Garcia-Dorado , Keller and Waller 2002), as well as our understanding of individual behaviour, particularly those involved in active avoidance of mating with kin (e.g. dispersal among mating groups) (Charlesworth and Charlesworth 1987, Townsend, Clark et al. 2010) and the consequences for infectious disease dynamics.

Inbreeding may be associated with a depressed immune response if the loss of genetic variability is within loci involved in pathogen defence (O'Brien and Evermann 1988). However, empirical data on inbreeding depression in pathogen response wild populations is very limited in wild populations (Spielman, Brook et al. 2004, Townsend, Clark et al. 2009). In song sparrows, cell-mediated immunity to a novel pathogen was lower in individuals with a higher inbreeding coefficient (Reid, Arcese et al. 2003). A similar result was found in American crows *(Corvus brachyrhynchos)*, where inbred crows were observed to mount weaker immune responses (Townsend, Clark et al. 2010). Inbreeding in remnant populations of the highly endangered Tasmanian devil *(Sarcophilus harrisii)* has been associated with low genetic diversity in their major histocompatibility complex (MHC), a diverse gene family that plays a crucial role in the vertebrate adaptive immune system and in autoimmunity (Sin, Dugdale et al. 2012), which is thought to increase their susceptibility to the devil

facial tumour disease (DFTD) (Siddle, Kreiss et al. 2007). Inbreeding has also been linked to higher rates of parasitism in a range of species (Coltman, Pilkington et al. 1999, Whiteman, Matson et al. 2006, Smallbone, van Oosterhout et al. 2016).

Three hypotheses have been proposed to explain correlations between inbreeding and fitness (David 1998). First, the 'general effect' hypothesis, which predicts that individuals with lower genome wide heterozygosity, due to inbreeding, will suffer fitness disadvantages when compared to individuals with higher heterozygosity (David 1998, Kardos, Allendorf et al. 2014). This may occur as inbreeding increases the expression of deleterious recessive alleles or through the reduction in heterozygotes in a population where heterozygosity conveys a fitness advantage (Keller and Waller 2002). The 'general effect' hypothesis assumes that inbreeding is homogenously spread across the genome rather than being restricted to particular loci (Slate, David et al. 2004). In order for the general effect hypothesis to be true, inbreeding (and therefore genome wide heterozygosity) should vary among individuals (Slate, David et al. 2004, Szulkin, Bierne et al. 2010, Kardos, Allendorf et al. 2014). Where this is the case, heterozygosity in the markers examined should be correlated with each other if they are to be a good proxy for genome wide heterozygosity; this is termed 'identity disequilibrium' (Weir and Cockerham 1973, Kardos, Allendorf et al. 2014).

There are alternatives to the 'general effect' hypothesis. The 'local effect' hypothesis (also known as associative overdominance) predicts that some proportion of the microsatellite marker loci are physically linked with loci that influence fitness and hence they are transmitted together. In this case, individuals who are heterozygous at particular marker loci tend to be heterozygous at the trait loci (Slate, David et al. 2004, Kardos, Allendorf et al. 2014). Where the state (heterozygous or homozygous) of a given loci, or subset of loci, is a better predictor of a fitness trait than multi-locus heterozygosity, this would support the local effect hypothesis (Annavi, Newman et al. 2014). Finally, the 'direct effect' hypothesis predicts that the microsatellite marker loci themselves are having a direct effect on the fitness trait being measured; this is thought to be unlikely as marker loci are thought to be

generally non-functional (Slate, David et al. 2004) but not impossible, with some microsatellites documented as having functional roles (Li, Korol et al. 2002, Annavi, Newman et al. 2014).

In terms of inbreeding and susceptibility to infection by members of the *Mycobacterium tuberculosis* (TB) complex, there is evidence from a range of species that host genotype has an influence on the outcome of TB infection (le Roex, van Helden et al. 2013). Within the field of human medicine, inbreeding depression has been associated with increased susceptibility to TB infection (Lyons, Frodsham et al. 2009). Inbred rabbits have been demonstrated to be more susceptible to TB infection than outbred rabbits under experimental infection conditions (Dorman, Hatem et al. 2004). Early experimental infection studies on inbred lines of mice suggested that TB resistance may be under relatively simple genetic control in mice, as different genetic lines of mice segregated distinctly into resistance and susceptible phenotypes, without the existence an intermediate level of resistance (Briles 2012). Inbreeding depression has been linked to increased susceptibility to *Mycobacterium bovis* (bovine TB) infection in African lions *(Panthera leo)* (Trinkel, Cooper et al. 2011). In wild boar *(Sus scrofa)*, a wildlife reservoir of bovine TB in Spain (Naranjo, Gortazar et al. 2008), genetic heterozygosity was an important predictor of both resistance to bovine TB infection and progression of disease (Acevedo-Whitehouse, Vicente et al. 2005).

In the UK, the European badger *(Meles meles)* is the principal wildlife reservoir of bovine TB, with an established role in transmission and persistence of the disease within cattle populations. The social structure and kin structure typical of high-density badger populations is known to result in non-random mating (Carpenter, Pope et al. 2005), which increases the possibility of mating between relatives (Szulkin and Sheldon 2008). However, to date, there has been no investigation of the potential role of inbreeding in bovine TB progression within badgers. Previous genetic studies badger populations suggest that extra-group mating is commonplace (Evans, Macdonald et al. 1989, Dugdale, Macdonald et al. 2007), with around half of all cubs fathered by males from other groups (Carpenter, Pope et al. 2005); this behaviour may have evolved to mitigate against inbreeding (Durrant and Hughes 2005, Annavi, Newman et al. 2014). In

order for such inbreeding avoidance strategies to evolve, it is assumed that there must be a fitness cost associated with inbreeding (Szulkin, Stopher et al. 2013) . This could be a reduction in fitness of an individual (e.g. females who mate with relatives have smaller litters), or of offspring (e.g. inbred offspring are smaller, more susceptible to predation or parasitism). Additionally, inbreeding is thought to be usually deleterious in species that are normally outbred, but less severe when inbreeding is part of the natural social system (Soulé 1987, Pertoldi, Loeschcke et al. 2001). Mating with a close relative may in fact enhance an individual's inclusive fitness even if inbreeding affects offspring fitness (Kokko, Ots et al. 2006), for example a female mating with a close male relative (e.g. a brother) gains the direct fitness benefit in terms of offspring production but also an indirect fitness benefit of improving her brothers mating success (Kokko, Ots et al. 2006). Early work on population genetics in a high density badger population suggested inbreeding rates were lower than those reported in other social mammals (Evans, Macdonald et al. 1989), however more recent, pedigree-based findings in a similar density population reported that 5% of matings were incestuous (Dugdale, Macdonald et al. 2007).

In this study, I use data from a long-term population study to investigate the relationship between an exposed badger's estimated inbreeding coefficient and their immune responses, as measured using a range of TB diagnostic test results. As inbreeding effects have been reported to differ between the sexes (Ebel and Phillips 2016) and between age classes (Charlesworth and Hughes 1996), I include these factors to allow for such variation. I predict that exposed individuals with higher inbreeding coefficients will be more likely to show signs of progressed TB infection (as indicated by an increased likelihood to test positive to an antibody test and an increased likelihood of becoming infectious). I also predict that the magnitude of the cell-mediated immune response to *M. bovis* infection will be lower for individuals with lower heterozygosity, due to compromised immune function.

4.3 Methods

Badger Sampling and TB Diagnostic Tests

All data used in these analyses were collected from the long-term capture-mark-recapture study at Woodchester Park in Gloucestershire. Badgers from this study population have been routinely trapped, up to four times a year, since 1976 (Delahay, Walker et al. 2013). Trapped badgers are brought back to a sampling facility, anaesthetised (for full details see Appendix) and a range of clinical samples taken (oesophageal and tracheal aspirates, faeces, urine, swabs of bite wounds or abscesses) for the detection of *M. bovis* by microbiological culture (Gallagher and Horwill 1977). Blood samples are collected and used for TB diagnostic testing. Between 1990 and 2005, the Brock ELISA antibody test (Goodger, Nolan et al. 1994) and the culture of clinical samples were the diagnostic tests used to assign TB status to individual badgers. From 2006 onwards, the Brock ELISA was replaced with the improved Stat-Pak antibody test (Chambers, Crawshaw et al. 2008) and the gamma interferon (IFN) test for T-cell responses to *M. bovis* (Dalley, Davé et al. 2008) were introduced (Delahay, Walker et al. 2013). The combination of diagnostic tests provides a biologically meaningful picture of the progression of disease within an individual badger (Tomlinson, Chambers et al. 2013). The cell-mediated response (as measured by the gamma interferon test) is likely to reflect the first line of defence to *M. bovis* exposure, whereas serological responses (as measured by the Brock ELISA and StatPak tests) take time to develop and hence these antibody tests are most sensitive in animals with progressed disease [39, 42]. Some individuals go on to become 'infectious', characterised by the potential excretion of *M. bovis* bacteria via various routes (Pritchard, Stuart et al. 1986, Graham, Smith et al. 2013). Culture from clinical samples is the only diagnostic approach that has been used throughout the entire study period.

Selection of exposed individuals

To select individuals from the capture database who could be considered to have been 'exposed' to *M. bovis*, individuals were identified who had tested positive to any TB diagnostic test at any point during their capture history. This

dataset was then restricted to individuals who had been caught in their first year (and hence were of known age) and for whom genotype data were available, from which an Individual Inbreeding Coefficient (IIC) could be estimated. The resultant dataset was comprised of 3712 capture events from 490 individual badgers. All the diagnostic tests employed have limitations in terms of their sensitivity and specificity, which have been explored in detail elsewhere (Drewe, Tomlinson et al. 2010, Buzdugan, Chambers et al. 2016). A recent multi-event model constructed for this population demonstrated that at a given sampling event, a truly uninfected individual has a 94% probability of testing negative to the gamma interferon and StatPak tests and on culture. Under these circumstances, the risk of wrongly classifying a truly uninfected badger as infected is low. We can therefore have some confidence that those individuals identified as 'exposed' in the current study are not truly uninfected individuals.

Genotyping:

On first capture, a hair sample was routinely taken from trapped badgers, and stored in 80% ethanol prior to DNA extraction and genotyping (Carpenter, Pope et al. 2005). All genotyping data reported in this study were generated by the team headed by T.Burke at the Molecular Ecology Lab, University of Sheffield. Genotyping has been routinely carried out on hair samples collected from 1990 until present, although only data up to 2011 were available for this analysis. Genotyping involved the use of 22 microsatellite markers, each with 4-7 alleles. This has been demonstrated to be a sufficient number of markers to identify deeper relationships between individuals e.g. differentiating between cousins and unrelated individuals (Goodnight and Queller 1999).

Definitions of 'inbreeding coefficient'

An individual's inbreeding coefficient can be inferred directly from a pedigree, where it refers to the amount of ancestry that is shared by the parents of an individual (Keller and Waller 2002). This measure is therefore determined completely by rates of breeding among relatives. However, in the absence of a pedigree, multi-locus heterozygosity (MLH) may be used as a proxy (Keller and Waller 2002, Jombart 2008). If an individual has related parents, it will have lower heterozygosity, as many loci will be identical by descent (IBD), with a

single gene copy having been passed onto both parents from a single common ancestor (Kardos, Allendorf et al. 2014). MLH can be calculated from a set of microsatellite markers and associated with a fitness trait of interest in an approach termed heterozygosity-fitness correlations (HFCs) (Slate, David et al. 2004). An individual's inbreeding coefficient (hereafter referred to as IIC) can also be estimated directly from microsatellite markers, where it is defined as the probability of an individual inheriting two identical alleles from a single ancestor. This is the approach used here to determine IIC values.

All statistical analyses in this chapter were conducted in R version 3.3.2 (R Core Development Team 2016). In order to identify any relationship between badger genetic profile and disease progression, a number of metrics were used:

1. Individual inbreeding coefficient (IIC): As a pedigree was not available from which to calculate an inbreeding coefficient for each individual (which in a pedigree, is the relatedness between its parents), the R package 'adegenet' was used (v 1.3-9.2) (Jombart 2008) to generate inbreeding estimates from the microsatellite marker data.

2. Multi-locus heterozygosity (MLH): Most studies looking for potential inbreeding depression use multi-locus heterozygosity (MLH) as their measure and are therefore looking for HFC's (heterozygosity-fitness correlations). A previous study looking at the effect of genotype on *M. bovis* infection risk in wild boar found an association between MLH and disease progression (Acevedo-Whitehouse, Vicente et al. 2005); Therefore MLH was used as an additional measure of badger genotype in order to compare results of the current study with those from the wild boar study. The MLH function in the 'inbreedR' package was used (Stoffel, Esser et al. 2016) to calculate MLH values.

3. In order to test for single-locus effects, separate consecutive models were run using each locus' status (homozygous vs heterozygous), instead of the MLH value and then compared with the MLH value model. Models were then constructed containing only the status of the 'important' (i.e. statistically significant in single models) loci, and compared to those in the marker-wide MLH model.

Identity Disequilibrium

To test for the presence of statistically significant identity disequilibrium (David 1998), the g2 statistic and its standard error for the badger genotypes were calculated (using the 'g2_microsats' function within the R package 'inbreedR' [51]). Following the methodology in a similar study (Harrison, York et al. 2013) a randomisation approach was used to quantify the correlation between inbreeding estimates and heterozygosity (Balloux, Amos et al. 2004) and to calculate the heterozygosity – heterozygosity calculation (HHC) (Balloux, Amos et al. 2004) between microsatellites (randomisation code provided by X. Harrison, *pers comm*).

Modelling immune responses

Factors relating to badger immune response (using a variety of diagnostic test responses) were investigated via three distinct analyses, all consisting of generalised linear mixed models constructed using the R package 'lme4' (v1.0-5) (Bates 2010).

1. **Gamma interferon response**

To investigate the effect of IIC on the magnitude of the initial gamma interferon response, a subset of the exposed individuals were selected that were interferon negative on their first capture with evidence of prior absence of infection (i.e. individuals who were previously Stat-Pak and culture negative). As the gamma interferon test has only been used on the population since mid-2006, this resulted in a dataset of 41 individuals. The response variable was the difference in gamma interferon responses to bovine and avian mycobacterial antigens (mean optical density of PPD-B (bovine antigen) response minus the mean optical density of the PPD-A (avian antigen) response), at the incident interferon positive capture event. A mixed effects model was used to investigate the effect of an individual's IIC on the magnitude of their initial gamma interferon response, with sex and age (years) included as fixed effects. Social group and capture year were included as random effects. Wald's chi squared tests were used to assess significance of fixed effects. Due to limitations in the number of observations, interaction terms were not included in this model. IIC values were standardised to mean = 0, SD = 0.5.

2. Likelihood of positive antibody test result

To investigate the effect of IIC on the likelihood of an exposed individual giving a positive result to an antibody test at a given sampling event, a model was constructed with test outcome as a binary response variable (1 = positive, 0 = negative). In both this model and the culture response model below, age, sex and IIC were included as fixed effects. IIC values were standardised to mean = 0, SD = 0.5. Two-way interactions between fixed effects were also included. Individual ID, social group and capture year were included as random effects.

3. Likelihood of culture positivity

To investigate the effect of IIC on the likelihood of an *M. bovis* positive culture result at a given sampling event, a model was constructed with culture result as a binary response variable (1 = positive, 0 = negative). Fixed and random effects were as described for the antibody response model.

4.4 Results

Individual Inbreeding Coefficients.

The mean Individual Inbreeding Coefficient (IIC) for 490 badgers from the Woodchester Park study population with evidence of exposure to *M. bovis* (1990 – 2011) was 0.18 (Fig 4.1). Figure 4.2 suggests that there is little spatial structure in inbreeding in the Woodchester Park population, with more inbred social groups appearing across the study area. It also appears that mean social group IIC is not obviously related to mean social group size, with higher inbreeding values appearing across a range of group sizes, although this was not formally tested.

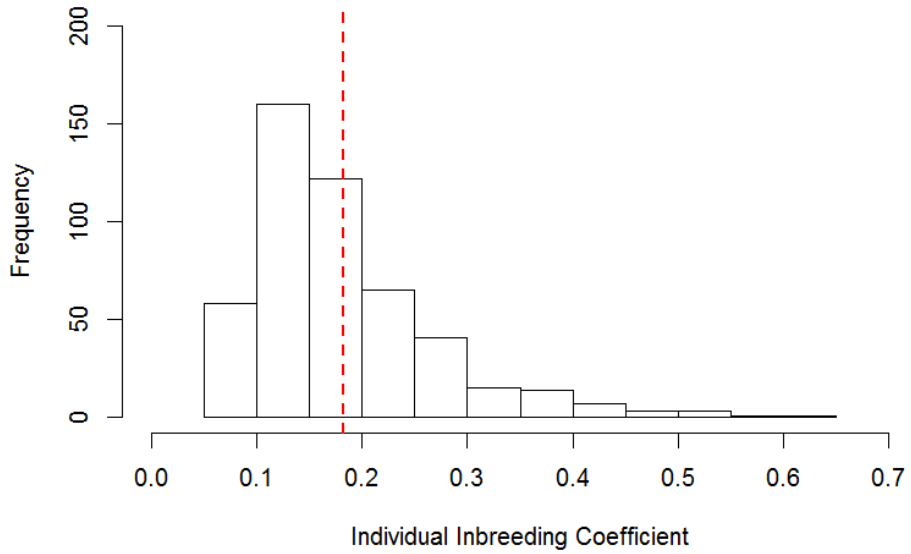


Fig 4.1 Histogram of estimated individual inbreeding coefficients (IIC) for 490 badgers from the Woodchester Park study population with evidence of exposure to *M. bovis* (1990 – 2011). The mean IIC value was 0.18, as indicated by the red line.

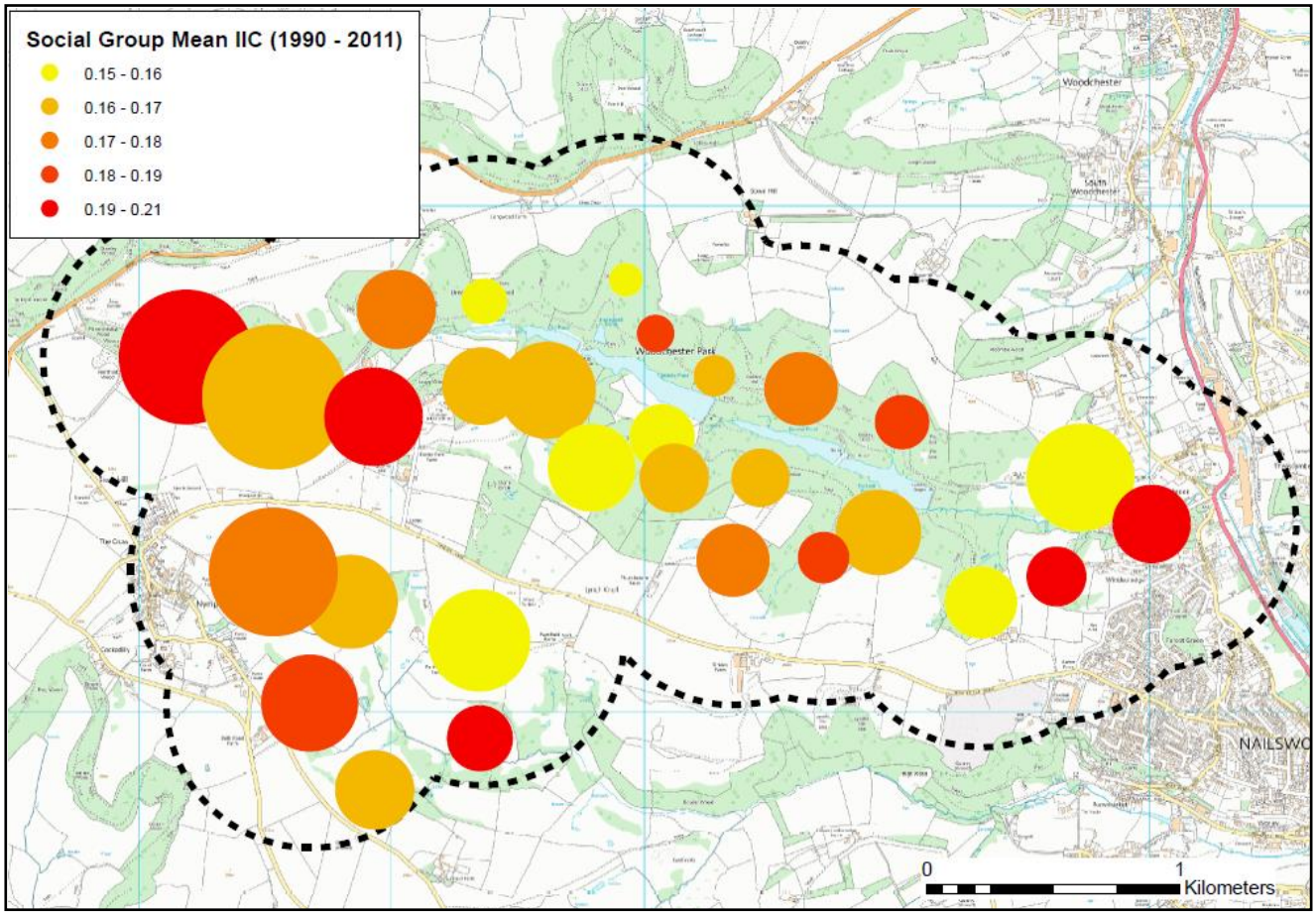


Fig 4.2 Average individual inbreeding coefficients (IIC) by social group for all genotyped badgers within the Woodchester Park badger population (1990-2011, n = 1534). The size of each circle represents the average number of group members in the social group during this time period.

Identity Disequilibrium

Weak but significant identity disequilibrium was detected in the genotype dataset of 1899 badgers from Woodchester Park (g2 = 0.005, SD = 0.0006, P = 0.01 based on 100 iterations). Additionally, a significant heterozygosity – heterozygosity correlation (HHC) was detected, consistent with the presence of identity disequilibrium (HHC = 0.07 (CI 0.03 – 0.11)). As heterozygosity between the marker loci was significantly correlated in this population, this suggests that the heterozygosity of the marker loci would reflect heterozygosity at unlinked, functionally important loci; a key requirement for the general effect hypothesis of inbreeding depression (Slate, David et al. 2004). Heterozygosity and inbreeding estimates were significantly negatively correlated, based on 100 iterations (r = -0.08 (CI: -0.10 to -0.07); this is in line with the expected

correlation given the number of microsatellites employed and the population structure (Balloux, Amos et al. 2004). This result suggests that the variation in heterozygosity among individuals is informative of their inbreeding level (Harrison, York et al. 2013).

Variation in Immune Responses

Gamma interferon responses

The magnitude of an individual's initial gamma interferon response was not related to their age (Wald's $\chi^2_{(1)} = 1.26$, $p = 0.27$), sex (Wald's $\chi^2_{(1)} = 1.43$, $p = 0.23$), or IIC (Wald's $\chi^2_{(1)} = 0.38$, $p = 0.54$).

Likelihood of positive antibody test result

The likelihood of an individual testing positive to an antibody test at a given sampling event was influenced by several predictors and their interactions, as determined by model averaging and information criteria

Table 4.1 Details of the top 12 regression models with a ΔAICc<6 predicting the odds of badgers giving a positive antibody test result at a given sampling event. Individual ID, social group and study year were included as random effects. Each row in the table indicates a model with a given set of variables (+). Degrees of freedom, ΔAICc, model weight and R^2 values are included for each model. Marginal R^2 (R^2M) represents the variance explained by fixed factors and Conditional R^2 (R^2C) represents the variance explained by both fixed and random factors (Barton 2011).

Model	Age (years)	Sex	IIC	Age (years):IIC	Sex:IIC	Sex:Age (years)	df	ΔAIC_C	ω_i	R^2C	R^2M
1	+		+	+			7	0.00	0.25	0.47	0.05
2	+	+	+	+		+	9	0.28	0.21	0.47	0.05
3	+	+	+	+			8	1.85	0.10	0.47	0.05
4	+	+	+	+	+	+	10	1.96	0.09	0.47	0.05
5	+	+				+	7	2.12	0.09	0.47	0.05
6	+	+	+			+	8	2.85	0.06	0.47	0.05
7	+						5	3.05	0.05	0.46	0.05
8	+	+	+	+	+		9	3.44	0.04	0.47	0.05
9	+		+				6	3.64	0.04	0.46	0.05
10	+	+	+		+	+	9	4.56	0.03	0.46	0.05
11	+	+					6	4.80	0.02	0.46	0.04
12	+	+	+				7	5.72	0.02	0.46	0.04

From Table 4.1 and Fig 4.3 it is evident that age is a consistently positive predictor of antibody test positivity in exposed badgers; this is consistent with the antibody tests used being indicative of progressed *M. bovis* infection (Goodger, Nolan et al. 1994, Chambers, Crawshaw et al. 2008) and hence more sensitive in individuals who have been infected for a longer time period. The interaction between sex and age was included in the top model set, with exposed male badgers having a higher probability of giving a positive antibody test result as they aged, however the effect was variable (as indicated by the

95% confidence intervals touching zero in Fig 4.3). The interaction between age and IIC was in the top four most supported models with a consistent directional effect (Fig 4.3), indicating that inbreeding effects on the risk of antibody positivity increased with age.

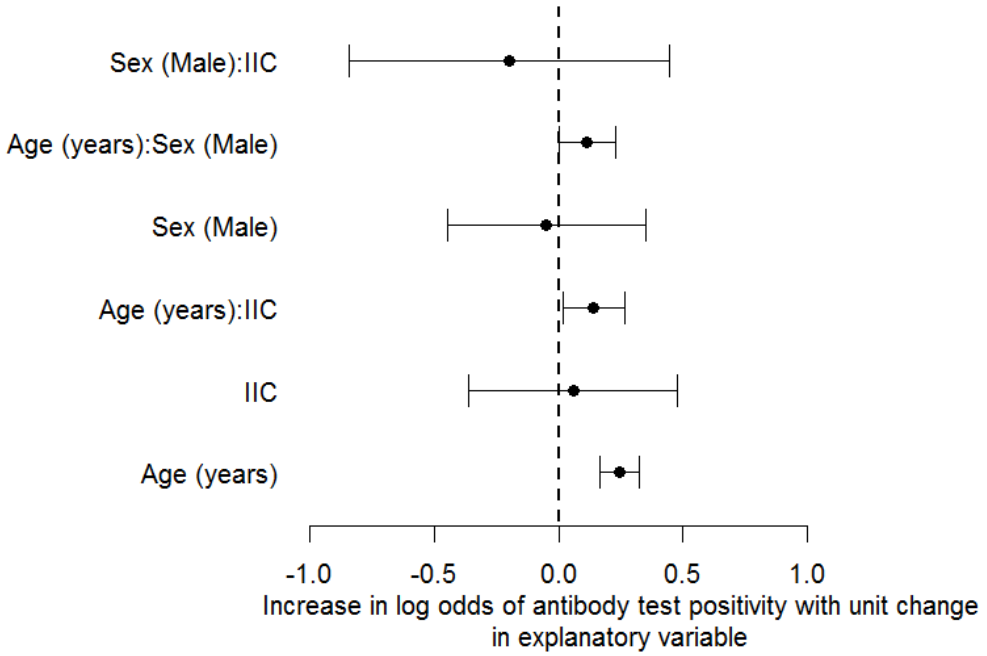


Fig 4.3 Factors affecting the risk of badgers in the Woodchester Park population testing positive to an antibody test at a given capture event (1990-2011). The average model coefficients (log odds) shown were calculated for variables included in the top model set (see Table 4.1). Bars indicate 95% confidence intervals. Model-averaged regression slopes are considered important if their confidence intervals do not span zero.

Figure 4.4 illustrates the impacts of inbreeding on the likelihood of antibody test positivity with age. In both male and female badgers, individuals with higher IIC values become more likely to test positive to the antibody tests as they age. Age related increases in risk are stronger for males than females, such that male badgers in the oldest age class are at a higher risk of testing positive to an antibody test than females in the same age class. Also, individuals with higher IIC values do not appear to be represented in the older age classes (see Fig 4.5) consistent with a survival cost to inbreeding, although this was not investigated further in the current study

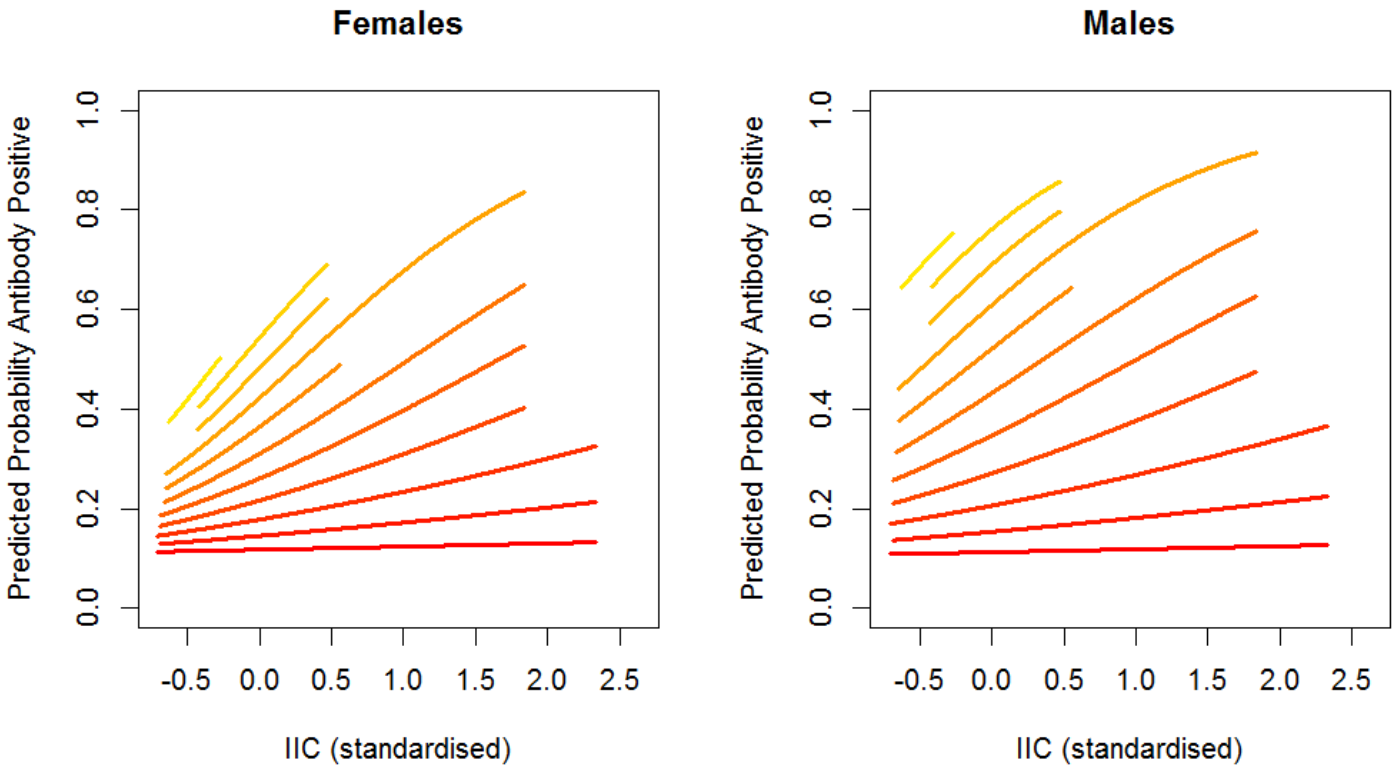


Fig 4.4 Changing relationship between IIC and the likelihood of an individual badger giving a positive result to an antibody test, shown by sex and age (colour gradient represents age, with darkest red = individuals less than one year, up to yellow representing individuals older than 10).

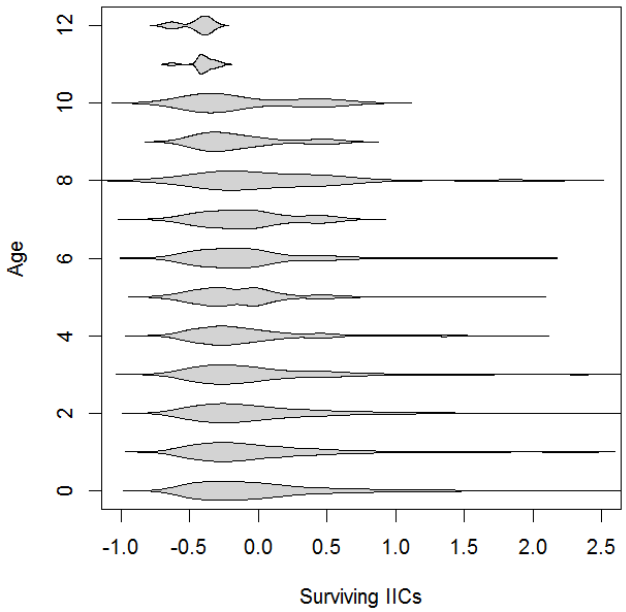


Fig 4.5 IIC values for individual badgers (Woodchester Park, 1990-2011) presented by age class.

Likelihood of positive culture result

The likelihood of an individual testing culture positive at a given sampling event was also influenced by several predictors and their interactions, however in most cases these effects were variable, as indicated by the confidence intervals spanning zero in Fig 4.6.

From Fig 4.6 and Table 4.2, it is evident that the risk of an *M. bovis* positive culture being obtained from an exposed badger at a given sampling event increased consistently with age. IIC was also a positive predictor of risk, appearing in the majority of the top model set (Table 4.2), however, the confidence intervals for the estimate span 0 (Fig 4.6), indicating some variability in this effect. The effects of inbreeding increased with age, however in contrast to the antibody response model, the estimate for this effect was variable and spanned 0. A positive culture was more likely to be obtained from a male badger (sex appears in the majority of the top model set; Table 4.2), again however, this effect was variable as evidenced by confidence intervals spanning 0.

Table 4.2 Details of the top 13 regression models with a ΔAICc<6 predicting the odds of an *M. bovis* culture positive sample being obtained from an exposed badger at a given sampling event. Individual ID, social group and study year were included as random effects. Each row in the table indicates a model with a given set of variables (+). Degrees of freedom, ΔAICc, model weight and R^2 values are included for each model. Marginal R^2 (R^2M) represents the variance explained by fixed factors and Conditional R^2 (R^2C) represents the variance explained by both fixed and random factors (Barton 2011).

Model	Age (years)	Sex	IIC	Age (years):IIC	Sex:IIC	Sex:Age (years)	df	ΔAIC_C	ω_i	R^2C	R^2M
1	+	+	+		+	+	9	0.00	0.18	0.70	0.16
2	+	+				+	7	0.50	0.14	0.86	0.15
3	+	+	+	+	+	+	10	0.84	0.12	0.86	0.16
4	+	+	+		+		8	0.95	0.11	0.76	0.03
5	+	+	+			+	8	1.17	0.10	0.07	<0.001
6	+	+	+	+	+		9	1.62	0.08	0.85	0.16
7	+	+					6	2.11	0.06	0.04	0.001
8	+	+	+	+		+	9	2.12	0.06	0.61	0.05
9	+	+	+				7	2.61	0.05	0.86	0.14
10	+	+	+	+			8	3.29	0.03	0.79	0.16
11	+						5	3.34	0.03	0.17	<0.001
12	+		+				6	3.86	0.03	0.48	0.009
13	+		+	+			7	4.37	0.02	0.57	0.02

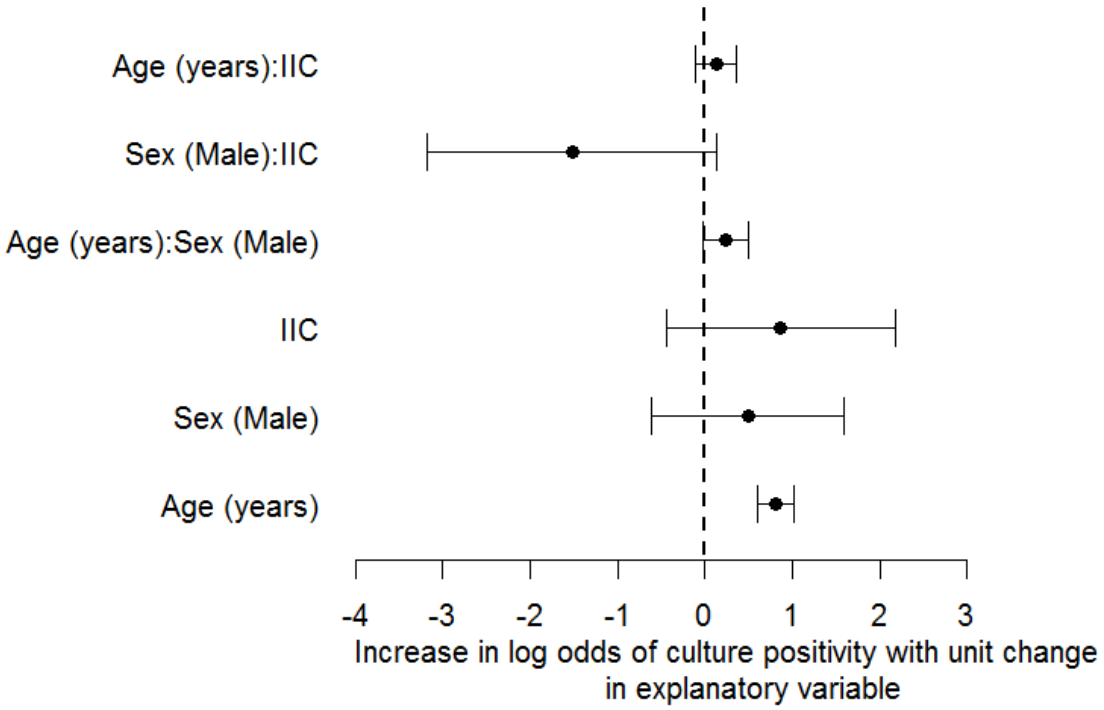


Fig 4.6 Factors affecting the risk of badgers in the Woodchester Park population testing culture positive at a given capture event (1990-2011). Average model coefficients (log odds) shown were calculated for variables included in the top model set (Table 4.2). Bars indicate 95% confidence intervals. Model-averaged regression slopes are considered important if their confidence intervals do not span zero.

Single locus effects

The single-locus models were better supported than the MLH models for both the antibody response model and the culture response model, as evidenced by lower AIC values, (antibody response: single-locus: AIC = 3417, MLH: AIC = 3430, culture response: single-locus: AIC =1461, MLH: AIC=1465). The separate models, in which each locus was included as a fixed effect, revealed that some loci may be more strongly associated with TB test status than others (see Table 4.3). The final set of models for each response variable, where only the loci and their interactive effects identified as significant in the individual loci models ($P < 0.05$ in Table 4.3) were included, were the best supported (indicated by lower AIC values than the MLH model, antibody response: important loci model AIC = 3368 vs 3420 for the MLH model, culture response: important loci model AIC = 1411 vs 1465 for the MLH model).

In several cases, the same loci are predictors of both the likelihood of a positive antibody response and the likelihood of culture positive (e.g. 1gm, 1gs, 1gxl, 2gxl, 2yl and 2ys). It is also interesting to note than in the majority of cases, the significant interaction identified is with age.

Table 4.3 P values from a series of individual models to assess the strength of associations between each locus and TB test status in badgers from Woodchester Park (1990-2011). In all cases Wald's chi squared tests were used to assess significance. Significant results ($P < 0.05$) are indicated in black bold, close to significant results (P between 0.05-0.09) are indicated in bold grey.

Marker Loci	Antibody Positive	Culture Positive
1bl	**0.01 (*sex)**	0.79
1bm	0.84	0.26
1bs	0.28	0.15
1bxl	**<0.001 (*age)**	0.31
1gl	0.12	0.07
1gm	**0.01 (*age)**	0.09 (*age)
1gs	**0.04 (*sex)**	**0.03 (*sex)**
1gxl	0.07 (*age)	0.09 (*age)
1ys	0.20	**<0.001 (*age)**
1yxl	**0.03 (*age)**	0.09 (* age)
2bs	0.32	0.09 (* age)
2bxl	0.61	0.56
2gl	0.41	0.77
2gs	0.76	0.39
2gxl	0.08	**0.04 (* age)**
2yl	**0.005 (*sex)**	**0.02 (* age)**
2ys	**<0.001 (*age)**	**<0.001 (*age)**
m1	0.90	0.09 (* age)
m10	**0.047 (*age)**	0.63
m12	**<0.001 (*age)**	**0.01**
m14	**0.05**	0.10 (* sex)
m15	0.06	0.10 (* age)

4.5 Discussion

The likelihood of positive responses to tests for *M. bovis* specific antibodies increased with the measure of inbreeding. This suggests that badgers with higher levels of inbreeding are less able to tolerate or contain *M. bovis* infection and are therefore more likely to subsequently exhibit progressed disease. Both the antibody tests used in this study (Brock ELISA and Brock TB StatPak) have been reported to be more sensitive in individuals with progressed or disseminated disease (Chambers, Crawshaw et al. 2008, Chambers, Waterhouse et al. 2009) and which may pose the greatest risk of transmitting infection (Chambers, Crawshaw et al. 2008). Further support for this interpretation is provided by the observation that inbreeding was a positive predictor of the likelihood of bacterial shedding being detected. These findings are consistent with work on wild boar which demonstrated that individuals with lower heterozygosity (indicative of inbreeding) had signs of more progressed disease (Acevedo-Whitehouse, Vicente et al. 2005). The authors inferred that individuals with lower heterozygosity were less successful at mounting an immune response capable of containing the *M. bovis* bacilli within single early stage granulomas (Acevedo-Whitehouse, Vicente et al. 2005). The pathogenesis of TB infection in badgers is known to be complex, in that not all badgers exposed to infection become diseased, with a proportion mounting a successful immune response. In some cases, lesions develop but these lesions remain dormant, such that the animal shows no clinical signs and is not infectious (Roper 2010), this may be considered a 'containment' or 'latent' phase (Gallagher and Clifton-Hadley 2000). It has been suggested that many badgers remain in this latent phase throughout their lives (Gallagher and Clifton-Hadley 2000). However, in a proportion of exposed badgers, the immune response is insufficient to contain, or 'wall off' the bacteria within lesions. The mycobacteria can then escape, spread to new body sites and the individual can become infectious, characterised by potentially large amounts of bacterial shedding through a range of routes (Gallagher and Clifton-Hadley 2000). The amount of bacteria shed by an infectious badger is related to how progressed its pathology is (Nolan 1991), hence individuals with evidence of more progressed disease are likely to be more important in the onward transmission of infection to susceptible individuals. Interestingly, one of the

potential stressors which has been suggested to result in badgers going from latent TB infection to active infection is senile deterioration (Gallagher and Clifton-Hadley 2000). The age-mediated inbreeding effect observed in this study could suggest that immune responses in inbred badgers deteriorate faster with age and hence they are more likely to convert from a latent to an active state of infection. From a management perspective, these individuals may be disproportionately important in transmitting infection to susceptible individuals within the badger population.

In the present study, the measure of inbreeding in badgers did not predict the magnitude of their initial gamma interferon responses to *M. bovis*. The gamma interferon response is considered to be essential for protection against tuberculosis infection (Cooper, Dalton et al. 1993) and is the first line immune response (Tomlinson 2013). The magnitude of the incident gamma interferon response in badgers has been identified as a potential predictor of the likelihood of disease progression, with consistently elevated responses observed in animals with evidence of advanced disease (Tomlinson, Chambers et al. 2015). Other factors may influence the magnitude of the initial gamma interferon response, for example the size of the infective dose (Lesellier, Corner et al. 2009) or the route of infection (Tomlinson, Chambers et al. 2015), in which case we would not expect any relationship with inbreeding.

The results of the current study are consistent with a cost of inbreeding for immune responses to *M. bovis* infection in badgers. Furthermore, they provide evidence for such inbreeding effects to be stronger in older animals. Natural selection is predicted to act more strongly on fitness traits which affect early survival or reproductive output than those affecting later life (Medawar 1952). As many individuals die before they reach old age through natural competition or mortality, there is little selection pressure for traits maintaining long-term viability. Mutations which detrimentally affect fitness in early life and hence affect an individual's subsequent reproductive success will be strongly selected against, whereas those affecting fitness in later life when an individual has already successfully reproduced may not be filtered out by natural selection (Medawar 1952, Hamilton 1966). This leads to the evolution of a life history where mortality increases and reproductive performance declines with age (Charlesworth and Hughes 1996). One evolutionary model of aging, or

'senescence', is 'mutation accumulation' where random, detrimental mutations, whose fitness effects are only seen late in life, fail to be purged from a population by natural selection and hence accumulate (Medawar 1952). Under the mutation accumulation model of aging, inbreeding related declines are predicted to increase with age, as I report in the current study. If sub-lethal effects of inbreeding occurred in early life, selection should act on them to purge them from the population. However, if the sub-lethal effects are only seen later in life, they may escape the selection filter. Within human medicine, inbreeding effects on the onset risk of late-acting diseases have been documented (Rudan, Rudan et al. 2003). However, age-specific inbreeding effects have rarely been demonstrated in wild populations (Reid, Arcese et al. 2003). One example is from a study of song sparrows *(Melospiza melodia)*in which annual reproductive success was related to individual inbreeding, interacting with individual age (Keller, Reid et al. 2008). As far as the authors are aware, this study represents the only other example drawn from a wild population.

I also noted that individuals with the highest inbreeding values in the study population were absent in the older age classes, which is consistent with their loss from the population. This is in line with the results of a number of other studies in which survival probabilities were lower for inbred individuals (Gjerde, Gunnes et al. 1983, Keller, Grant et al. 2002, Townsend, Clark et al. 2009). The potential impact of inbreeding on survival, and interactions with disease status within this system are an area worthy of further investigation. The mean IIC value presented here (0.18) is much higher than that reported from a similar density badger population (Annavi, Newman et al. 2014), where only 5% of individuals were considered to be inbred (based on a cut off IIC < 0.125 considered to be 'outbred''). However, it is important to note that the IIC calculation in this previous study was made from pedigree data rather than directly from microsatellite data as has been done in the present study. In some circumstances, inbreeding estimates from pedigrees may only be weakly correlated with marker-based estimates (Robinson, Simmons et al. 2013). Where a pedigree is available, an individual's IIC is the same thing as the relatedness between their parents; this relies on the availability of a deep, detailed pedigree. The IIC elicited directly from microsatellite data is defined as

the probability that, at a given locus, two identical alleles have been inherited from a common ancestor(Jombart 2008).

As previously suggested in badgers (Beirne, Waring et al. 2016), the results of this study indicate that immune responses of male badgers deteriorate with age to a greater extent than those of female badgers (Graham, Smith et al. 2013); i.e. they experience stronger 'immunosenescence'. Male badgers are known to experience more rapid disease progression and higher mortality from *M. bovis* infection relative to females (Graham, Smith et al. 2013, Tomlinson, Chambers et al. 2013); in the present study there are no exposed male badgers in the oldest age classes. A previous study of the Woodchester population found no evidence for differences in immunosenescence rates between male and female badgers, as measured by mean telomere lengths and rates of telomere attrition over time (Beirne, Waring et al. 2016). However, it was noted that within-individual variation in immunosenescence was greater than that observed at the population level, suggesting selective loss of individuals from the population who had shorter telomere lengths or whose telomeres degraded faster (Beirne, Waring et al. 2016). This previous study did not take into account host genotype; it would be of interest to incorporate individual heterozygosity in order to see if inbreeding is linked to faster rates of immunosenescence, which would be a possible mechanism for the observed age-related inbreeding effect on disease progression in the current study.

The presence of significant identity disequilibrium in the current study population is consistent with the 'general effect' hypothesis in which neutral marker loci are correlated with genome-wide heterozygosity. However, testing each marker separately also found some evidence of single-locus associative effects on bTB progression (see Table 4.3), consistent with the 'local effect' hypothesis and the best supported models for both responses contained only the loci identified as important by the individual models. These findings are consistent with the observation of genetic resistance to bovine TB in wild boar where only single-locus effects predicted TB progression, with several of the single loci being identified as mapping to regions of the genome with known immune function (Acevedo-Whitehouse, Vicente et al. 2005). However, rather than an observed fitness cost being due either to genome wide inbreeding depression (where large numbers of recessive genes have a small effect) or

due to local effects (where microsatellite markers are linked by chance to a functional gene which has a greater effect on immune response), it has been suggested that these explanations represent two ends of a spectrum, with true populations appearing somewhere along the scale (Balloux, Amos et al. 2004).

Building on these results, the availability of the badger pedigree for this population in the near future will allow us to investigate fitness costs in terms of immune responses in the offspring of parents with low heterozygosity. Fitness costs due to parental genetic effects may be important, whereby it is the offspring of inbred parents who bear the fitness cost, rather than the inbred individuals themselves (Annavi, Newman et al. 2014). Although it is not possible to consider them in the present analyses as parentage data is not currently available, these results suggest that it would be an avenue worth pursuing, and the combination of parentage data alongside microsatellite based estimates of inbreeding (Townsend, Clark et al. 2009) would add value to the current study. The current study did not investigate the relationship between susceptibility to becoming infected with *M. bovis* and inbreeding, as has been demonstrated previously in wild boar (Acevedo-Whitehouse, Vicente et al. 2005). As the likelihood of a badger becoming infected with *M. bovis* is affected by a range of factors, including early life environment (Benton, Delahay et al. 2016), social group composition (Tomlinson, Chambers et al. 2013, Benton, Delahay et al. 2016), position in the social network (Weber, Carter et al. 2013), individual movement behaviour (Woodroffe, Donnelly et al. 2009), age and sex (Graham, Smith et al. 2013), these would need to be carefully taken into account in order to investigate whether inbreeding coefficients provided an additional explanation of susceptibility to infection.

Variation in immunogenetic profiles amongst badgers may play an important role in M. bovis transmission and persistence within the social group and may potentially scale up to population level effects. The results of this study also highlight the possibility that single-locus effects may be important determinants of *M. bovis* infection outcomes in badgers. If single loci are powerful predictors of TB progression, as has been found in wild boar (Amos and Acevedo-Whitehouse 2009), then this could have important implications for disease management . For example, genotyping badger populations in areas where TB is not yet established may allow better predictions of rates of spread

within the badger population. This has implications for our understanding of TB transmission within badger populations. Additionally, management of badger populations through culling may alter the host genetic population structure in as yet unknown ways, potentially increasing levels of inbreeding through reducing the population density or decreasing inbreeding as surviving individuals range more widely (Riordan, Delahay et al. 2011). Further these results highlight the importance of considering the role that host genotype plays on disease outcomes, an area which until recently has been largely overlooked (Allen, Minozzi et al. 2010). Understanding how host genetics influence pathogen outcomes can help inform epidemiological models of disease spread (Hendricks, Epstein et al. 2017) and, as recently demonstrated in the highly endangered Tasmanian devil, can help to predict population level responses to pathogens and inform conservation interventions such as translocations and reintroductions (Hendricks, Epstein et al. 2017).

The above findings are consistent with age-specific inbreeding depression in a wild mammal; a phenomenon rarely documented in wild populations. They also lend support to the mutation accumulation model of senescence. A useful next step to build on the findings of the current study would be to start to explore potential mechanistic links between inbreeding and *M. bovis* progression in badgers; for example, do inbred badgers experience more rapid immunosenescence than outbred badgers, are inbred badgers in poorer body condition making them more susceptible to disease progression? The links between inbreeding, individual condition and disease progression have rarely been studied in wild populations (Townsend, Clark et al. 2010), making this an area of considerable interest.

CHAPTER 5: Data Chapter

Changing spatial patterns of bovine TB infections in a high-density badger population

5.1 Abstract

The structure of host populations will influence disease distribution and persistence but increases in prevalence inevitably cause reductions in disease clustering, making it difficult to identify demographic drivers. The European badger is a wildlife reservoir of bovine tuberculosis (caused by infection with *Mycobacterium bovis*) in the UK. Current understanding suggests that social structuring in moderate to high density badger populations promotes the spatial clustering of infection. Culling, as a disease management strategy, can disrupt this social structure and hence risk the spread of disease. In this chapter I test whether the spatial arrangement of *M. bovis* infection has significantly changed over a 20 year period in an unmanaged, high density population of badgers in the south west of England. During the first decade of the long-term study, *M. bovis* infection remained spatially clustered. However, during the second decade, the spatial distribution of infection became more widespread, characterised by a reduction in disease clustering among individual hosts and at the social group level. This occurred against a background of increasing TB prevalence in the population. These findings reveal changes in epidemiological trends occurring over relatively long time periods and challenge the view that *M. bovis* distribution necessarily remains stable, and tightly clustered, in badger populations which are not subject to management interventions such as culling. I suggest that natural perturbation events occur and might be responsible for these changes in infection distribution. Spatial and temporal patterns of disease prevalence and persistence observed in the field should be used to inform ecological and epidemiological models, helping to predict the outcome of management interventions.

5.2 Introduction

Empirical data on the spatial and temporal distribution and persistence of pathogens provide invaluable information for transmission models that attempt to predict the progression and extent of disease epidemics (Gaudart, Rebaudet et al. 2013) and inform management interventions (Miller and Conner 2005), for example determining the required intensity of cull (Fulford, Roberts et al. 2002), or the duration of spatially targeted vaccination programmes (Keeling and White 2010). Management interventions, particularly the culling of reservoir hosts, can disrupt disease dynamics (Woodroffe, Donnelly et al. 2005, Jenkins, Woodroffe et al. 2007, Delahay, Smith et al. 2008) with potentially counter-productive effects on disease prevalence and distribution. Any attempt to predict the impacts of management interventions, requires a baseline understanding of how ecological and behavioural processes may drive disease dynamics.

Bovine tuberculosis in UK cattle is continuing to increase in incidence and geographical distribution (Godfray, Donnelly et al. 2013). Infection in European badgers (*Meles meles*) has been implicated in the persistence of the disease in cattle since the 1970's (Gallagher, Muirhead et al. 1976). Experimental evidence for reservoir status of the badger comes from the Randomised Badger Culling Trial (RBCT: Krebs et al. 1997). It has long been suggested that badger social structure is itself a barrier to disease transmission between social groups (Overend 1980), the disruption of which could result in increased local disease transmission. Field studies have shown that in the wake of culling the stable social structure of badger populations is disrupted, resulting in a more mobile residual population (Cheeseman, Jones et al. 1981, Woodroffe, Donnelly et al. 2006), with potentially higher potential for contact with individuals from whom they would previously have been socially isolated . Hence clustering of *M. bovis* infection in badgers was observed to decrease in culled populations (Jenkins, Woodroffe et al. 2007). Recent analyses suggest that even small scale social perturbations result in measurable behavioural changes among survivors, which may be associated with increased disease transmission (Bielby, Donnelly et al. 2014).

Whilst much consideration has been given to the impact of population reduction on *M. bovis* distribution in badgers, fewer studies have considered the dynamics of *M. bovis* infection in unmanaged populations. The generally supported hypothesis, based both on data from long term epidemiological studies (Delahay, Langton et al. 2000)and shorter observations of badger social groups subject to population management (Woodroffe, Donnelly et al. 2005, Jenkins, Woodroffe et al. 2007) is that TB remains stably clustered in unmanaged, naturally infected badger populations. This finding is primarily based on a long-term study of an undisturbed population of European badgers (*Meles meles*) at Woodchester Park in south-west England. Badgers have been trapped, sampled and routinely tested for *M. bovis* infection in the study area since the mid-1980s, resulting in three decades of longitudinal information on host demography and disease epidemiology. Diagnostic test results from trapping data up to 1996 indicated that *M. bovis* infection was aggregated in social groups in the west of the Woodchester Park study area. Temporal trends of infection were not synchronised between neighbouring social groups, consistent with low rates of transmission and infected hosts distributed in 'stable persistent foci' (Delahay, Langton et al. 2000). With the benefit of a further decade of data collection at this site, I am now able to investigate whether this pattern has persisted in the long term. A localised phenomenon of immigration and depopulation occurred in four of the social groups in the northern part of the study area in the late 1990s. In these groups cub recruitment fell, immigrant adults were trapped at a higher rate than elsewhere in the study area and were more likely to go on to exhibit progressive disease (Delahay *et al*., unpublished). Since these observations were made, the spatial and temporal distribution of disease has not been formally examined at the population level.

Here I describe the spatial and temporal distribution of *M. bovis* infection in the Woodchester Park badger population over a 20 year period. In particular, I consider whether the spatial dynamics of *M. bovis* infection in this population have changed over time as a result of natural ecological processes.

5.3 Methods

Badger Sampling, Disease Status & Social Group Allocation

All data used in these analyses were collected from the long-term trapping and sampling study at Woodchester Park in Gloucestershire. Badgers from this study population have been routinely trapped, up to four times a year, since 1976 (for full details see Appendix). Data from 1985 – 2005 were included in the main analyses. Data from live trapping and *post mortem* records were included in these analyses. Between 1985 and 2005, the Brock ELISA & culture of clinical samples (Rogers, Cheeseman et al. 1997) were the diagnostic tests used to assign TB status to individual badgers. From 2006 onwards, the Brock ELISA was abandoned due to doubts over its performance and poor correlation with other test results and new diagnostic tests were introduced (Delahay, Walker et al. 2013). For this reason, data from 2006 onwards were not included in these analyses. As many badgers were trapped on multiple occasions during a calendar year, each was assigned to one social group and one infection status category per year. Rules used to assign badgers to social groups where they were caught in multiple groups within a year were as described previously (Vicente, Delahay et al. 2007). If it was not possible to assign a badger to a social group based on these criteria, they were excluded from the analysis, however this was only the case for less than 1% of capture events. As the spatial extent of the study area changed between 1985 and 2005, only social groups falling into a defined 'core' area, which was consistently trapped throughout the study, were included in analyses. The infection status of each individual animal in a given year was assigned according to a one-way progression scheme based on the results of the two diagnostic methods (Gallagher, Monies et al. 1998, Graham, Smith et al. 2013). Where badgers were caught multiple times per year, they were assigned to their most progressed status in that year. In this way, each individual trapped in a given year was given an infection status 'score' (Table 5.1). An average TB index was calculated for each social group by year (replicating the approach used in (Delahay, Langton et al. 2000)).

Table 5.1 Categorisation of infection status of individual badgers in the Woodchester Park study area between 1985 - 2005, based on a one way disease progression scheme (Gallagher, Monies et al. 1998, Graham, Smith et al. 2013).

Infection status	Brock ELISA test result	Culture of clinical samples	Assigned individual 'index' score
Negative	Negative	Negative	0
Seropositive	Positive	Negative	1
Single Site Excretor	Positive or Negative	Positive culture from one body site only (regardless of number of positive cultures)	2
Multi-Site Excretor	Positive or Negative	Positive culture obtained from 1+ body site over trapping history	3

Visualisation of Infection Distribution

The spatial configuration of *M. bovis* infection during the period of interest was visualised using kernel heatmaps, each representing a 7 year period (1985-1991, 1992-1998 and 1999-2005). This allowed visual examination of spatial patterns in *M. bovis* infection in GIS outputs (ESRI 2013, ArcGIS Desktop, Release 10.2, Redlands, CA: Environmental Systems Research Institute). The 'kernel density' tool was used to produce kernel maps of individual infection scores. For any given feature this tool calculates the density of other similar features in the surrounding neighbourhood as specified by the search radius size. Here, features were individual badgers, weighted by increasing values of

TB score. A raster surface is produced, where the density at each output raster cell is calculated by adding the values of all kernel surfaces where they overlay the raster cell centre.

Spatio-Temporal Structure

The methodology employed in the earlier spatio-temporal analysis (Delahay, Langton et al. 2000) was reproduced and extended it to more recent time periods, using three dimensional spatio-temporal variograms constructed using the 'gstat' package (Pebesma 2004) in software R version 3.0.2 (R Core Team, 2013). As in the original study (Delahay, Langton et al. 2000), these variograms plotted the residuals from a model where the response variable was the log transformed TB index of each badger social group in each year and the fitted factors were year and social group. The variograms plot half the averaged squared differences between the paired residuals against spatial and temporal distance. Changes in spatial autocorrelation over the time period of interest were examined using linear mixed effects models, constructed for each year which tested for the presence of different spatial autocorrelation structures. These were compared with null models in which no spatial autocorrelation function was present using AIC comparison, with the lowest AIC values indicating the best supported model (Burnham and Anderson 2004).

Clustering of Infection

Clustering analyses followed previously established methods (Jenkins, Woodroffe et al. 2007) which are expected to be robust to changes in badger density. Badger location was tied to the geographic co-ordinates of the main sett (the principal focus of breeding and social behaviour (Roper 2010)) in the social group that the animal was assigned to in each year. Euclidean (smallest direct) distances between main sett locations were calculated using the 'SpatialTools' package in R software v 3.0.2 (R Core Development Team 2013).The following information was calculated for trapping data for each year from 1985 – 2005:

- **Infected – Infected Nearest Neighbour Distance:** The distance to the nearest neighbouring 'infected' badger (i.e. those that had tested

positive to any diagnostic test at some point in their previous trapping history) from each 'infected' badger.

- **Uninfected – Infected Nearest Neighbour Distance:** The distance to the nearest neighbouring 'infected' badger (i.e. those that had tested positive to any diagnostic test at some point in their previous trapping history) from each 'uninfected' badger (i.e. those that had not up to that point tested positive to a diagnostic test).

For each year of the study, the log ratio of the difference between the distance from infected badgers to the nearest infected badger, and the distance from uninfected badgers to the nearest infected badger was calculated. Spatial clustering is indicated when the relative distance to the nearest infected badger is shorter for infected than for uninfected badgers. One metre was added to all distances to avoid infinite ratios where distances were 0 (indicating that the nearest infected badger was in the same social group and therefore assigned to the same spatial location). Changes in clustering patterns are linked by definition to changes in prevalence of infection: very high prevalence will necessarily yield an un-clustered distribution. In order to determine whether the observed changes in clustering scores were independent of changes in prevalence, a permutation testing procedure was used. In this approach, for each year, infection status (i.e. 'positive' or 'negative') was randomly shuffled among spatial locations of trapped badgers and the clustering coefficient calculated as described above. This procedure was carried out 10,000 times in order to give a null distribution of the clustering coefficient. The observed clustering coefficient from the true infection status was then compared to this null distribution and the percentile of the null distribution at which the true clustering value fell was recorded. The true value was deemed to be significantly different from the null distribution where the observation lay in the upper or lower 2.5% of the permutated null distribution. To visualise this, the Z score (the number of standard deviations of the true value from the mean of the null distribution) was calculated and plotted for each year.

To investigate whether the variation in clustering observed was related to spatial location, social groups were assigned to geographic zones within the study area (Fig 5.6). The clustering score (as defined above) was calculated for

each zone independently. A general linear model was used to test the influence of spatial zone and year on clustering score, with an interaction term included.

5.4 Results

Between 1985 and 2005, 11,349 *post mortem* and/or trapping records were collected from 2408 individual badgers. Population prevalence increased significantly with study year ($F_{(19,20)} = 34.7$, $P<0.05$) but was not significantly associated with contemporary estimates of population size ($F_{(18,19)} = 3.4$, $P=0.08$) (see Fig 5.1a) and there was no evidence of an interaction between year and estimated population size ($F_{(17,18)}=0.003$, $P=0.9$). The total number of badgers captured in each infection category by year is shown in Figure 5.1b.

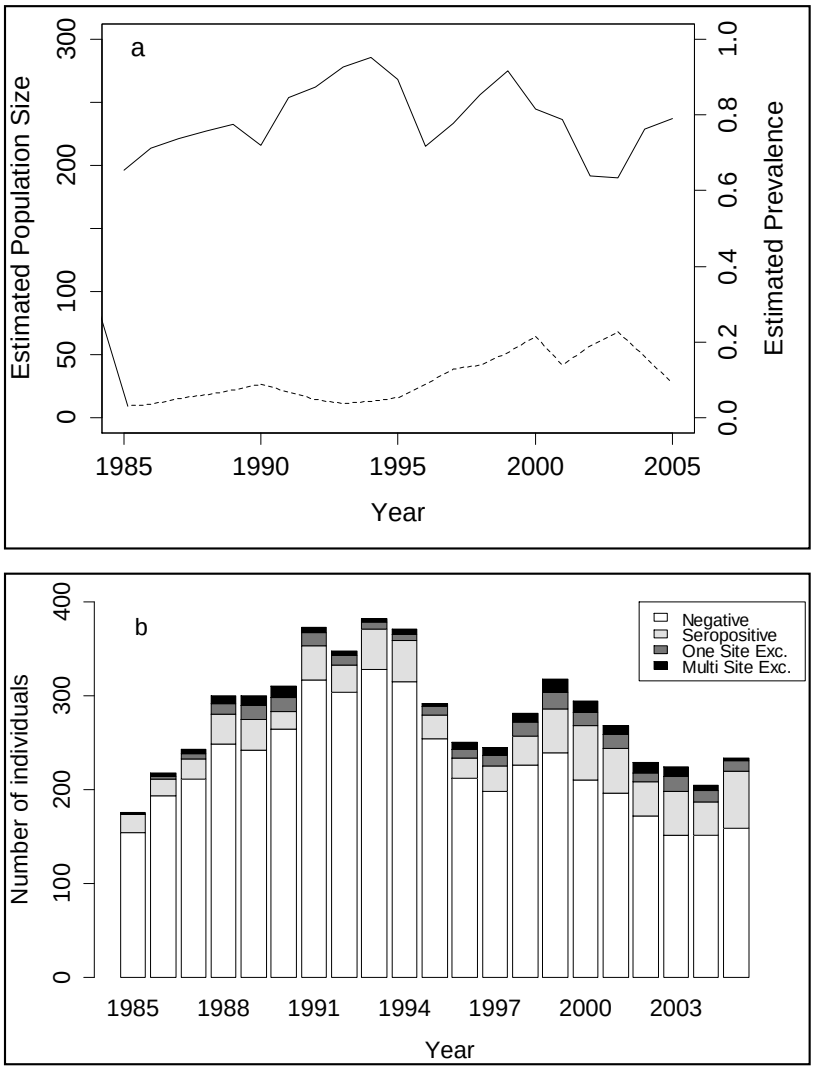
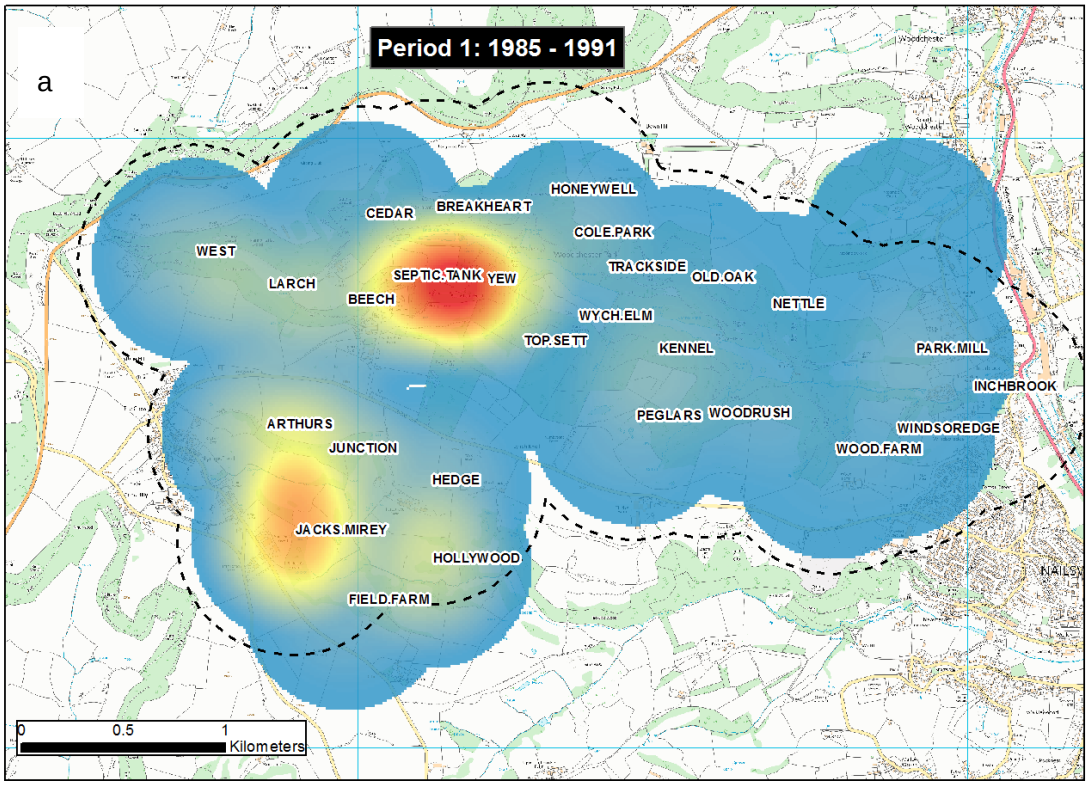


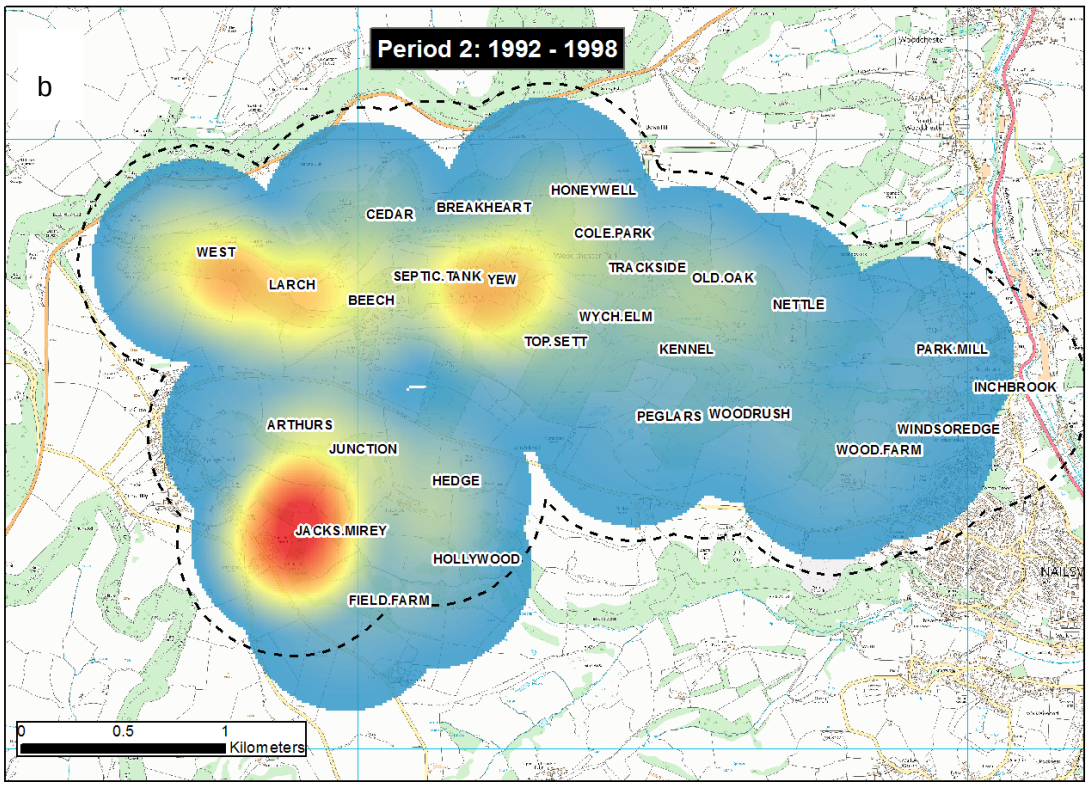
Fig 5.1 (a) The estimated trends in TB prevalence (dotted line; (Walker 2012)) and population size (solid line; (McDonald 2014)) for the Woodchester Park badger population. (b) The total number of badgers captured in the study area each year and the frequency in each TB infection class.

Visualisation of Infection Distribution

The kernel maps in Figures 5.2 (a-c) show the distribution of test positive animals in the study period over three different time periods. As these maps each cumulatively represent a time-span of 7 years, individual test positive animals will in some cases contribute multiple times to the observed output.

The spatial distribution of *M. bovis* infection in this study population between 1985 and 1998, as described previously (Delahay, Langton et al. 2000) is illustrated in Figs 5.2a and b. Test positive badgers were spatially clustered in the western region of the study area with very few in the eastern section. During Period 2 (1992-1998) infected cases start to appear in the central part of the study area, around Honeywell, Cole Park and Old Oak setts although the hotspot of infection centred on Jacks Mirey sett persists. Between 1999 and 2005, infection spread throughout the study area with the previously infection free eastern area, 'warming up' considerably. The previous hotspots of infection in the western and southern regions also persisted.


a
Period 1: 1985 - 1991
HONEYWELL
CEDAR
BREAKHEART
COLE.PARK
WEST
TRACKSIDE
LARCH
SEPTIC.TANK
YEW
OLD.OAK
BEECH
NETTLE
WYCH.ELM
TOP.SETT
KENNEL
PARK.MILL
INCHBROOK
PEGLARS
WOODRUSH
ARTHURS
WINDSOREDGE
JUNCTION
WOOD.FARM
HEDGE
JACKS.MIREY
HOLLYWOOD
FIELD.FARM
0
0.5
1
Kilometers


b
Period 2: 1992 - 1998
HONEYWELL
CEDAR
BREAKHEART
COLE.PARK
WEST
TRACKSIDE
LARCH
SEPTIC.TANK
YEW
OLD.OAK
BEECH
NETTLE
WYCH.ELM
TOP.SETT
KENNEL
PARK.MILL
INCHBROOK
PEGLARS
WOODRUSH
ARTHURS
WINDSOREDGE
JUNCTION
WOOD.FARM
HEDGE
JACKS.MIREY
HOLLYWOOD
FIELD.FARM
0
0.5
1
Kilometers

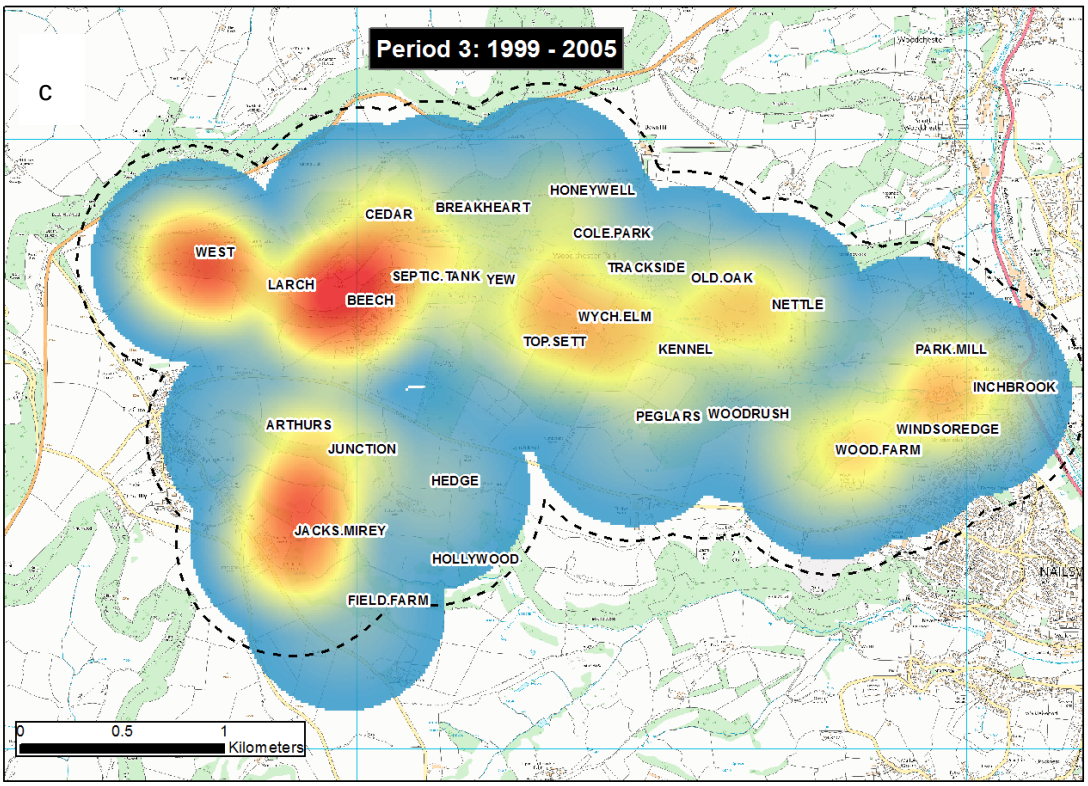


Fig 5.2 *(a – c)* Kernel density maps of the spatial distribution of TB test positive badgers at Woodchester Park between a) 1985-1991, b) 1992-1998 and c) 1999-2005. Warmer colours indicate more test positive individuals present or the presence of individuals at a more progressed disease state.

Spatio-Temporal Structure

Variograms displaying spatio-temporal autocorrelations in TB index are similar in the two study periods (Figure 5.3), with dissimilarity increasing rapidly with distance in space and time. However, the second period shows less dissimilarity at medium and long distances and at increasing time-lags; this is consistent with *M. bovis* infection becoming less clustered, yielding a pattern of greater similarity in TB index scores amongst social groups across the whole population.

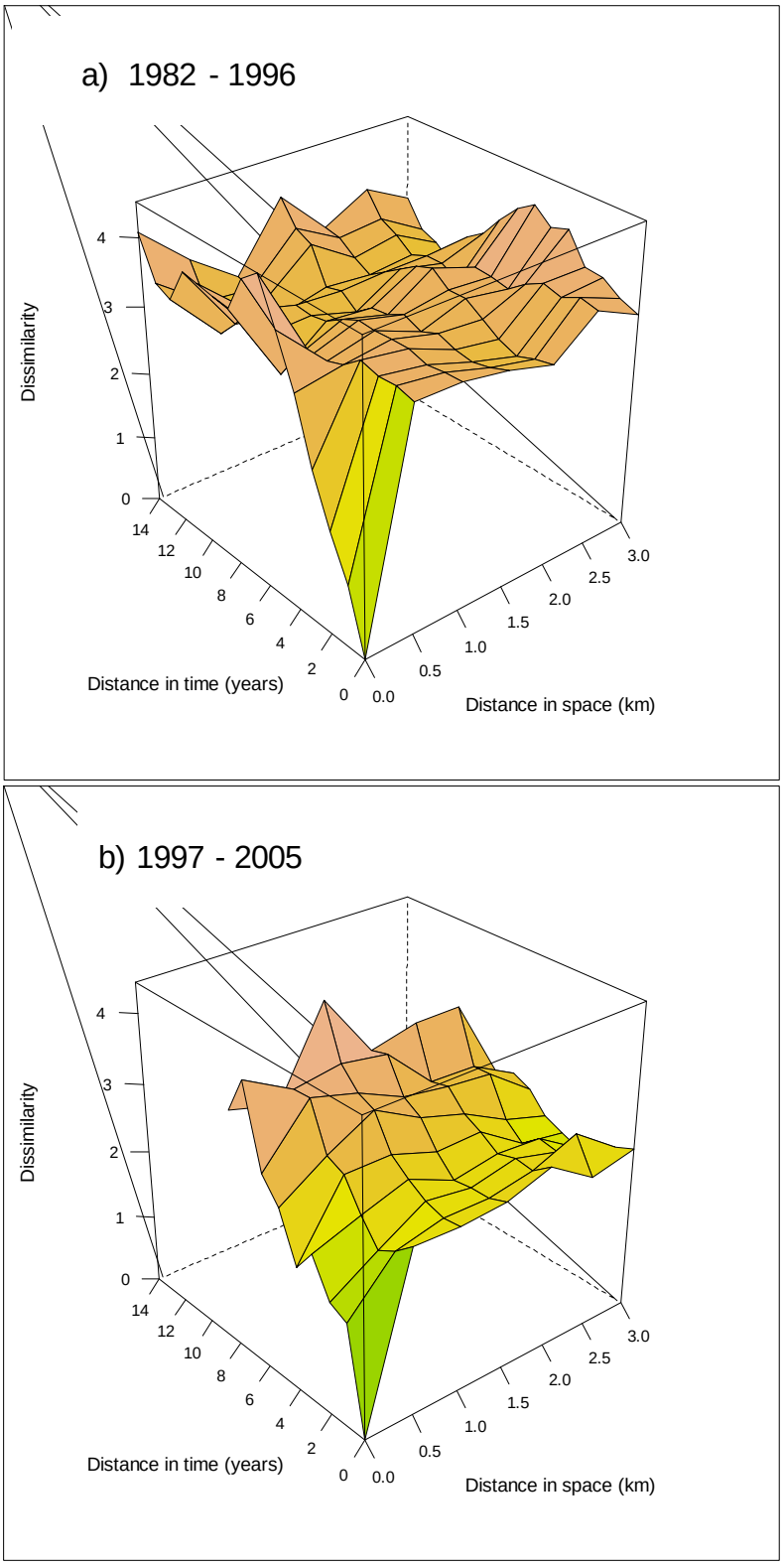


Fig 5.3 Spatio-temporal variograms of dissimilarity in TB index between social groups separated in space and time

Prior to 1998, significant spatial autocorrelation was not detected, however between 1998 and 2000 spatial autocorrelation was found in all three years at significant or close to significant levels (1998, likelihood ratio statistic = 3.08, P = 0.08; 1999, likelihood ratio statistic = 6.41, P = 0.01; 2000, likelihood ratio statistic = 3.36, P =0.07). No significant spatial autocorrelation was identified between 2001 and 2005 (see Table 5.2). However, when correction for multiple testing was applied based on Benjamini False Discovery Rate (Benjamini and Hochberg 1995), no significant spatial autocorrelation was noted in any year.

Table 5.2 Results of spatial autocorrelation testing by year, using linear mixed effects models for TB index data (1985 – 2005). Where significant, or close to significant spatial autocorrelation structures were identified, these are highlighted in bold.

Year	Spatial Autocorrelation Kernel	Non-corrected P value: spatial models vs null model	Corrected P value (spatial models vs null model)
1985	All NS	1	1
1986	All NS	1	1
1987	All NS	0.52 – 0.76	1
1988	All NS	0.55 - 1	1
1989	All NS	0.13 – 0.20	0.72 – 0.87
1990	Gaussian	0.07	0.73 – 0.74
1991	All NS	0.23 - 0.40	1
1992	All NS	0.94 - 1	1
1993	All NS	1	1
1994	All NS	1	1
1995	All NS	0.68 - 1	1
1996	All NS	1	1
1997	All NS	0.98 - 1	1
1998	Gaussian	0.08	0.73
1999	Gaussian	0.01	0.73
2000	Exponential	0.07	0.73
2001	All NS	1	1
2002	All NS	1	1
2003	All NS	1	1
2004	All NS	0.40 - 1	1
2005	All NS	0.21 – 0.43	1

Clustering of Infection

A general trend of reducing clustering of infected badgers was observed over the study period, as Z scores moved closer to the mean of the null distribution (Fig 5.4). The observed reduction in clustering is therefore greater than that expected by chance.

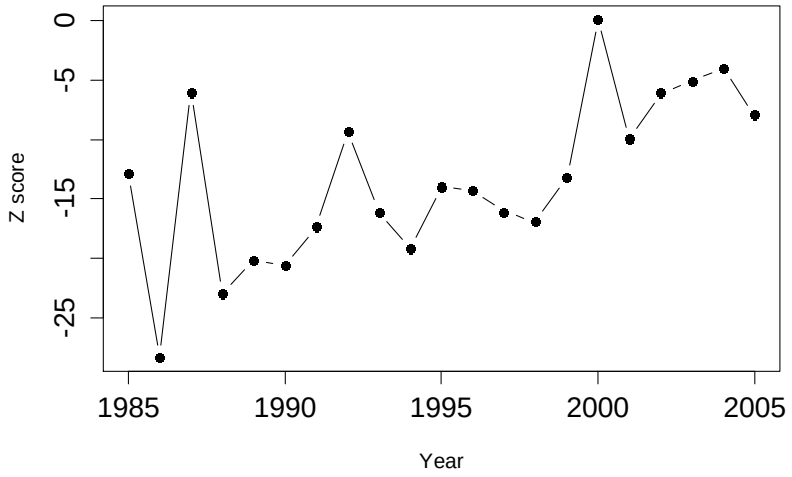


Fig 5.4 The trend in dispersion of TB test positive badgers in the Woodchester Park population as revealed by Z-scores. The plot identifies a gradual increase in dispersion (reduction in clustering) of infection through time, independent of expectations due to changes in prevalence. The Z-score on the y-axis describes how many standard deviations the observed clustering index lies from the mean of a null distribution created by shuffling spatial positions of infected and uninfected badgers. A Z-score of zero indicates random dispersion of infected cases.

When variation in the extent to which *M. bovis* infections are spatially clustered within the Woodchester population was investigated, the top model identified through model averaging included an interaction between year and location, and was more than 2 AIC units lower than the null model (Table 5.3). However, Table 5.4 indicates that in most cases parameter estimates span 0, indicating that effects are not consistent and directional. The exception to this is the interaction between location and year in the east of the study area (see Fig 5.5), where clustering is weakening more rapidly than in the reference zone (South). This could be due to infected individuals moving into social groups in the east of the study area from outside at a higher rate than in the southern social groups, resulting in a less clustered disease distribution. Alternatively, this trend could be indicative of a higher incidence of *M. bovis* infection over time in the eastern

social groups, potentially indicating that the transmission of *M. bovis* infection between social groups is happening more frequently in the eastern social groups.

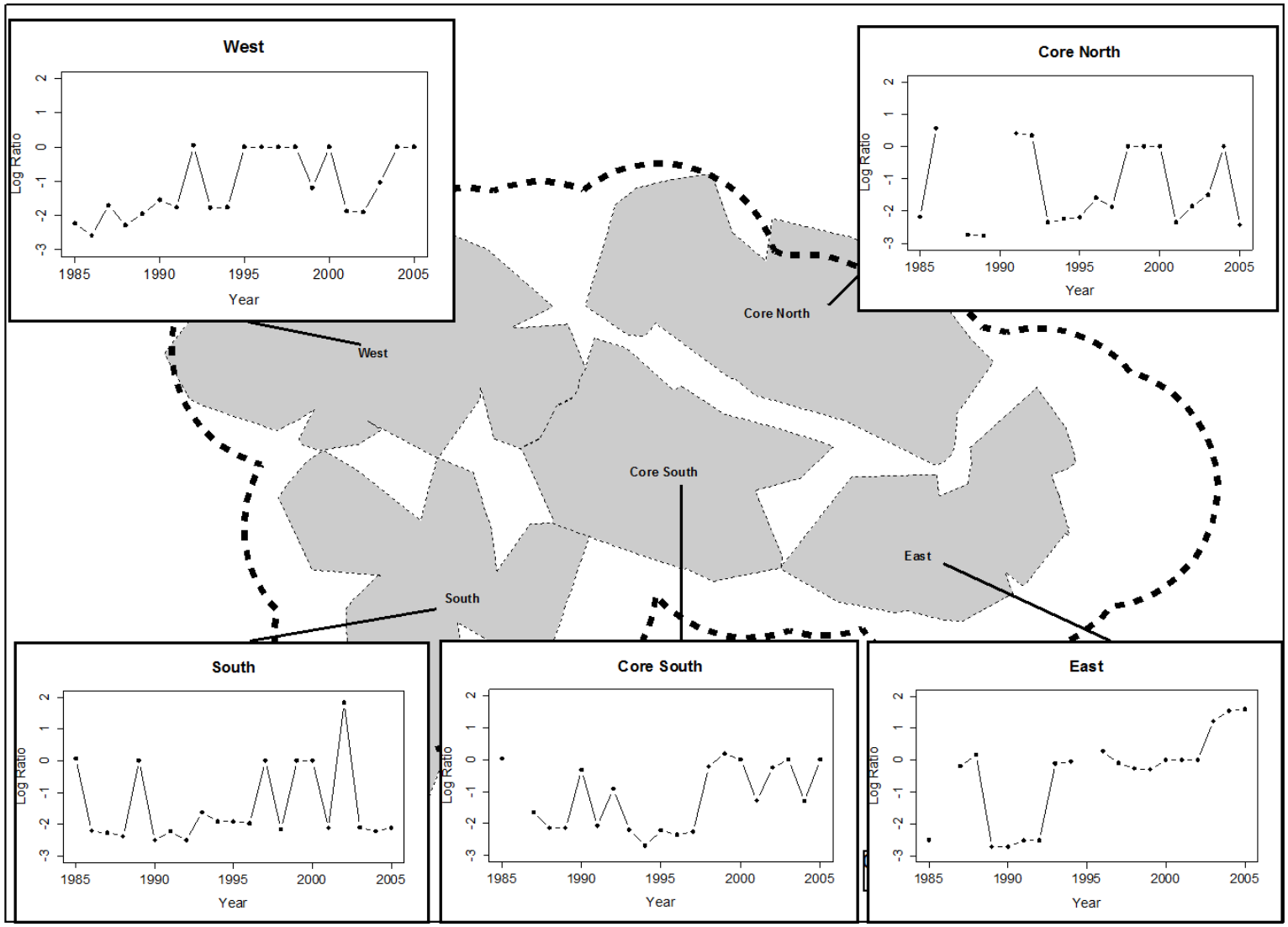
Table 5.3 Details of top models with a ΔAICc<6 predicting the clustering of *M. bovis* infections in the Woodchester population in a given year. Year, location and their interaction were included as fixed effects. Each row in the table indicates a model, with a + indicating the inclusion of a given variable. Degrees of freedom, ΔAICc, model weight and adjusted R^2 values are also included for each of the top models, with the null model included for comparison.

Model	Location	Year	Location: Year	df	ΔAIC_C	ω_i	Adjusted R^2
1	+	+	+	11	0.00	0.37	0.21
2		+		3	0.20	0.33	0.12
3	+	+		7	0.38	0.30	0.16
Null model				2	11.97	0.001	N/A

Table 5.4 Average model coefficients and relative importance of variables included in the top model set (AICc of ≤6) explaining variation in the extent to which *M. bovis* infections are spatially clustered within the Woodchester population. Parameter names with brackets show the effect of that parameter category in relation to the reference category (Location = South). The intercept value indicates the coefficient estimate where location is 'South' and the year is '1985'. Parameters highlighted in bold are those with confidence intervals that do not span zero, indicating a consistent directional effect on disease clustering.

	Coefficient Estimate	Std. Error	2.5%	97.5%	Relative Importance
Intercept	-112.1	68.5	-247.5	23.4	
Location (Core North)	23.0	85.4	-146.6	192.6	0.67
Location (Core South)	-37.6	89.3	-214.8	139.5	0.67
Location (East)	-151.8	161.5	-469.4	165.9	0.67
Location (West)	-66.6	100.5	-265.2	132.1	0.67
Year	0.06	0.03	-0.01	0.12	1
Core North : Year	-0.02	0.06	-0.13	0.09	0.37
Core South : Year	0.03	0.06	-0.08	0.15	0.37
East : Year	**0.14**	**0.06**	**0.03**	**0.25**	**0.37**
West : Year	0.06	0.05	-0.05	0.17	0.37

Fig 5.5 Fluctuations in disease clustering scores in different zones of the study area. Missing points indicate that no infected individuals were trapped in that zone in a given year, hence the clustering coefficient could not be calculated.



5.5 Discussion

Demographic and behavioural changes in badger populations which have been subjected to population management have been widely demonstrated (Carter, Delahay et al. 2007) (Riordan, Delahay et al. 2011), with associated implications for the distribution of *M. bovis* infection (Jenkins, Woodroffe et al. 2007). However, using data from a long term study of an unmanaged population of badgers, I found that the distribution of *M. bovis* infection can change dramatically even in the absence of management. This demonstrates that the accepted understanding of *M. bovis* infection distribution in unmanaged badger populations as stable and persistent (Delahay, Langton et al. 2000), with limited spread between social groups, may not represent the full picture. Analysis of two decades of data on *M. bovis* distribution in an unmanaged badger population suggests a more dynamic picture, where the spatial distribution of infection and the degree to which infection is clustered, can alter substantially, as a result of natural processes.

A reduction in, or lack of spatial clustering has been noted in other studies of *M. bovis* distribution in badger populations (Olea-Popelka, Griffin et al. 2003, Jenkins, Woodroffe et al. 2007); however this has always previously been in culled populations, in which the social structure is seriously disrupted. Social disruption caused by culling increases the home range of surviving individuals, which is hypothesized to result in increased contact rates between individuals and elevated opportunities for disease transmission (Riordan, Delahay et al. 2011). Additionally, it has been suggested that the stress caused at an individual level by this social disruption may result in immunosuppression and enhanced disease expression (Macdonald, Riordan et al. 2006, Riordan, Delahay et al. 2011). Badgers which survive culling operations may be more likely to move between social groups (Tuyttens, Macdonald et al. 2000, Riordan, Delahay et al. 2011). Territories of social groups are more likely to overlap in populations which have been culled (Tuyttens, Delahay et al. 2000). In the undisturbed population at Woodchester Park, which has not been subject to culling, it is interesting to note a breakdown in disease clustering. However, the rate of decline in clustering observed in this undisturbed population was lower than that observed in populations which had been subjected to culling

operations. Clustering (as measured by the percentage difference between the nearest neighbor distance between infected badgers and the nearest neighbor distance from uninfected badgers to the nearest infected badger) was predicted to decline by 15% with each annual cull (Jenkins, Woodroffe et al. 2007), compared to a 3% reduction per annum in Woodchester Park. However, my finding of a reduction in disease clustering in an unmanaged population provides important baseline information on natural disease dynamics, from which to assess the impact of population management.

It is feasible that natural ecological events could also generate behavioral and demographic changes similar to those observed in post- culled badger populations: increased movement between groups, increased ranging behavior, increased contact rates between members of different social groups, and increased territory overlap. Recently, it has been suggested that perturbation effects may occur when even a small number of individuals, or even a single individual is removed from a social group (Bielby, Donnelly et al. 2014). The natural loss of individuals from social groups, for example through natural, stochastic mortality, might therefore be enough to disrupt the stability of TB levels within a social group. In other wildlife disease systems, such as Chronic Wasting Disease (CWD) in deer and *M. bovis* infection in white-tailed deer, age and sex specific prevalence trends have been noted (O'Brien, Schmitt et al. 2002, Miller and Conner 2005, Grear, Samuel et al. 2006). These findings illustrate how natural changes in the demographic structure of a population can lead to changes in observed prevalence patterns. Demographic changes in the population could result in increased movement, which has been linked to increased transmission, both in this system (Rogers, Delahay et al. 1998, Vicente, Delahay et al. 2007) and other wildlife disease scenarios (Cross, Lloyd-Smith et al. 2004, Clements, Hygnstrom et al. 2011). Badger social groups occasionally merge together, which may be in response to availability of mates or resources within territories (Robertson, Palphramand et al. 2015). Potentially infectious contacts between members of different social groups can occur during extra-group mating; this is thought to be favoured where within-group genetic relatedness is higher and may represent an inbreeding avoidance strategy (Annavi, Newman et al. 2014). Food availability may affect the size of an individual's home range, seasonal variation in individual home range has

been observed and food availability is considered to be the driver of this (Palphramand, Newton-Cross et al. 2007). In dry years when food is scarcer, badgers may range more widely in search of food which may increase their contact rate with members of other social groups, affording opportunities for disease transmission. This effect has been noted in other systems, for example .*M. bovis* infection in African buffalo where drought conditions favoured population mixing, potentially enhancing disease transmission (Cross, Lloyd-Smith et al. 2004).

In the second time period considered (1997 – 2005), badger social groups were more homogenous in terms of their infection status. Again, this is consistent with a breakdown in clustering and may be explained by increased mixing between badger social groups during this period. This period is coincident with the identification of a localised phenomenon of immigration and depopulation in four contiguous social groups in the 'Core North' zone of the study area in the late 1990s (Delahay et al., unpublished). In these groups cub recruitment fell and immigrant adults were trapped at a higher rate than elsewhere in the study area and were more likely to go on to exhibit progressive disease. The proportion of immigrants that were associated with these unstable groups was particularly high after 1999, and coincided with a decline in cub numbers and an increase in the prevalence of M. bovis excretors (Delahay et al., unpublished). It may be that the social disruption caused by this unusual but natural social perturbation resulted in a change in spatio-temporal dynamics at this time, either through enhancing disease transmission within these disrupted groups or potentially through seeding this part of the population (which had previously not had high levels of TB infection) with disease.

More complete elucidation of the demographic and epidemiological processes that have yielded the observed spread of infection through this population, will require molecular epidemiological approaches (Benton, Delahay et al. 2014). For example, if *M. bovis* infection has spread spatially from the west to the east of the study area over the period of interest (Fig 2), ancestral state reconstruction (Cunningham, Omland et al. 1998) of genotyped *M. bovis* isolates may indicate that 'older' strains of the bacteria are isolated from the west of the study area and eastern strains are derived from these. An

alternative to this process of spread is that infection may have been seeded into the population on multiple occasions from outside the core study area. The introduction of new infection from outside the study population, or the immigration of infected individuals into the population, may explain why clustering broke down more quickly in the eastern edge of the study area (Fig 5). Whole genome sequence data are available for a group of *M. bovis* isolates collected from badgers in this population, which will allow us further insight into transmission dynamics (see Chapter 6).

In conclusion, I have demonstrated that even in an undisturbed population, the spatial dispersion of this chronic disease is dynamic, and that a reduction in disease clustering, as has been previously observed in culled populations (Woodroffe, Donnelly et al. 2005, Jenkins, Woodroffe et al. 2007), can occur as a result of natural ecological processes. It is important to note that the magnitude of reduction in disease clustering observed here, under natural conditions, is lower than that observed previously in response to culling operations (Jenkins, Woodroffe et al. 2007). However, I have shown that even in undisturbed badger populations, patterns of TB prevalence, distribution and dispersion are dynamic, and it is recommended that these dynamics be considered during the design of bTB management programmes. Beyond this context, these findings add to the weight of evidence generated from a variety of wildlife disease systems, highlighting the importance of taking into account the impact of natural demographic changes on prevalence patterns, in order to refine management strategies and allow more accurate forecasting of disease trends (Miller and Conner 2005).

CHAPTER 6: Data Chapter

CHAPTER 6: Genetic evidence of demographic changes within an unmanaged badger population; implications for TB transmission

6.1 Abstract

Host population structure and dispersal patterns are known to influence disease transmission, hence fluctuations in these patterns may have epidemiological consequences. Here I characterise temporal fluctuations in the genetic population structure of an intensively studied, unmanaged, high density population of European badgers (*Meles meles)*; the key wildlife reservoir of bovine tuberculosis, *Mycobacterium bovis*, in the UK and Republic of Ireland. Using a combination of observational data from two decades of a capture-mark-recapture programme, alongside genetic metrics obtained from microsatellite analysis, I demonstrate the presence of fine-scale genetic population structure and highlight heterogeneity in demographic trends, with important implications for *M. bovis* incidence. Examination of the genetic structure suggests that the population has undergone a period of demographic flux, characterised by a loss of genetic isolation by distance between badger social groups, which I suggest has resulted in a change in the spatial distribution of *M. bovis* infection within the population.

6.2 Introduction

Social structure, host movement and individual behaviour are important determinants of disease transmission in wildlife populations (Delahay, Smith et al. 2008). Genetic sampling of individuals is increasingly being employed to provide valuable information on these phenomena which may otherwise be difficult or impossible to capture (Huck, Frantz et al. 2008, Banks and Peakall 2012). Monitoring population genetics over time can help to identify changes in movement patterns and demographic structure, either in response to natural events or management interventions, which may have important implications for disease dynamics (Pope, Butlin et al. 2007). The genetic structure of populations is not always consistent with the geographical proximity of individuals. A seemingly continuous population of individuals may actually be genetically structured due to unidentified barriers to gene flow (Evanno, Regnaut et al. 2005).

Dispersal of individuals is a key ecological process, impacting on conservation genetics and population demography (Robertson, Chilvers et al. 2006, Harrison, York et al. 2014), and is a key factor in the epidemiology of infectious disease (Hess 1996). Dispersal has been shown to vary between the sexes in a wide range of species (Pusey 1987, Goudet, Perrin et al. 2002, Banks and Peakall 2012, Harrison, York et al. 2014), and amongst individuals of different age classes (Harris, Caillaud et al. 2009) and social positions (Bekoff 1977, Ekman, Eggers et al. 2002). Heterogeneity in movement patterns can have important consequences for disease transmission, for example a few dispersing individuals may be responsible for the majority of pathogen spread across a landscape (Smith, Rand et al. 1996). Understanding the drivers of individual dispersal therefore has important consequences for pathogen transmission. Demographic or ecological change may lead to spatial and temporal variation in dispersal within species (Banks and Peakall 2012), however such processes are only likely to be captured using longitudinal datasets. Due to the logistical constraints in collecting field data, dispersal may be inferred entirely from genetic analyses (Banks and Peakall 2012, Harrison, York et al. 2014). Where observational data are available, they are often collected over a short-time period and may therefore represent transient rather than long-term patterns (Harrison, York et al. 2014). If observational data are only available from a

short time period, genetic approaches are considered to better represent these long-term patterns in dispersal (Goudet, Perrin et al. 2002, Banks and Peakall 2012, Harrison, York et al. 2014). In the current study the availability of detailed data from a longitudinal capture-mark-recapture study provides the rare opportunity to compare direct, observational data with genetic data from the same population (Winters and Waser 2003).

The European badger *(Meles meles)* is the key wildlife reservoir of bovine TB in the UK and the Republic of Ireland (Delahay, De Leeuw et al. 2002, Abernethy, Upton et al. 2013). In badger populations dispersal is considered to be determined by an interaction of individual, social and environmental factors (Frantz, San et al. 2010). Sex-biased dispersal in both directions has been documented (female (Woodroffe, Macdonald et al. 1995, Tuyttens, Delahay et al. 2000) and male biased (Kruuk and Parish 1987, Cheeseman, Cresswell et al. 1988, Rogers, Delahay et al. 1998, Roper, Ostler et al. 2003)). Population density is thought to be important in determining dispersal patterns, with delayed dispersal predicted when all suitable surrounding habitat is occupied (Frantz, San et al. 2010). Many individuals do not disperse from their natal social group (Roper 2010), resulting in spatial clustering of related individuals and hence pronounced genetic structure (Pope, Domingo-Roura et al. 2006). Where dispersal does occur, it is generally to a neighbouring social group, although longer distance movements crossing several territories have been documented (Roper 2010).The removal of individuals from the population, for example through badger culling to attempt to control TB transmission to cattle, has been documented to result in detectable changes in the genetic population structure, as surviving individuals range more widely and encounter previously spatially remote individuals (Pope, Butlin et al. 2007). This increase in badger movement has been linked to increases in *M. bovis* infection in badgers and cattle (Carter, Delahay et al. 2007) and is associated with a decrease in spatial clustering of infection in badgers (Donnelly, Woodroffe et al. 2003, Donnelly, Woodroffe et al. 2005, Donnelly, Wei et al. 2007, Jenkins, Woodroffe et al. 2007).

Chapter 5 of this thesis described a temporal change in the spatial distribution of *M. bovis* infection in the Woodchester Park badger population. Initially (1982-1996), *M. bovis* infection remained spatially clustered, primarily within badger

social groups in the south and west of the study area. However, subsequently (1997-2005), infection became more widespread, characterised by a reduction in the clustering of infected hosts and greater similarity in measures of TB infection at the social group level. As a result cases of *M. bovis* infection were no longer restricted to the western region of the study area and a new hotspot appeared in the previously low prevalence eastern region. Increases in movement between social groups, ranging behavior, contact rates between members of different social groups, and territory overlap may have led to the observed reduction in *M. bovis* clustering (Chapter 5).

A meta-analysis of genetic data from European badger populations across Europe demonstrated that genetic isolation by distance (i.e. spatially proximate individuals tend to be genetically closer than spatially distant individuals) is the norm over a range of spatial scales (Pope, Domingo-Roura et al. 2006). However, this meta-analysis focused generally on a wider spatial scale than considered in the current study (500m – 30 km) and did not investigate within population changes in genetic structure over time. The current chapter aims to characterise the fine-scale genetic population structure in the Woodchester Park population over a two decade period and, in combination with detailed observational data, to consider whether this structure has changed over time. Genetic and observational data were used in tandem to consider whether dispersal patterns had changed over time and whether this was consistent with temporal demographic change driving the observed change in TB spatial distribution described in Chapter 5.

6.3 Methods

Badger Sampling

All data used in these analyses were collected from the long-term capture-mark-recapture and epidemiological sampling study of a wild badger population at Woodchester Park in Gloucestershire (for full details see Appendix). As limited genotype data were available in 1990 and 1991, and genotypes from 2011 to the present were not available, only the period 1992-2011 was considered. Analyses were restricted to adult badgers (i.e. > 1 year old) in order to control for fluctuating cub numbers between years. This was necessary because if the proportion of cubs in the population increased in a particular year, the population would appear to be more genetically differentiated as, in most cases, cubs remain in their natal group in their first year of life.

Host Genotyping

Genotyping of hair samples was carried out on samples obtained from trapped badgers from 1992 until 2011. On first capture, a hair sample was routinely taken from each trapped badger, and stored in 80% ethanol before being submitted for DNA extraction and genotyping (Carpenter, Pope et al. 2005). All genotyping data were generated by the Molecular Ecology Lab, University of Sheffield. Genotyping data were derived from 22 microsatellite markers, each with 4-7 alleles.

The MicroDrop Programme (Wang and Rosenberg 2012) was used to impute missing microsatellite data. Deviations from Hardy-Weinberg equilibrium for each of the 22 microsatellite markers were tested on the imputed dataset using the hwtest function in ‘adegenet’ and none were identified. The Bartlett test of homogeneity in the same package was used to confirm homogeneity of variance among loci ($P = 0.78$).

Host Genetic Population Structure

Population structure in the Woodchester badger genotypes was analysed using the open-source population genetics programme ‘STRUCTURE’ (v2.3.4) (Pritchard, Stephens et al. 2000), which uses a Bayesian clustering method to identify linkage disequilibrium in unlinked loci caused by true population

structure. The aim was to identify the smallest number of clusters of genotypes (K) within the population that captured the major structuring in the data and then to assign individuals to these clusters (Evanno, Regnaut et al. 2005). Previous studies on badger population genetics have shown that extra-group mating is commonplace (Carpenter, Pope et al. 2005), therefore at the small spatial scale of Woodchester Park pronounced genetic population structure with strong divisions between clusters was not expected. For this reason individuals were allowed to be 'admixed' such that they could belong to a mixture of clusters rather than being exclusively assigned to one cluster. Allele frequencies were allowed to be correlated to account for correlations between the markers used as this has been demonstrated to provide more accurate inference of population structure (Falush, Stephens et al. 2003). Social group was not included as a prior as the aim was to understand the population structure based purely on the genetic information available. STRUCTURE simulation runs were carried out where K was set to 2-9 (20 iterations at each value of K), with a burn-in period of 10,000 followed by 30,000 MCMC reps. Results of this simulation were fed into the online STRUCTURE HARVESTER programme (Earl and vonHoldt 2012) which assesses likelihood values over different values of K from STRUCTURE analyses in order to detect the value of K that best fits the data (Pritchard, Stephens et al. 2000); this confirmed the optimum value of K to be 6. A Q matrix was generated for each genotyped individual indicating the proportional membership of that individual to each of the 6 genetic clusters. Individual 'i' has inherited some fraction of their genome from ancestors in each genetic cluster; the Q matrix provides the posterior mean estimates of these proportions.

Genetic cluster membership at the population level was visualised over time by assigning adult badgers trapped in the population in a given year to the cluster with which they had the highest proportional membership. For each year, this data was visualised in ArcGIS 10.2 in order to examine the spatial arrangement of genetic cluster membership of adult badgers, based on their social group of residence in that year. Individuals were assigned to one social group per year based on previously established assignment rules (Vicente, Delahay et al. 2007). Proportional cluster membership by social group was calculated as the proportion of social group members assigned to each of the 6 genotype

clusters. The Whittaker's beta diversity index (an ecological measure of pairwise community differentiation) (Whittaker 1960) was calculated for the displayed years in order to illustrate the extent of genetic differentiation between social groups in a given year. This analysis was restricted to social groups where more than two genotyped individuals had been trapped in a given year. To assess the significance of the Whittaker's indices in each year, a permutation procedure was conducted on the cluster assignment data to generate a null distribution in each year as follows; for the genotyped adults in a given year, dominant cluster assignments were shuffled 1000 times. From these randomised data, Whittaker's index was calculated, thus generating a null distribution of Whittaker's indices if there was no genetic differentiation between social groups in a given year. The true Whittaker's index observed from the unshuffled data was then compared to this null distribution and significance assessed (p = number of null Whittaker's index values that were greater than the observed value / 1000).

Testing for Isolation by Distance

To test whether spatially closer social groups were genetically closer (isolation by distance; IBD) in each of the years under consideration (1992 – 2011), a pairwise genetic distance matrix was constructed from the badger microsatellite data using the R package 'adegenet'(Jombart 2008). A corresponding spatial distance matrix was produced based on the XY co-ordinates of the social groups that the genotyped individuals trapped in a given year were assigned to using the R package 'ecodist' (Goslee and Urban 2007). A Mantel test permutation procedure (which generates a null distribution by randomising the matrices) was used to test for significant isolation by distance using 1000 permutations in all cases. The mean genetic distance between social groups was calculated for each year. Also, a Standardised Major Axis (SMA) regression was carried out for the spatial and genetic data in each year, using the R package 'smatr' (Warton, Duursma et al. 2012). SMA regression is favoured over standard least squares regression because it does not assign cause and effect in the relationship between genetic and spatial distance, and accepts the presence of error in both variables (Harper 2014). The spatial distances between social groups were regressed against the corresponding

genetic distances between those groups in order to obtain an intercept and slope of the regression line for each year between 1992 and 2011.

Inter-individual Heterogeneity in Probability of Moving Social Groups

To determine whether changes in genetic population structure were driven by temporal changes in individual dispersal behaviour, variation in the likelihood of an individual dispersing from their birth social group was related to the year of study and demographic class. Firstly, an analysis was conducted using a mixed effects model on the capture history data only. The probability of an individual moving in a given year was modelled as a binary variable (1 indicating that, at the given capture event, that individual had moved from the social group it was previously trapped in, and 0 indicating that it had not moved since its last capture). Year (as a factor), badger age (years) and sex were included as fixed effects, with individual ID and birth social group included as random effects. Analyses investigating heterogeneity in the likelihood of moving were restricted to capture records from adult badgers of known age (i.e. those that had first been caught as cubs in the study), which had been caught more than once (n = 3016 capture records). The number of times an individual had been caught in a given year was included as a fixed effect to account for inter-individual differences in the numbers of captures. Birth social group was assigned as the first social group a badger was caught in during its first year of life.

To investigate whether the probability of moving in a given year was related to the population size, a Pearson's correlation test was performed on the annual estimates of movement probability from the above model, and the number of individuals caught in the population.

Heterogeneity in Destinations of Group-Moving Individuals

In the second part of the analysis, the above dataset was restricted to only individuals who had moved from their birth social group (n = 590) in order to investigate heterogeneity in the destinations of moving individuals. Firstly, the capture data was examined for inter-annual differences in the spatial scale of dispersal events (measured as the straight line distance between an individual's birth group and the social group that they had moved into). A mixed effects

model was constructed including year, age and sex as fixed effects, with individual ID and birth social group included as random effects.

The above dataset was then further restricted to badgers for whom a genotype was available (n = 509). Corrected individual assignment indices (Alc); a traditional genetic metric of dispersal (Goudet, Perrin et al. 2002, Pope, Butlin et al. 2007) were calculated for each badger-capture event in the dataset using the R package 'hierfstat'(Goudet 2005). The Alc for each individual was calculated by determining the likelihood of an individual's genotype occurring at a given location, based on the allele frequencies in the population. The assignment index value was then corrected using the social group mean. In the current context, an individual more likely than average to be assigned to the social group (i.e. more likely to be a resident born into the social group) will have a positive Alc value, whereas an individual less likely than average to be assigned to the social group will have a negative Alc value (i.e. more likely to be an immigrant to the social group). A mixed effects model was constructed where the response variable was the variance between an individual's Alc value in its birth group to its Alc value in the social group it had moved to. A positive variance would suggest that an individual had moved into a similar or more related social group, whereas a negative variance would suggest that an individual had dispersed into a less related social group. Age and sex were included as fixed effects along with year, with individual ID and birth social group included as random effects.

Social Group Population Structure and *M. bovis* incidence

A previous study on this population had demonstrated that badger social groups diminishing in size were at a higher risk of containing new cases of infection than stable or expanding social groups (Vicente, Delahay et al. 2007). To investigate this further, and with the benefit of the availability of the host genetic data, a similar analytical approach was used on the current dataset, which covered a different temporal period (original study; 1989-2004, current analyses; 1992-2011). As in the original study, analyses were restricted to social groups with no resident cases of *M. bovis* in the previous year (i.e. susceptible groups, where no individuals who had tested positive to a TB diagnostic test were resident). A social group was classed as an 'incident' case

if at least one *M. bovis* positive individual was detected in any subsequent year, based on the diagnostic test procedure used at the time (for details of diagnostic test regime, see Appendix) (Vicente, Delahay et al. 2007). In contrast to the previous study, cubs were excluded from the social group size calculations. This was in order to focus on social group size changes driven by immigration / emigration of adult individuals or retention of adults in their natal group rather than by cub recruitment. Social groups were categorised based on their change in adult group size from the previous year and the change in the mean adult social group AIc (see Table 6.1). Once social groups had been categorised for each year, as described above, a linear mixed effects model was constructed using the 'lmer' function in the R package 'lme4'. The response variable was whether the social group included an incident case in that year (binomial: 1 or 0), with category (Table 6.1) as the explanatory variable. Year and Social Group were included as random effects. *Post-hoc* testing was carried out to identify any significant differences in incidence risk between categories using the 'mcposthoc.fnc' function in the 'LMERConvenienceFunctions' package (Tremblay and Ransijn 2015).

Table 6.1 Categories of social groups based on group size and assignment index trends.

Category	Adult Group Size Trend & AIc Trend	Suggested interpretation
“Expanding -Falling”	Adult group size increases, mean AIc decreases	Group size increase, driven by adult immigration into group
“Shrinking -Falling”	Adult group size decreases, mean AIc decreases	Dispersal of adults born in the group
“Stable - Falling”	Adult group size stable, mean AIc decreases	Dispersal of adults born in the group compensated by arrival of immigrant adults resulting in stable group size
“Expanding - Rising”	Adult group size increases, mean AIc increases	Group growing in size, driven by retention of adults born in the group (natal philopatry)
“Shrinking - Rising”	Adult group size decreases, mean AIc increases	Group shrinking driven by immigrant adults leaving group
“Stable - Rising”	Adult group size stable, mean AIc increases	Immigrant adults leaving compensated by adults born in the group remaining

6.4 Results

Host Genetic Population Structure

Permutation analysis indicated that, in the majority of years, the observed Whittaker's index was significantly larger than if there were no true genetic differentiation between social groups ($p < 0.01$). The exceptional years were 1996, 1999, 2007 and 2009, where the observed Whittaker's index was not significantly different from the null distribution in that year (1996; $p = 0.06$, 1999; $p = 0.13$, 2007; $p = 0.9$, 2009; $p = 0.69$). Fig 6.1 illustrates that badger genotypes were generally clustered within social groups in 1992, with social groups more genetically differentiated from each other, as indicated by the low number of 'mixed' social groups (where badger genotypes from multiple clusters are represented) and the high Whittaker's index Z score. This is consistent with badgers from the same genotype cluster (i.e. more genetically similar, related badgers) residing together in social groups, resulting in genetically differentiated groups. In contrast, in 1999, = more genotype mixing between social groups is evident, with most groups comprising multiple badger genotypes and a lower Whittaker's index Z score, consistent with lower genetic differentiation between social groups. In the final map showing the spatial distribution of badger genotypes in 2011, a return towards the 1992 picture is seen, with a lower number of 'mixed' social groups and the higher Whittaker's index Z score.

The Whittaker's index Z score from the whole time period (Fig 6.2) shows that initially in 1992 the observed genetic differentiation between social groups is very different from the null distribution, but between 1993 and 1999 it is generally close to the null distribution. This suggests that there is weaker genetic differentiation between social groups from 1993 to 1999, consistent with higher levels of inter-group mixing. In the early 2000's genetic differentiation appears to re-establish, consistent with lower levels of inter-group mixing, although from the mid 2000's until 2010 genetic differentiation is low or non-existent, consistent with a rise in levels of inter-group mixing before re-establishing to 1992 levels by 2011.

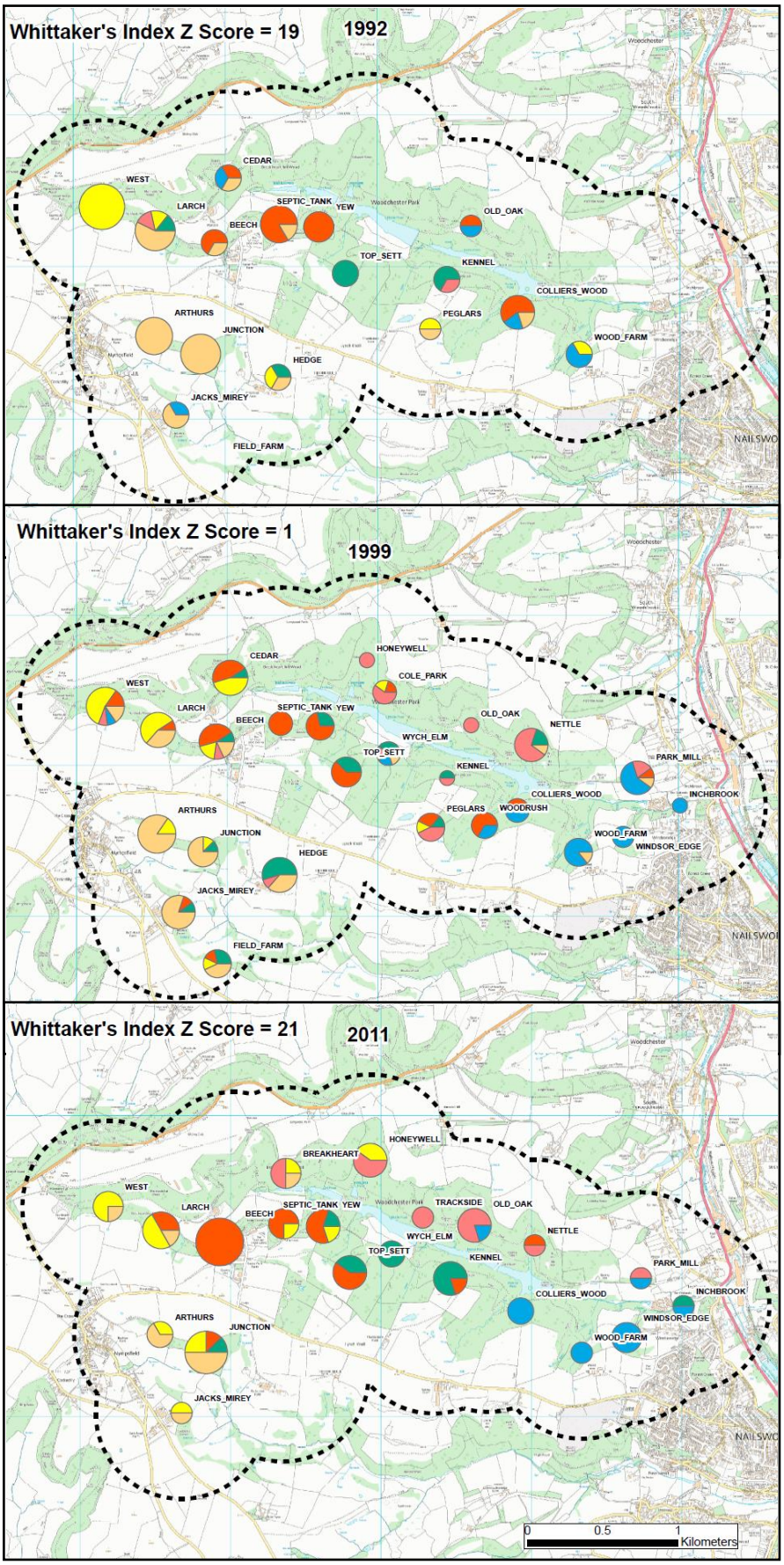


Fig 6.1 Genotype clustering at the social group level at Woodchester Park, indicating proportional membership to each of the six genotype clusters elicited from the STRUCTURE analysis in 1992, 1999 and 2011. Whittaker's diversity index is shown, with higher values

indicating more differentiation among communities. Many individuals were strongly assigned to a single genetic cluster, which has been suggested to be indicative of the presence of true population structure (Pritchard, Stephens et al. 2000).

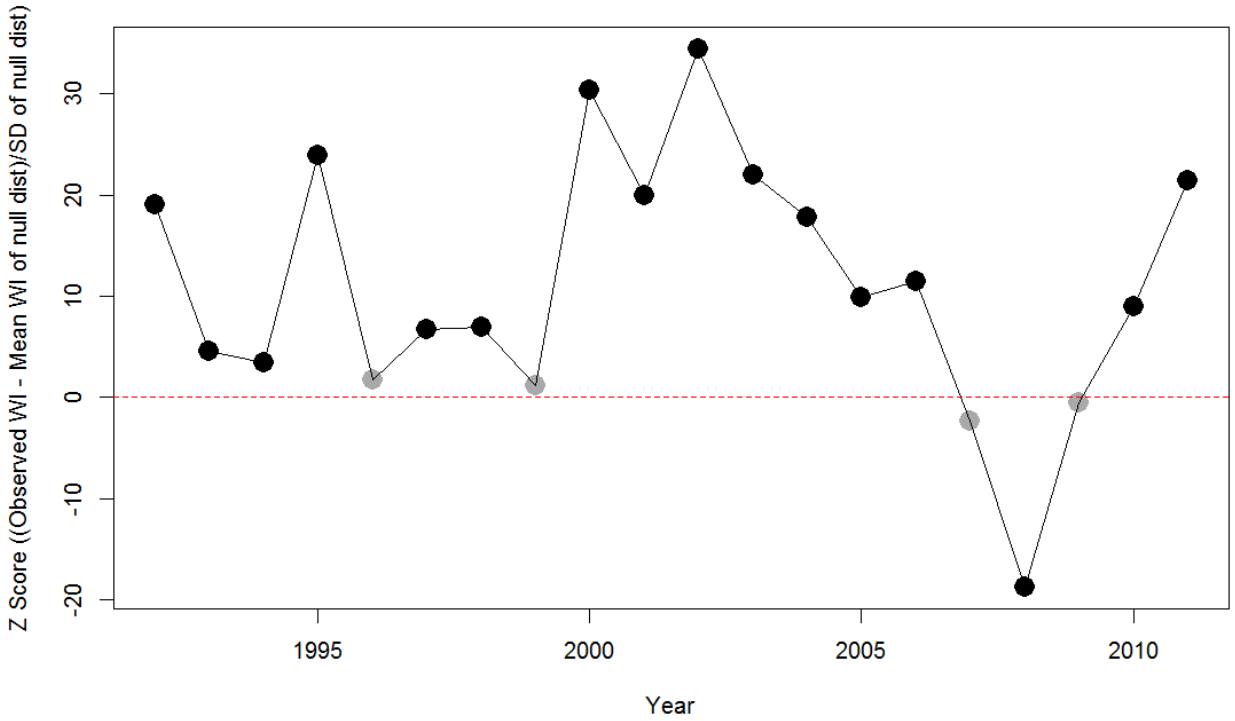


Fig 6.2 The trend in genetic differentiation between social groups in the Woodchester Park population as revealed by Z-scores. The plot reveals temporal differences in the extent of genetic differentiation, independent of expectations due to changes in population size. The Z-score on the y-axis describes how many standard deviations the observed clustering index lies from the mean of a null distribution created by shuffling genetic cluster assignments among genotyped badgers. A Z-score of zero indicates social groups are not significantly genetically differentiated from one another.

Isolation by Distance

Significant Isolation by Distance (IBD) was noted in 1992 (Table 6.2), indicating that spatially proximate social groups were more genetically similar to one another. This is consistent with extra-group mating primarily occurring between neighbouring social groups and dispersing individuals generally moving to spatially proximate groups. However, between 1993 and 2008, IBD is non-significant, indicating that spatially proximate social groups are no more likely to be genetically similar than spatially distant social groups. IBD then becomes statistically significant again from 2009 to 2011 inclusive.

Table 6.2 Results of isolation by distance analyses by year, intercept and slope values based on SMA regression, p values based on Mantel isolation by distance test procedure in 'adegenet'. Years with significant IBD are indicated in bold.

Year	Intercept (from SMA)	Slope of IBD (from SMA)	Significance of IBD (from Mantel)
1992	**0.31**	0.00012	**0.02 ***
1993	0.61	-0.00001	0.77
1994	0.64	-0.00012	0.69
1995	0.58	0.00001	0.59
1996	0.50	-0.00008	0.62
1997	0.57	-0.00008	0.93
1998	0.27	0.0001	0.32
1999	0.25	0.00009	0.08
2000	0.27	0.00009	0.10
2001	0.29	0.00009	0.47
2002	0.59	-0.00009	0.74
2003	0.61	-0.0001	0.61
2004	0.32	0.0001	0.44
2005	0.37	0.00009	0.10
2006	0.34	0.0001	0.16
2007	0.35	0.00009	0.08
2008	0.33	0.0001	0.11
2009	**0.32**	**0.0001**	**0.02 ***
2010	**0.32**	**0.00007**	**< 0.001 ***

2011	**0.32**	**0.0001**	**0.004 ***

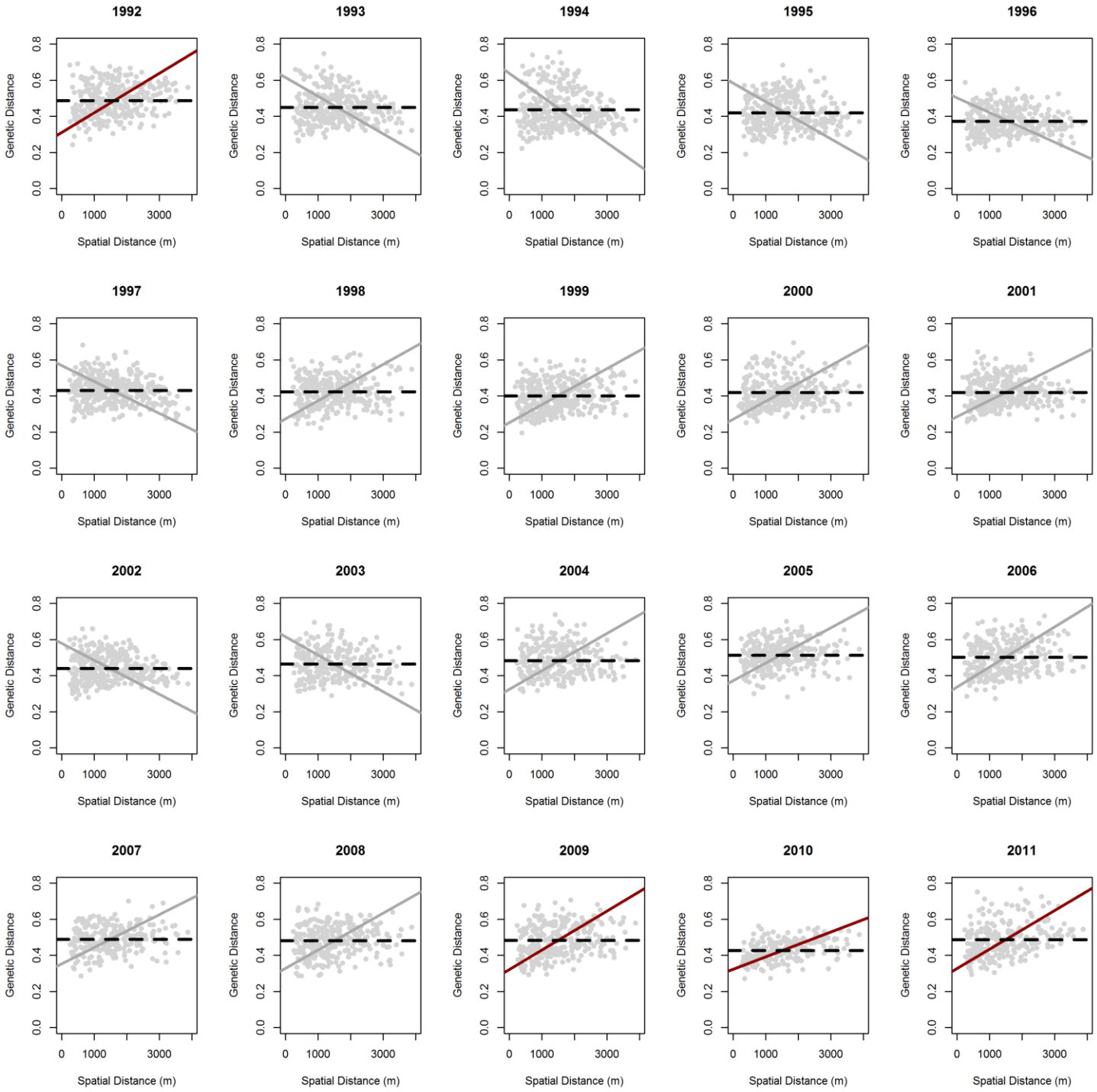


Fig 6.3 Genetic IBD amongst Woodchester badger social groups between 1992 and 2011. Raw data are displayed in grey. Significant IBD regression lines are indicated in red. Where IBD was not supported in a given year, a grey line indicates the estimated intercept from the SMA regression between genetic and spatial distances between social groups. In all years, the black dashed line indicates the mean genetic distance between social groups in that year.

Inter-individual Heterogeneity in Probability of Moving Social Groups

The probability of an individual having moved social groups in a given year was higher for males (Wald's $\chi^2_{(1)} = 44.8$, $p < 0.001$). Overall, the predicted probability of a male badger moving social groups in a given year was double that of a female badger (male predicted probability of moving social groups in a given year = 22%, female = 11%). Dispersal probabilities varied significantly among years (Wald's $\chi^2_{(19)} = 43.2$, $p = 0.001$). The probability of an individual moving social groups in a given year was unrelated to its age (Wald's $\chi^2_{(1)} = 0.02$, $p = 0.89$). In many years, the average probability of a badger moving social group was approximately 10% or less (see Fig 6.4). However, peaks in movement rates were observed from 1997 to 1998 and from 2003 to 2009, where the predicted probability of inter-group movement was up to 20%. The annual estimates of the predicted probability of moving were significantly negatively correlated to population size, such that movement was more likely at lower population densities (Pearson's correlation coefficient = -0.47, t= - 2.25, p = 0.03, see Fig 6.4).

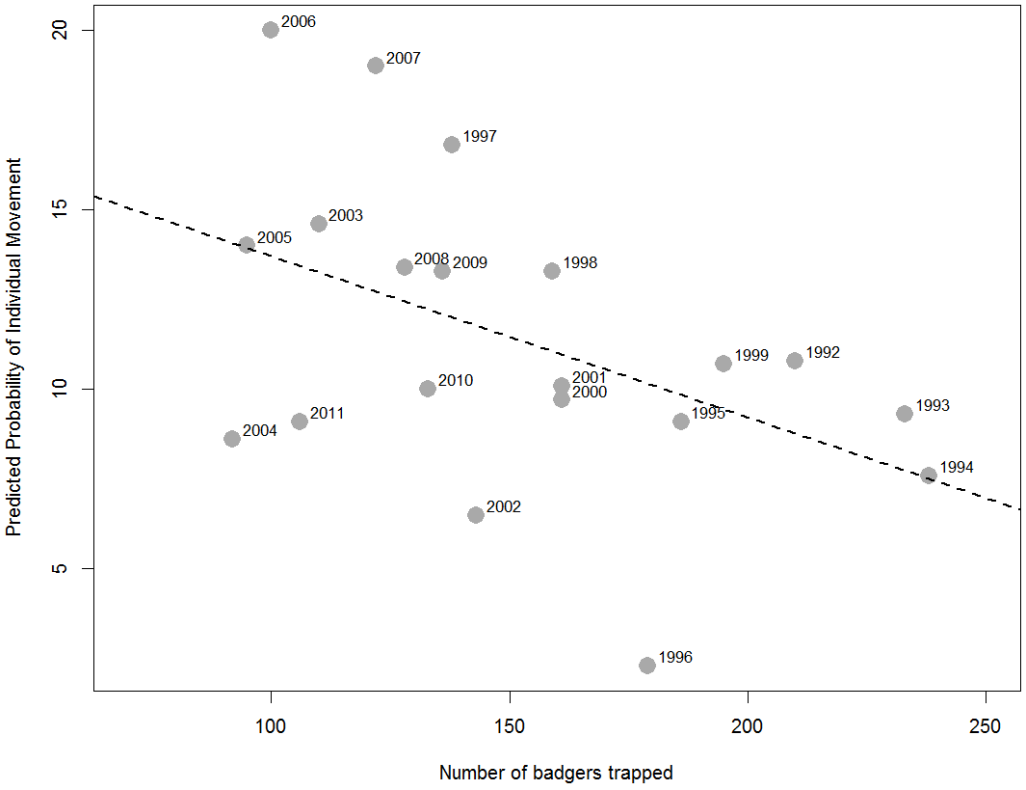


Fig 6.4 Predicted probability of an individual badger moving social groups in a given year at different population densities.

Heterogeneity in Destinations of Group-Moving Individuals

The spatial distances moved by individuals dispersing from a social group did not vary significantly between years (Wald's $\chi^2_{(19)} = 17.4$, $p = 0.56$). Males tended to move to social groups that were spatially further from their natal group than females, although this difference was not statistically significant (Wald's $\chi^2_{(1)} = 2.8$, $p = 0.10$). The distance moved by badgers was unrelated to their age at the time of movement (Wald's $\chi^2_{(1)} = 0.11$, $p = 0.9$).

In the later years of the time period considered, badgers tended to move to social groups that they were less related to than in earlier years (Wald's $\chi^2_{(1)} = 10.5$, $p = 0.001$, see Fig 6.5). However, the average genetic distance between social groups remained similar throughout this period (as indicated by the dashed black lines in Fig 6.3) which suggests that this effect was not due to changing genetic population structure. Neither age (Wald's $\chi^2_{(1)} = 0.67$, $p = 0.41$) nor sex (Wald's $\chi^2_{(1)} = 0.21$, $p = 0.65$) predicted the difference in relatedness between the destination and natal social group. These results suggest that until 1999, badgers that left their natal social group tended to move into social groups where there were related badgers already resident. As extra-group mating is thought to primarily take place between neighbouring social groups (Carpenter, Pope et al. 2005, Roper 2010) this suggests that badgers are generally moving into spatially proximate groups. However, from 2000 onwards this appeared to change, with badgers tending to move into social groups comprised of increasingly unrelated residents.

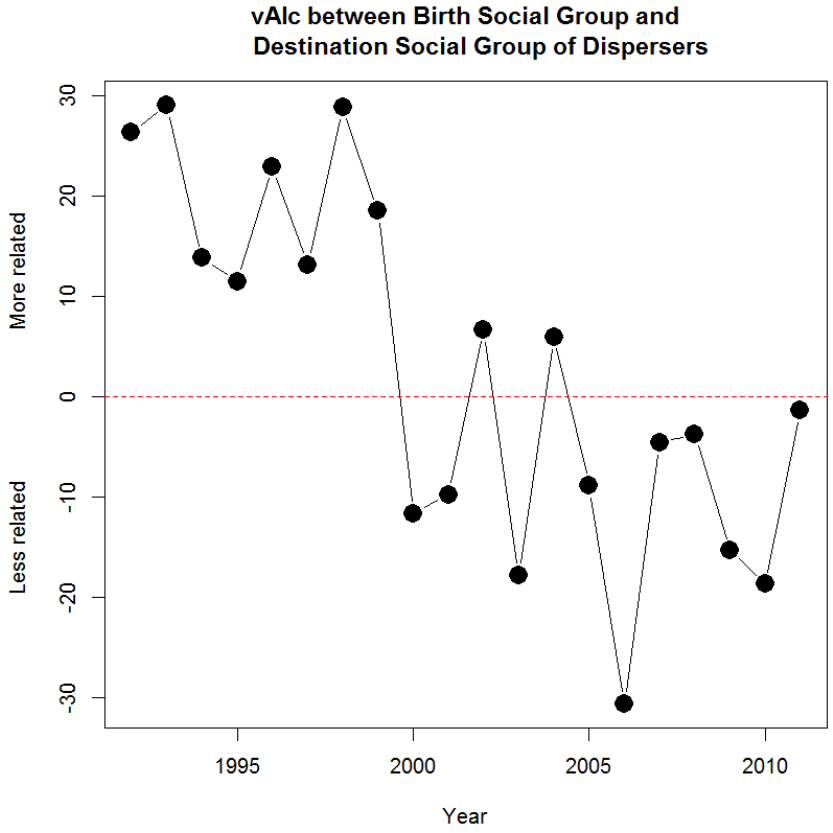


Fig 6.5 Annual estimates of variance between a badger's natal assignment index (vAIc) to its assignment index in the social group it has moved into. A negative assignment index is consistent with movement into a social group to which an individual is less genetically related.

Social Group Population Structure and *M. bovis* incidence

The likelihood of a badger social group becoming an 'incident' case was significantly higher for social groups in the 'Stable – Falling ' category (see Table 6.1) when compared to the 'Expanding -Falling' and 'Stable - Rising' categories (SF vs EF; $z_{(33)} = 2.59$, p =0.009, SF vs SR; $z_{(33)}=2.17$, p =0.03, marginal $R^2 = 0.16$, conditional $R^2 =0.53$).

Social groups in the 'Stable-Falling' category experienced the highest risk of subsequently having incident cases of TB (Fig 6.6). As illustrated in Fig 6.7, the membership flux may be greater in these social groups than that within the 'Expanding-Falling' and 'Stable-Rising' groups as a higher proportion of the social group is likely to be of immigrant origin. However, in contrast to the previous study no evidence was found of enhanced incidence risk in shrinking social groups; although the probability estimates for the shrinking group categories were higher than for the expanding group categories (see Fig 6.6; ShF vs ExF and ShR vs ExR), these differences were not significant (as indicated by overlapping error bars in Fig 6.6).

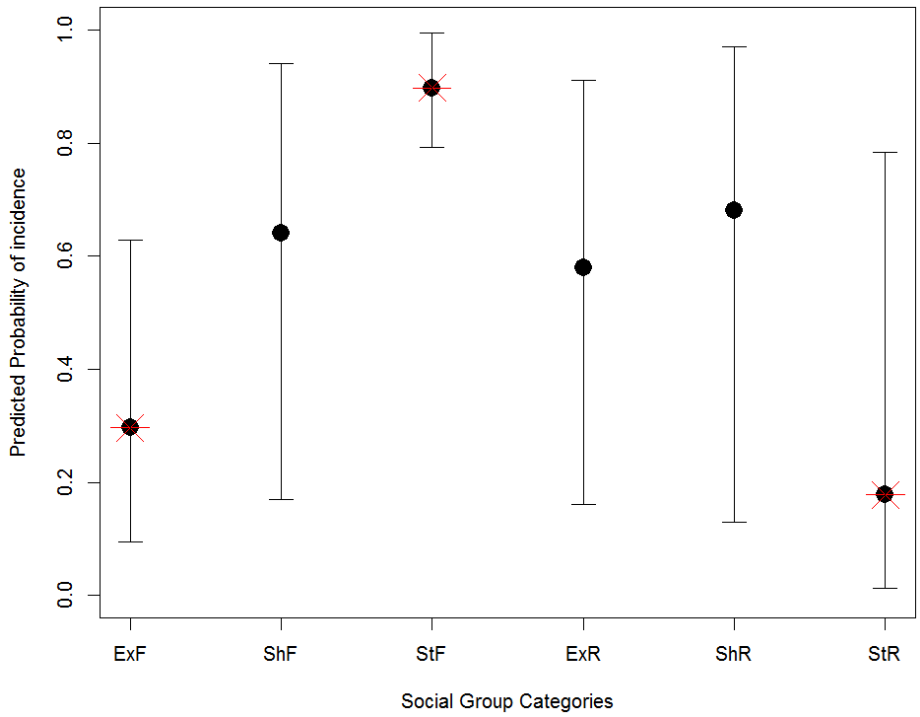


Fig 6.6 Predicted probability of social group having an incident case of TB detected in a subsequent year, based on demographic category (for full details of categories see Table 6.1; ExF = 'Expanding–Falling', ShF = 'Shrinking–Falling', StF = 'Stable-Falling', ExR = 'Expanding-Rising', ShR = 'Shrinking-Rising' and StR = 'Stable-Rising'). Standard errors are displayed. Red

markers indicate significant differences at the more conservation p value threshold to reflect correction for multiple testing, using the 'mcposthoc.fnc' function in LMERConvenienceFunctions package (incidence risk significantly higher for 'StF' category compared to 'ExF; and 'StR').

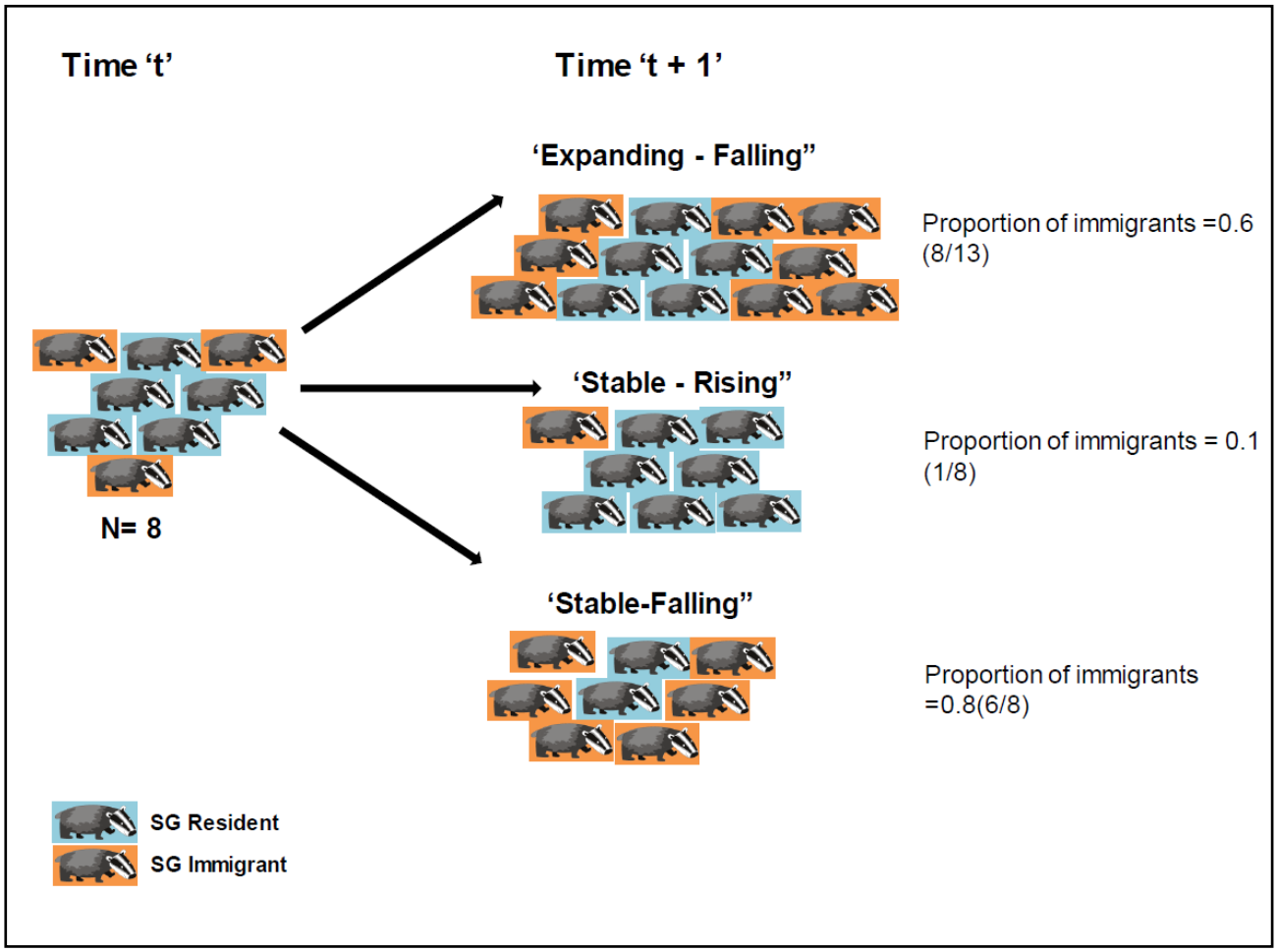


Fig 6.7 Schematic for suggested drivers of observed enhanced incidence risk in "Stable – Falling" category.

6.5 Discussion

The present study has identified temporal changes in the genetic population structure within this high-density badger population suggesting that this unmanaged population has undergone a period of demographic 'flux'. During the period of flux, social groups which were previously relatively isolated appear to have become linked by inter-group movements and potentially inter-group mating events, resulting in the population becoming increasingly genetically homogenous. Initially, spatially closer social groups were also genetically closer, as predicted under models of restricted dispersal across short distances (Pope, Domingo-Roura et al. 2006). However, for a period of years in the middle of the study period, this was not the case, suggesting that inter-group movement and mating events were no more likely between spatially proximate social groups than between spatially distant social groups. These changes are concurrent with a period of rising *M. bovis* prevalence and incidence in the population (Delahay, Walker et al. 2013) where the previously spatially clustered pattern of infection became more diffuse and social groups became more homogenous in their disease status (Chapter 5). In the final years of the study period, genetic isolation by distance re-established, consistent with a return to the former situation. A previous study which used only a small number of individual genotypes from the Woodchester Park population (N = 20) found no evidence of genetic isolation by distance. This result was attributed to the small spatial scale of the study area (Pope, Domingo-Roura et al. 2006). However, using a larger number of genotypes from this population and considering the population over a two-decade time period, the present study has demonstrated the presence of fine-scale genetic population structure, at a very restricted spatial scale and further, that this population structure can fluctuate over time, potentially in response to changes in movement patterns and population density.

Visualising the genetic population structure allows us to gain a different perspective on possible opportunities for *M. bovis* transmission amongst badger social groups. For example, forays into other social group territories to seek mating opportunities may be ephemeral and are unlikely to be detected by observational data or by delineating social territory boundaries. Shared matings between social groups may be driven more by the composition of the social

groups, for example the availability of breeding age females, rather than simply the spatially proximity to their social group. This is supported by observational data; in two studies, one of which was based on the Woodchester population, male badgers had a tendency to move into social groups with more females, and females tended to move into groups with more males (non-significant result in Woodchester study) (Macdonald, Newman et al. 2008, Robertson, Palphramand et al. 2015). The decision of whether or not an individual decides to leave their natal group is likely to be influenced by a range of factors (Lawson Handley and Perrin 2007), potentially related to current social position, the sex-ratio of the group and mate availability or potentially an individual's relatedness to other group members which may favour dispersal as an inbreeding avoidance strategy (Moore and Ali 1984).

Movement is known to be a key driver of TB incidence in badger populations, with years of higher movement rates tending to be followed by years with an increased TB incidence (Rogers, Delahay et al. 1998). Where badger population density is reduced by management interventions such as culling, this favours inter-group movement amongst the surviving individuals (Riordan, Delahay et al. 2011). The present study shows that the probability of inter-group movement can vary over time in an unmanaged population, with lower population densities favouring increased movement. There are a number of drivers which may explain this relationship. Mate availability is thought to be a key driver of dispersal in high-density badger populations, as reproduction within the social group is usually monopolised in a given year by a small number of residents (Roper 2010). It is possible that mate availability within social groups may be limited at lower population densities; therefore individuals may be more likely to seek mating opportunities outside their resident group. Territory defence may be less intense at lower population densities, as fewer badgers are available to share the cost of maintaining territorial boundaries (Roper 2010), which may result in more inter-group movement. In the Woodchester Park population, it has been noted that the delineated territories identified annually (Delahay, Brown et al. 2000) became less distinct in the latter years of the present study, suggesting greater territorial overlap, however this phenomenon has yet to be formally investigated. It is interesting to note that a number of 'super-groups' comprised of merged neighbouring groups,

formed from 2000 onwards whereas they were largely absent prior to this point. A recent analysis suggested that these merging events tended to occur between groups of differing sex ratios, potentially suggesting a response to mate availability, however this effect was not statistically significant (Robertson, Palphramand et al. 2015). As the distance moved by dispersing badgers does not vary between years, it appears that rather than badgers moving spatially further in later years, there have been changes in the connections between social groups, such that groups which traditionally were strongly linked by inter-group movement have become uncoupled and new links between social groups which previously lacked gene flow have formed. Further consideration of what drives dispersal and gene flow between pairs of social groups, for example taking into account sex-ratios and levels of inbreeding, would be required in order to further understand the drivers of temporal changes in dispersal patterns.

In the present study, male badgers were both more likely to move social groups and also tended to disperse further than female badgers, as has been previously documented for badger populations (Kruuk and Parish 1987, Cheeseman, Cresswell et al. 1988, Rogers, Delahay et al. 1998, Roper, Ostler et al. 2003, Pope, Domingo-Roura et al. 2006). As well as bringing individuals from previously spatially separated social groups together, the process of moving social groups is likely to carry enhanced *M. bovis* transmission risk as moving badgers are more likely to be on the receiving end of aggressive encounters as they trespass into a neighbouring territory (Macdonald, Newman et al. 2008). Additionally, movement imposed physiological stress may suppress the immune response, such that moving badgers are particularly vulnerable to acquiring infection or that latent infection becomes activated (Gallagher and Clifton-Hadley 2000). Male badgers may therefore play a more important role in the transmission of *M. bovis* infection amongst badger social groups than female badgers.

As well as changes in genetic population structure being driven by changes in inter-group movement, extra-group mating, which is known to be commonplace in badgers (Carpenter, Pope et al. 2005, Annavi, Newman et al. 2014), can result in a ‘gamete-dispersal’ effect. This involves dispersal of an individual’s genes over a greater distance than the individual itself permanently disperses

over (Winters and Waser 2003). Temporal changes in the proportion of mating which occurred outside the social group may also have contributed to the observed changes in genetic population structure, and the future availability of a pedigree for this population will help to determine whether such changes have occurred.

The risk of a badger social group experiencing an incident case of *M. bovis* has previously been linked to changes in group size, with shrinking social groups at the highest risk, suggesting that the process of group size reduction had the most influence on disease dynamics (Vicente, Delahay et al. 2007). By incorporating genetic data with observational data, the present study found that the highest incidence risk was associated with social groups where the group size was stable but the social group membership was unstable, potentially driven by the replacement of residents by immigrants. These results suggest that it may be the proportion of immigrants moving into a social group and the associated effect on the social group contact network which carries an enhanced transmission risk, rather than any absolute change in group size or the number of immigrants into a group (Fig 6.8). It is likely that the higher the proportion of immigrant members joining a social group, the more impact there is on the existing within-group contact network. This disruption may result in more within group mixing, resulting in enhanced incidence risk. An extension of the present study would be to determine who is at risk of becoming an incident case; the immigrant animals themselves, potentially as a result of the movement related stressors described above, as has been previously documented in badgers (Rogers, Delahay et al. 1998, Woodroffe, Donnelly et al. 2009) or the individuals already resident in the receiving social group.

The present study suggests that temporal changes in the genetic population structure of the Woodchester Park badger population are consistent with a demographic change in movement and mating behaviour. These processes may potentially be a response to a fall in population density, and may have influenced the incidence and spatial arrangement of *M. bovis* infections. However, an alternative explanation is that the decrease in spatial clustering of *M. bovis* infections (Chapter 5) and increasing incidence observed is due to the seeding of infection into the population from either the surrounding, un-sampled badger population or from the local cattle population. Examination of whole

genome sequences from both cattle and badgers will allow further investigation of the respective roles of local cattle and resident badgers in spatio-temporal variations in *M. bovis* infection in the Woodchester Park badger population.

The present study has demonstrated the existence of fine-scale genetic population structure and temporal variation in its magnitude within a high-density badger population. Through coupling genetic and detailed observational data, this study has identified a period of demographic change in this unmanaged population with implications for *M. bovis* transmission. As management interventions such as culling have been noted to result in changes in genetic population structure, related to movement patterns, it is valuable to be able to quantify the changes that can occur within a unmanaged system as a result of natural demographic processes. With increasing interest in using genetic methods to help understand badger population structure and disease transmission, particularly at the moving front of the current endemic area, this study provides a useful demonstration of what host genotype data can tell us about population structuring and movement patterns in this system. In terms of its wider application, this study adds to a limited body of work in which host genetic structure has been used to understand pathogen spread related to host movement at a restricted spatial scale (Cullingham, Merrill et al. 2011, Mazé-Guilmo, Blanchet et al. 2016) rather than as a larger scale landscape genetics approach (Biek and Real 2010) looking at host population structure across regions (Lee, Ruell et al. 2012), countries (Lang and Blanchong 2012, Vander Wal, Edye et al. 2013) or continents (Streicker, Winternitz et al. 2016). In some cases, researchers have also incorporated pathogen genetic information to investigate how host genetic population structure drives pathogen genetic population structure (Lee, Ruell et al. 2012, Streicker, Winternitz et al. 2016). This is the approach that will be taken in Chapter 7; combining whole genome sequence data from a group of *M. bovis* isolates isolated from the Woodchester badger population with host genotype data.

CHAPTER 7: Data Chapter

Transmission pathways of bovine tuberculosis revealed by Whole Genome Sequencing

7.1 Abstract

The advent of whole genome sequencing (WGS) of pathogens allows the study of fine-scale variation in transmission pathways, across space and through time. Here, I employ WGS on *Mycobacterium bovis* isolates from a well-studied reservoir population of badgers in order to resolve fine-scale transmission dynamics within the population. I demonstrate that, in comparison to traditional genotyping approaches, WGS adds much higher resolution and allows us to address epidemiological questions that traditional approaches could not resolve. I use the sequence data to suggest that infection has spread spatially across the study population over the period of a decade and that badger to badger transmission is occurring. Further, I combine genotype data from both the badgers themselves with the WGS data from the *M. bovis* isolates to demonstrate that host population genetic structure influences pathogen population genetic structure in this population. I find support for the importance of kin structure in *M. bovis* transmission within badger social groups but suggest that inbreeding avoidance may influence contact rates and subsequent disease transmission outside of the social group. These results provide an encouraging indication that WGS technologies will have much to add to our understanding of bTB transmission within wildlife populations.

7.2 Introduction

The advent of whole genome sequencing (WGS) of pathogens allows the study of fine-scale variation in transmission pathways, across space and through time. Such approaches have been used powerfully in understanding the transmission dynamics of a range of human, wildlife and livestock pathogens (Benton, Delahay et al. 2014), including HIV (Henn, Boutwell et al. 2012), influenza (Ghedin, Sengamalay et al. 2005, Holmes, Ghedin et al. 2005), Ebola (Gire, Goba et al. 2014) and TB (Gardy, Johnston et al. 2011, Bryant, Schürch et al. 2013, Roetzer, Diel et al. 2013). In many cases, WGS has uncovered diversity between pathogen strains which traditional methods were not able to elucidate (Bryant, Schürch et al. 2013, Roetzer, Diel et al. 2013) and hence has proved to be a significant step forward in understanding transmission dynamics, particularly at a fine scale.

As discussed in Chapter 5, the spatial distribution of *M. bovis* infections in the Woodchester Park badger population has changed over the course of the 30+ year study. Infection was initially restricted to the western part of the study area, however over time infections started to be detected in the east of the area. The recent availability of a group of whole-genome-sequenced *M. bovis* isolates collected over a ten year period provides a valuable opportunity to look more closely at fine-scale transmission dynamics during this time and allow us to determine whether the change in spatial distribution was due to the introduction of new infections from an outside source (e.g. the local cattle population) or whether it represents spatial spread from west to east via transmission across the badger population.

The typing methods traditionally used to categorize strains of *Mycobacterium bovis* (the causative agent of bovine TB) are spoligotyping (spacer-oligonucleotide typing) and VNTR (Variable Number Tandem Repeat) typing, both of which are based on small genomic regions that are generally evolving at a higher rate than the rest of the genome (Joshi, Harris et al. 2012). Such methods are therefore potentially more useful for differentiating between organisms at a coarser evolutionary scale than detecting finer scale intra-specific variation (Joshi, Harris et al. 2012). The *M. tuberculosis* complex, of which *M. bovis* is a member, is highly clonal and therefore in this case, the

spoligotype pattern can be more confidently used as a proxy for the evolutionary history of the complete genome of the cell (Frothingham 1995, Smith, Gordon et al. 2006). Spoligotyping and VNTR typing have had valuable contributions as molecular tools in our understanding of *M. bovis* epidemiology; helping to infer transmission events between livestock populations (Smith, Gordon et al. 2006, Munyeme, Rigouts et al. 2009), identify probable infection sources (Duarte, Domingos et al. 2008) suggest the existence of wildlife reservoirs (Pavlik, Dvorska et al. 2002, Santos, Correia-Neves et al. 2009) and have also had some value in inferring cross-species transmission events (Serraino, Marchetti et al. 1999, Woodroffe, Donnelly et al. 2005, Duarte, Domingos et al. 2010, Cunha, Matos et al. 2012). Within cattle populations in the UK, *M. bovis* spoligotypes are highly geographically clustered, with spoligotypes having a typical 'home range'(Smith, Gordon et al. 2006). The appearance of a non-typical spoligotype outside its characteristic home range has been used to demonstrate the need for pre-movement testing of cattle to prevent the movement of infected animals and further spread of *M. bovis* infection between cattle populations (Smith, Gordon et al. 2006).

Despite these valuable applications, spoligotyping and VNTR typing are limited in their ability to infer finer-scale transmission dynamics e.g. inferring direction of transmission between local populations (Smith, Gordon et al. 2006). Epidemiologists wishing to study these finer scale dynamics in pathogens with very little variation between strains, such as *M. bovis,* will require a typing method that is able to detect small differences between isolates. Where discrimination between isolates is not possible using conventional methods, whole genome sequencing (WGS) may be the only tool suitable for looking at fine-scale transmission dynamics (Benton, Delahay et al. 2014). The exceptionally high level of genetic resolution achievable through sequencing the whole genome of an organism means that even sequencing a restricted number of isolates can reveal a wealth of epidemiologically valuable information (Biek, O'Hare et al. 2012). Where access to long term studies is possible, a 'phylodynamic' approach (Grenfell, Pybus et al. 2004) of overlaying pathogen phylogenies onto well documented epidemiological systems is advocated.

The identification of shared genotypes of *M. bovis* between wildlife and livestock populations has been used to infer transmission in a number of contexts (Lisle,

Yates et al. 1995, Serraino, Marchetti et al. 1999, Woodroffe, Donnelly et al. 2005, Duarte, Domingos et al. 2010, Cunha, Matos et al. 2012, Buddle, Lisle et al. 2015, Glaser, Carstensen et al. 2016). In the UK, where badgers are the principal wildlife reservoir of *M. bovis,* with a key role in transmission to cattle, marked spatial correlation of *M. bovis* spoligotypes of cattle and badger isolates has been documented (Woodroffe, Donnelly et al. 2005). These data have been used to support the hypothesis that transmission occurs between the two host species with the caveat that they cannot be used to infer to relative importance of cattle to badger vs badger to cattle transmission (Smith, Gordon et al. 2006). More recently whole genome sequencing of a limited number of *M. bovis* isolates from badgers and cattle provided evidence for recent transmission events between the two hosts, with the suggestion that more extensive sampling could allow for the quantification of the extent and direction of transmission between badgers and cattle (Biek, O'Hare et al. 2012).

As well as inferring transmission dynamics between species, the overlaying of data on pathogen strain diversity onto ecological information could be used in wildlife populations to assess transmission rates in relation to within - population structure (e.g. social groups, herds etc.) (Benton, Delahay et al. 2014). In the case of the European badger the prevailing social structure in high density populations has been associated with the clustering of infection within social groups (Delahay, Langton et al. 2000, Delahay, Langton et al. 2000). Disruption of this social structure, as observed following culling, leads to a reduction in this clustering, as surviving individuals range more widely (Jenkins, Woodroffe et al. 2007). Further information on the role of social behaviour in the spread of infection may be achievable by investigating the genetic diversity of *M. bovis* strains in badger populations. In Chapter 3 I highlighted the importance of badger social group and specifically kin structure in predicting infection risk in badger cubs. If social structure acts as a barrier to disease spread then we would expect the degree of genetic similarity amongst *M. bovis* strains within badger social groups to be greater than that observed between social groups. Within social groups, kin structure may also influence transmission rates between social group members (Benton, Delahay et al. 2016). Examining the genetic distances between *M. bovis* strains of related and

unrelated social group members may shed further light on the influence of relatedness structure on *M. bovis* transmission within badger populations.

In my final data chapter, I use the pathogen genetic data, derived from whole genome sequencing of a group of *M. bovis* isolates collected from the Woodchester badger population, to demonstrate the added value of WGS over traditional genotyping approaches in understanding fine-scale TB transmission dynamics within a single badger population. I demonstrate how the availability of such sequence data allows epidemiological questions to be addressed by specifically using the data to look for evidence of spatial spread of infection across the population. Finally, I use the pathogen genetic data in tandem with the host genetic data to consider how host kin structure, both within and outside badger social groups, affects pathogen population structure.

7.3 Methods

Badger Sampling

All data used in these analyses were collected from the long-term trapping and sampling study at Woodchester Park in Gloucestershire. Badgers from this study population have been routinely trapped, up to four times a year, since 1976. At first capture, badgers are tattooed with a permanent ID code. Trapped badgers are brought back to a sampling facility, anaesthetised (for full details see Appendix) and a range of clinical samples are taken (oesophageal aspirates, tracheal aspirates, faeces, urine, swabs of bite wounds or burst abscesses) from which *M. bovis* may be cultured. Cultures, which have been routinely isolated from routine clinical sampling of captured badgers at Woodchester, have been archived since 1990.

Host Genotyping

On first capture, a hair sample is routinely taken from trapped badgers. This is stored in a tube with 80% ethanol before being submitted for genotyping. DNA extraction and genotyping procedures were as described previously (Carpenter, Pope et al. 2005). All genotyping data reported in this study were generated by the team headed by T.Burke at the Molecular Ecology Lab, University of Sheffield. Genotyping of hair samples from this population took place routinely from 1990 until present, however only genotype data up to 2011 were available for this analysis. 22 microsatellite markers were used, each with 4-7 alleles.

The MicroDrop Programme (Wang and Rosenberg 2012) was used to impute missing data in the microsatellite data set. Deviations from Hardy-Weinberg equilibrium for each of the 22 microsatellite markers were tested on the MicroDrop-corrected dataset using the hwtest function in the R package ‘adegenet’ (Jombart 2008); none were identified. The Bartlett test of homogeneity in the same package to confirm homogeneity of variance among loci ($P = 0.78$).

Microsatellite data were used from individuals for whom pathogen sequence data was available (see below). This resulted in a dataset of 66 individuals for whom host and pathogen genetic information was available.

Pathogen Genotyping

Spoligotyping (spacer-oligonucleotide typing) was carried out routinely on *M. bovis* isolates from Woodchester Park from 1990 onwards; VNTR (Variable Number Tandem Repeat) typing began in 2000 but was not routinely carried out until 2002 onwards. Spoligotyping and VNTR genotyping was carried out by Noel Smith and the TB Genotyping Group, APHA.

As part of a collaborative pilot study with the University of Glasgow, 230 archived *M. bovis* isolates cultured from samples taken from Woodchester badgers between 2000 - 2010 were regrown. Re-culturing and DNA extraction were carried out at the Agri-Food and Biosciences Institute in Northern Ireland (AFBINI). The extracted DNA was sequenced at the Glasgow Polyomics facility using an Illumina MiSeq platform. Bioinformatics, estimation of the molecular clock rate and construction of the initial phylogenetic tree of the Woodchester badger isolates were carried out by Hannah Trewby; for details see (Trewby 2016). For each isolate, 503 high quality single nucleotide polymorphisms (SNPs) were identified (Trewby 2016). The program PhyML v3.0 was used to generate a maximum likelihood phylogeny for the concatenated sequences, using the HKY nucleotide substitution model and 1000 non-parametric bootstraps to assess node support (Trewby 2016). The phylogeny generated previously (Trewby 2016) was visualised using the R package 'ape' (Paradis, Claude et al. 2004). Trees display the number of SNP (single nucleotide polymorphisms) between isolates. Analysis was restricted to 163 'high-quality' isolates as determined during bioinformatics procedures carried out previously (Crispell 2017).

Pathogen Population Structure

Lineages of the sequenced isolates were previously assigned, based on a threshold of 10 SNP differences between lineages and bootstrap support of >99% (Trewby 2016). From this tree, maps were produced in ArcGIS 10.2 to visualise the spatial extent and persistence of the different *M. bovis* lineages identified from the phylogeny. Previous analysis describing cattle-badger transmission clusters within Woodchester Park were also included in the visualisation (Crispell 2017). These clusters were identified as follows; the 163 isolates from the Woodchester badgers described above and 81 isolates

collected from dairy and beef cattle herds between 1997 and 2012, within a 15km buffer of the study area were used to construct a maximum likelihood phylogenetic tree and rooted against a reference sequence – AF2122/97 (Crispell 2017). Potential inter-species transmission events were identified where clusters of highly genetically related isolates from badgers and cattle were noted. The direction of transmission was inferred based on which host species was closest to the root of the clade. If a clade contained closely related isolates from badgers and cattle, and an isolate from a cow was closest to the root of the clade, this was taken as an indicator of a potential transmission event from cattle to badgers. Full details are here (Crispell 2017).

In Chapter 5 it was suggested that the spatial distribution of *M. bovis* infections in Woodchester Park was consistent with *M. bovis* infection being introduced initially into badgers in the western social groups of the Woodchester study area and subsequent spread to badgers in the eastern social groups. To test for this, a phylogenetic discrete traits analysis was carried out to look for a genetic signal of west to east spread (Lemey, Rambaut et al. 2009). This analysis was restricted to high-quality isolates from Lineage 1. Only 1 isolate per individual was included as there is currently no capacity to include individual as a random effect within the below analyses. Where more than 1 isolate was available for an individual, the earliest isolate was used (i.e. from the earliest available capture event). If multiple isolates were all obtained at a single capture event, an isolate was randomly selected. This resulted in a dataset of 66 isolates for inclusion in the analyses. Despite low evolutionary rates within these data, previous work using tip-date randomisations has shown that the sequences contain sufficient temporal-genetic signal to allow the use of Bayesian phylogenetics analysis to estimate molecular clock rates (Grenfell, Pybus et al. 2004, Trewby 2016). In the current analysis a discrete traits analysis was carried out using the Bayesian phylogenetics programme 'BEAST' (Drummond, Suchard et al. 2012). The 'BEAGLE' library was used to allow more computationally efficient likelihood calculation (Ayres, Darling et al. 2011). Data preparation was carried out in the programme 'BEAUti' (Drummond, Suchard et al. 2012). Two independent MCMC chains were run for each analysis and log files were evaluated in 'Tracer' (Rambaut and Drummond 2007) to assess convergence, both by visually examining the posterior trace, and by checking

that Estimated Sample Size values were greater than 200 for each parameter. 'TreeAnnotator' (Rambaut and Drummond 2013) was used to summarise the maximum clade credibility tree from the distribution of posterior trees generated during the first analysis (see below) after removing a burnin of 10% of MCMC steps. The programme 'FigTree' was used to display the maximum clade credibility tree (Rambaut 2007). Prior settings used for the discrete traits analysis were consistent with those used previously (Trewby 2016) for the same samples, as follows. The uncorrelated log normal relaxed clock model (Drummond, Ho et al. 2006) was used with a normally distributed prior, truncated at zero, with a mean of 0.2 SNPs per year (based on the evolutionary rate determined from a different set of genotyped *M. bovis* isolates from Northern Ireland (Trewby 2016, Trewby, Wright et al. 2016) and a wide standard deviation of 10 SNPs per genome per year. The HKY model of nucleotide substitution was used (Hasegawa, Kishino et al. 1985, Yang 1994) to model the variation in rates at which nucleotides replace each other during DNA evolution. The coalescent Bayesian skyline demographic model, which allows population size to be estimated at different points, was also used as a flexible demographic prior (Drummond, Nicholls et al. 2002, Drummond, Rambaut et al. 2005): setting a more restrictive demographic prior (e.g. exponential growth) might have imposed an improper constraint on the estimation of the phylogeny, as coalescence rate and population size are interlinked (Ho and Shapiro 2011). Bayesian skyline reconstruction (Drummond and Rambaut 2007) was carried out in 'Tracer' to visualise the effective population size over time by assessing the coalescence pattern of the posterior trees. As the coalescence rate of an infectious disease is thought to be driven primarily by new transmissions (i.e incidence) and is also indirectly linked to the number of infected individuals (i.e prevalence) (Volz, Pond et al. 2009, Frost and Volz 2010), the Bayesian skyline plot was visualised alongside population level estimates of prevalence and incidence in the Woodchester badgers over the same temporal period (Walker 2012).

In the first BEAST analysis Bayesian Stochastic Search Variable Selection (BSSVS) (Lemey, Rambaut et al. 2009) was used to identify which transition states (west to east, with eastern isolates having western ancestors; or east to west, with western isolates having eastern ancestors) were significantly

supported. The MCMC process was set to 500 million iterations, sampling every 50,000th iteration, with a 10% burnin removed. The branches of this tree were then coloured based on the estimated trait of the most recent common ancestor of the branch (is the isolate derived from an 'eastern' ancestor or a 'western' ancestor), with posterior values displayed at each node to show how well supported this ancestral state is. 'SPREAD' (Bielejec, Rambaut et al. 2011) was used to calculate Bayes Factor from the east to west and west to east transition rates generated in order to determine whether the rates were well supported (a Bayes Factor of greater than 10 indicates moderate support for a given transition rate, with Bayes Factor greater than 100 indicating strong support (Lemey, Rambaut et al. 2009)).

In a separate BEAST analysis, State Change Count Reconstruction was used to estimate of the rate of transition between the two states (west to east vs east to west (Minin and Suchard 2008)). The MCMC process was set to 200 million iterations, sampling every 20,000th iteration. Markov Jump Counts were reported, giving estimated counts of the number of transitions between east and west.

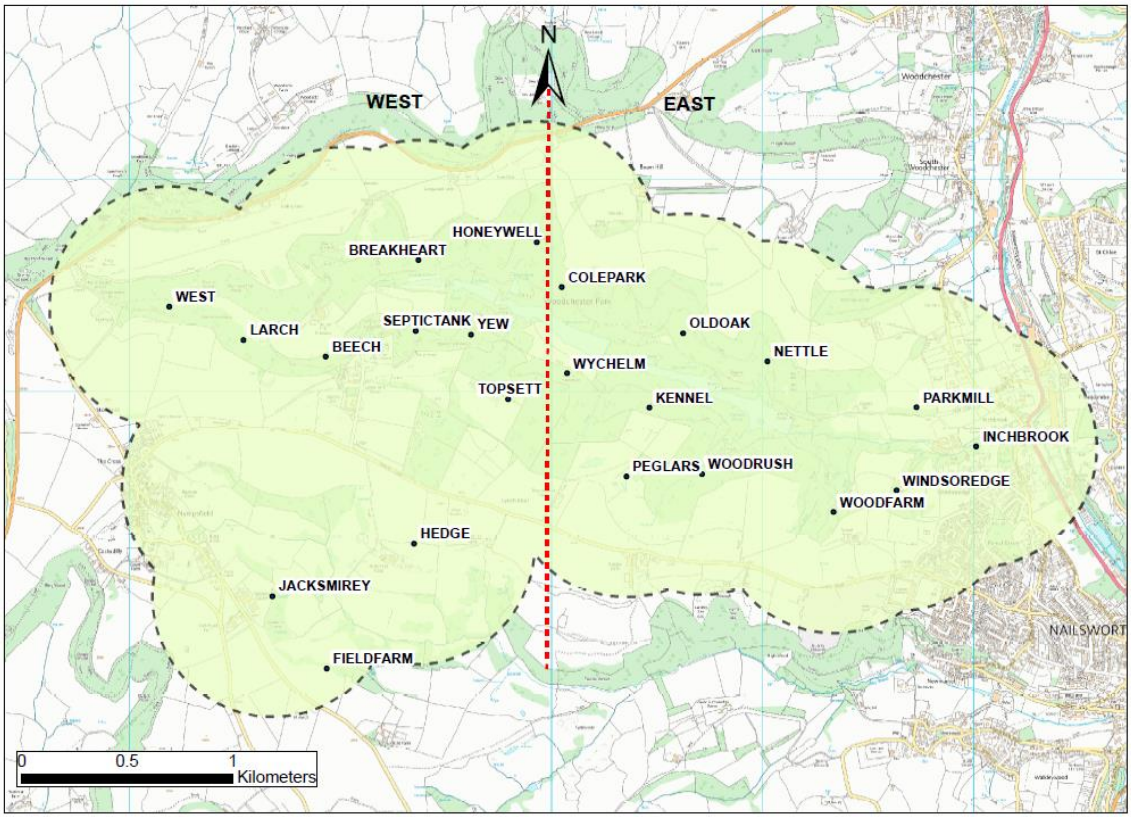


Fig 7.1 Spatial division of social groups from west and east of the Woodchester Park study area for whom high-quality Lineage 1 *M. bovis* isolate sequences were available.

Isolate Genetic Distance, Social Group & Kin Structure

The impact on genetic distance between *M. bovis* isolates of within-individual variation, social group membership and kin structure was analysed using permutation testing. Genetic distances between all available pairs of high-quality *M. bovis* isolates were included as the response variable in this analysis. Isolates were grouped into categories based on relatedness between hosts (as estimated using the rxy estimator within the R package 'Demerelate'(Kraemer, Gerlach et al. 2013)) and social group membership as detailed in Table 7.1. An individual's main social group was defined as the social group that the badger had spent the most time in. This definition was used as 'main' social group was found to be the best spatial predictor of genetic distance between isolates (Crispell 2017). The dataset was restricted to only include pairwise comparisons from isolates collected within 3 years of each other, in order to look only at pairwise comparisons between individuals who were likely to have been in the population at the same time. This is a conservative estimate, as the average life expectancy of adult badgers in the Woodchester population is 4.5 years. As the response variable was a pairwise distance, a permutation procedure was used to account for the non-independence of samples (i.e. the genetic distance between a pair of hosts and the genetic distance between their *M. bovis* isolates are non-independent, as both distance values are collected from the same pair of individuals). To test whether related badgers in the same social group had more similar *M. bovis* isolates than non-related badgers within the same social group, a subset of the *M. bovis* genetic distance dataset was created which was restricted to those two categories ('Related In SG' and 'NonRelated In SG', see Table 7.1). *M. bovis* genetic distances were then shuffled across the badgers and the average genetic distance between isolates in the two categories calculated. This process was repeated 10,000 times in order to generate a null distribution. The observed 'true' distance between the mean of the two categories was then compared to this null distribution and significance assessed as follows ((p = number of null mean genetic distance values that were greater than the observed value / 1000). This process was repeated to test for significant genetic differences between *M. bovis* isolates from related and non-related badgers outside of an individual badgers social group ('Related Out SG and 'NonRelated Out SG', see Table 7.1). Finally,

genetic differences between isolates from the same individual were compared with those from related individuals in the same social group ('Related In SG' and 'Same Individual'; see Table 7.1).

Table 7.1 Categories to describe social group membership and relatedness relationship between hosts

Category	Description
NonRel Out SG	Isolate is from an **unrelated** individual from **outside** the focal individual's main assigned social group
Related Out SG	Isolate is from a **related** individual from **outside** the focal individual's main assigned social group
NonRel In SG	Isolate is from an **unrelated** individual from **within** the focal individual's main assigned social group
Related In SG	Isolate is from a **related** individual from **within** the focal individual's main assigned social group
Same Individual	A different *M. bovis* isolate (from a different body site or different sampling occasion) from the focal individual

7.4 Results

Pathogen Genotyping

The dominant spoligotype within the genotyped isolates from the Woodchester badger population is Spoligotype 17 (Fig 7.2) which is also the genotype typically found in cattle associated with the Woodchester location (Trewby 2016). VNTR typing adds a little more resolution (Fig 7.3); breaking Spoligotype 17 down into two VNTR types, although the majority are of one VNTR type 17:a. In contrast, the sequence data available from the whole genome sequencing allow the dominant spoligotype 17 to be broken up into multiple lineages (see Fig 7.4) and also allow pairwise distances between isolates to be estimated (in terms of differing SNP profiles at 503 high-quality SNP locations).

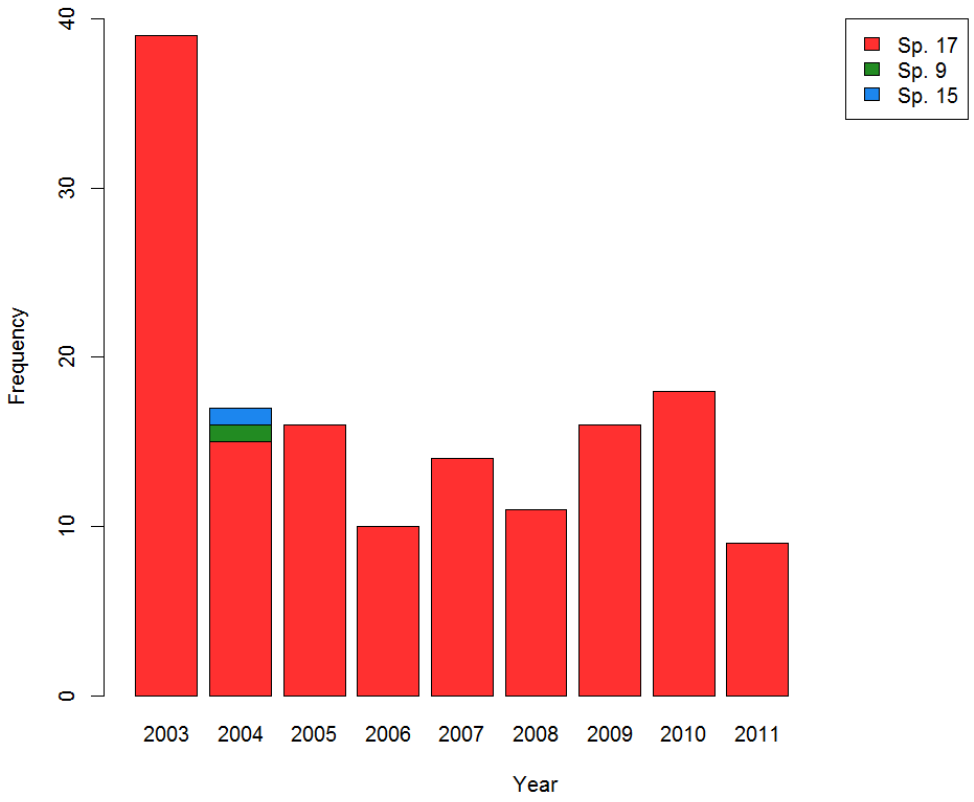


Fig 7.2 Spoligotypes of 163 high-quality sequenced *M. bovis* isolates

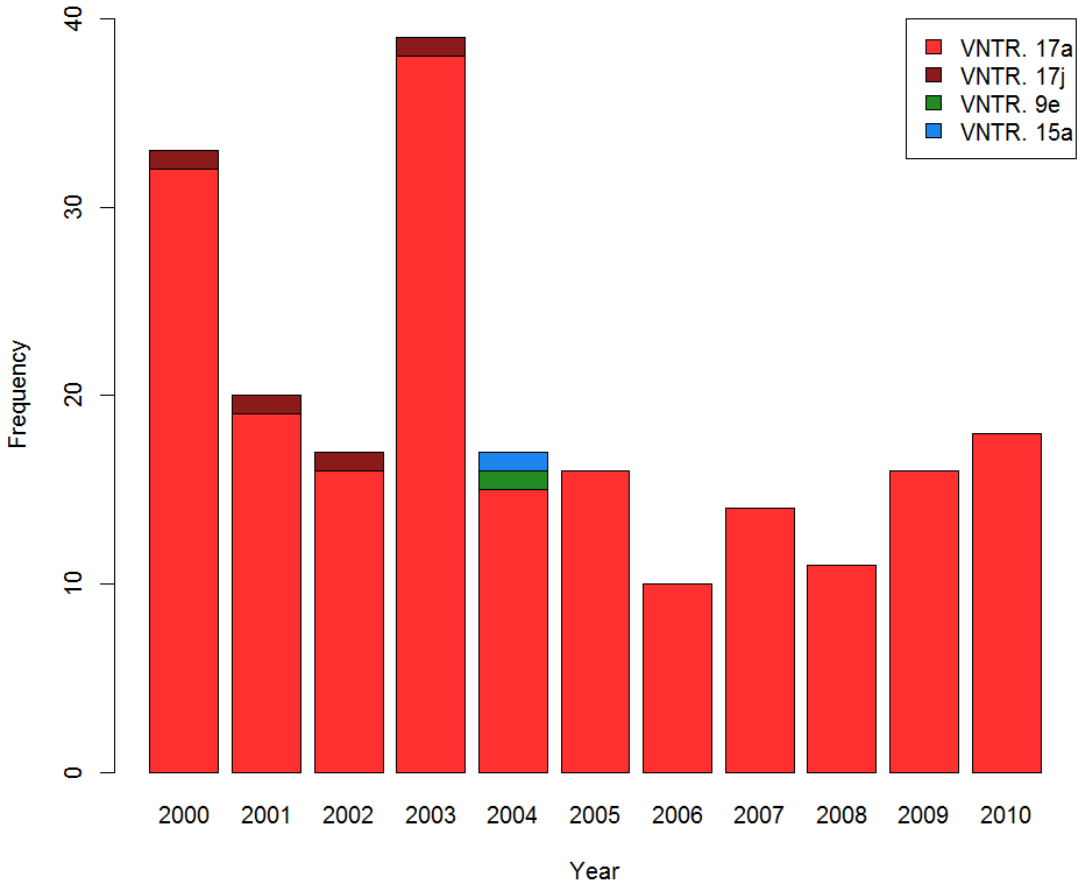


Fig 7.3 VNTR types of high quality *M. bovis* isolates

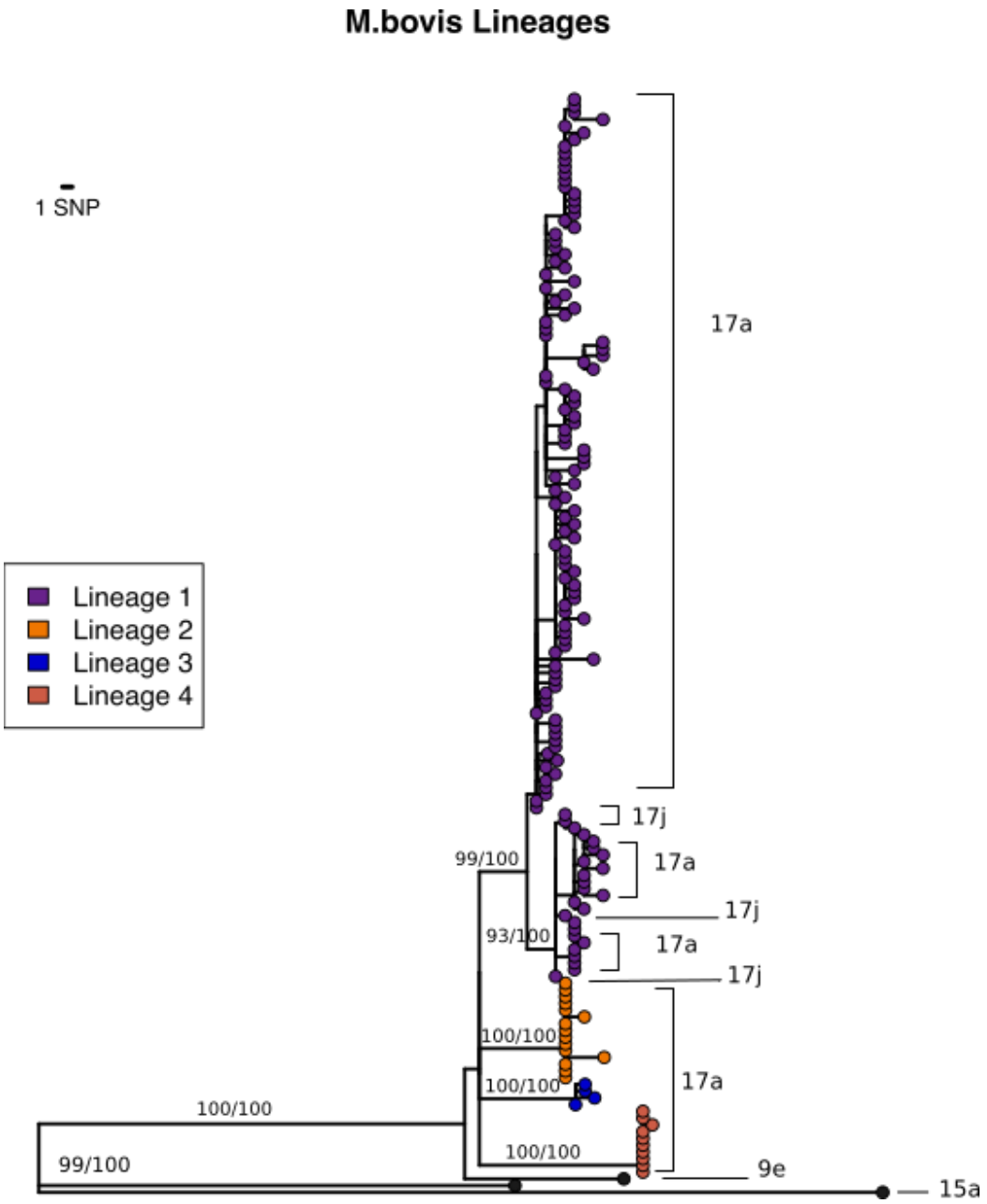


Fig 7.4 Phylogenetic tree of 163 high quality sequenced M. bovis isolates based on the number of SNP differences between isolates. Node colours indicate which lineage the isolate has been assigned to. The tree is rooted on the non Spoligotype 17 isolates, indicated in black. Lineages were assigned based on differences of 10 SNPs or greater between isolates (lineage assignments by Hannah Trewby, see (Trewby 2016)). VNTR genotypes are displayed. Left-hand value indicates percentage bootstrap support from the Maximum Likelihood phylogeny, and the right-hand value shows posterior probability of the node in the Bayesian phylogeny (Trewby 2016).

Lineage 1 (Fig 7.5) represents a cluster of spoligotype 17 *M. bovis* isolates which appears to be spatially widespread across the study area and has been present over the entire time period from which the sequenced samples were isolated. Over half of the isolates (71/131 high quality isolates available for this lineage) have been included in badger-cattle transmission clusters. The majority of these isolates were assigned to badger-cattle transmission clusters where the direction of transmission was thought to be from badgers to cattle (fewer isolates sourced from the cattle; cattle isolates nested within isolates from badgers) (Crispell 2017). However, 46% of the Lineage 1 badger isolates have not been included in any badger-cattle transmission cluster.

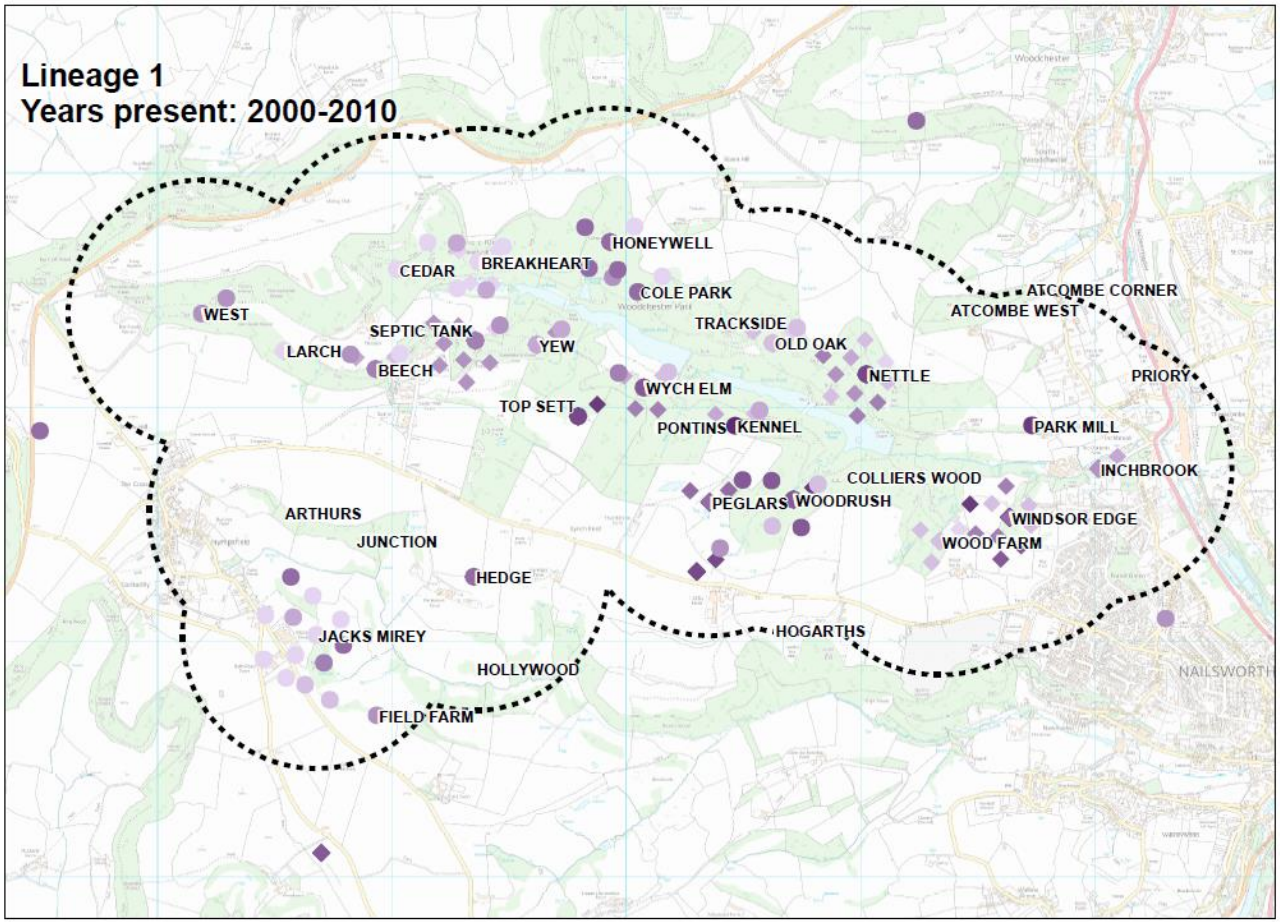


Fig 7.5 Spatial and temporal distribution of *M. bovis* isolates from Woodchester badgers assigned to Lineage 1 (Trewby 2016). Symbol colour indicates the year when the isolate was collected from a badger, with more recently collected isolates marked darker. Diamonds indicate isolates that have been assigned to a badger-cattle transmission cluster (full details available (Crispell 2017)), circles indicate unassigned isolates which may be representative of badger – badger transmission. Markers have been dispersed in ArcGIS 10.2 to aid visualisation of points collected from individuals in the same social group.

Lineage 2 (Fig 7.6) also represents a group of spoligotype 17 isolates which were present in the badger population in a number of years. The majority of the isolates of this lineage are linked to a single badger – cattle transmission cluster, where transmission from cattle to badgers has been inferred (cattle isolates present towards the root of the cluster, suggesting outbreak originated in cattle population (Crispell 2017)). However, there are a couple of isolates in this lineage which were not placed in this badger-cattle transmission cluster.

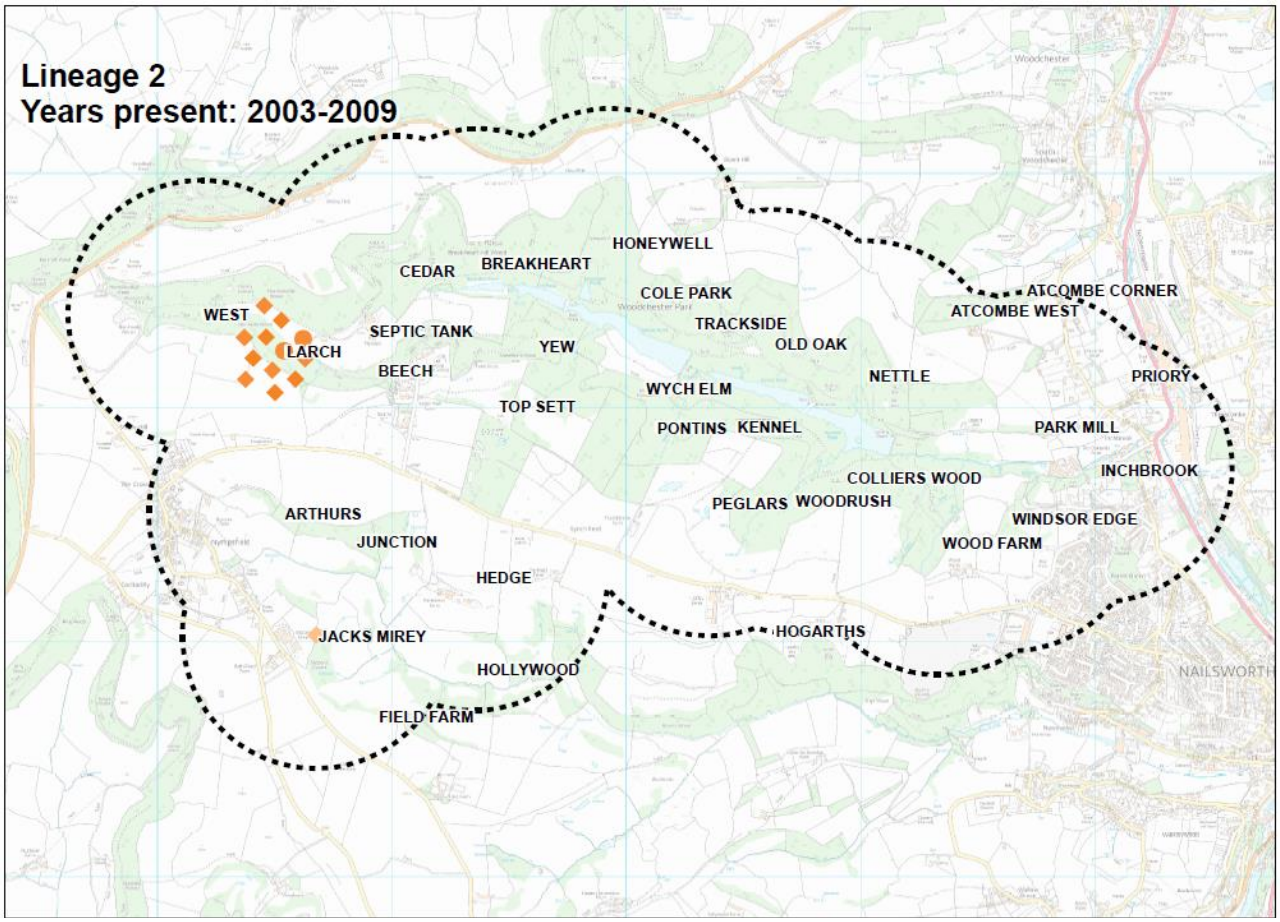


Fig 7.6 Spatial and temporal distribution of *M. bovis* isolates from Woodchester badgers assigned to Lineage 2 (Trewby 2016). Symbol colour indicates the year at which the isolate was collected from a badger, with more recently collected isolates marked darker. Diamonds indicate isolates that have been assigned to a badger-cattle transmission cluster (full details available (Crispell 2017)), circles indicate unassigned isolates which may be representative of badger – badger transmission. Markers have been dispersed to aid visualisation of points collected from individuals in the same social group.

Four Lineage 3 *M. bovis* isolates (Fig 7.7) were detected in a single year of the study (2000). All isolates were from a single badger (X090) that was resident in Nettle social group in that year. The isolates were linked to a badger-cattle transmission cluster in which cattle to badger transmission was inferred (Crispell 2017).

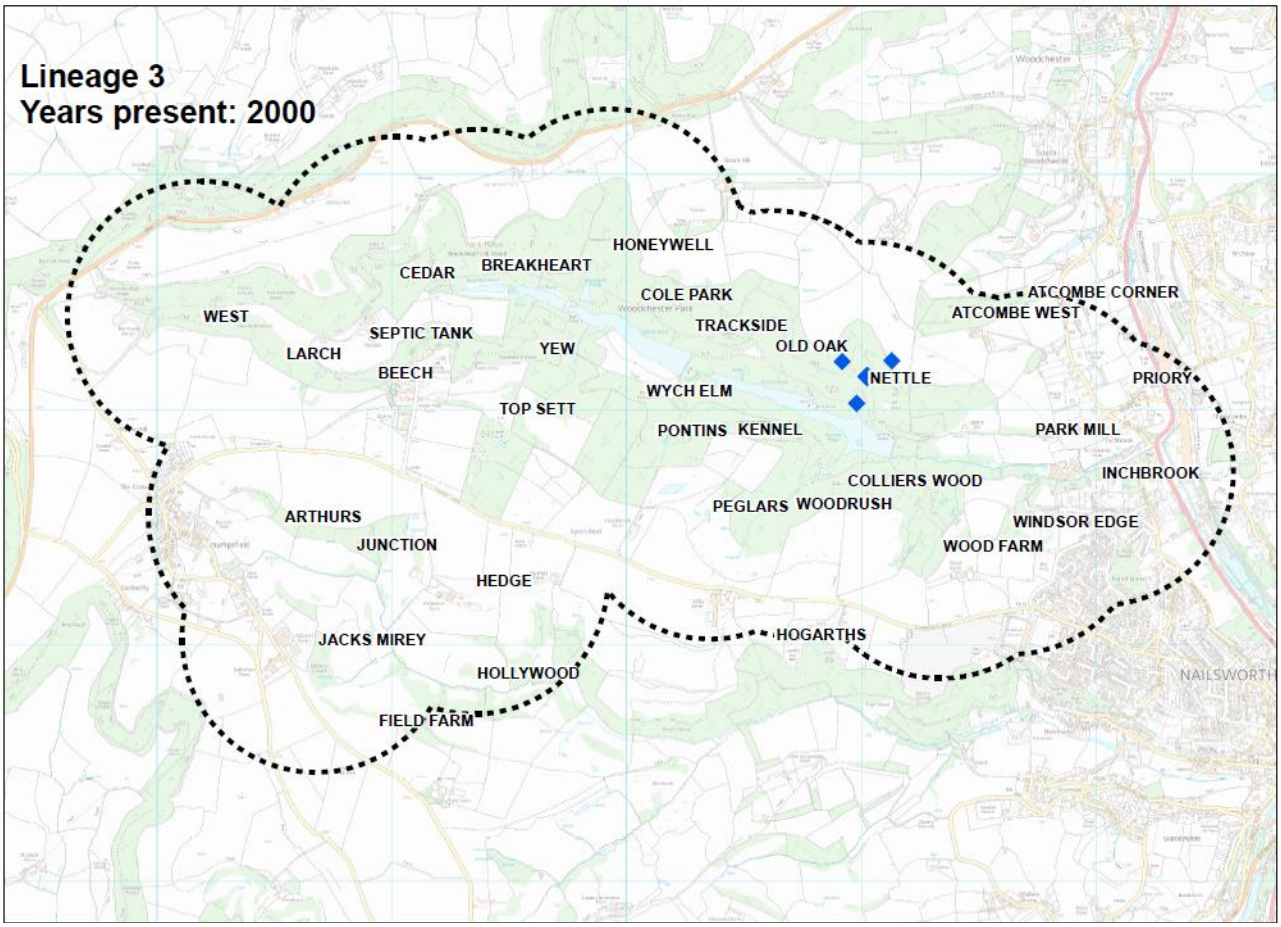


Fig 7.7 Spatial and temporal distribution of *M. bovis* isolates from Woodchester badgers assigned to Lineage 3 (Trewby 2016). Diamonds indicate isolates that have been assigned to a badger-cattle transmission cluster (full details available (Crispell 2017)). Markers have been dispersed to aid visualisation of points collected from individuals in the same social group.

Nine of the ten *M. bovis* isolates from Lineage 4 have been assigned to a badger-cattle transmission cluster where the direction of transmission is thought to be from cattle to badgers (cattle isolates present towards the root of the cluster, suggesting outbreak originated in cattle population (Crispell 2017)).The only isolate not assigned to this transmission cluster appears in badger U049 in Yew social group; this isolate is somewhat spatially isolated and does not fit with the rest of the transmission cluster.

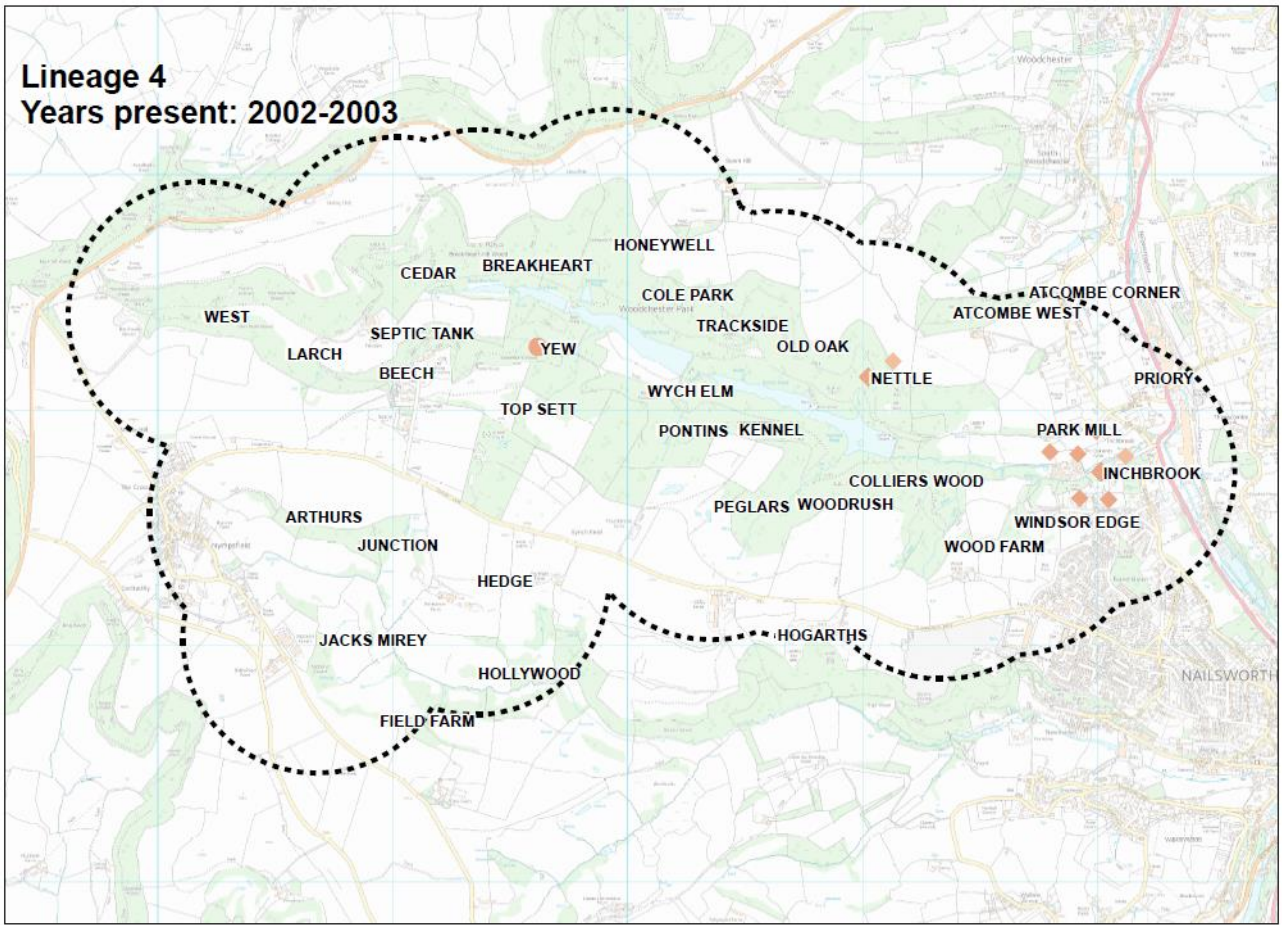


Fig 7.8 Spatial and temporal distribution of *M. bovis* isolates from Woodchester badgers assigned to Lineage 4 (Trewby 2016). Symbol colour indicates the year at which the isolate was collected from a badger, with more recently collected isolates marked darker. Diamonds indicate isolates that have been assigned to a badger-cattle transmission cluster (full details available (Crispell 2017)), circles indicate unassigned isolates which may be representative of badger – badger transmission. Markers have been dispersed to aid visualisation of points collected from individuals in the same social group.

The west to east transition rate was higher than the east to west transition rate (west to east: mean rate = 1.51, HPD intervals 0.045 – 3.744, east to west: mean rate = 0.43, HPD intervals 0.09 – 1.34). ESS values in both BEAST runs were >200 indicating good convergence. The west to east transition rate was

well supported (Bayes Factor = 1603) whereas the east to west transition rate was not (Bayes Factor <3). This suggests that there is strong support for west to east transitions (i.e. eastern isolates having western ancestors) but no support for east to west transitions (western isolates having eastern ancestors) and is consistent with *M. bovis* infection spatially spreading west to east across the study area (Fig 7.9).

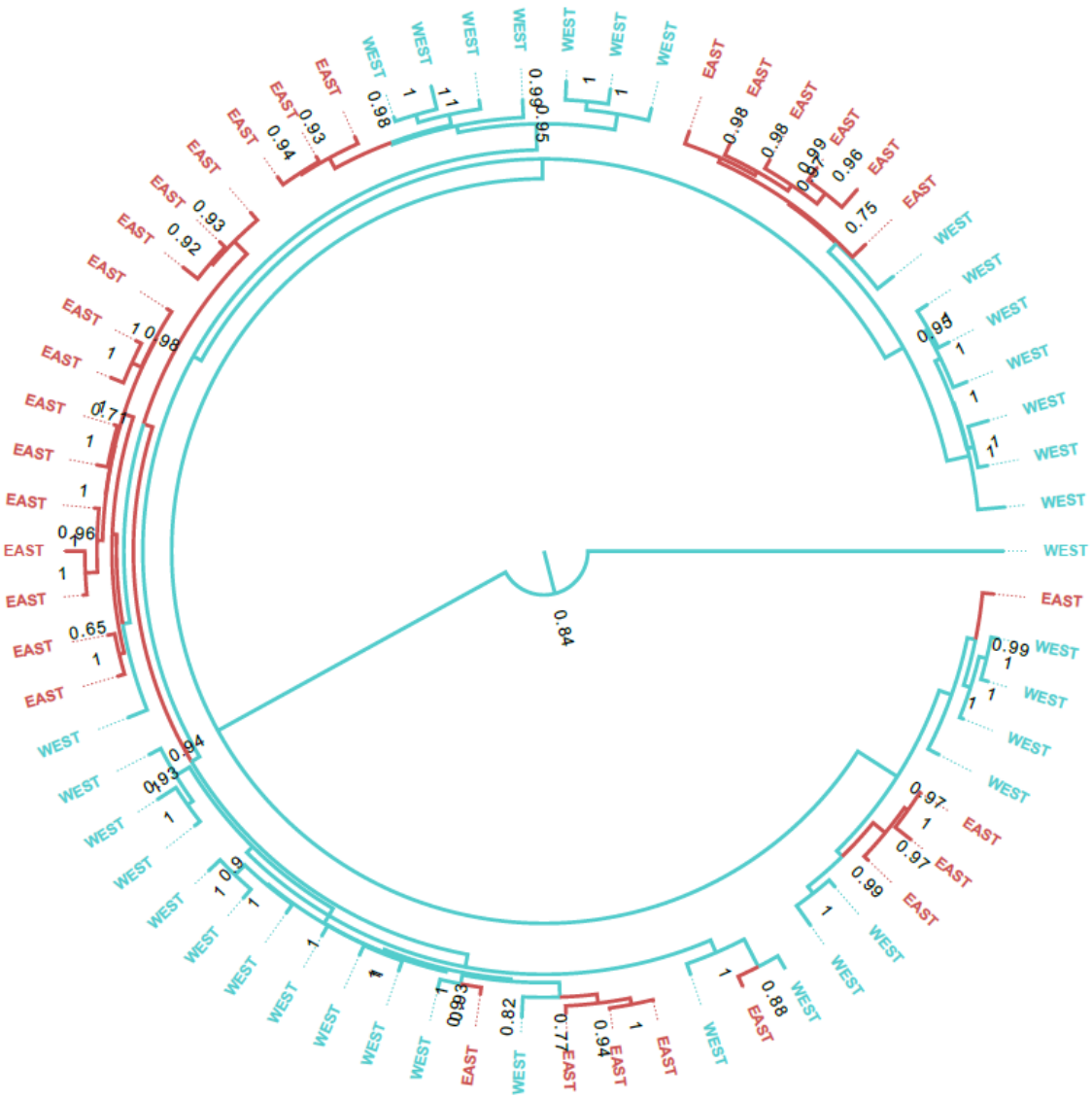


Fig 7.9 Bayesian phylogenetic tree built using BEAST of high quality Lineage 1 *M. bovis* isolates. Branches are coloured based on ancestral trait reconstruction, with support values for the ancestral state indicated. A value of 1 indicates that the assigned ancestral state (east or west) was chosen in all the trees generated within the posterior distribution.

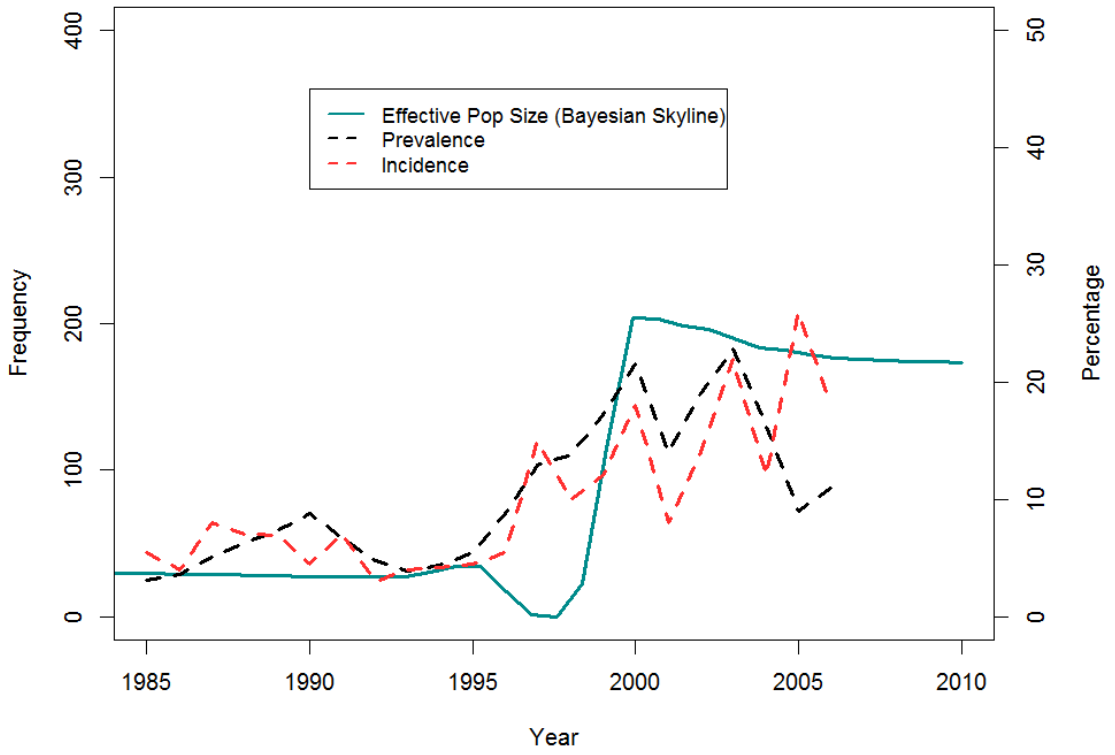


Fig 7.10 Bayesian skyline plot of effective population size over time (blue line) with population level estimates of incidence (red dashed line) and prevalence (black dashed line) generated from diagnostic test results from the Woodchester population (Walker 2012).

From Fig 7.10 there appears to be good agreement between the effective population size estimate from the BEAST analysis and the prevalence and incidence data. This is consistent with transmission of *M. bovis* in the Woodchester population increasing in the late 90's, peaking in 2000 before flattening off.

Isolate Genetic Distance, Social Group & Kin Structure

Genetic distances between *M. bovis* isolates were significantly lower from pairs of related individuals within the same social group than from pairs of unrelated individuals within the same social group (see Fig 7.11a). Isolates from unrelated group members differed by an average of 6 more SNPs when compared to isolates from related group members (mean difference = 7.5, $p < 0.001$). This is consistent with a kin-association model of *M. bovis* transmission within badger social groups. However, if a related badger was not in the same social group, genetic distances between isolates were smaller between unrelated group members (mean difference between Related Out SG category vs NonRel Out SG = -2.5, $p < 0.001$, Fig 7.11b). Isolates from within the same individual had significantly shorter genetic distances than isolates from related individuals in the same social group, although this distance was only 2 SNPs on average. ($p < 0.001$, see Fig 7.12). Significant differences between all categories are visualised in Fig 7.13.

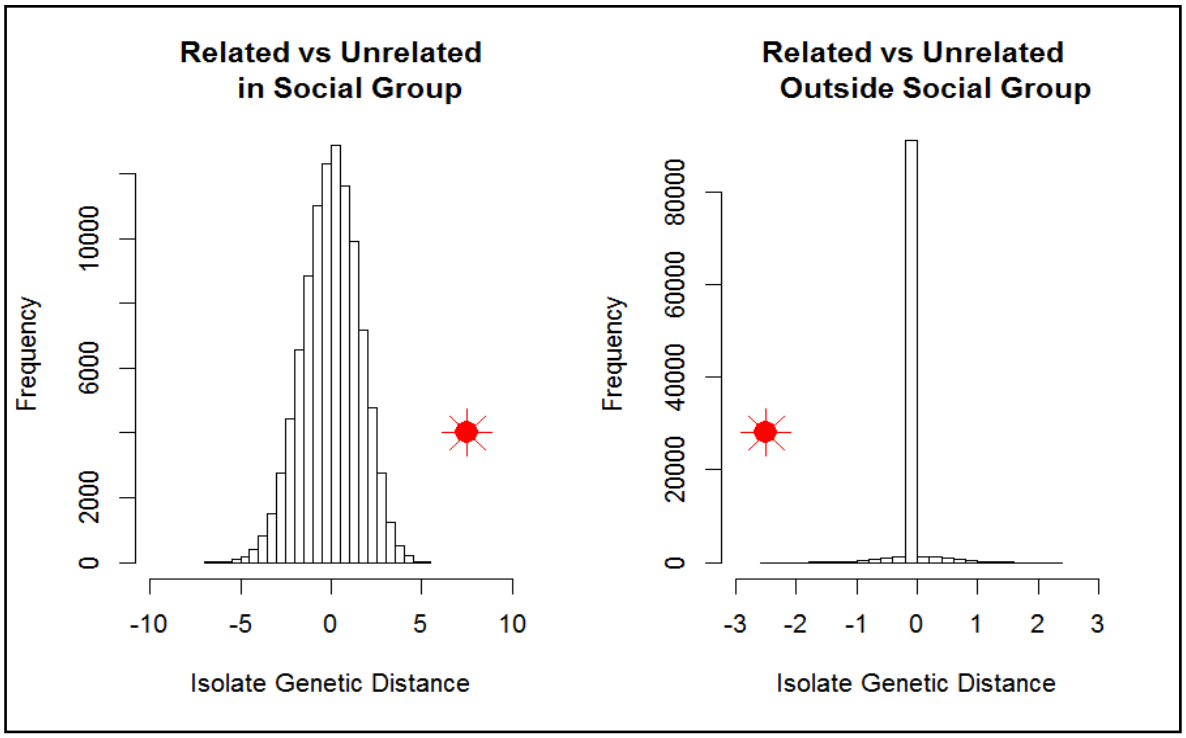


Fig 7.11 Permutation results for comparisons of mean genetic distance between *M. bovis* isolates from related and unrelated badgers a) inside the same social group and b) in different social groups. The null distribution based on shuffling isolate genetic distances across the categories is shown; in both cases the observed genetic distance between categories (indicated in red) is significantly different from the null distribution ($p < 0.01$ in both cases).

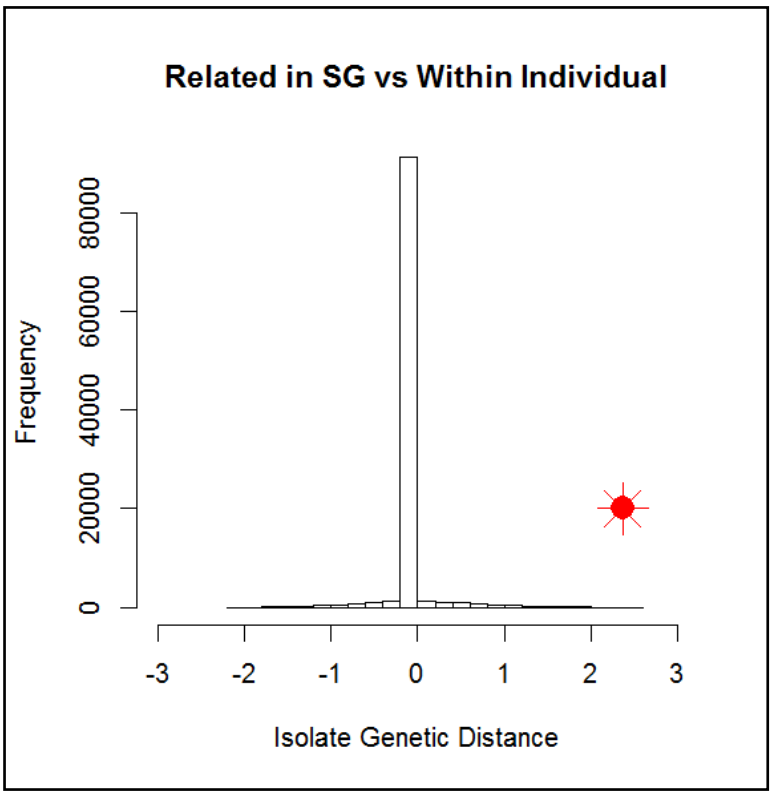


Fig 7.12 Permutation results for comparisons of mean genetic distance between *M. bovis* isolates from related badgers within the same social group and isolates taken from the same individual. The null distribution based on shuffling isolate genetic distances across the categories is shown; the observed genetic distance between categories (indicated in red) is significantly different from the null distribution ($p < 0.01$).

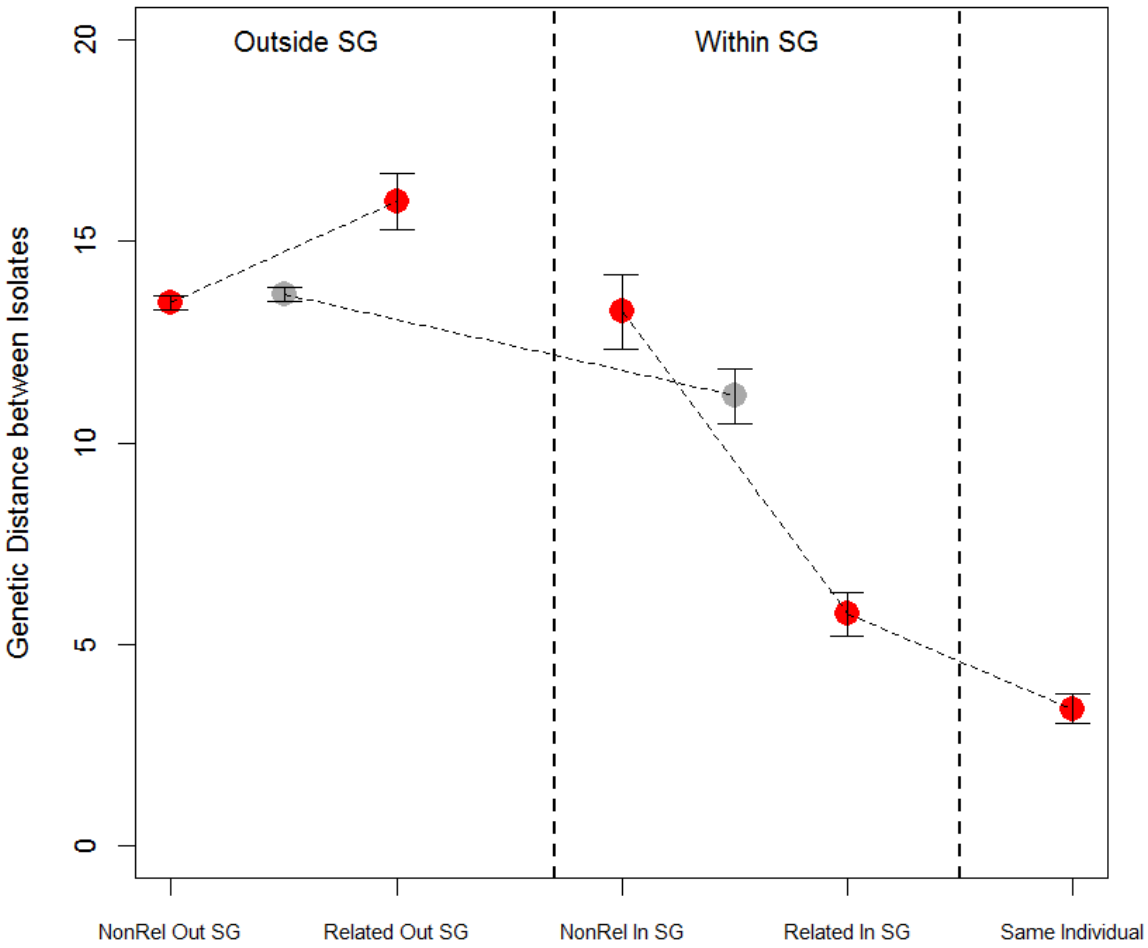


Fig 7.13 Mean genetic distances between *M. bovis* isolates obtained from related and unrelated individuals. Isolates are further grouped based on whether they are members of the same social group or not. Red points indicate mean values based on social group membership and relatedness, arrows indicate standard errors. Grey points indicate the average genetic distances for individuals in different social groups as compared to isolates from individuals in the same social groups. Significant differences between categories are indicated by dashed lines (as determined by permutation testing).

7.5 Discussion

These results describe one of the first applications of whole genome sequencing technology to look at *M. bovis* transmisison within a wildlife population. Examination of the WGS data supports the spatial spread of a dominant lineage of *M. bovis* from west to east across a high density badger population over a 10 year period. This suggests that the change in *M. bovis* spatial arrangement captured in Chapter 5, where infection distribution was initially clumped and spatially restricted to the western region of the study area and subsequently was more spatially dispersed, is primarily driven by transmission from badger social groups in the west of the study area to those in the east rather than seeding of infection from the surrounding badger of cattle population. Changes in the genetic population structure and badger movement patterns (as evidenced in Chapter 6) are potential drivers of this spatial spread, resulting in a change from spatially stable infection loci in the west to a more widespread infection distribution across the study area, as characterised in Chapter 5. There are other potential explanations; it may be that infection prevalence reached a critical level in the western social groups and subsequently spilled over into neighbouring groups, driving the west to east spatial spread observed. Alternatively, stochastic events, such as the movement of certain individuals transmitting infection between social groups, or breeding encounters between individuals from different groups resulting in transmission contacts, rather than any population wide demographic change. However, given the evidence that movement patterns and population mixing patterns have significantly changed, I would suggest that these are the most likely explanation for the observed change in spatial distribution of *M. bovis* infection within this study population.

The spoligotype and VNTR data available for the *M. bovis* isolates from the badger population revealed a near-monoculture of a single spoligotype; spoligotype 17. Although VNTR typing added a little more resolution; in terms of looking at fine-scale transmission dynamics within the badger population, the options are limited. In contrast, WGS revealed the presence of multiple lineages and further, allowed analysis of pairwise distances between isolates to investigate transmission dynamics within badger social groups. The availability of whole genome sequence data from the neighbouring cattle population

(Crispell 2017) suggests that there have been multiple occasions where *M. bovis* has been transmitted from the cattle population to the badgers within the study area, however on the whole, these lineages have failed to establish.

WGS revealed the presence of a dominant lineage, Lineage 1, which appeared to be spatially widespread across the study area and was present over the entire time period from which the sequenced samples were isolated. Over half of these isolates were assigned to badger-cattle transmission clusters where the direction of transmission was thought to be from badgers to cattle. However, just under half of the Lineage 1 badger isolates were not included in any badger-cattle transmission cluster. These could be examples of badger-badger transmission events; it has been suggested that the low divergence between many of the badger samples in this lineage may be consistent with direct transmission of this lineage within the badger population (Trewby 2016). It is important to note however that by comparison to the badger infections, *M. bovis* infections from the surrounding cattle population are likely to be considerably under sampled, as only one isolate per herd breakdown is sent for genotyping. It may therefore be the case that spillover transmission of the cattle-derived infections has gone undetected within the badger population. Although other lineages were noted in the badger population, these were generally only found in a small number of badgers and were generally linked to transmission from the cattle population which failed to establish within the badger population. Lineage 3 and 4 were introduced into the east of the study area and therefore may have contributed to the decrease in spatial clustering of *M. bovis* infections noted particularly in the eastern region of the study area, which I characterised in Chapter 5..

The high prevalence of Lineage 1 raises an intriguing possibility that this dominant lineage within the badger population is a 'badger-adapted' genotype, while the introductions from the cattle population are 'cattle-adapted' and this is the reason that they have failed to persist within the badger population. However, given the low effective population size of *M. bovis* in Great Britain (Smith, Gordon et al. 2006) (implying that selective pressures are likely to play a lesser role in the evolution of these bacteria compared to genetic drift), in addition to the very slow evolutionary rate of *M. bovis* in general, it has been suggested that the evolution of badger- or cattle-adapted lineages would be

surprising over the temporal and evolutionary scales covered by these sequences (H. Trewby, *pers comms*). A plausible explanation for the establishment of Lineage 1 as the dominant lineage within the badger population in this study is that it is due to stochastic events, for example a critical number of badgers becoming infected with this lineage to allow onward transmission. The distribution of the different lineages within the Spoligotype 17 group of isolates echoes that of the distribution of spoligotypes; with one major dominant lineage and occasional detection of other lineages. Whether the original source of the Lineage 1 infection in the badgers was from cattle or from the unsampled badger population surrounding the study area is not known; it may be of value to sequence some of the archived *M. bovis* isolates from the Woodchester badgers from 1990 -2000 to shed further light on this.

Through applying whole genome sequencing to a well characterised host-pathogen system this study has demonstrated the value of the added resolution afforded by whole genome sequencing in capturing spatial spread of *M. bovis* infection across a wildlife reservoir population. Within the immediate context of bovine tuberculosis transmission in the UK, whole genome sequencing is likely to be a useful tool in determining the proximate causes of new cases of bovine tuberculosis infection in UK cattle at the fringes of the endemic areas, helping to distinguish whether infection is seeded from livestock movements or the presence of infected wildlife (Benton, Delahay et al. 2014). Combining whole genome sequencing with epidemiology is a promising approach as a means of conducting in-depth investigations of *M. bovis* transmission dynamics, adding resolution and helping to distinguish between local persistance or introduction (Crispell, Zadoks et al. 2017). Theoretically, the discrete traits analysis approach used in this study could be used to assess the relative transmission rates from cattle to badgers and badgers to cattle, as has been attempted recently to assess interspecies transmission rates of *M. bovis* between possums and cattle in New Zealand (Crispell, Zadoks et al. 2017). However, this relies heavily on the availability of a balanced sampled of isolates from each host species, and estimates of inter-species transmission are sensitive to the sampling strategy used (Crispell, Zadoks et al. 2017). If one population is undersampled (De Maio, Wu et al. 2016), or if populations are not sampled evenly with reference to the underlying infection prevalence (as would often be

the case where infection is present in a wildlife reservoir), the method is likely to overestimate the prevalence/occurrence of transitions from the over-sampled to the undersampled population, thus generating biased estimates of transition rates between populations ((Crispell, Zadoks et al. 2017), H. Trewby, *pers comms*). Temporal sampling biases can also influence inter-species transmission estimates; if there is dense sampling of wildlife in the early part of the time period, followed by later dense sampling of livestock, this has been demonstrated to result in the dominant direction of spread being estimated as wildlife to livestock (Crispell, Zadoks et al. 2017). Other analytical options exist for estimating inter-species transmission rates however which can better deal with such biases, for example (De Maio, Wu et al. 2016). It should be noted also that there is no whole genome sequence information on the *M. bovis* badger population surrounding the Woodchester Park population. Woodchester Park is located in a high-incidence area of cattle TB and, given the observed prevalence in the study population it is highly likely that the surrounding badger population is also infected to some unknown extent. Although there are modelling approaches which can be employed to deal with an unsampled potential reservoir of disease (De Maio, Wu et al. 2016),(J. Crispell, *pers comm)*, it would be valuable to collect *M. bovis* isolates from the surrounding badger population in order to gain a more complete picture of the transmission processes operating within the Woodchester population, between the Woodchester population and the surrounding badger population, and between the badger and local cattle population.

In addition to helping us understand the changing spatial distribution of *M. bovis* infection within this study population, the availability of the sequence data has allowed us to examine the importance of kin structure and social group structure on disease transmission in an alternative way to that employed in Chapter 3. Shorter genetic distances between *M. bovis* isolates from related individuals within the same social group, as compared to isolates from unrelated individual within the same social group, supports the importance of kin structure in transmission of *M. bovis* within badger social groups (Benton, Delahay et al. 2016). Outside the social group however genetic distances between isolates were shorter between unrelated badgers when compared to related badgers outside the social group. This is consistent with recent findings from social

network analysis in this population, where contact rates with related badgers outside of the group were lower than those with unrelated badgers outside the group; this has been suggested to be an inbreeding avoidance mechanism. (Steward 2017). This result adds further weight to the proposal that where a local population is partitioned into small groups there is not a uniform transmission rate but rather two; within groups and between groups (Loehle 1995). The results described in this chapter support the importance of kin structure in disease transmission within badger social groups (as explored in detail in Chapter 3), consistent with related individuals having higher contact rates and living in closer proximity (Loehle 1995) within the social group. However, they also suggest that outside the social group, individuals may actively avoid contact with relatives and hence transmission may be more likely between unrelated individuals. In order for such inbreeding avoidance behaviours to evolve, kin recognition must be possible and inbreeding depression must be severe enough to incur some fitness cost (Pusey and Wolf 1996). Whether either of these conditions can be met in badgers is unknown, however the results in Chapter 4 suggest a potential survival or fitness cost of inbreeding. This is an area that warrents further investigation and the future availability of the pedigree for this population will be of great value in this.

These results have important implications for our understanding of the complexity and heterogeneity in disease transmission rates in socially structured populations. Behaviours such as inbreeding avoidance, which may vary in strength depending on the demographic parameters in a given year could have important implications for transmission of disease between social units and hence favour spatial spread of a pathogen across a landscape. Although further work is required to get a more complete picture in the current system, these results provide an encouraging indication that whole genome sequencing technologies will have much to add to our understanding of fine-scale transmission dynamics within wildlife disease systems.

CHAPTER 8: General Discussion and Synthesis

8.1 Overview

In this final chapter, I will discuss how the research in this thesis has contributed to three general areas: spatial epidemiology; the role of social and kin structure on pathogen transmission; and the impact of host genotype on infection outcomes. Suggestions for further areas of research are also discussed.

8.2 Spatial Epidemiology

Incidence and prevalence of disease are dynamics processes through time. When new pockets of disease emerge, it is important to understand whether they are driven by spatial spread of existing infection across a population or are a result of a new introduction of infection. If it is determined that spatial spread has occurred, understanding the drivers of such spread can help to predict future disease dynamics.

8.2.1 Capturing change in pathogen distribution

One of the motivating factors behind the research proposal for this thesis was a particular epidemiological question, based on an observation - specifically that the spatial distribution of *M. bovis* infection within the Woodchester Park badger population had changed over time. This was an untested hypothesis based solely on observations from within the research group. A principal aim of this thesis was to test rigorously whether there was empirical evidence of a spatial change and further to suggest possible reasons for the change. Finding an answer to this question was important as our existing understanding of how *M. bovis* infection is distributed in badger populations at the outset of this thesis was that it tended to remain highly spatially clustered with limited evidence of spread between neighbouring social groups (Delahay, Langton et al. 2000) in the absence of population management. Changes in the spatial distribution of *M. bovis* infections as a result of culling badger populations have been well documented however (Woodroffe, Donnelly et al. 2005, Woodroffe, Donnelly et al. 2006, Jenkins, Woodroffe et al. 2007). Through analysing spatial trends in

M. bovis distribution over a two decade period, I was able to demonstrate that the previously clustered, stable spatial distribution of *M. bovis* infection within the Woodchester Park population had broken down, characterised by a decrease in spatial clustering (robust to background changes in *M. bovis* infection prevalence at the population level) with badger social groups becoming more homogenous in their infection profiles (Chapter 5). These findings revealed changes in epidemiological trends occurring over relatively long time periods and also challenged the view that *M. bovis* distribution necessarily remains stable, and tightly clustered, in unmanaged badger populations. However, the rate of the breakdown of spatial clustering observed in this unmanaged population is well below that noted as a result of culling operations (Jenkins, Woodroffe et al. 2007). This suggests that although stable foci of *M. bovis* infection can destabilise in the absence of population management, the process is far more rapid when it occurs as a result of culling induced social perturbation (Jenkins, Woodroffe et al. 2007).

8.2.2 Genetic Population Structure, Movement and TB transmission

Once it had been established that a real change had been detected in the spatial distribution of *M. bovis* infections in the study population, attention turned to the potential drivers of that change. The movement of badgers is known to be an important predictor of *M. bovis* incidence risk, with years of high inter-group movement often followed by years of elevated incidence (Rogers, Delahay et al. 1998). However, previous investigations have relied entirely on capture data to infer inter-social group movement (Vicente, Delahay et al. 2007) and to calculate population level movement metrics (Rogers, Delahay et al. 1998). By characterising host genetic population structure in tandem with the capture database, I was able to demonstrate that, where initially badger social groups were genetically isolated by distance (i.e. spatially closer badger social groups were also genetically more similar), this effect disappeared for a subsequent period, consistent with more widespread mixing of individuals in the population (Chapter 6). This was supported by the observation that there was interannual variation in the probability of an individual badger moving social groups, with movement favoured during years of lower population density. This finding has key management implications, as lowering population density through culling is therefore likely to increase the individual probability of a

surviving badger moving social groups. Increases in ranging behaviour and inter-social group movement have previously been documented in response to culling interventions (Tuyttens, Macdonald et al. 2000, Carter, Delahay et al. 2007, Riordan, Delahay et al. 2011, Bielby, Donnelly et al. 2014). When considering unmanaged populations, a relationship between population density and movement has been suggested based on comparing movement probabilities from populations of varying densities (Woodroffe, Macdonald et al. 1995), although some studies report no effect of density on the likelihood of badger dispersal (Macdonald, Newman et al. 2008). The negative relationship between population density and movement probability is in line with the 'social fence' hypothesis (Hestbeck 1982) which predicts low dispersal in high-density populations, as a result of an increased individual cost of dispersal.

Returning to the observation motivating this section of the thesis, I have been able to use both genetic and observational data in combination to suggest that the Woodchester study population has undergone a period of demographic flux, potentially prompted by a fall in population density, favoring inter-group movement and transmission of *M. bovis* infection between social groups which were previously socially isolated, with limited interactions between group members. However, there are a few important points to note. Firstly, the extent to which the Woodchester badger population is truly 'unmanaged' should be considered carefully, as there have been ad hoc reports of illegal culling of badgers at the periphery of the study population (D. Delahay, P. Spyvee, *pers comm*). Secondly, the Woodchester study site is not a closed system; rather it is surrounded by a wider badger population about whom we have little information on *M. bovis* prevalence. Thirdly, *M. bovis* can be transmitted from badgers to cattle and the Woodchester study population is close to a number of cattle herds, many of which have experienced TB breakdowns at various points over the course of the study. Therefore an alternative explanation for the observed change in spatial distribution of *M. bovis* infections, characterised in Chapter 5, is that it is due to the seeding of infection from an outside source, such as local cattle or adjacent badger populations. In Chapter 7 I used genomic data from a sample of *M. bovis* isolates to visualise the multiple occasions where infection from badgers to cattle has been inferred but apparently not established within the badger population (Crispell 2017). Further,

I was able to demonstrate evidence of spatial spread of a dominant lineage of *M. bovis* across the Woodchester badger population, moving from social groups in the western region where infection has long been documented, into the eastern social groups which were consistently low prevalence for an extended period (Delahay, Langton et al. 2000).

8.2.3 Management Implications

Whole genome sequencing has only very recently been applied to examining transmission dynamics of bacterial pathogens in wildlife (Kamath, Foster et al. 2016) and hence the results presented in this thesis represent one of the very first applications of this technology to *M. bovis* transmission dynamics in wildlife (Biek, O'Hare et al. 2012, Glaser, Carstensen et al. 2016, Crispell, Zadoks et al. 2017). There is considerable interest amongst policy makers and scientists as to what WGS technologies can reveal about *M. bovis* transmission dynamics. It may provide insights into transmission between badgers and cattle and within each of the host populations, with suggestions that it may be possible to quantify the role that badgers play in the persistence of TB infection in cattle herds. However, it has been suggested that the most valuable approach to best make use of the data provided by these new technologies will be to integrate WGS data with existing epidemiological data (Trewby, Wright et al. 2016, Crispell, Zadoks et al. 2017). In this way, more bespoke 'farm-scale' management interventions may be possible which take into account the heterogeneity of transmission routes, including the role of wildlife, in different contexts. Given the slow rate at with the *M. bovis* bacteria accumulates mutations (Biek, O'Hare et al. 2012, Trewby, Wright et al. 2016, Crispell, Zadoks et al. 2017), it is thought to be unlikely that even WGS of *M. bovis* isolates will be able to uncover transmission links between individual animals, however it is likely to be useful to look at transmission dynamics between cattle herds (Crispell, Zadoks et al. 2017) or potentially badger social groups. Through using the newly available WGS data from the Woodchester Park badger isolates in tandem with the extensive epidemiological and ecological data available, I have been able to address a particular epidemiological question based on an observation of changing spatial distribution of disease. This highlights the benefits of a multi- faceted approach, incorporating traditional spatial epidemiology to capture changing distributions of infection,

using longitudinal datasets to detect demographic trends and relationships, coupled with genetic and genomic approaches to capture changes in host population structure and to track the pathogen itself through space and time. My findings also demonstrate the potential of WGS technologies to reveal transmission dynamics of *M. bovis* at a very restricted spatial scale (Biek, O'Hare et al. 2012).This is in contrast to traditional *M. bovis* genotyping approaches which have largely been used to infer long-distance transmission events from inter-regional differences in strain types (Smith, Dale et al. 2003). There is increasing interest in using WGS approaches to inform TB management at the farm-level scale, for example to trace the source of outbreaks, to identify the local introduction of infection or local persistence (Crispell, Zadoks et al. 2017), or to suggest where local wildlife may act as a reservoir of infection.

8.3 The Impact of Social Structure and Kin Structure on Pathogen Transmission

It is well established that social structure influences pathogen transmission dynamics in host populations (Loehle 1995). In particular, social structure has been found to favour the persistence of chronic diseases such as TB (Cross, Lloyd-Smith et al. 2005). Levels of pathogen infection are expected to be higher for social species compared to solitary species (Ezenwa 2004) and for individuals in larger social groups (Brown and Brown 1986), as contact rates between individuals are higher in social situations, facilitating pathogen transmission (Altizer, Nunn et al. 2003, Ezenwa, Ghai et al. 2016).

8.3.1 Social Group Size and TB transmission

The relationship between host density and parasite transmission (Anderson, May et al. 1992) has been extensively investigated. Within badger social groups, transmission of *M. bovis* is not considered to be density dependent (Cheeseman, Wilesmith et al. 1989) with research findings suggesting higher prevalence (Woodroffe, Donnelly et al. 2009) and incidence (Vicente, Delahay et al. 2007) in smaller or shrinking social groups. The research described in Chapter 3 confirms that cubs born into larger social groups are at a lower risk of testing TB positive in their first year of life, suggesting a dilution effect consistent with the result previously reported in badgers (Woodroffe, Donnelly et al. 2009).

This is consistent with the 'anti-parasite' hypothesis of group-living, where enhanced pathogen resistance or tolerance has been associated with group living (Ezenwa, Ghai et al. 2016). By considering the number of test positive and test negative residents in a badger social group, I was able to demonstrate that the risk of a badger cub testing positive within its first year of life is positively related to the number of test positive adults in the group, but negatively related to the number of test negative adults. It is important to note that these are not independent factors; groups with a large number of test negative adult residents are unlikely to also have a large number of test positive adult residents, as badger social groups tend to typically contain four to eight adults (Roper 2010). A badger cub born into a large social group predominantly comprised of test negative adults may therefore be at a lower risk of acquiring infection than a badger cub born into a smaller social group comprised of predominantly test positive adults, even though the badger in the larger social group is mixing with a larger number of individuals.

8.3.2 Parallels with TB transmission in human contexts

The most risky scenario for a badger cub is likely to be where cubs are born into large badger social groups with a high prevalence of TB infection, where there is likely to be lots of mixing with infected residents and further, they may be in closer proximity to other group members due to a higher number of animals sharing the limited sett space. This is comparable to our understanding of TB transmission within human populations. Transmission of TB in humans is considered to be most likely where infected persons are in close contact with others in confined spaces, with overcrowding further increasing the risk of transmission (Beggs, Noakes et al. 2003). In humans, TB mortality rates have been directly related to the number of people living in a house in a specific population (Elender, Bentham et al. 1998), even where there is no relationship between overcrowding and TB mortality at a regional scale. Even though prolonged close contact is thought to generally be necessary for transmission of TB amongst humans (Antunes and Waldman 2001), there are reports of rapid transmission events occurring, for example during long haul flights (Kenyon , Valway et al. 1996) and other scenarios (Houk, Baker et al. 1968, Nardell, Keegan et al. 1991) in which infected persons are in close contact with others in confined spaces.

Returning to the badger context, it is possible that rather than badger social group size itself being important in determining infection risks for cubs, an interplay of several factors may be involved. These might include the proportion of the group who are already infected, the number of individuals living in the sett and the available living space in that sett which determines the extent to which individuals are crowded together. If, as in humans, rapid transmission of TB is possible between individual badgers in shared confined spaces (Houk, Baker et al. 1968, Kenyon , Valway et al. 1996) then the extended periods of time badgers spend underground in the sett may represent high-risk periods for transmission (Corner, Murphy et al. 2011, Ní Bhuachalla, Corner et al. 2014). The underground environment of the badger sett, where airflow is restricted, humidity is high and temperature is relatively stable (Roper 2010) may favour the transmission of TB amongst badgers. Radio tracking studies suggest that badgers tend to rest in underground sleep chambers singly or in pairs, although up to five individuals have been recorded sharing a single chamber (Kowalczyk, Zalewski et al. 2004). In the winter, badgers tend to occupy a single sleeping chamber (Butler and Roper 1996, Roper, Ostler et al. 2001, Kowalczyk, Zalewski et al. 2004). If chamber sharing were more common in larger social groups, due to constraints on available space, this would suggest that inter-individual transmission risks would be higher for badgers in large social groups where a substantial proportion of the group are infected. There is currently very little published work on den-sharing in badgers (Butler and Roper 1996, Roper, Ostler et al. 2001, Kowalczyk, Zalewski et al. 2004) and none on whether den-sharing behaviour changes with social group size and this may be a valuable area of future investigation. Bite-wounding is known to occur between members of the same social group as well as being a territory defence behaviour (Stewart, Ellwood et al. 1997), has been associated with more rapid progression of clinical TB in badgers (Gallagher and Clifton-Hadley 2000), and is more frequent at higher population densities (Macdonald, Harmsen et al. 2004). However, whether the probability of being bitten is higher in larger social groups is currently unknown.

8.3.3 Kin Structure and Pathogen Transmission

In addition to uncovering heterogeneities in cub infection risk related to group size and composition, I also used host genotype data to show that kin structure

within badger social groups adds further complexity to transmission dynamics. There is limited published work on the impact of kin structure on pathogen transmission dynamics (Grear, Samuel et al. 2010, Dharmarajan, Beasley et al. 2012, Vander Wal, Edye et al. 2013) and none which focuses on badgers and TB. The research described in Chapter 3 demonstrates that badger cubs have higher infection risks in social groups containing resident infectious adults, and these risks are exaggerated when cubs and infectious adults are closely related. These results were supported by analysis of the whole genome sequences of *M. bovis* described in Chapter 7, with isolates cultured from relatives within the same social group being more genetically similar than those from non-relatives within the same social group. Interestingly, although *M. bovis* isolates from individuals resident in the same social group were more genetically similar than those from individuals resident in different social groups, consistent with badger social structure limiting transmission between individuals in different social groups, when kin structure was accounted for, the interpretation of this relationship changed. Pairwise comparisons of isolates from unrelated badgers revealed no significant genetic distance between isolates from individuals within the same social group and individuals in different social groups. This suggests that transmission between unrelated badgers is just as likely regardless of social group membership, which is surprising. Outside the social group, isolates from unrelated badgers were more genetically similar than isolates from related badgers. This is consistent with badgers avoiding relatives that reside in other social groups, potentially as an inbreeding avoidance strategy. This result is consistent with recent data generated from social network analysis of this population, indicating that badgers spend less time with relatives outside of their group (Steward 2017).

Overall, the results of these studies on heterogeneity in cub infection risk and the comparison of the WGS isolates of *M. bovis* lend strong support to the importance of kin structure in disease transmission in badger populations. The results suggest that within the social group itself, kin structure produces heterogeneity in contact patterns, with the result that being related to an infected individual confers an additional risk of acquiring infection. Interestingly, this conclusion contrasts somewhat with recent findings using proximity collars in this population which suggested that, with social groups, adult badgers do not

spend more time with relatives than non-relatives (Steward 2017). It should be noted however, that as the collars can only be deployed on adult badgers, the data generated only represents patterns of contact amongst adults whereas the research covered in Chapter 3 of this thesis focuses on infection risk in cubs. It is likely that cubs mix preferentially with kin during early life, particularly their mothers (Roper 2010), potentially generating heterogeneity in infection risk depending on the infection status of relatives. Additionally, the proximity collar data only describes spatial proximity, registering when adult badgers come within half a metre of one another (Steward 2017) rather than describing particular behaviours which may enhance transmission risks. Studies of human TB have revealed a strong relationship between transmission risk and spatial distance between individuals (Houk, Baker et al. 1968, Beggs, Noakes et al. 2003, Pantelic, Sze-To et al. 2009). For example in a report of rapid TB transmission during a long haul flight, those seated closest to the index case were much more likely to acquire infection (Kenyon , Valway et al. 1996) than those seated elsewhere in the aeroplane compartment. It may be the case that behaviours which bring badgers into very close contact, such as grooming, which has been linked to enhanced risk of acquiring TB in meerkats (Drewe 2010) are preferentially performed between relatives. In badgers the primary route of *M. bovis* infection is the lower respiratory tract following inhalation of small infectious aerosol particles (Corner, Murphy et al. 2011), therefore behaviours which bring individuals into very close proximity are likely to incur a higher transmission risk. Although a wealth of observational work has been carried out on badger behaviour (Roper 2010), very little has taken account of kin-structure (Dugdale, Ellwood et al. 2010), hence at this stage it is only possible to speculate as to the mechanism of the observed effects.

8.3.4 Management Implications

These findings have key management applications as they help us to explain the observation of a herd-immunity effect in badgers following low-intensity vaccination campaigns (Carter, Chambers et al. 2012). They also highlight the heterogeneity in early life infection risk experienced by badger cubs, which has important implications for modelling TB transmission. From a wider perspective, they have highlighted the roles of kinship and kin-association as important (Ezenwa, Ghai et al. 2016), and often cryptic, drivers of disease

transmission in social mammals. Even in large, open populations, kin association can result in unexpected pathogen transmission outcomes.

8.4 Host Genotype and Pathogen Exposure Outcomes

As I have explored above, social structure and kin structure can influence exposure risk to *M. bovis* in badger cubs. The risks of exposure can vary based on social group membership, kin structure (Benton, Delahay et al. 2016) and individual behaviour illustrated for example by movement between social groups (Rogers, Delahay et al. 1998, Woodroffe, Donnelly et al. 2009) or position within a social network (Drewe 2010, Weber, Carter et al. 2013). However as well as heterogeneity in exposure risk, there is also heterogeneity in the outcome of exposure. This is the case in human TB, where it is estimated that over a third of the world's population has been infected with *M. tuberculosis* but only a minority of those individuals will ever go on to develop clinical disease or 'active' tuberculosis (Mack, Migliori et al. 2009). In badgers, the pathogenesis of TB infection is known to be complex, as not all exposed badgers become diseased, with a proportion mounting a successful immune response. In some cases, lesions develop but may remain dormant, such that the animal shows no clinical signs and is not infectious (Roper 2010), which may be considered a 'containment' or 'latent' phase (Gallagher and Clifton-Hadley 2000). It has been suggested that many or even the majority of infected badgers remain in a latent phase throughout their lives (Gallagher and Clifton-Hadley 2000, Murphy, Gormley et al. 2010, Corner, Murphy et al. 2011). However, in a proportion of exposed badgers, the immune response is insufficient to contain, or 'wall off' the bacteria within lesions. The mycobacteria can then escape, spread to new body sites and the individual can become infectious, characterised by bacterial shedding through a range of routes (Gallagher and Clifton-Hadley 2000). The amount of bacteria shed by an infectious badger is related to the progression of pathology (Nolan 1991), and so individuals with evidence of more progressed disease are likely to be more important in the onward transmission of infection to susceptible individuals.

What factors determine whether an exposed badger goes on to develop clinical TB? There are a number of possibilities. It may be related to the size of the infective dose received (Dean, Rhodes et al. 2005), or to the route of

transmission (Gallagher and Clifton-Hadley 2000, Gavier-Widén, Cooke et al. 2009), the condition of the animal at the time of exposure, environmental factors or the genotype of the *M. bovis* strain (Aguilar León, Zumárraga et al. 2009). Another possibility is that host genotype influences the outcome of exposure to *M. bovis*. Although this has been explored in other wildlife species (Dorman, Hatem et al. 2004, Acevedo-Whitehouse, Vicente et al. 2005, Trinkel, Cooper et al. 2011) and livestock (Allen, Minozzi et al. 2010, Brotherstone, White et al. 2010, Vordermeier, Ameni et al. 2012), no published work exists on host genotype and outcomes of TB exposure in badgers. In Chapter 4 I used host genotype data to investigate how individual inbreeding coefficients influenced the likelihood of badgers exposed to *M. bovis* exhibiting progressed disease (as measured by excretion of *M. bovis* bacilli and the result of a serological test known to be more sensitive in individuals with progressed infection (Chambers, Crawshaw et al. 2008)). Exposed badgers with higher inbreeding coefficients were more likely to test positive to an antibody test at a given capture event (indicative of progressed disease). The impacts of inbreeding also became stronger with age. I also found evidence of single locus effects predicting the likelihood of an exposed badger becoming culture positive (indicative of infectiousness). This is the first demonstration of a link between badger genotype and TB progression and only one other published study from wildlife links genetic profile to TB progression (Acevedo-Whitehouse, Vicente et al. 2005). I also found potential evidence of a survival cost to inbreeding, as I observed that the most inbred individuals were not present in the most advanced age classes. Although this did not form part of the formal analysis in Chapter 4, it is an avenue that should be explored in future research.

8.4.1 Management Implications

The above findings contribute to our understanding in number of areas. Firstly, from an evolutionary perspective, the observed age-mediated inbreeding relationship lends support to the mutation accumulation hypothesis of senescence. This has rarely been demonstrated in natural populations (Keller, Reid et al. 2008). Secondly, they suggest further heterogeneity in transmission risk, with the outcome of exposure to *M. bovis* varying between badgers of different genetic profiles. This has implications for our understanding of TB transmission in badger populations. Additionally, management of badger

populations through culling may alter the host genetic population structure in as yet unknown ways, potentially increasing levels of inbreeding through reducing population density or decreasing inbreeding as surviving individuals range more widely (Riordan, Delahay et al. 2011). Finally, if there are true fitness costs to inbreeding in badgers, as my findings suggest, then we may expect inbreeding avoidance mechanisms to have evolved, such as dispersal (Charlesworth and Charlesworth 1987, Townsend, Clark et al. 2010) from social groups or extra-group mating (Carpenter, Pope et al. 2005, Annavi, Newman et al. 2014), both of which are known to occur in badgers. However, if the fitness costs of inbreeding are not seen until later life, when badgers have already successfully reproduced then they may not be filtered out by natural selection (Medawar 1952, Hamilton 1966). Although there is further work to be done to more extensively explore the relationship between inbreeding, *M. bovis* progression and host survival, the results presented in this thesis provide an encouraging starting point.

8.5 Conclusions

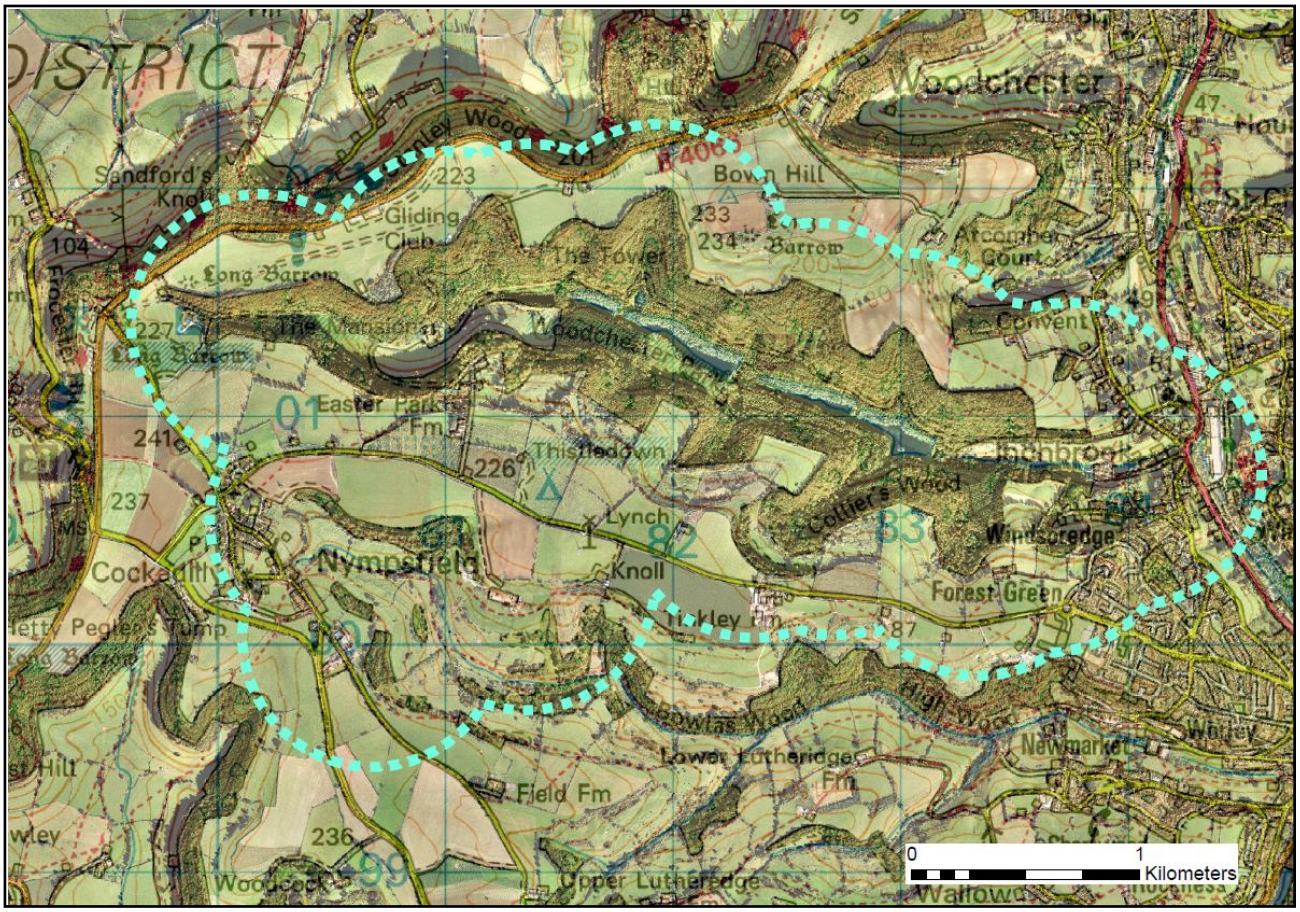
In the same way as using multiple diagnostic tests can improve the likelihood of detecting infection, using multiple approaches to capture contact patterns amongst individuals may improve the probability of detecting epidemiologically important interactions. The data chapters in this thesis have included a wide array of analytical approaches from the fields of spatial epidemiology, genetics and genomics to reveal fine-scale transmission dynamics within a single badger population and to reveal further heterogeneity in transmission risk incorporating kin structure and host genotype. Through the availability of a detailed dataset generated from a valuable long-term study of a badger population, I have shown how traditional epidemiological approaches and lines of enquiry can complement more recently available data, such as that generated from whole genome sequencing. Currently, the potential of these technologies is vast but to a large extent, untested in bacterial pathogens of wildlife (Kamath, Foster et al. 2016). I have been able to demonstrate that, when coupled with ecological, genetic and epidemiological data from well-studied systems, they can be usefully applied to epidemiological questions and as such are likely to play an important future role in our understanding of pathogen transmission, within and between wildlife and livestock populations.

Further, the results of this thesis have contributed to our understanding of the epidemiology of bTB in badgers. A population of badgers has experienced spatial spread of bTB infection, in the absence of social perturbation driven by management interventions. This breakdown in disease clustering followed a period of demographic flux, as detected by changes in badger movement patterns and host genetic population structure. There is powerful evidence of kin structure in badger social groups strongly influencing transmission risk to resident cubs, despite an overall dilution effect of group size. I have presented evidence of links between genotype and the progression of bTB infection in badgers, however further work is required to determine whether genome wide of loci-specific homozygosity is generating this effect. As the challenge of managing bTB transmission between badgers and livestock populations continues in the UK and other contexts, these results will contribute to our understanding of fine-scale transmission dynamics within badger populations and have provided new insights into the role of badger genotype on infection outcomes.

Appendix: Badger Trapping and Sampling Regime at Woodchester Park

The research included in this thesis is entirely based on data collected from the long-term study of a badger population at Woodchester Park, Gloucestershire, UK. The study area consists of a central wooded valley, with a small number of lakes, surrounded by farmland (Delahay, Carter et al. 2006). Cattle are grazed in the core of the study area which is approximately 7 km^2 in area (See Map 1). The location of the study area falls within the designated 'High Risk' endemic area for bovine tuberculosis in cattle (DEFRA 2017).

Map 1 Location of the Woodchester Park study area, with core study area delineated in blue (OS Grid Ref: SO 80907 01377; Woodchester Mansion)



Long Term Population Study

The study of this population of badgers naturally infected with *M. bovis* first started in 1975, following the discovery a few years earlier of a badger found dead from generalised bovine tuberculosis on a farm in Gloucestershire which was experiencing a TB outbreak in its cattle herd (Muirhead and Burns 1974). The objective of the study was to provide epidemiological and ecological data to inform TB control policy (Delahay, Walker et al. 2013). It is the only study of its kind, detailing the natural dynamics of *M. bovis* infection in live badgers and is also one of the longest running population studies on any wild mammal (Delahay, Walker et al. 2013). The study has been funded since its inception by the UK Government, specifically Defra (Department for Environment, Food and Rural Affairs, formerly MAFF; Ministry of Agriculture, Fisheries and Food). Between 1990 – 2011 (the period on which this thesis primarily focuses, due to the availability of badger genotyping data), there were 10,578 live capture events of 2003 individual badgers, with each being caught an average of five times over the period of study (min = 1, max = 42). Approximately 93% of individuals were first caught as cubs, and hence were of known age.

Trapping

Badger traps are deployed at all active main setts four times a year. A closed season of February to April inclusive is observed to avoid the capture of heavily lactating females or highly dependent cubs during this period (Tomlinson 2013). Traps are pre-baited with peanuts for about four days, before being set to catch for two consecutive nights. Badgers that are trapped on the first night are held in the sampling facility overnight following general anaesthesia and sampling, to avoid recapture. Traps are checked early in the morning, and additionally from December through to January, late at night, in order to release any adult females with evidence of pregnancy or lactation. Trapped badgers are transferred to individual holding cages, which are labelled with the sett of capture, and transported to the holding facility where they are placed on a metal rack, arranged in social groups to minimise the potential for infection transfer between groups (Tomlinson 2013).

Image 1 Badger trap as used in routine trapping operations at Woodchester Park (photo credit, Amy Griffiths)

Sampling

Until 2001, captured badgers were anaesthetized using ketamine hydrochloride (Vetalar™ V, Pharmacia and Upjohn, UK) alone (Mackintosh, MacArthur et al. 1976), but subsequently this has been used in combination with medetomidine hydrochloride (Domitor®, Pfizer, UK) and butorphanol tartrate (Torbugesic®, Fort Dodge Animal Health Ltd, UK) (De Leeuw, Forrester et al. 2004). On first capture each badger is tattooed on the belly with a unique identifying code. At each capture badger sex and weight, and the location of the trap were recorded (Delahay, Walker et al. 2013). A variety of other measurements are taken (neck circumference and body length, body temperature), along with qualitative assessments of body condition, and tooth wear which can be used to estimate age (Delahay, Walker et al. 2011). Differentiation between cubs and adults is made on the basis of cubs having a smaller body size, narrower head, bright and silky pelage, and the presence of completely unworn and unstained teeth (Delahay, Walker et al. 2013). The year of birth is assigned as the year of capture for cubs but is unknown for animals first caught as adults. When the morphometrics described above have been recorded, along with the clinical sampling detailed below, badgers are returned to the holding cage from which

they had been removed and placed in lateral recumbency for recovery (Tomlinson 2013). Once badgers have recovered from anaesthesia, they are offered oral rehydration solution (Lectade Small Animal; Elanco Companion Animal Health, Basingstoke, UK) in drinking troughs fixed within the holding cages. Badgers from the first night of capture are held overnight and released the following morning. Badgers from the second night of capture are released later the same day. Release is at the sett of capture, close to sett entrances (Tomlinson 2013).

Diagnostic testing for *M. bovis* infection

A range of clinical samples are taken from the anaesthetised badgers in order to diagnose *M. bovis* infection. Blood is taken from the jugular vein and used for serological testing as follows. Between 1985 and 2005, the Brock ELISA (Rogers, Cheeseman et al. 1997) was the serological test used to assign TB status to individual badgers. However, from 2006 onwards, the Brock ELISA was abandoned due to doubts over its performance and poor correlation with other test results (Delahay, Walker et al. 2013). From 2006 onwards, the Brock ELISA was replaced with the improved Stat-Pak antibody test (Chambers, Crawshaw et al. 2008) and the gamma interferon (IFN) test for T-cell responses to *M. bovis* was introduced (Dalley, Davé et al. 2008). The combination of diagnostic tests used provides a biologically meaningful picture of the progression of disease within an individual (Tomlinson 2013). It is thought that the cell-mediated response (as measured by the gamma interferon test) is the first line response to *M. bovis* exposure, whereas the serological response (as measured by the ELISA test and StatPak) takes time to develop as infection progresses (Tomlinson 2013). A range of clinical samples are taken are also taken for the attempted culture of *M. bovis*; these include oesophageal and tracheal aspirates, swabs of open bite wounds or ruptured lymph node abscesses, urine and faeces (Tomlinson 2013).

Delineating social territories

Badger social group territories are mapped annually using a bait-marking technique (Delahay, Brown et al. 2000). In the spring of each year, active badger main setts are fed bait comprised of peanuts, syrup and indigestible coloured beads. Each sett is fed with a different colour of bead such that on

surveying the study area after a period of feeding the marked bait, the coloured beads are visible in the badger latrines. In this way, the territorial boundaries between the badger social groups can be delineated and visualised using ArcGIS.

References

Abernethy, D., P. Upton, I. Higgins, G. McGrath, A. Goodchild, S. Rolfe, J. Broughan, S. Downs, R. Clifton-Hadley and F. Menzies (2013). "Bovine tuberculosis trends in the UK and the Republic of Ireland, 1995–2010." Veterinary Record: vetrec-2012-100969.

Acevedo-Whitehouse, K. and A. A. Cunningham (2006). "Is MHC enough for understanding wildlife immunogenetics?" Trends Ecol Evol **21**(8): 433-438.

Acevedo-Whitehouse, K., J. Vicente, C. Gortazar, U. Hofle, I. G. Fernandez-De-Mera and W. Amos (2005). "Genetic resistance to bovine tuberculosis in the Iberian wild boar." Molecular Ecology **14**(10): 3209-3217.

Achtman, M. (2001). "A Phylogenetic Perspective on Molecular Epidemiology." Molecular Medical Microbiology **1-3**: 485-509.

Achtman, M., A. van der Ende, P. Zhu, I. S. Koroleva, B. Kusecek, G. Morelli, I. G. Schuurman, N. Brieske, K. Zurth, N. N. Kostyukova and A. E. Platonov (2001). "Molecular epidemiology of serogroup a meningitis in Moscow, 1969 to 1997." Emerg Infect Dis **7**(3): 420-427.

Aguilar León, D., M. Zumárraga, R. Jiménez Oropeza, A. Gioffré, A. Bernardelli, H. Orozco Estévez, A. Cataldi and R. Hernández Pando (2009). "*Mycobacterium bovis* with different genotypes and from different hosts induce dissimilar immunopathological lesions in a mouse model of tuberculosis." Clinical & Experimental Immunology **157**(1): 139-147.

Alexander, R. D. (1974). "The evolution of social behavior." Annual review of ecology and systematics **5**(1): 325-383.

Allen, A. R., G. Minozzi, E. J. Glass, R. A. Skuce, S. W. J. McDowell, J. A. Woolliams and S. C. Bishop (2010). "Bovine tuberculosis: the genetic basis of host susceptibility." Proceedings of the Royal Society B: Biological Sciences **277**(1695): 2737-2745.

Altizer, S., R. Bartel and B. A. Han (2011). "Animal migration and infectious disease risk." science **331**(6015): 296-302.

Altizer, S., C. L. Nunn, P. H. Thrall, J. L. Gittleman, J. Antonovics, A. A. Cunningham, A. P. Dobson, V. Ezenwa, K. E. Jones and A. B. Pedersen (2003). "Social organization and parasite risk in mammals: integrating theory and empirical studies." Annual Review of Ecology, Evolution, and Systematics **34**(1): 517-547.

Amos, W. and K. Acevedo-Whitehouse (2009). "A new test for genotype–fitness associations reveals a single microsatellite allele that strongly predicts the nature of tuberculosis infections in wild boar." Molecular Ecology Resources **9**(4): 1102-1111.

Anderson, R. and W. Trewhella (1985). "Population dynamics of the badger (*Meles meles*) and the epidemiology of bovine tuberculosis (*Mycobacterium bovis*)." Philosophical Transactions of the Royal Society of London. B, Biological Sciences **310**(1145): 327-381.

Anderson, R. M., R. M. May and B. Anderson (1992). Infectious diseases of humans: dynamics and control, Wiley Online Library.

Angulo, E. and B. Juan (2007). "Towards a unique and transmissible vaccine against myxomatosis and rabbit haemorrhagic disease for rabbit populations." Wildlife Research **34**: 567-577.

Annavi, G., C. Newman, C. D. Buesching, D. W. Macdonald, T. Burke and H. L. Dugdale (2014). "Heterozygosity–fitness correlations in a wild mammal population: accounting for parental and environmental effects." Ecology and Evolution **4**(12): 2594-2609.

Annavi, G., C. Newman, H. L. Dugdale, C. D. Buesching, Y. W. Sin, T. Burke and D. W. Macdonald (2014). "Neighbouring-group composition and within-group relatedness drive extra-group paternity rate in the European badger (*Meles meles*)." Journal of Evolutionary Biology **27**(10): 2191-2203.

Annavi, G., C. Newman, H. L. Dugdale, C. D. Buesching, Y. W. Sin, T. Burke and D. W. Macdonald (2014). "Neighbouring-group composition and within-group relatedness drive extra-group paternity rate in the European badger (*Meles meles*)." J Evol Biol **27**(10): 2191-2203.

Antunes, J. L. F. and E. A. Waldman (2001). "The impact of AIDS, immigration and housing overcrowding on tuberculosis deaths in São Paulo, Brazil, 1994–1998." Social Science & Medicine **52**(7): 1071-1080.

Archie, E. A., G. Luikart and V. O. Ezenwa (2009). "Infecting epidemiology with genetics: a new frontier in disease ecology." Trends in Ecology & Evolution **24**(1): 21-30.

Ardia, D. R., H. K. Parmentier and L. A. Vogel (2011). "The role of constraints and limitation in driving individual variation in immune response." Functional Ecology **25**(1): 61-73.

Artois, M., R. Bengis, R. J. Delahay, M.-J. Duchêne, J. P. Duff, E. Ferroglio, C. Gortazar, M. R. Hutchings, R. A. Kock and F. A. Leighton (2009). Wildlife disease surveillance and monitoring. Management of disease in wild mammals, Springer**:** 187-213.

Ayres, D. L., A. Darling, D. J. Zwickl, P. Beerli, M. T. Holder, P. O. Lewis, J. P. Huelsenbeck, F. Ronquist, D. L. Swofford and M. P. Cummings (2011). "BEAGLE: an application programming interface and high-performance computing library for statistical phylogenetics." Systematic biology: syr100.

Baker, M. (2010). "Next-generation sequencing: adjusting to data overload." Nat Meth **7**(7): 495-499.

Ball, F., D. Mollison and G. Scalia-Tomba (1997). "Epidemics with two levels of mixing." The Annals of Applied Probability: 46-89.

Balloux, F., W. Amos and T. Coulson (2004). "Does heterozygosity estimate inbreeding in real populations?" Molecular Ecology **13**(10): 3021-3031.

Banks, S. C. and R. Peakall (2012). "Genetic spatial autocorrelation can readily detect sex-biased dispersal." Molecular ecology **21**(9): 2092-2105.

Barker, E. N., A. C. Darby, C. R. Helps, I. R. Peters, K. J. Heesom, C. J. Arthur, B. Crossett, M. A. Hughes, A. D. Radford and S. Tasker "Molecular characterization of the uncultivatable hemotropic bacterium *Mycoplasma haemofelis*." Vet Res **42**(1): 83.

Baron, H., S. Hummel and B. Herrmann (1996). "*Mycobacterium tuberculosis* complex DNA in ancient human bones." Journal of Archaeological Science **23**(5): 667-671.

Barreiro, L. B., L. Tailleux, A. A. Pai, B. Gicquel, J. C. Marioni and Y. Gilad (2012). "Deciphering the genetic architecture of variation in the immune response to *Mycobacterium tuberculosis* infection." Proceedings of the National Academy of Sciences **109**(4): 1204-1209.

Barton, K. (2011). "MuMIn: Multi-model inference. R package version 1.0. 0." Vienna, Austria: R Foundation for Statistical Computing. See http://CRAN. R-project. org/package= MuMIn.

Bates, D. M. (2010). "lme4: Mixed-effects modeling with R." URL http://lme4. r-forge. r-project. org/book.

Beggs, C., C. Noakes, P. Sleigh, L. Fletcher and K. Siddiqi (2003). "The transmission of tuberculosis in confined spaces: an analytical review of alternative epidemiological models." The international journal of tuberculosis and lung disease **7**(11): 1015-1026.

Begon, M., S. M. Hazel, D. Baxby, K. Bown, R. Cavanagh, J. Chantrey, T. Jones and M. Bennett (1999). Transmission dynamics of a zoonotic pathogen within and between wildlife host species.

Beirne, C., L. Waring, R. A. McDonald, R. Delahay and A. Young (2016). Age-related declines in immune response in a wild mammal are unrelated to immune cell telomere length. Proc. R. Soc. B, The Royal Society.

Bekoff, M. (1977). "Mammalian dispersal and the ontogeny of individual behavioral phenotypes." American Naturalist: 715-732.

Benjamini, Y. and Y. Hochberg (1995). "Controlling the false discovery rate: a practical and powerful approach to multiple testing." Journal of the royal statistical society. Series B (Methodological): 289-300.

Benton, C., R. Delahay, H. Trewby and D. Hodgson (2014). "What has molecular epidemiology ever done for wildlife disease research? Past contributions and future directions." European Journal of Wildlife Research: 1-16.

Benton, C. H., R. J. Delahay, A. Robertson, R. A. McDonald, A. J. Wilson, T. A. Burke and D. Hodgson (2016). Blood thicker than water: kinship, disease

prevalence and group size drive divergent patterns of infection risk in a social mammal. Proc. R. Soc. B, The Royal Society.

Biek, R., A. O'Hare, D. Wright, T. Mallon, C. McCormick, R. J. Orton, S. McDowell, H. Trewby, R. A. Skuce and R. R. Kao (2012). "Whole Genome Sequencing Reveals Local Transmission Patterns of *Mycobacterium bovis* in Sympatric Cattle and Badger Populations." PLoS Pathog **8**(11): e1003008.

Biek, R. and L. A. Real (2010). "The landscape genetics of infectious disease emergence and spread." Mol Ecol **19**(17): 3515-3531.

Bielby, J., C. A. Donnelly, L. C. Pope, T. Burke and R. Woodroffe (2014). "Badger responses to small-scale culling may compromise targeted control of bovine tuberculosis." Proceedings of the National Academy of Sciences **111**(25): 9193-9198.

Bielejec, F., A. Rambaut, M. A. Suchard and P. Lemey (2011). "SPREAD: spatial phylogenetic reconstruction of evolutionary dynamics." Bioinformatics **27**(20): 2910-2912.

Blackwood, J. C., D. G. Streicker, S. Altizer and P. Rohani (2013). "Resolving the roles of immunity, pathogenesis, and immigration for rabies persistence in vampire bats." Proc Natl Acad Sci U S A **110**(51): 20837-20842.

Blanchong, J. A., S. J. Robinson, M. D. Samuel and J. T. Foster (2016). "Application of genetics and genomics to wildlife epidemiology." The Journal of Wildlife Management **80**(4): 593-608.

Blanchong, J. A., M. D. Samuel, K. T. Scribner, B. V. Weckworth, J. A. Langenberg and K. B. Filcek (2008). "Landscape genetics and the spatial distribution of chronic wasting disease." Biology Letters **4**(1): 130-133.

Blehert, D. S. (2011). "Population Genetics of Geomyces destructans from Whole Genome Sequences " US Geological Survey.

Blouin, M., M. Parsons, V. Lacaille and S. Lotz (1996). "Use of microsatellite loci to classify individuals by relatedness." Molecular Ecology **5**(3): 393-401.

Böhm, M., M. R. Hutchings and P. C. L. White (2009). "Contact Networks in a Wildlife-Livestock Host Community: Identifying High-Risk Individuals in the Transmission of Bovine TB among Badgers and Cattle." PLoS ONE **4**(4): e5016.

Bourne, F., C. Donnelly, D. Cox, G. Gettinby, J. McInerney, I. Morrison and R. Woodroffe (2007). Bovine TB: the scientific evidence. Final report of the independent scientific group on cattle TB, London, UK: Defra. See h ttp://archive. defra. gov. uk/foodfarm/farmanimal/diseases/atoz/tb/isg/report/final_report. pdf.

Boyce, W. M., M. E. Weisenberger, M. C. T. Penedo and C. K. Johnson (2011). "Wildlife translocation: the conservation implications of pathogen exposure and genetic heterozygosity." BMC ecology **11**(1): 5.

Boyd, S. D. (2013). "Diagnostic applications of high-throughput DNA sequencing." Annu Rev Pathol **8**: 381-410.

Briles, D. E. (2012). Genetic control of the susceptibility to bacterial infection, Springer Science & Business Media.

Brotherstone, S., I. M. S. White, M. Coffey, S. H. Downs, A. P. Mitchell, R. S. Clifton-Hadley, S. J. More, M. Good and J. A. Woolliams (2010). "Evidence of genetic resistance of cattle to infection with Mycobacterium bovis." Journal of Dairy Science **93**(3): 1234-1242.

Brown, C. R. and M. B. Brown (1986). "Ectoparasitism as a cost of coloniality in cliff swallows (Hirundo pyrrhonota)." Ecology **67**(5): 1206-1218.

Brown, G. E. and J. A. Brown (1996). "Does kin-biased territorial behavior increase kin-biased foraging in juvenile salmonids?" Behavioral Ecology **7**(1): 24-29.

Bryant, J. M., A. C. Schürch, H. van Deutekom, S. R. Harris, J. L. de Beer, V. de Jager, K. Kremer, S. A. van Hijum, R. J. Siezen and M. Borgdorff (2013). "Inferring patient to patient transmission of *Mycobacterium tuberculosis* from whole genome sequencing data." BMC infectious diseases **13**(1): 110.

Buckee, C. O., M. Recker, E. R. Watkins and S. Gupta (2011). "Role of stochastic processes in maintaining discrete strain structure in antigenically diverse pathogen populations." Proc Natl Acad Sci U S A **108**(37): 15504-15509.

Buddle, B., G. Lisle, J. Griffin and S. Hutchings (2015). "Epidemiology, diagnostics, and management of tuberculosis in domestic cattle and deer in New Zealand in the face of a wildlife reservoir." N Z Vet J **63**.

Bull, J. J. (1994). "Perspective: Virulence." Evolution **48**(5): 1423-1437.

Burnham, K. P. and D. R. Anderson (2004). "Multimodel Inference: Understanding AIC and BIC in Model Selection." Sociological Methods & Research **33**(2): 261-304.

Butler, J. and T. Roper (1996). "Ectoparasites and sett use in European badgers." Animal Behaviour **52**(3): 621-629.

Buzdugan, S., M. Chambers, R. Delahay and J. Drewe (2016). "Diagnosis of tuberculosis in groups of badgers: an exploration of the impact of trapping efficiency, infection prevalence and the use of multiple tests." Epidemiology and infection **144**(08): 1717-1727.

Byrnes, E. J., Ill, W. Li, Y. Lewit, H. Ma, K. Voelz, P. Ren, D. A. Carter, V. Chaturvedi, R. J. Bildfell, R. C. May and J. Heitman (2010). "Emergence and Pathogenicity of Highly Virulent *Cryptococcus gattii* Genotypes in the Northwest United States." PLoS Pathog **6**(4): e1000850.

Carey, A. J., P. Timms, G. Rawlinson, J. Brumm, K. Nilsson, J. M. Harris and K. W. Beagley (2010). "A Multi-Subunit Chlamydial Vaccine Induces Antibody and Cell-Mediated Immunity in Immunized Koalas (*Phascolarctos cinereus*):

Comparison of Three Different Adjuvants." American Journal of Reproductive Immunology **63**(2): 161-172.

Carpenter, P. J., L. C. Pope, C. Greig, D. A. Dawson, L. M. Rogers, K. Erven, G. J. Wilson, R. J. Delahay, C. L. Cheeseman and T. Burke (2005). "Mating system of the Eurasian badger, *Meles meles*, in a high density population." Molecular Ecology **14**(1): 273-284.

Carter, S. P., M. A. Chambers, S. P. Rushton, M. D. F. Shirley, P. Schuchert, S. Pietravalle, A. Murray, F. Rogers, G. Gettinby, G. C. Smith, R. J. Delahay, R. G. Hewinson and R. A. McDonald (2012). "BCG Vaccination Reduces Risk of Tuberculosis Infection in Vaccinated Badgers and Unvaccinated Badger Cubs." PloS one **7**(12): e49833.

Carter, S. P., R. J. Delahay, G. C. Smith, D. W. Macdonald, P. Riordan, T. R. Etherington, E. R. Pimley, N. J. Walker and C. L. Cheeseman (2007). "Culling-induced social perturbation in Eurasian badgers *Meles meles* and the management of TB in cattle: an analysis of a critical problem in applied ecology." Proceedings of the Royal Society B: Biological Sciences **274**(1626): 2769-2777.

Castro-Prieto, A., B. Wachter, J. Melzheimer, S. Thalwitzer, H. Hofer and S. Sommer (2012). "Immunogenetic Variation and Differential Pathogen Exposure in Free-Ranging Cheetahs across Namibian Farmlands." PloS one **7**(11): e49129.

Cauchemez, S., A. Bhattarai, T. L. Marchbanks, R. P. Fagan, S. Ostroff, N. M. Ferguson, D. Swerdlow, S. V. Sodha, M. E. Moll and F. J. Angulo (2011). "Role of social networks in shaping disease transmission during a community outbreak of 2009 H1N1 pandemic influenza." Proceedings of the National Academy of Sciences **108**(7): 2825-2830.

Chambers, M. A., T. Crawshaw, S. Waterhouse, R. Delahay, R. G. Hewinson and K. P. Lyashchenko (2008). "Validation of the BrockTB Stat-Pak assay for detection of tuberculosis in Eurasian badgers (*Meles meles*) and influence of disease severity on diagnostic accuracy." Journal of clinical microbiology **46**(4): 1498-1500.

Chambers, M. A., T. Crawshaw, S. Waterhouse, R. Delahay, R. G. Hewinson and K. P. Lyashchenko (2008). "Validation of the BrockTB stat-pak assay for detection of tuberculosis in Eurasian badgers (*Meles meles*) and influence of disease severity on diagnostic accuracy." J Clin Microbiol **46**.

Chambers, M. A., F. Rogers, R. J. Delahay, S. Lesellier, R. Ashford, D. Dalley, S. Gowtage, D. Davé, S. Palmer, J. Brewer, T. Crawshaw, R. Clifton-Hadley, S. Carter, C. Cheeseman, C. Hanks, A. Murray, K. Palphramand, S. Pietravalle, G. C. Smith, A. Tomlinson, N. J. Walker, G. J. Wilson, L. A. L. Corner, S. P. Rushton, M. D. F. Shirley, G. Gettinby, R. A. McDonald and R. G. Hewinson (2010). "Bacillus Calmette-Guérin vaccination reduces the severity and progression of tuberculosis in badgers." Proceedings of the Royal Society B: Biological Sciences.

Chambers, M. A., S. Waterhouse, K. Lyashchenko, R. Delahay, R. Sayers and R. G. Hewinson (2009). "Performance of TB immunodiagnostic tests in Eurasian badgers (Meles meles) of different ages and the influence of duration of infection on serological sensitivity." BMC Veterinary Research **5**(1): 42.

Chandra, R. K. (1996). "Nutrition, immunity and infection: from basic knowledge of dietary manipulation of immune responses to practical application of ameliorating suffering and improving survival." Proceedings of the National Academy of Sciences **93**(25): 14304-14307.

Chang, S., Z. Ding, S.-t. Yang, Y.-w. Gao, X.-h. Zou, T.-c. Wang and X.-z. Xia (2007). "Study on the histopathology of cats inoculated with H5N1 subtype high pathogenic avian influenza virus originated from tigers." Chinese Journal of Virology **23**(6): 477-480.

Chantrey, J., T. D. Dale, J. M. Read, S. White, F. Whitfield, D. Jones, C. J. McInnes and M. Begon (2014). "European red squirrel population dynamics driven by squirrelpox at a gray squirrel invasion interface." Ecology and evolution **4**(19): 3788-3799.

Chapman, D. A., A. C. Darby, M. Da Silva, C. Upton, A. D. Radford and L. K. Dixon (2011). "Genomic analysis of highly virulent Georgia 2007/1 isolate of African swine fever virus." Emerging Infectious Diseases **17**(4): 599-605.

Charlesworth, B. and K. A. Hughes (1996). "Age-specific inbreeding depression and components of genetic variance in relation to the evolution of senescence." Proceedings of the National Academy of Sciences **93**(12): 6140-6145.

Charlesworth, D. and B. Charlesworth (1987). "Inbreeding depression and its evolutionary consequences." Annual review of ecology and systematics **18**(1): 237-268.

Cheeseman, C., G. Jones, J. Gallagher and P. Mallinson (1981). "The population structure, density and prevalence of tuberculosis (*Mycobacterium bovis*) in badgers (*Meles meles*) from four areas in south-west England." Journal of Applied Ecology: 795-804.

Cheeseman, C., J. Wilesmith and F. Stuart (1989). "Tuberculosis: the disease and its epidemiology in the badger, a review." Epidemiology and infection **103**(01): 113-125.

Cheeseman, C. L., W. J. Cresswell, S. Harris and P. J. Mallinson (1988). "Comparison of dispersal and other movements in two Badger (*Meles meles*) populations." Mammal Review **18**(1): 51-59.

Cheeseman, C. L., J. W. Wilesmith, F. A. Stuart and P. J. Mallinson (1988). "Dynamics of tuberculosis in a naturally infected Badger population." Mammal Review **18**(1): 61-72.

Clements, G. M., S. E. Hygnstrom, J. M. Gilsdorf, D. M. Baasch, M. J. Clements and K. C. Vercauteren (2011). "Movements of white-tailed deer in riparian habitat: Implications for infectious diseases." The Journal of Wildlife Management **75**(6): 1436-1442.

Clifton-Hadley, R., J. Wilesmith and F. Stuart (1993). "Mycobacterium bovis in the European badger *(Meles meles)*: epidemiological findings in tuberculous badgers from a naturally infected population." Epidemiology and infection **111**(01): 9-19.

Coltman, D. W., J. G. Pilkington, J. A. Smith and J. M. Pemberton (1999). "Parasite-mediated selection against inbred Soay sheep in a free-living, island population." Evolution: 1259-1267.

Cook, V. J., S. J. Sun, J. Tapia, S. Q. Muth, D. F. Argüello, B. L. Lewis, R. B. Rothenberg and P. D. McElroy (2007). "Transmission Network Analysis in Tuberculosis Contact Investigations." Journal of Infectious Diseases **196**(10): 1517-1527.

Cooper, A. M., D. K. Dalton, T. A. Stewart, J. Griffin, D. Russell and I. Orme (1993). "Disseminated tuberculosis in interferon gamma gene-disrupted mice." Journal of Experimental Medicine **178**(6): 2243-2247.

Corner, L. A., D. Murphy and E. Gormley (2011). "*Mycobacterium bovis* infection in the Eurasian badger (*Meles meles*): the disease, pathogenesis, epidemiology and control." J Comp Pathol **144**(1): 1-24.

Cortey, M., K. Bertran, J. Toskano, N. Majo and R. Dolz (2012). "Phylogeographic distribution of very virulent infectious bursal disease virus isolates in the Iberian Peninsula." Avian Pathol **41**(3): 277-284.

Côté, I. M. and R. Poulinb (1995). "Parasitism and group size in social animals: a meta-analysis." Behavioral Ecology **6**(2): 159-165.

Cowled, B. D., M. P. Ward, S. W. Laffan, F. Galea, M. G. Garner, A. J. MacDonald, I. Marsh, P. Muellner, K. Negus, S. Quasim, A. P. Woolnough and S. D. Sarre (2012). "Integrating survey and molecular approaches to better understand wildlife disease ecology." PloS one **7**(10): e46310-e46310.

Crispell, J. (2017). Using Whole Genome Sequencing to investigate the inter-species transmission dynamics of *Mycobacterium bovis* PhD, University of Glasgow.

Crispell, J., R. N. Zadoks, S. R. Harris, B. Paterson, D. M. Collins, G. W. de-Lisle, P. Livingstone, M. A. Neill, R. Biek, S. J. Lycett, R. R. Kao and M. Price-Carter (2017). "Using whole genome sequencing to investigate transmission in a multi-host system: bovine tuberculosis in New Zealand." BMC Genomics **18**(1): 180.

Cross, P. C., J. O. Lloyd-Smith, J. A. Bowers, C. T. Hay, M. Hofmeyr and W. M. Getz (2004). Integrating association data and disease dynamics in a social ungulate: bovine tuberculosis in African buffalo in the Kruger National Park. Annales Zoologici Fennici, JSTOR.

Cross, P. C., J. O. Lloyd-Smith, P. L. F. Johnson and W. M. Getz (2005). "Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations." Ecology Letters **8**(6): 587-595.

Cullingham, C. I., E. H. Merrill, M. J. Pybus, T. K. Bollinger, G. A. Wilson and D. W. Coltman (2011). "Broad and fine-scale genetic analysis of white-tailed deer populations: estimating the relative risk of chronic wasting disease spread." Evolutionary Applications **4**(1): 116-131.

Cunha, M. V., F. Matos, A. Canto, T. Albuquerque, J. R. Alberto, J. M. Aranha, M. Vieira-Pinto and A. Botelho (2012). "Implications and challenges of tuberculosis in wildlife ungulates in Portugal: A molecular epidemiology perspective." Research in Veterinary Science **92**(2): 225-235.

Cunningham, C. W., K. E. Omland and T. H. Oakley (1998). "Reconstructing ancestral character states: a critical reappraisal." Trends in Ecology & Evolution **13**(9): 361-366.

Daboussi, M. J. (1997). Fungal transposable elements and genome evolution. Evolution and Impact of Transposable Elements. P. Capy, Springer Netherlands. **6:** 253-260.

Dalley, D., D. Davé, S. Lesellier, S. Palmer, T. Crawshaw, R. G. Hewinson and M. Chambers (2008). "Development and evaluation of a gamma-interferon assay for tuberculosis in badgers (*Meles meles*)." Tuberculosis **88**(3): 235-243.

Darwin, C. (1876). The effects of cross and self fertilisation in the vegetable kingdom, J. Murray.

Daszak, P., A. A. Cunningham and A. D. Hyatt (2000). "Emerging Infectious Diseases of Wildlife - Threats to Biodiversity and Human Health." Science **287**: 443-449.

Daszak, P., G. M. Tabor, A. M. Kilpatrick, J. Epstein and R. Plowright (2004). "Conservation medicine and a new agenda for emerging diseases." Ann N Y Acad Sci **1026**: 1-11.

David, P. (1998). "Heterozygosity–fitness correlations: new perspectives on old problems." Heredity **80**(5): 531-537.

de Carvalho, L. M., L. B. Santos, N. R. Faria and W. de Castro Silveira (2013). "Phylogeography of foot-and-mouth disease virus serotype O in Ecuador." Infect Genet Evol **13**: 76-88.

De Leeuw, A., G. Forrester, P. Spyvee, M. Brash and R. Delahay (2004). "Experimental comparison of ketamine with a combination of ketamine, butorphanol and medetomidine for general anaesthesia of the Eurasian badger (*Meles meles*)." The Veterinary Journal **167**(2): 186-193.

De Maio, N., C.-H. Wu and D. J. Wilson (2016). "SCOTTI: Efficient Reconstruction of Transmission within Outbreaks with the Structured Coalescent." PLOS Computational Biology **12**(9): e1005130.

Dean, G. S., S. G. Rhodes, M. Coad, A. O. Whelan, P. J. Cockle, D. J. Clifford, R. G. Hewinson and H. M. Vordermeier (2005). "Minimum infective dose of *Mycobacterium bovis* in cattle." Infect Immun **73**(10): 6467-6471.

DEFRA. (2017, May 2017). "TB Risk Area Map." Retrieved May 2017, 2017, from http://www.tbhub.co.uk/tb-policy/england/.

Delahay, R., J. Brown, P. Mallinson, P. Spyvee, D. Handoll, L. Rogers and C. Cheeseman (2000). "The use of marked bait in studies of the territorial organization of the European badger (*Meles meles*)." Mammal Review **30**(2): 73-87.

Delahay, R., A. De Leeuw, A. Barlow, R. Clifton-Hadley and C. Cheeseman (2002). "The status of *Mycobacterium bovis* infection in UK wild mammals: a review." The Veterinary Journal **164**(2): 90-105.

Delahay, R. J., S. P. Carter, G. J. Forrester, A. Mitchell and C. L. Cheeseman (2006). "Habitat correlates of group size, bodyweight and reproductive performance in a high-density Eurasian badger (*Meles meles*) population." Journal of Zoology **270**(3): 437-447.

Delahay, R. J., S. Langton, G. C. Smith, R. S. Clifton-Hadley and C. L. Cheeseman (2000). "The spatio-temporal distribution of *Mycobacterium bovis* (bovine tuberculosis) infection in a high-density badger population." Journal of Animal Ecology **69**(3): 428-441.

Delahay, R. J., S. Langton, G. C. Smith, R. S. Clifton-hadley and C. L. Cheeseman (2000). "The spatio-temporal distribution of Mycobacterium bovis(bovine tuberculosis) infection in a high-density badger population." J Anim Ecol **69**.

Delahay, R. J., G. C. Smith and M. R. Hutchings (2008). Management of Disease in Wild Mammals, Springer.

Delahay, R. J., N. Walker, M. R. Gunn, C. Christie, G. J. Wilson, C. L. Cheeseman and R. A. McDonald (2011). "Using lifetime tooth-wear scores to predict age in wild Eurasian badgers: performance of a predictive model." Journal of Zoology **284**: 183-191.

Delahay, R. J., N. Walker, G. C. Smith, D. Wilkinson, R. S. Clifton-Hadley, C. L. Cheeseman, A. J. Tomlinson and M. A. Chambers (2013). "Long-term temporal trends and estimated transmission rates for *Mycobacterium bovis* infection in an undisturbed high-density badger (*Meles meles*) population." Epidemiol Infect **141**(7): 1445-1456.

Delahay, R. J., N. Walker, G. S. Smith, D. Wilkinson, R. S. Clifton-Hadley, C. L. Cheeseman, A. J. Tomlinson and M. A. Chambers (2013). "Long-term temporal trends and estimated transmission rates for *Mycobacterium bovis* infection in an undisturbed high-density badger (*Meles meles*) population." Epidemiology & Infection **141**(Special Issue 07): 1445-1456.

Dharmarajan, G., J. C. Beasley, J. A. Fike, E. A. Raizman, C. C. Wu, R. M. Pogranichniy and O. E. Rhodes Jr (2012). "Effects of kin-structure on disease dynamics in raccoons (*Procyon lotor*) inhabiting a fragmented landscape." Basic and Applied Ecology **13**(6): 560-567.

Doll, S. R. (1980). "The epidemiology of cancer." Cancer **45**(10): 2475-2485.

Domingo, E. and J. J. Holland (1997). "RNA virus mutations and fitness for survival." Annu Rev Microbiol **51**: 151-178.

Donnelly, C. A., G. Wei, W. T. Johnston, D. R. Cox, R. Woodroffe, F. J. Bourne, C. L. Cheeseman, R. S. Clifton-Hadley, G. Gettinby, P. Gilks, H. E. Jenkins, A. M. Le Fevre, J. P. McInerney and W. I. Morrison (2007). "Impacts of widespread badger culling on cattle tuberculosis: concluding analyses from a large-scale field trial." International Journal of Infectious Diseases **11**(4): 300-308.

Donnelly, C. A., R. Woodroffe, D. Cox, F. J. Bourne, C. Cheeseman, R. S. Clifton-Hadley, G. Wei, G. Gettinby, P. Gilks and H. Jenkins (2005). "Positive and negative effects of widespread badger culling on tuberculosis in cattle." Nature **439**(7078): 843-846.

Donnelly, C. A., R. Woodroffe, D. Cox, J. Bourne, G. Gettinby, A. M. Le Fevre, J. P. McInerney and W. I. Morrison (2003). "Impact of localized badger culling on tuberculosis incidence in British cattle." Nature **426**(6968): 834-837.

Donnelly, C. A., R. Woodroffe, D. R. Cox, J. Bourne, G. Gettinby, A. M. Le Fevre, J. P. McInerney and W. I. Morrison (2003). "Impact of localized badger culling on tuberculosis incidence in British cattle." Nature **426**(6968): 834-837.

Doolittle, W. F. and O. Zhaxybayeva (2010). "Metagenomics and the units of biological organization." BioScience **60**(2): 102-112.

Dorman, S. E., C. L. Hatem, S. Tyagi, K. Aird, J. Lopez-Molina, M. L. Pitt, B. C. Zook, A. M. Dannenberg, Jr., W. R. Bishai and Y. C. Manabe (2004). "Susceptibility to tuberculosis: clues from studies with inbred and outbred New Zealand White rabbits." Infect Immun **72**(3): 1700-1705.

Drewe, J., H. O'Connor, N. Weber, R. McDonald and R. Delahay (2013). "Patterns of direct and indirect contact between cattle and badgers naturally infected with tuberculosis." Epidemiology and infection **141**(07): 1467-1475.

Drewe, J. A. (2010). "Who infects whom? Social networks and tuberculosis transmission in wild meerkats." Proceedings of the Royal Society B: Biological Sciences **277**(1681): 633-642.

Drewe, J. A., A. J. Tomlinson, N. J. Walker and R. J. Delahay (2010). "Diagnostic Accuracy and Optimal Use of Three Tests for Tuberculosis in Live Badgers." PloS one **5**(6).

Drummond, A. J., S. Y. Ho, M. J. Phillips and A. Rambaut (2006). "Relaxed phylogenetics and dating with confidence." PLoS Biol **4**(5): e88.

Drummond, A. J., G. K. Nicholls, A. G. Rodrigo and W. Solomon (2002). "Estimating mutation parameters, population history and genealogy simultaneously from temporally spaced sequence data." Genetics **161**(3): 1307-1320.

Drummond, A. J. and A. Rambaut (2007). "BEAST: Bayesian Evolutionary Analysis by Sampling Trees." BMC Evol Biol **7**.

Drummond, A. J., A. Rambaut, B. Shapiro and O. G. Pybus (2005). "Bayesian coalescent inference of past population dynamics from molecular sequences." Molecular biology and evolution **22**(5): 1185-1192.

Drummond, A. J., M. A. Suchard, D. Xie and A. Rambaut (2012). "Bayesian Phylogenetics with BEAUti and the BEAST 1.7." Molecular Biology and Evolution **29**(8): 1969-1973.

Duarte, E. L., M. Domingos, A. Amado and A. Botelho (2008). "Spoligotype diversity of *Mycobacterium bovis* and *Mycobacterium caprae* animal isolates." Veterinary Microbiology **130**(3-4): 415-421.

Duarte, E. L., M. Domingos, A. Amado, M. Cunha and A. Botelho (2010). "MIRU-VNTR typing adds discriminatory value to groups of *Mycobacterium bovis* and *Mycobacterium caprae* strains defined by spoligotyping." Veterinary microbiology **143**(2): 299-306.

Duffield, B. and D. Young (1985). "Survival of *Mycobacterium bovis* in defined environmental conditions." Veterinary Microbiology **10**(2): 193-197.

Dugdale, H. L., S. A. Ellwood and D. W. Macdonald (2010). "Alloparental behaviour and long-term costs of mothers tolerating other members of the group in a plurally breeding mammal." Animal Behaviour **80**(4): 721-735.

Dugdale, H. L., D. W. Macdonald, L. C. Pope and T. Burke (2007). "Polygynandry, extra-group paternity and multiple-paternity litters in European badger (*Meles meles*) social groups." Molecular Ecology **16**(24): 5294-5306.

Dugdale, H. L., D. W. Macdonald, L. C. Pope, P. J. Johnson and T. Burke (2008). "Reproductive skew and relatedness in social groups of European badgers, *Meles meles*." Molecular Ecology **17**(7): 1815-1827.

Durrant, K. L. and J. M. Hughes (2005). "Differing Rates of Extra-Group Paternity between Two Populations of the Australian Magpie (*Gymnorhina tibicen*)." Behavioral Ecology and Sociobiology **57**(6): 536-545.

Earl, D. and B. vonHoldt (2012). "STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method." Conservation Genetics Resources **4**(2): 359-361.

Ebel, E. R. and P. C. Phillips (2016). "Intrinsic differences between males and females determine sex-specific consequences of inbreeding." BMC evolutionary biology **16**(1): 1.

Ekman, J., S. Eggers and M. Griesser (2002). "Fighting to stay: the role of sibling rivalry for delayed dispersal." Animal Behaviour **64**(3): 453-459.

El-Sayed, N. M., P. J. Myler, G. Blandin, M. Berriman, J. Crabtree, G. Aggarwal, E. Caler, H. Renauld, E. A. Worthey, C. Hertz-Fowler, E. Ghedin, C. Peacock, D. C. Bartholomeu, B. J. Haas, A.-N. Tran, J. R. Wortman, U. C. M. Alsmark, S. Angiuoli, A. Anupama, J. Badger, F. Bringaud, E. Cadag, J. M. Carlton, G. C. Cerqueira, T. Creasy, A. L. Delcher, A. Djikeng, T. M. Embley, C. Hauser, A. C. Ivens, S. K. Kummerfeld, J. B. Pereira-Leal, D. Nilsson, J. Peterson, S. L. Salzberg, J. Shallom, J. C. Silva, J. Sundaram, S. Westenberger, O. White, S.

E. Melville, J. E. Donelson, B. Andersson, K. D. Stuart and N. Hall (2005). "Comparative Genomics of Trypanosomatid Parasitic Protozoa." Science **309**(5733): 404-409.

Elender, F., G. Bentham and I. Langford (1998). "Tuberculosis mortality in England and Wales during 1982–1992: its association with poverty, ethnicity and AIDS." Social science & medicine **46**(6): 673-681.

Elliott, P. and D. Wartenberg (2004). "Spatial epidemiology: current approaches and future challenges." Environmental health perspectives: 998-1006.

Evanno, G., S. Regnaut and J. Goudet (2005). "Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study." Molecular ecology **14**(8): 2611-2620.

Evans, P., D. Macdonald and C. Cheeseman (1989). "Social structure of the Eurasian badger (*Meles meles*): genetic evidence." Journal of Zoology **218**(4): 587-595.

Eyre, D. W., M. L. Cule, D. J. Wilson, D. Griffiths, A. Vaughan, L. O'Connor, C. L. C. Ip, T. Golubchik, E. M. Batty, J. M. Finney, D. H. Wyllie, X. Didelot, P. Piazza, R. Bowden, K. E. Dingle, R. M. Harding, D. W. Crook, M. H. Wilcox, T. E. A. Peto and A. S. Walker (2013). "Diverse Sources of *C. difficile* Infection Identified on Whole-Genome Sequencing." New England Journal of Medicine **369**(13): 1195-1205.

Ezenwa, V. O. (2004). "Host social behavior and parasitic infection: a multifactorial approach." Behavioral Ecology **15**(3): 446-454.

Ezenwa, V. O., R. R. Ghai, A. F. McKay and A. E. Williams (2016). "Group living and pathogen infection revisited." Current Opinion in Behavioral Sciences **12**: 66-72.

Falush, D., M. Stephens and J. K. Pritchard (2003). "Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies." Genetics **164**(4): 1567-1587.

Fell, R. J., C. D. Buesching and D. W. Macdonald (2006). "The social integration of European badger (*Meles meles*) cubs into their natal group." Behaviour **143**(6): 683-700.

Feng, Y. and L. Xiao (2011). "Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis." Clinical Microbiology Reviews **24**(1): 110-140.

Fierer, N., M. Breitbart, J. Nulton, P. Salamon, C. Lozupone, R. Jones, M. Robeson, R. A. Edwards, B. Felts, S. Rayhawk, R. Knight, F. Rohwer and R. B. Jackson (2007). "Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil." Applied and Environmental Microbiology **73**(21): 7059-7066.

Foster, J. T., S. M. Beckstrom-Sternberg, T. Pearson, J. S. Beckstrom-Sternberg, P. S. G. Chain, F. F. Roberto, J. Hnath, T. Brettin and P. Keim (2009). "Whole-Genome-Based Phylogeny and Divergence of the Genus Brucella." Journal of Bacteriology **191**(8): 2864-2870.

Frantz, A., E. D. L. San, L. Pope and T. Burke (2010). "Using genetic methods to investigate dispersal in two badger (*Meles meles*) populations with different ecological characteristics." Heredity **104**(5): 493-501.

Frost, S. D. and E. M. Volz (2010). "Viral phylodynamics and the search for an 'effective number of infections'." Philosophical Transactions of the Royal Society of London B: Biological Sciences **365**(1548): 1879-1890.

Frothingham, R. (1995). "Differentiation of strains in *Mycobacterium tuberculosis* complex by DNA sequence polymorphisms, including rapid identification of M. bovis BCG." J Clin Microbiol **33**(4): 840-844.

Fulford, G. R., M. G. Roberts and J. A. P. Heesterbeek (2002). "The Metapopulation Dynamics of an Infectious Disease: Tuberculosis in Possums." Theoretical Population Biology **61**(1): 15-29.

Gallagher, J. and R. Clifton-Hadley (2000). "Tuberculosis in badgers; a review of the disease and its significance for other animals." Research in veterinary science **69**(3): 203-217.

Gallagher, J. and D. Horwill (1977). "A selective oleic acid albumin agar medium for the cultivation of *Mycobacterium bovis*." Journal of Hygiene **79**(01): 155-160.

Gallagher, J., R. Monies, M. Gavier-Widen and B. Rule (1998). "Role of infected, non-diseased badgers in the pathogenesis of tuberculosis in the badger." The Veterinary Record **142**(26): 710-714.

Gallagher, J., R. Muirhead and K. Burn (1976). "Tuberculosis in wild badgers (*Meles meles*) in Gloucestershire: pathology." Veterinary Record **98**(1): 9-14.

Gallagher, J., R. Muirhead and K. Burn (1976). "Tuberculosis in wild badgers (*Meles meles*) in Gloucestershire: pathology." The Veterinary Record **98**(1): 9-14.

Gallagher, J. and J. Nelson (1979). "Cause of ill health and natural death in badgers in Gloucestershire." Tuberculosis **10**(12): 14-16.

Gardy, J. L., J. C. Johnston, S. J. H. Sui, V. J. Cook, L. Shah, E. Brodkin, S. Rempel, R. Moore, Y. Zhao, R. Holt, R. Varhol, I. Birol, M. Lem, M. K. Sharma, K. Elwood, S. J. M. Jones, F. S. L. Brinkman, R. C. Brunham and P. Tang (2011). "Whole-Genome Sequencing and Social-Network Analysis of a Tuberculosis Outbreak." New England Journal of Medicine **364**(8): 730-739.

Gaudart, J., S. Rebaudet, R. Barrais, J. Boncy, B. Faucher, M. Piarroux, R. Magloire, G. Thimothe and R. Piarroux (2013). "Spatio-temporal dynamics of cholera during the first year of the epidemic in Haiti." PLoS neglected tropical diseases **7**(4): e2145.

Gavier-Widén, D., M. M. Cooke, J. Gallagher, M. A. Chambers and C. Gortázar (2009). "A review of infection of wildlife hosts with *Mycobacterium bovis* and the diagnostic difficulties of the 'no visible lesion' presentation." New Zealand Veterinary Journal **57**(3): 122-131.

Geue, L., S. Schares, C. Schnick, J. Kliemt, A. Beckert, C. Freuling, F. J. Conraths, B. Hoffmann, R. Zanoni, D. Marston, L. McElhinney, N. Johnson, A. R. Fooks, N. Tordo and T. Müller (2008). "Genetic characterisation of attenuated SAD rabies virus strains used for oral vaccination of wildlife." Vaccine **26**(26): 3227-3235.

Ghedin, E., N. A. Sengamalay, M. Shumway, J. Zaborsky, T. Feldblyum, V. Subbu, D. J. Spiro, J. Sitz, H. Koo, P. Bolotov, D. Dernovoy, T. Tatusova, Y. Bao, K. St George, J. Taylor, D. J. Lipman, C. M. Fraser, J. K. Taubenberger and S. L. Salzberg (2005). "Large-scale sequencing of human influenza reveals the dynamic nature of viral genome evolution." Nature **437**(7062): 1162-1166.

Gire, S. K., A. Goba, K. G. Andersen, R. S. G. Sealfon, D. J. Park, L. Kanneh, S. Jalloh, M. Momoh, M. Fullah, G. Dudas, S. Wohl, L. M. Moses, N. L. Yozwiak, S. Winnicki, C. B. Matranga, C. M. Malboeuf, J. Qu, A. D. Gladden, S. F. Schaffner, X. Yang, P.-P. Jiang, M. Nekoui, A. Colubri, M. R. Coomber, M. Fonnie, A. Moigboi, M. Gbakie, F. K. Kamara, V. Tucker, E. Konuwa, S. Saffa, J. Sellu, A. A. Jalloh, A. Kovoma, J. Koninga, I. Mustapha, K. Kargbo, M. Foday, M. Yillah, F. Kanneh, W. Robert, J. L. B. Massally, S. B. Chapman, J. Bochicchio, C. Murphy, C. Nusbaum, S. Young, B. W. Birren, D. S. Grant, J. S. Scheiffelin, E. S. Lander, C. Happi, S. M. Gevao, A. Gnirke, A. Rambaut, R. F. Garry, S. H. Khan and P. C. Sabeti (2014). "Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak." Science **345**(6202): 1369-1372.

Gjerde, B., K. Gunnes and T. Gjedrem (1983). "Effect of inbreeding on survival and growth in rainbow trout." Aquaculture **34**(3): 327-332.

Glaser, L., M. Carstensen, S. Shaw, S. Robbe-Austerman, A. Wunschmann and D. Grear (2016). "Descriptive epidemiology and whole genome sequencing analysis for an outbreak of bovine tuberculosis in beef cattle and white-tailed deer in northwestern Minnesota." PLoS One **11**.

Godfray, H. C. J., C. A. Donnelly, R. R. Kao, D. W. Macdonald, R. A. McDonald, G. Petrokofsky, J. L. N. Wood, R. Woodroffe, D. B. Young and A. R. McLean (2013). "A restatement of the natural science evidence base relevant to the control of bovine tuberculosis in Great Britain†." Proceedings of the Royal Society B: Biological Sciences **280**(1768).

Goodchild, A. V., G. H. Watkins, A. R. Sayers, J. R. Jones and R. S. Clifton-Hadley (2012). "Geographical association between the genotype of bovine tuberculosis in found dead badgers and in cattle herds." Vet Rec **170**(10): 259.

Goodger, J., A. Nolan, W. Russell, D. Dalley, C. Thorns, F. Stuart, P. Croston and D. Newell (1994). "Serodiagnosis of *Mycobacterium bovis* infection in badgers: development of an indirect ELISA using a 25 kDa antigen." The Veterinary Record **135**(4): 82-85.

Goodger, J., A. Nolan, W. P. Russell, D. J. Dalley, C. J. Thorns, F. A. Stuart, P. Croston and D. G. Newell (1994). "Serodiagnosis of *Mycobacterium bovis* infection in badgers: development of an indirect ELISA using a 25 kDa antigen." Vet Rec **135**.

Goodnight, K. F. and D. C. Queller (1999). "Computer software for performing likelihood tests of pedigree relationship using genetic markers." Molecular Ecology **8**(7): 1231-1234.

Gortázar, C., E. Ferroglio, U. Höfle, K. Frölich and J. Vicente (2007). "Diseases shared between wildlife and livestock: a European perspective." European Journal of Wildlife Research **53**(4): 241-256.

Goslee, S. C. and D. L. Urban (2007). "The ecodist package for dissimilarity-based analysis of ecological data." Journal of Statistical Software **22**(7): 1-19.

Goudet, J. (2005). "Hierfstat, a package for R to compute and test hierarchical F-statistics." Molecular Ecology Notes **5**(1): 184-186.

Goudet, J., N. Perrin and P. Waser (2002). "Tests for sex-biased dispersal using bi-parentally inherited genetic markers." Molecular Ecology **11**(6): 1103-1114.

Grad, Y. H., M. Lipsitch, M. Feldgarden, H. M. Arachchi, G. C. Cerqueira, M. FitzGerald, P. Godfrey, B. J. Haas, C. I. Murphy, C. Russ, S. Sykes, B. J. Walker, J. R. Wortman, S. Young, Q. Zeng, A. Abouelleil, J. Bochicchio, S. Chauvin, T. DeSmet, S. Gujja, C. McCowan, A. Montmayeur, S. Steelman, J. Frimodt-Møller, A. M. Petersen, C. Struve, K. A. Krogfelt, E. Bingen, F.-X. Weill, E. S. Lander, C. Nusbaum, B. W. Birren, D. T. Hung and W. P. Hanage (2012). "Genomic epidemiology of the Escherichia coli O104:H4 outbreaks in Europe, 2011." Proceedings of the National Academy of Sciences **109**(8): 3065-3070.

Graham, J., G. Smith, R. Delahay, T. Bailey, R. McDonald and D. Hodgson (2013). "Multi-state modelling reveals sex-dependent transmission, progression and severity of tuberculosis in wild badgers." Epidemiology and infection **141**(07): 1429-1436.

Graham, J., G. C. Smith, R. J. Delahay, T. Bailey, R. A. McDonald and D. Hodgson (2013). "Multi-state modelling reveals sex-dependent transmission, progression and severity of tuberculosis in wild badgers." Epidemiol Infect **141**(7): 1429-1436.

Grear, D. A., M. D. Samuel, J. A. Langenberg and D. Keane (2006). "Demographic patterns and harvest vulnerability of chronic wasting disease infected white-tailed deer in Wisconsin." Journal of Wildlife Management **70**(2): 546-553.

Grear, D. A., M. D. Samuel, K. T. Scribner, B. V. Weckworth and J. A. Langenberg (2010). "Influence of genetic relatedness and spatial proximity on chronic wasting disease infection among female white-tailed deer." Journal of Applied Ecology **47**(3): 532-540.

Grenfell, B. T., O. G. Pybus, J. R. Gog, J. L. Wood, J. M. Daly, J. A. Mumford and E. C. Holmes (2004). "Unifying the epidemiological and evolutionary dynamics of pathogens." Science **303**(5656): 327-332.

Grenfell, B. T., O. G. Pybus, J. R. Gog, J. L. N. Wood, J. M. Daly, J. A. Mumford and E. C. Holmes (2004). "Unifying the Epidemiological and Evolutionary Dynamics of Pathogens." Science **303**(5656): 327-332.

Griffin, A. S., S. A. West and A. Buckling (2004). "Cooperation and competition in pathogenic bacteria." Nature **430**(7003): 1024-1027.

Griffin, J. F. T. and C. G. Mackintosh (2000). "Tuberculosis in Deer: Perceptions, Problems and Progress." The Veterinary Journal **160**(3): 202-219.

Hacker, J. and E. Carniel (2001). "Ecological fitness, genomic islands and bacterial pathogenicity." EMBO reports **2**(5): 376-381.

Hamede, R., S. Lachish, K. Belov, G. Woods, A. Kreiss, A. M. Pearse, B. Lazenby, M. Jones and H. McCallum (2012). "Reduced effect of Tasmanian devil facial tumor disease at the disease front." Conserv Biol **26**(1): 124-134.

Hamede, R. K., J. Bashford, H. McCallum and M. Jones (2009). "Contact networks in a wild Tasmanian devil (*Sarcophilus harrisii*) population: using social network analysis to reveal seasonal variability in social behaviour and its implications for transmission of devil facial tumour disease." Ecology Letters **12**(11): 1147-1157.

Hamede, R. K., A.-M. Pearse, K. Swift, L. A. Barmuta, E. P. Murchison and M. E. Jones (2015). "Transmissible cancer in Tasmanian devils: localized lineage replacement and host population response." Proceedings of the Royal Society B: Biological Sciences **282**(1814).

Hamilton, W. D. (1966). "The moulding of senescence by natural selection." Journal of theoretical biology **12**(1): 12-45.

Harper, W. V. (2014). "Reduced Major Axis Regression: Teaching Alternatives to Least Squares."

Harris, S. R., E. J. P. Cartwright, M. E. Török, M. T. G. Holden, N. M. Brown, A. L. Ogilvy-Stuart, M. J. Ellington, M. A. Quail, S. D. Bentley, J. Parkhill and S. J. Peacock (2013). "Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study." The Lancet Infectious Diseases **13**(2): 130-136.

Harris, T. R., D. Caillaud, C. A. Chapman and L. Vigilant (2009). "Neither genetic nor observational data alone are sufficient for understanding sex-biased dispersal in a social-group-living species." Molecular Ecology **18**(8): 1777-1790.

Harrison, X., J. York, D. Cram and A. Young (2013). "Extra-group mating increases inbreeding risk in a cooperatively breeding bird." Molecular ecology **22**(22): 5700-5715.

Harrison, X. A., J. E. York and A. J. Young (2014). "Population genetic structure and direct observations reveal sex-reversed patterns of dispersal in a cooperative bird." Molecular ecology **23**(23): 5740-5755.

Harwood, V. J., J. Whitlock and V. Withington (2000). "Classification of Antibiotic Resistance Patterns of Indicator Bacteria by Discriminant Analysis: Use in Predicting the Source of Fecal Contamination in Subtropical Waters." Applied and Environmental Microbiology **66**(9): 3698-3704.

Hasegawa, M., H. Kishino and T.-a. Yano (1985). "Dating of the human-ape splitting by a molecular clock of mitochondrial DNA." Journal of molecular evolution **22**(2): 160-174.

Hatchwell, B. J. (2010). "Cryptic Kin Selection: Kin Structure in Vertebrate Populations and Opportunities for Kin-Directed Cooperation." Ethology **116**(3): 203-216.

Haydon, D. T., S. Cleaveland, L. H. Taylor and M. K. Laurenson (2002). "Identifying reservoirs of infection: a conceptual and practical challenge." Emerg Infect Dis **8**(12): 1468-1473.

Hedrick, P. (2011). Genetics of populations, Jones & Bartlett Learning.

Hedrick, P. W. and A. Garcia-Dorado "Understanding Inbreeding Depression, Purging, and Genetic Rescue." Trends in Ecology & Evolution **31**(12): 940-952.

Hemelaar, J., E. Gouws, P. D. Ghys and S. Osmanov (2011). "Global trends in molecular epidemiology of HIV-1 during 2000-2007." AIDS **25**(5): 679-689.

Henderson, M. and B. Okamura (2004). "The phylogeography of salmonid proliferative kidney disease in Europe and North America." Proceedings of the Royal Society of London. Series B: Biological Sciences **271**(1549): 1729-1736.

Hendricks, S., B. Epstein, B. Schönfeld, C. Wiench, R. Hamede, M. Jones, A. Storfer and P. Hohenlohe (2017). "Conservation implications of limited genetic diversity and population structure in Tasmanian devils (Sarcophilus harrisii)." Conservation Genetics: 1-6.

Henn, M. R., C. L. Boutwell, P. Charlebois, N. J. Lennon, K. A. Power, A. R. Macalalad, A. M. Berlin, C. M. Malboeuf, E. M. Ryan and S. Gnerre (2012). "Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection." PLoS pathogens **8**(3): e1002529.

Hess, G. (1996). "Disease in metapopulation models: implications for conservation." Ecology **77**(5): 1617-1632.

Hestbeck, J. B. (1982). "Population Regulation of Cyclic Mammals: The Social Fence Hypothesis." Oikos **39**(2): 157-163.

Ho, S. Y. and B. Shapiro (2011). "Skyline-plot methods for estimating demographic history from nucleotide sequences." Molecular ecology resources **11**(3): 423-434.

Hoffmann, C., N. Minkah, J. Leipzig, G. Wang, M. Q. Arens, P. Tebas and F. D. Bushman (2007). "DNA bar coding and pyrosequencing to identify rare HIV drug resistance mutations." Nucleic Acids Research **35**(13): e91.

Holman, P. J., J. E. Carroll, R. Pugh and D. S. Davis (2011). "Molecular detection of *Babesia bovis* and *Babseia bigemina* in white-tailed deer *(Odocoileus virginianus)* from Tom Green County in central Texas." Veterinary Parasitology **177**: 298-304.

Holmes, E. C., E. Ghedin, N. Miller, J. Taylor, Y. Bao, K. St George, B. T. Grenfell, S. L. Salzberg, C. M. Fraser and D. J. Lipman (2005). "Whole-genome analysis of human influenza A virus reveals multiple persistent lineages and reassortment among recent H3N2 viruses." PLoS biology **3**(9): e300.

Hostnik, P., E. Picard-Meyer, D. Rihtaric, I. Toplak and F. Cliquet (2014). "Vaccine-induced Rabies in a Red Fox (*Vulpes vulpes*): Isolation of Vaccine Virus in Brain Tissue and Salivary Glands." J Wildl Dis.

Houk, V. N., J. H. Baker, K. Sorensen and D. C. Kent (1968). "The epidemiology of tuberculosis infection in a closed environment." Arch Environ Health **16**(1): 26-35.

Hu, M., N. Wang, Z. H. Pan, C. P. Lu and Y. J. Liu (2012). "Identity and virulence properties of *Aeromonas* isolates from diseased fish, healthy controls and water environment in China." Letters in Applied Microbiology **55**(3): 224-233.

Huck, M., A. C. Frantz, D. A. Dawson, T. Burke and T. J. Roper (2008). "Low genetic variability, female-biased dispersal and high movement rates in an urban population of Eurasian badgers *Meles meles*." Journal of Animal Ecology **77**(5): 905-915.

Hudson, P. J., A. Rizzoli, B. T. Grenfell, J. Heesterbeek and A. P. Dobson (2002). Ecology of wildlife diseases, Oxford University Press.

Hyndman, R. J. and Y. Khandakar (2007). Automatic time series for forecasting: the forecast package for R, Monash University, Department of Econometrics and Business Statistics.

Ip, H. S., M. R. Wiley, R. Long, G. Palacios, V. Shearn-Bochsler and C. A. Whitehouse (2014). "Identification and characterization of Highlands J virus from a Mississippi sandhill crane using unbiased next-generation sequencing." Journal of Virological Methods **206**(0): 42-45.

Jenkins, H. E., R. Woodroffe, C. A. Donnelly, D. R. Cox, W. T. Johnston, F. J. Bourne, C. L. Cheeseman, R. S. Clifton-Hadley, G. Gettinby, P. Gilks, R. G. Hewinson, J. P. McInerney and W. I. Morrison (2007). "Effects of culling on spatial associations of *Mycobacterium bovis* infections in badgers and cattle." Journal of Applied Ecology **44**(5): 897-908.

Jombart, T. (2008). "adegenet: a R package for the multivariate analysis of genetic markers." Bioinformatics **24**(11): 1403-1405.

Jombart, T., A. Cori, X. Didelot, S. Cauchemez, C. Fraser and N. Ferguson (2014). "Bayesian Reconstruction of Disease Outbreaks by Combining Epidemiologic and Genomic Data." PLoS Comput Biol **10**(1): e1003457.

Jones, K. E., N. G. Patel, M. A. Levy, A. Storeygard, D. Balk, J. L. Gittleman and P. Daszak (2008). "Global trends in emerging infectious diseases." Nature **451**(7181): 990-993.

Joshi, D., N. B. Harris, R. Waters, T. Thacker, B. Mathema, B. Krieswirth and S. Sreevatsan (2012). "Single Nucleotide Polymorphisms in the *Mycobacterium*

bovis Genome Resolve Phylogenetic Relationships." Journal of Clinical Microbiology **50**(12): 3853-3861.

Kamath, P. L., J. T. Foster, K. P. Drees, G. Luikart, C. Quance, N. J. Anderson, P. R. Clarke, E. K. Cole, M. L. Drew, W. H. Edwards, J. C. Rhyan, J. J. Treanor, R. L. Wallen, P. J. White, S. Robbe-Austerman and P. C. Cross (2016). "Genomics reveals historic and contemporary transmission dynamics of a bacterial disease among wildlife and livestock." Nature Communications **7**.

Kao, R. R., D. T. Haydon, S. J. Lycett and P. R. Murcia (2014). "Supersize me: how whole-genome sequencing and big data are transforming epidemiology." Trends in Microbiology **22**(5): 282-291.

Kardos, M., F. W. Allendorf and G. Luikart (2014). "Evaluating the role of inbreeding depression in heterozygosity-fitness correlations: how useful are tests for identity disequilibrium?" Molecular Ecology Resources **14**(3): 519-530.

Keeling, M. J. and P. J. White (2010). "Targeting vaccination against novel infections: risk, age and spatial structure for pandemic influenza in Great Britain." Journal of The Royal Society Interface: rsif20100474.

Keim, P., K. L. Smith, C. Keys, H. Takahashi, T. Kurata and A. Kaufmann (2001). "Molecular Investigation of the Aum Shinrikyo Anthrax Release in Kameido, Japan." Journal of Clinical Microbiology **39**(12): 4566-4567.

Keller, L., J. Reid and P. Arcese (2008). "Testing evolutionary models of senescence in a natural population: age and inbreeding effects on fitness components in song sparrows." Proceedings of the Royal Society of London B: Biological Sciences **275**(1635): 597-604.

Keller, L. F., P. R. Grant, B. R. Grant, K. Petren and J. Merilä (2002). "Environmental Conditions Affect the Magnitude of Inbreeding Depression in Survival of Darwin's Finches " Evolution **56**(6): 1229-1239.

Keller, L. F. and D. M. Waller (2002). "Inbreeding effects in wild populations." Trends in Ecology & Evolution **17**(5): 230-241.

Kenyon , T. A., S. E. Valway , W. W. Ihle , I. M. Onorato and K. G. Castro (1996). "Transmission of Multidrug-Resistant *Mycobacterium tuberculosis* during a Long Airplane Flight." New England Journal of Medicine **334**(15): 933-938.

Kokko, H., I. Ots and T. Tregenza (2006). "When not to avoid inbreeding." Evolution **60**(3): 467-475.

Koopmans, M., J. Vinjé, M. de Wit, I. Leenen, W. van der Poel and Y. van Duynhoven (2000). "Molecular Epidemiology of Human Enteric Caliciviruses in The Netherlands." Journal of Infectious Diseases **181**(Supplement 2): S262-S269.

Köser, C. U., M. J. Ellington, E. J. P. Cartwright, S. H. Gillespie, N. M. Brown, M. Farrington, M. T. G. Holden, G. Dougan, S. D. Bentley, J. Parkhill and S. J. Peacock (2012). "Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology." PLoS Pathogens **8**(8): e1002824.

Kovats, R., D. Campbell-Lendrum, A. McMichel, A. Woodward and J. S. H. Cox (2001). "Early effects of climate change: do they include changes in vector-borne disease?" Philosophical Transactions of the Royal Society B: Biological Sciences **356**(1411): 1057-1068.

Kowalczyk, R., A. Zalewski and B. Jędrzejewska (2004). "Seasonal and spatial pattern of shelter use by badgers (*Meles meles*) in Białowieża Primeval Forest (Poland)." Acta Theriologica **49**(1): 75-92.

Kraemer, P., G. Gerlach, M. P. Kraemer and M. Suggests (2013). "Package 'Demerelate'."

Krebs, J., R. Anderson, T. Clutton-Brock, D. Young, C. Donnelly, S. Frost and R. Woodroffe (1997). "Bovine Tuberculosis in Cattle and Badgers."

Kruuk, H. and T. Parish (1987). "Changes in the size of groups and ranges of the European badger (*Meles meles*) in an area in Scotland." The Journal of Animal Ecology: 351-364.

Kumarasamy, K. K., M. A. Toleman, T. R. Walsh, J. Bagaria, F. Butt, R. Balakrishnan, U. Chaudhary, M. Doumith, C. G. Giske, S. Irfan, P. Krishnan, A. V. Kumar, S. Maharjan, S. Mushtaq, T. Noorie, D. L. Paterson, A. Pearson, C. Perry, R. Pike, B. Rao, U. Ray, J. B. Sarma, M. Sharma, E. Sheridan, M. A. Thirunarayan, J. Turton, S. Upadhyay, M. Warner, W. Welfare, D. M. Livermore and N. Woodford (2010). "Emergence of a new antibiotic resistance mechanism in India, Pakistan, and the UK: a molecular, biological, and epidemiological study." The Lancet Infectious Diseases **10**(9): 597-602.

Lang, K. R. and J. A. Blanchong (2012). "Population genetic structure of white-tailed deer: Understanding risk of chronic wasting disease spread." The Journal of Wildlife Management **76**(4): 832-840.

Lawson Handley, L. and N. Perrin (2007). "Advances in our understanding of mammalian sex-biased dispersal." Molecular Ecology **16**(8): 1559-1578.

le Roex, N., P. D. van Helden, A. P. Koets and E. G. Hoal (2013). "Bovine TB in livestock and wildlife: what's in the genes?" Physiol Genomics **45**(15): 631-637.

Lee, J. S., E. W. Ruell, E. E. Boydston, L. M. Lyren, R. S. Alonso, J. L. Troyer, K. R. Crooks and S. Vandewoude (2012). "Gene flow and pathogen transmission among bobcats (*Lynx rufus*) in a fragmented urban landscape." Mol Ecol **21**(7): 1617-1631.

Lemey, P., A. Rambaut, A. J. Drummond and M. A. Suchard (2009). "Bayesian phylogeography finds its roots." PLoS Comput Biol **5**.

Lemey, P., A. Rambaut, A. J. Drummond and M. A. Suchard (2009). "Bayesian phylogeography finds its roots." PLoS Comput Biol **5**(9): e1000520.

Leopardi, S., D. Blake and S. J. Puechmaille (2015). "White-Nose Syndrome fungus introduced from Europe to North America." Current Biology **25**(6): R217-R219.

Lesellier, S., L. Corner, E. Costello, K. Lyashchenko, R. Greenwald, J. Esfandiari, M. Singh, R. G. Hewinson, M. Chambers and E. Gormley (2009). "Immunological responses and protective immunity in BCG vaccinated badgers following endobronchial infection with Mycobacterium bovis." Vaccine **27**(3): 402-409.

Li, Y. C., A. B. Korol, T. Fahima, A. Beiles and E. Nevo (2002). "Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review." Molecular ecology **11**(12): 2453-2465.

Lisle, G. W., G. F. Yates, D. M. Collins, R. W. MacKenzie, K. B. Crews and R. Walker (1995). "A study of bovine tuberculosis in domestic animals and wildlife in the MacKenzie Basin and surrounding areas using DNA fingerprinting." N Z Vet J **43**.

Little, T., P. Naylor and J. Wilesmith (1982). "Laboratory study of Mycobacterium bovis infection in badgers and calves." The Veterinary Record **111**(24): 550-557.

Liu, C., C. J. Graber, M. Karr, B. A. Diep, L. Basuino, B. S. Schwartz, M. C. Enright, S. J. O'Hanlon, J. C. Thomas, F. Perdreau-Remington, S. Gordon, H. Gunthorpe, R. Jacobs, P. Jensen, G. Leoung, J. S. Rumack and H. F. Chambers (2008). "A Population-Based Study of the Incidence and Molecular Epidemiology of Methicillin-Resistant *Staphylococcus aureus* Disease in San Francisco, 2004–2005." Clinical Infectious Diseases **46**(11): 1637-1646.

Lloyd-Smith, J. O., P. C. Cross, C. J. Briggs, M. Daugherty, W. M. Getz, J. Latto, M. S. Sanchez, A. B. Smith and A. Swei (2005). "Should we expect population thresholds for wildlife disease?" Trends Ecol Evol **20**(9): 511-519.

Lloyd-Smith, J. O., S. J. Schreiber, P. E. Kopp and W. M. Getz (2005). "Superspreading and the effect of individual variation on disease emergence." Nature **438**(7066): 355-359.

Lochmiller, R. L., M. R. Vestey and J. C. Boren (1993). "Relationship between protein nutritional status and immunocompetence in northern bobwhite chicks." The Auk: 503-510.

Loehle, C. (1995). "Social barriers to pathogen transmission in wild animal populations." Ecology **76**(2): 326-335.

Lyons, E. J., A. J. Frodsham, L. Zhang, A. V. S. Hill and W. Amos (2009). Consanguinity and susceptibility to infectious diseases in humans.

Maan, S., M. Belaganahalli, N. Maan and P. C. Mertens (2013). Whole Genome Sequencing Strategies and Development of Orbivirus Sequence Database: Implications for Novel dsRNA Virus Detection. Biotechnology: Prospects and Applications. R. K. Salar, S. K. Gahlawat, P. Siwach and J. S. Duhan, Springer India**:** 237-255.

Macdonald, D. W., B. J. Harmsen, P. J. Johnson and C. Newman (2004). "Increasing frequency of bite wounds with increasing population density in Eurasian badgers, *Meles meles*." Animal Behaviour **67**(4): 745-751.

Macdonald, D. W., C. Newman, C. D. Buesching and P. J. Johnson (2008). "Male-biased movement in a high-density population of the Eurasian badger (*Meles meles*)." Journal of Mammalogy **89**(5): 1077-1086.

Macdonald, D. W., P. Riordan and F. Mathews (2006). "Biological hurdles to the control of TB in cattle: A test of two hypotheses concerning wildlife to explain the failure of control." Biological Conservation **131**(2): 268-286.

Macdonald, G. (1952). "The analysis of equilibrium in malaria." Trop Dis Bull **49**(9): 813-829.

Mack, U., G. B. Migliori, M. Sester, H. L. Rieder, S. Ehlers, D. Goletti, A. Bossink, K. Magdorf, C. Hölscher, B. Kampmann, S. M. Arend, A. Detjen, G. Bothamley, J. P. Zellweger, H. Milburn, R. Diel, P. Ravn, F. Cobelens, P. J. Cardona, B. Kan, I. Solovic, R. Duarte and D. M. Cirillo (2009). "LTBI: latent tuberculosis infection or lasting immune responses to *M. tuberculosis*? A TBNET consensus statement." European Respiratory Journal **33**(5): 956-973.

Mackintosh, C., J. MacArthur, T. Little and P. Stuart (1976). "The immobilization of the badger (*Meles meles*)." British Veterinary Journal (UK).

Mackintosh, C. G., T. Qureshi, K. Waldrup, R. E. Labes, K. G. Dodds and J. F. Griffin (2000). "Genetic resistance to experimental infection with *Mycobacterium bovis* in red deer (*Cervus elaphus*)." Infect Immun **68**(3): 1620-1625.

Mazé-Guilmo, E., S. Blanchet, K. D. McCoy and G. Loot (2016). "Host dispersal as the driver of parasite genetic structure: a paradigm lost?" Ecology letters **19**(3): 336-347.

McCallum, H., N. Barlow and J. Hone (2001). "How should pathogen transmission be modelled?" Trends in Ecology & Evolution **16**(6): 295-300.

McDonald, J. L. (2014). Disease and demography in the Woodchester Park badger population PhD, University of Exeter, Penryn Campus.

McDonald, J. L., T. Bailey, R. J. Delahay, R. A. McDonald, G. C. Smith and D. J. Hodgson (2016). "Demographic buffering and compensatory recruitment promotes the persistence of disease in a wildlife population." Ecology letters.

Medawar, P. B. (1952). An unsolved problem of biology, College.

Medini, D., D. Serruto, J. Parkhill, D. A. Relman, C. Donati, R. Moxon, S. Falkow and R. Rappuoli (2008). "Microbiology in the post-genomic era." Nat Rev Microbiol **6**(6): 419-430.

Meli, M. L., V. Cattori, F. Martínez, G. López, A. Vargas, M. A. Simón, I. Zorrilla, A. Muñoz, F. Palomares, J. V. López-Bao, J. Pastor, R. Tandon, B. Willi, R. Hofmann-Lehmann and H. Lutz (2009). "Feline Leukemia Virus and Other Pathogens as Important Threats to the Survival of the Critically Endangered Iberian Lynx (*Lynx pardinus*)." PloS one **4**(3): e4744.

Metzker, M. L. (2010). "Sequencing technologies [mdash] the next generation." Nat Rev Genet **11**(1): 31-46.

Millan, J. and I. Blasco-Costa (2012). "Molecular evidence of shared hookworm *Ancylostoma tubaeforme* haplotypes between the critically endangered Iberian lynx and sympatric domestic cats." Veterinary Parasitology **186**: 518-522.

Miller, M. W. and M. M. Conner (2005). "Epidemiology of chronic wasting disease in free-ranging mule deer: spatial, temporal, and demographic influences on observed prevalence patterns." Journal of Wildlife Diseases **41**(2): 275-290.

Minin, V. N. and M. A. Suchard (2008). "Fast, accurate and simulation-free stochastic mapping." Philosophical Transactions of the Royal Society B: Biological Sciences **363**(1512): 3985-3995.

Möller, L., L. Beheregaray, S. Allen and R. Harcourt (2006). "Association patterns and kinship in female Indo-Pacific bottlenose dolphins (*Tursiops aduncus*) of southeastern Australia." Behavioral Ecology and Sociobiology **61**(1): 109-117.

Monot, M., N. Honore, T. Garnier, N. Zidane, D. Sherafi, A. Paniz-Mondolfi, M. Matsuoka, G. M. Taylor, H. D. Donoghue, A. Bouwman, S. Mays, C. Watson, D. Lockwood, A. Khamispour, Y. Dowlati, S. Jianping, T. H. Rea, L. Vera-Cabrera, M. M. Stefani, S. Banu, M. Macdonald, B. R. Sapkota, J. S. Spencer, J. Thomas, K. Harshman, P. Singh, P. Busso, A. Gattiker, J. Rougemont, P. J. Brennan and S. T. Cole (2009). "Comparative genomic and phylogeographic analysis of *Mycobacterium leprae*." Nat Genet **41**(12): 1282-1289.

Moore, J. and R. Ali (1984). "Are dispersal and inbreeding avoidance related?" Animal Behaviour **32**(1): 94-112.

Muirhead, R. and K. Burns (1974). "Tuberculosis in wild badgers in Gloucestershire: epidemiology." Veterinary Record **95**(24): 552-555.

Munyeme, M., L. Rigouts, I. C. Shamputa, J. B. Muma, M. Tryland, E. Skjerve and B. Djonne (2009). "Isolation and characterization of *Mycobacterium bovis* strains from indigenous Zambian cattle using Spacer oligonucleotide typing technique." Bmc Microbiology **9**.

Murphy, D., E. Gormley, E. Costello, D. O'Meara and L. A. Corner (2010). "The prevalence and distribution of Mycobacterium bovis infection in European badgers (*Meles meles*) as determined by enhanced post mortem examination and bacteriological culture." Res Vet Sci **88**(1): 1-5.

Naranjo, V., C. Gortazar, J. Vicente and J. De La Fuente (2008). "Evidence of the role of European wild boar as a reservoir of *Mycobacterium tuberculosis* complex." Veterinary microbiology **127**(1): 1-9.

Nardell, E. A., J. Keegan, S. A. Cheney and S. C. Etkind (1991). "Airborne infection. Theoretical limits of protection achievable by building ventilation." Am Rev Respir Dis **144**(2): 302-306.

Neel, C. (2010). "Molecular Epidemiology of Simian Immunodeficiency Virus Infection in Wild-Living Gorillas." Journal of Virology **84**(3): 1464-1476.

Ní Bhuachalla, D., L. A. Corner, S. J. More and E. Gormley (2014). "The role of badgers in the epidemiology of *Mycobacterium bovis* infection (tuberculosis) in cattle in the United Kingdom and the Republic of Ireland: current perspectives on control strategies."

Nieberding, C. M., M. C. Durette-Desset, A. Vanderpoorten, J. C. Casanova, A. Ribas, V. Deffontaine, C. Feliu, S. Morand, R. Libois and J. R. Michaux (2008). "Geography and host biogeography matter for understanding the phylogeography of a parasite." Molecular Phylogenetics and Evolution **47**(2): 538-554.

Nolan, A. (1991). An investigation of the development of specific antibody responses of badgers (Meles meles) to infection with Mycobacterium boviswith reference to the pathogenesis and epidemiology of the disease. Brunel University, Department of Biology and Biochemistry.

Nunes, M. R., N. R. Faria, H. B. Vasconcelos, D. B. Medeiros, C. P. Silva de Lima, V. L. Carvalho, E. V. Pinto da Silva, J. F. Cardoso, E. C. Sousa, K. N. Nunes, S. G. Rodrigues, A. B. Abecasis, M. A. Suchard, P. Lemey and P. F. Vasconcelos (2012). "Phylogeography of dengue virus serotype 4, Brazil, 2010-2011." Emerg Infect Dis **18**(11): 1858-1864.

O'Brien, S. J. and J. F. Evermann (1988). "Interactive influence of infectious disease and genetic diversity in natural populations." Trends in Ecology & Evolution **3**(10): 254-259.

O'Brien, D. J., S. M. Schmitt, J. S. Fierke, S. A. Hogle, S. R. Winterstein, T. M. Cooley, W. E. Moritz, K. L. Diegel, S. D. Fitzgerald and D. E. Berry (2002). "Epidemiology of *Mycobacterium bovis* in free-ranging white-tailed deer, Michigan, USA, 1995–2000." Preventive veterinary medicine **54**(1): 47-63.

O'brien, R. M. (2007). "A caution regarding rules of thumb for variance inflation factors." Quality & Quantity **41**(5): 673-690.

Olea-Popelka, F. J., J. M. Griffin, J. D. Collins, G. McGrath and S. W. Martin (2003). "Bovine tuberculosis in badgers in four areas in Ireland: does tuberculosis cluster?" Preventive Veterinary Medicine **59**(1–2): 103-111.

Ostfeld, R. S., G. E. Glass and F. Keesing (2005). "Spatial epidemiology: an emerging (or re-emerging) discipline." Trends in ecology & evolution **20**(6): 328-336.

Overend, E. D. (1980). "Badgers and TB—does gassing spread the disease." Oryx **15**(04): 338-340.

Padgett, D. A. and R. Glaser (2003). "How stress influences the immune response." Trends in Immunology **24**(8): 444-448.

Paillot, R., A. C. Darby, C. Robinson, N. L. Wright, K. F. Steward, E. Anderson, K. Webb, M. T. Holden, A. Efstratiou, K. Broughton, K. A. Jolley, S. L. Priestnall, M. C. Marotti Campi, M. A. Hughes, A. Radford, K. Erles and A. S. Waller (2010). "Identification of three novel superantigen-encoding genes in Streptococcus equi subsp. zooepidemicus, szeF, szeN, and szeP." Infect Immun **78**(11): 4817-4827.

Palphramand, K., G. Newton-Cross and P. L. White (2007). "Spatial organization and behaviour of badgers (*Meles meles*) in a moderate-density population." Behavioral Ecology and Sociobiology **61**(3): 401-413.

Pantelic, J., G. N. Sze-To, K. W. Tham, C. Y. Chao and Y. C. Khoo (2009). "Personalized ventilation as a control measure for airborne transmissible disease spread." J R Soc Interface **6 Suppl 6**: S715-726.

Paradis, E., J. Claude and K. Strimmer (2004). "APE: analyses of phylogenetics and evolution in R language." Bioinformatics **20**(2): 289-290.

Paull, S. H., S. Song, K. M. McClure, L. C. Sackett, A. M. Kilpatrick and P. T. J. Johnson (2011). "From superspreaders to disease hotspots: linking transmission across hosts and space." Frontiers in Ecology and the Environment **10**(2): 75-82.

Pavlik, I., L. Dvorska, M. Bartos, I. Parmova, I. Melicharek, A. Jesenska, M. Havelkova, M. Slosarek, I. Putova and G. Martin (2002). "Molecular epidemiology of bovine tuberculosis in the Czech Republic and Slovakia in the period 1965-2001 studied by spoligotyping." VETERINARNI MEDICINA-PRAHA- **47**(7): 181-194.

Pebesma, E. J. (2004). "Multivariable geostatistics in S: the gstat package." Computers & Geosciences **30**(7): 683-691.

Perez de Rosas, A. R., E. L. Segura and B. A. Garcia (2011). "Molecular phylogeography of the Chagas disease vector *Triatoma infestans* in Argentina." Heredity **107**(1): 71-79.

Perry, S., J. H. Manson, L. Muniz, J. Gros-Louis and L. Vigilant (2008). "Kin-biased social behaviour in wild adult female white-faced capuchins, *Cebus capucinus*." Animal Behaviour **76**(1): 187-199.

Pertoldi, C., V. Loeschcke, A. B. Madsen, E. Randi and N. Mucci (2001). "Effects of habitat fragmentation on the Eurasian badger (*Meles meles*) subpopulations in Denmark." Hystrix, the Italian Journal of Mammalogy **12**(1).

Plowright, R. K., P. Eby, P. J. Hudson, I. L. Smith, D. Westcott, W. L. Bryden, D. Middleton, P. A. Reid, R. A. McFarlane and G. Martin (2015). Ecological dynamics of emerging bat virus spillover. Proc. R. Soc. B, The Royal Society.

Pop, M. and S. L. Salzberg (2008). "Bioinformatics challenges of new sequencing technology." Trends in Genetics **24**(3): 142-149.

Pope, L., X. Domingo-Roura, K. Erven and T. Burke (2006). "Isolation by distance and gene flow in the Eurasian badger (*Meles meles*) at both a local and broad scale." Molecular Ecology **15**(2): 371-386.

Pope, L. C., R. K. Butlin, G. J. Wilson, R. Woodroffe, K. Erven, C. M. Conyers, T. Franklin, R. J. Delahay, C. L. Cheeseman and T. Burke (2007). "Genetic evidence that culling increases badger movement: implications for the spread of bovine tuberculosis." Molecular Ecology **16**(23): 4919-4929.

Price, L. B., M. Stegger, H. Hasman, M. Aziz, J. Larsen, P. S. Andersen, T. Pearson, A. E. Waters, J. T. Foster, J. Schupp, J. Gillece, E. Driebe, C. M. Liu, B. Springer, I. Zdovc, A. Battisti, A. Franco, J. Żmudzki, S. Schwarz, P. Butaye, E. Jouy, C. Pomba, M. C. Porrero, R. Ruimy, T. C. Smith, D. A. Robinson, J. S. Weese, C. S. Arriola, F. Yu, F. Laurent, P. Keim, R. Skov and F. M. Aarestrup (2012). "*Staphylococcus aureus* CC398: Host Adaptation and Emergence of Methicillin Resistance in Livestock." mBio **3**(1).

Pritchard, D., F. Stuart, J. Wilesmith, C. Cheeseman, J. Brewer, R. Bode and P. Sayers (1986). "Tuberculosis in East Sussex: III. Comparison of post-mortem and clinical methods for the diagnosis of tuberculosis in badgers." Journal of Hygiene **97**(01): 27-36.

Pritchard, J. K., M. Stephens and P. Donnelly (2000). "Inference of Population Structure Using Multilocus Genotype Data." Genetics **155**(2): 945-959.

Pusey, A. and M. Wolf (1996). "Inbreeding avoidance in animals." Trends in Ecology & Evolution **11**(5): 201-206.

Pusey, A. E. (1987). "Sex-biased dispersal and inbreeding avoidance in birds and mammals." Trends in ecology & evolution **2**(10): 295-299.

Queller, D. C. and K. F. Goodnight (1989). "Estimating relatedness using genetic markers." Evolution: 258-275.

R Core Development Team (2013). R: a language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing.

R Core Development Team (2016). R: a language and environment for statistical computing. R Development Core Team, Vienna.

Rambaut, A. (2007). "FigTree, a graphical viewer of phylogenetic trees." See http://tree. bio. ed. ac. uk/software/figtree.

Rambaut, A. and A. Drummond (2007). Tracer v1. 4. Molecular Evolution, Phylogenetics and Epidemiology.

Rambaut, A. and A. Drummond (2013). "TreeAnnotator v1. 7.0." Available as part of the BEAST package at http://beast. bio. ed. ac. uk.

Ramsey, D., N. Spencer, P. Caley, M. Efford, K. Hansen, M. Lam and D. Cooper (2002). "The effects of reducing population density on contact rates between brushtail possums: implications for transmission of bovine tuberculosis." Journal of Applied Ecology **39**(5): 806-818.

Rankin, D. J., K. Bargum and H. Kokko (2007). "The tragedy of the commons in evolutionary biology." Trends in Ecology & Evolution **22**(12): 643-651.

Rasko, D. A., P. L. Worsham, T. G. Abshire, S. T. Stanley, J. D. Bannan, M. R. Wilson, R. J. Langham, R. S. Decker, L. Jiang, T. D. Read, A. M. Phillippy, S. L. Salzberg, M. Pop, M. N. Van Ert, L. J. Kenefic, P. S. Keim, C. M. Fraser-Liggett and J. Ravel (2011). "*Bacillus anthracis* comparative genome analysis in support of the Amerithrax investigation." Proceedings of the National Academy of Sciences of the United States of America **108**(12): 5027-5032.

Recker, M., O. G. Pybus, S. Nee and S. Gupta (2007). "The generation of influenza outbreaks by a network of host immune responses against a limited set of antigenic types." Proceedings of the National Academy of Sciences **104**(18): 7711-7716.

Reid, J. M., P. Arcese and L. F. Keller (2003). "Inbreeding depresses immune response in song sparrows (*Melospiza melodia*): direct and inter–generational effects." Proceedings of the Royal Society of London. Series B: Biological Sciences **270**(1529): 2151-2157.

Relman, D. A. (2011). "Microbial genomics and infectious diseases." N Engl J Med **365**(4): 347-357.

Richards, S. A. (2005). "Testing ecological theory using the information-theoretic approach: examples and cautionary results." Ecology **86**(10): 2805-2814.

Richards, S. A., M. J. Whittingham and P. A. Stephens (2011). "Model selection and model averaging in behavioural ecology: the utility of the IT-AIC framework." Behavioral Ecology and Sociobiology **65**(1): 77-89.

Riordan, P., R. J. Delahay, C. Cheeseman, P. J. Johnson and D. W. Macdonald (2011). "Culling-Induced Changes in Badger *(Meles meles)* Behaviour, Social Organisation and the Epidemiology of Bovine Tuberculosis." PLoS ONE **6**(12): e28904.

Robertson, A., K. L. Palphramand, S. P. Carter and R. J. Delahay (2015). "Group size correlates with territory size in European badgers: implications for the resource dispersion hypothesis?" Oikos **124**(4): 507-514.

Robertson, B. C., B. L. Chilvers, P. J. Duignan, I. S. Wilkinson and N. J. Gemmell (2006). "Dispersal of breeding, adult male *Phocarctos hookeri*: implications for disease transmission, population management and species recovery." Biological Conservation **127**(2): 227-236.

Robinson, S. P., L. W. Simmons and W. J. Kennington (2013). "Estimating relatedness and inbreeding using molecular markers and pedigrees: the effect of demographic history." Molecular Ecology **22**(23): 5779-5792.

Roetzer, A., R. Diel, T. A. Kohl, C. Rückert, U. Nübel, J. Blom, T. Wirth, S. Jaenicke, S. Schuback, S. Rüsch-Gerdes, P. Supply, J. Kalinowski and S. Niemann (2013). "Whole Genome Sequencing versus Traditional Genotyping for Investigation of a *Mycobacterium tuberculosis* Outbreak: A Longitudinal Molecular Epidemiological Study." PLoS Med **10**(2): e1001387.

Rogers, L., R. Delahay, C. Cheeseman, S. Langton, G. Smith and R. Clifton-Hadley (1998). "Movement of badgers (*Meles meles*) in a high–density population: individual, population and disease effects." Proceedings of the Royal Society of London. Series B: Biological Sciences **265**(1403): 1269-1276.

Rogers, L. M., C. L. Cheeseman, P. J. Mallinson and R. Clifton-Hadley (1997). "The demography of a high-density badger (*Meles meles*) population in the west of England." Journal of Zoology **242**(4): 705-728.

Roper, T., J. Ostler and L. Conradt (2003). "The process of dispersal in badgers *Meles meles*." Mammal Review **33**(3-4): 314-318.

Roper, T., J. Ostler, T. Schmid and S. Christian (2001). "Sett use in European badgers *Meles meles*." Behaviour **138**(2): 173-187.

Roper, T. J. (2010). Badger, Collins London.

Rudan, I., D. Rudan, H. Campbell, A. Carothers, A. Wright, N. Smolej-Narancic, B. Janicijevic, L. Jin, R. Chakraborty and R. Deka (2003). "Inbreeding and risk of late onset complex disease." Journal of Medical Genetics **40**(12): 925-932.

Russell, A. F. and B. J. Hatchwell (2001). "Experimental evidence for kin-biased helping in a cooperatively breeding vertebrate." Proceedings of the Royal Society of London B: Biological Sciences **268**(1481): 2169-2174.

Sainsbury, A. W., P. Nettleton, J. Gilray and J. Gurnell (2000). "Grey squirrels have high seroprevalence to a parapoxvirus associated with deaths in red squirrels." Animal Conservation **3**(3): 229-233.

Salathé, M. and J. H. Jones (2010). "Dynamics and control of diseases in networks with community structure." PLoS Comput Biol **6**(4): e1000736.

Salyer, S. J., T. R. Gillespie, I. B. Rwego, C. A. Chapman and T. L. Goldberg (2012). "Epidemiology and Molecular Relationships of *Cryptosporidium* spp. in People, Primates, and Livestock from Western Uganda." Plos Neglected Tropical Diseases **6**(4).

Santolaya, M. E., M. L. O'Ryan, M. T. Valenzuela, V. Prado, R. Vergara, A. Munoz, D. Toneatto, G. Grana, H. Wang, R. Clemens and P. M. Dull (2012). "Immunogenicity and tolerability of a multicomponent meningococcal serogroup B (4CMenB) vaccine in healthy adolescents in Chile: a phase 2b/3 randomised, observer-blind, placebo-controlled study." Lancet **379**(9816): 617-624.

Santos, N., M. Correia-Neves, S. Ghebremichael, G. Källenius, S. B. Svenson and V. Almeida (2009). "Epidemiology of *Mycobacterium bovis* infection in wild boar (*Sus scrofa*) from Portugal." Journal of Wildlife Diseases **45**(4): 1048-1061.

Schino, G. (2001). "Grooming, competition and social rank among female primates: a meta-analysis." Animal Behaviour **62**(2): 265-271.

Seib, K. L., G. Dougan and R. Rappuoli (2009). "The Key Role of Genomics in Modern Vaccine and Drug Design for Emerging Infectious Diseases." PLoS Genet **5**(10): e1000612.

Serraino, A., G. Marchetti, V. Sanguinetti, M. C. Rossi, R. G. Zanoni, L. Catozzi, A. Bandera, W. Dini, W. Mignone and F. Franzetti (1999). "Monitoring of transmission of tuberculosis between wild boars and cattle: genotypical analysis of strains by molecular epidemiology techniques." Journal of Clinical Microbiology **37**(9): 2766-2771.

Serruto, D., L. Serino, V. Masignani and M. Pizza (2009). "Genome-based approaches to develop vaccines against bacterial pathogens." Vaccine **27**(25-26): 3245-3250.

Sevi, A., G. Annicchiarico, M. Albenzio, L. Taibi, A. Muscio and S. Dell'Aquila (2001). "Effects of solar radiation and feeding time on behavior, immune response and production of lactating ewes under high ambient temperature." Journal of Dairy Science **84**(3): 629-640.

Siddle, H. V., A. Kreiss, M. D. B. Eldridge, E. Noonan, C. J. Clarke, S. Pyecroft, G. M. Woods and K. Belov (2007). "Transmission of a fatal clonal tumor by biting occurs due to depleted MHC diversity in a threatened carnivorous marsupial." Proceedings of the National Academy of Sciences **104**(41): 16221-16226.

Siddle, H. V., J. Marzec, Y. Cheng, M. Jones and K. Belov (2010). "MHC gene copy number variation in Tasmanian devils: implications for the spread of a contagious cancer." Proc Biol Sci **277**(1690): 2001-2006.

Sin, Y. W., H. L. Dugdale, C. Newman, D. W. Macdonald and T. Burke (2012). "MHC class II genes in the European badger (*Meles meles*): characterization, patterns of variation, and transcription analysis." Immunogenetics **64**(4): 313-327.

Slate, J., P. David, K. G. Dodds, B. A. Veenvliet, B. C. Glass, T. E. Broad and J. C. McEwan (2004). "Understanding the relationship between the inbreeding coefficient and multilocus heterozygosity: theoretical expectations and empirical data." Heredity (Edinb) **93**(3): 255-265.

Smallbone, W., C. van Oosterhout and J. Cable (2016). "The effects of inbreeding on disease susceptibility: *Gyrodactylus turnbulli* infection of guppies, *Poecilia reticulata*." Experimental Parasitology **167**: 32-37.

Smith, N. H., J. Dale, J. Inwald, S. Palmer, S. V. Gordon, R. G. Hewinson and J. M. Smith (2003). "The population structure of *Mycobacterium bovis* in Great Britain: clonal expansion." Proc Natl Acad Sci U S A **100**(25): 15271-15275.

Smith, N. H., S. V. Gordon, R. de la Rua-Domenech, R. S. Clifton-Hadley and R. G. Hewinson (2006). "Bottlenecks and broomsticks: the molecular evolution of *Mycobacterium bovis*." Nature Reviews Microbiology **4**: 670-681.

Smith, R. P., P. W. Rand, E. H. Lacombe, S. R. Morris, D. W. Holmes and D. A. Caporale (1996). "Role of Bird Migration in the Long-Distance Dispersal of *Ixodes dammini*, the Vector of Lyme Disease." Journal of Infectious Diseases **174**(1): 221-224.

Snitkin, E. S., A. M. Zelazny, P. J. Thomas, F. Stock, D. K. Henderson, T. N. Palmore and J. A. Segre (2012). "Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing." Sci Transl Med **4**(148): 148ra116.

Snow, M. (2011). "The contribution of molecular epidemiology to the understanding and control of viral diseases of salmonid aquaculture." Veterinary Research **42**(1): 56.

Soulé, M. E. (1987). Viable populations for conservation, Cambridge university press.

Spibey, N., V. J. McCabe, N. M. Greenwood, S. C. Jack, D. Sutton and L. van der Waart (2012). "Novel bivalent vectored vaccine for control of myxomatosis and rabbit haemorrhagic disease." Veterinary Record.

Spielman, D., B. W. Brook, D. A. Briscoe and R. Frankham (2004). "Does Inbreeding and Loss of Genetic Diversity Decrease Disease Resistance?" Conservation Genetics **5**(4): 439-448.

Starkman, S. E., D. M. MacDonald, J. C. M. Lewis, E. C. Holmes and P. Simmonds (2003). "Geographic and species association of hepatitis B virus genotypes in non-human primates." Virology **314**: 381-393.

Steward, L. C. (2017). Badger social networks and their implications for disease transmission. PhD PhD, University of Exeter, Penryn Campus.

Stewart, P., S. Ellwood and D. Macdonald (1997). "Remote video-surveillance of wildlife–an introduction from experience with the European badger *Meles meles*." Mammal Review **27**(4): 185-204.

Stoffel, M. A., M. Esser, M. Kardos, E. Humble, H. Nichols, P. David and J. I. Hoffman (2016). "inbreedR: an R package for the analysis of inbreeding based on genetic markers." Methods in Ecology and Evolution: n/a-n/a.

Streicker, D. G., P. Lemey, A. Velasco-Villa and C. E. Rupprecht (2012). "Rates of Viral Evolution Are Linked to Host Geography in Bat Rabies." PLoS Pathog **8**(5): e1002720.

Streicker, D. G., J. C. Winternitz, D. A. Satterfield, R. E. Condori-Condori, A. Broos, C. Tello, S. Recuenco, A. Velasco-Villa, S. Altizer and W. Valderrama (2016). "Host–pathogen evolutionary signatures reveal dynamics and future invasions of vampire bat rabies." Proceedings of the National Academy of Sciences: 201606587.

Szentiks, C. A., K. Tsangaras, B. Abendroth, M. Scheuch, M. D. Stenglein, P. Wohlsein, F. Heeger, R. Höveler, W. Chen, W. Sun, A. Damiani, V. Nikolin, A. D. Gruber, M. Grobbel, D. Kalthoff, D. Höper, G. Á. Czirják, J. DeRisi, C. J. Mazzoni, A. Schüle, A. Aue, M. L. East, H. Hofer, M. Beer, N. Osterrieder and A. D. Greenwood (2014). "Polar Bear Encephalitis: Establishment of a Comprehensive Next-generation Pathogen Analysis Pipeline for Captive and Free-living Wildlife." Journal of Comparative Pathology **150**(4): 474-488.

Szulkin, M., N. Bierne and P. David (2010). "Heterozygosity-Fitness Correlations: A time for reappraisal." Evolution **64**(5): 1202-1217.

Szulkin, M. and B. C. Sheldon (2008). "Dispersal as a means of inbreeding avoidance in a wild bird population." Proceedings of the Royal Society B: Biological Sciences **275**(1635): 703-711.

Szulkin, M., K. V. Stopher, J. M. Pemberton and J. M. Reid (2013). "Inbreeding avoidance, tolerance, or preference in animals?" Trends in Ecology & Evolution **28**(4): 205-211.

Tettelin, H., D. Riley, C. Cattuto and D. Medini (2008). "Comparative genomics: the bacterial pan-genome." Curr Opin Microbiol **11**(5): 472-477.

Tibayrenc, M. (1998). "Beyond Strain Typing and Molecular Epidemiology: Integrated Genetic Epidemiology of Infectious Diseases." Parasitology Today **14**(8): 323-329.

Tomlinson, A., M. Chambers, G. Wilson, R. McDonald and R. Delahay (2013). "Sex-related heterogeneity in the life-history correlates of *Mycobacterium bovis* infection in European badgers (*Meles meles*)." Transboundary and emerging diseases **60**(s1): 37-45.

Tomlinson, A. J. (2013). Life-history correlates of Myobacterium bovis infection in individual Eurasion badgers (*Meles meles*), University of Liverpool.

Tomlinson, A. J., M. A. Chambers, S. P. Carter, G. J. Wilson, G. C. Smith, R. A. McDonald and R. J. Delahay (2013). "Heterogeneity in the risk of Mycobacterium bovis infection in European badger (*Meles meles*) cubs." Epidemiol Infect **141**(7): 1458-1466.

Tomlinson, A. J., M. A. Chambers, R. A. McDonald and R. J. Delahay (2015). "Association of quantitative interferon-γ responses with the progression of naturally acquired *Mycobacterium bovis* infection in wild European badgers (*Meles meles*)." Immunology **144**(2): 263-270.

Tompkins, D. M., A. M. Dunn, M. J. Smith and S. Telfer (2011). "Wildlife diseases: from individuals to ecosystems." Journal of Animal Ecology **80**(1): 19-38.

Tonellato, P. J., J. M. Crawford, M. S. Boguski and J. E. Saffitz (2011). "A National Agenda for the Future of Pathology in Personalized Medicine: Report of the Proceedings of a Meeting at the Banbury Conference Center on Genome-Era Pathology, Precision Diagnostics, and Preemptive Care: A Stakeholder Summit." American Journal of Clinical Pathology **135**(5): 668-672.

Townsend, A. K., A. B. Clark, K. J. McGowan, E. L. Buckles, A. D. Miller and I. J. Lovette (2009). "Disease-mediated inbreeding depression in a large, open population of cooperative crows." Proceedings of the Royal Society B: Biological Sciences **276**(1664): 2057-2064.

Townsend, A. K., A. B. Clark, K. J. McGowan, A. D. Miller and E. L. Buckles (2010). "Condition, innate immunity and disease mortality of inbred crows." Proceedings of the Royal Society of London B: Biological Sciences **277**(1695): 2875-2883.

Tremblay, A. and J. Ransijn (2015). "LMERConvenienceFunctions: A suite of functions to back-fit fixed effects and forward-fit random effects, as well as other miscellaneous functions. R package version 2.5. 2013." cran. r-project. org/web/packages/LMERConvenienceFunctions/index. html. Accessed at **15**.

Trewby, H. (2016). The genetic and spatial epidemiology of bovine tuberculosis in the UK: from molecular typing to bacterial whole genome sequencing, University of Glasgow.

Trewby, H., D. Wright, E. L. Breadon, S. J. Lycett, T. R. Mallon and C. McCormick (2016). "Use of bacterial whole-genome sequencing to investigate local persistence and spread in bovine tuberculosis." Epidemics Elsevier BV **14**.

Trinkel, M., D. Cooper, C. Packer and R. Slotow (2011). "Inbreeding depression increases susceptibility to bovine tuberculosis in lions: an experimental test using an inbred-outbred contrast through translocation." J Wildl Dis **47**(3): 494-500.

Tuyttens, F., R. Delahay, D. Macdonald, C. Cheeseman, B. Long and C. Donnelly (2000). "Spatial perturbation caused by a badger (*Meles meles*) culling operation: implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*)." Journal of Animal Ecology **69**(5): 815-828.

Tuyttens, F. A. M., R. J. Delahay, D. W. Macdonald, C. L. Cheeseman, B. Long and C. A. Donnelly (2000). "Spatial perturbation caused by a badger (*Meles meles*) culling operation: implications for the function of territoriality and the control of bovine tuberculosis (*Mycobacterium bovis*)." Journal of Animal Ecology **69**(5): 815-828.

Tuyttens, F. A. M., D. W. Macdonald, L. M. Rogers, C. L. Cheeseman and A. W. Roddam (2000). "Comparative study on the consequences of culling badgers (Meles meles) on biometrics, population dynamics and movement." Journal of Animal Ecology **69**(4): 567-580.

Vander Wal, E., I. Edye, P. C. Paquet, D. W. Coltman, E. Bayne, R. K. Brook and J. A. Andrés (2013). "Juxtaposition between host population structures: implications for disease transmission in a sympatric cervid community." Evolutionary Applications **6**(7): 1001-1011.

Vicente, J., R. J. Delahay, N. J. Walker and C. L. Cheeseman (2007). "Social organization and movement influence the incidence of bovine tuberculosis in an undisturbed high-density badger *Meles meles* population." J Anim Ecol **76**.

Vicente, J., R. J. Delahay, N. J. Walker and C. L. Cheeseman (2007). "Social Organization and Movement Influence the Incidence of Bovine Tuberculosis in an Undisturbed High-Density Badger *Meles meles* Population." Journal of Animal Ecology **76**(2): 348-360.

Vineis, P. and F. Perera (2007). "Molecular epidemiology and biomarkers in etiologic cancer research: the new in light of the old." Cancer Epidemiol Biomarkers Prev **16**(10): 1954-1965.

Vogler, A. J., F. Chan, D. M. Wagner, P. Roumagnac, J. Lee, R. Nera, M. Eppinger, J. Ravel, L. Rahalison, B. W. Rasoamanana, S. M. Beckstrom-Sternberg, M. Achtman, S. Chanteau and P. Keim (2011). "Phylogeography and Molecular Epidemiology of *Yersinia pestis* in Madagascar." PLoS Negl Trop Dis **5**(9): e1319.

Volz, E. M., S. L. K. Pond, M. J. Ward, A. J. L. Brown and S. D. Frost (2009). "Phylodynamics of infectious disease epidemics." Genetics **183**(4): 1421-1430.

Vordermeier, M., G. Ameni, S. Berg, R. Bishop, B. D. Robertson, A. Aseffa, R. G. Hewinson and D. B. Young (2012). "The influence of cattle breed on susceptibility to bovine tuberculosis in Ethiopia." Comp Immunol Microbiol Infect Dis **35**(3): 227-232.

Walker, N. J. (2012). The analysis of epidemiological data on bovine tuberculosis in a wild badger population: an investigation of approaches to statistical modelling PhD, University of Bristol.

Walker, T. M., C. L. C. Ip, R. H. Harrell, J. T. Evans, G. Kapatai, M. J. Dedicoat, D. W. Eyre, D. J. Wilson, P. M. Hawkey, D. W. Crook, J. Parkhill, D. Harris, A. S. Walker, R. Bowden, P. Monk, E. G. Smith and T. E. A. Peto (2013). "Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study." The Lancet Infectious Diseases **13**(2): 137-146.

Wang, C. and N. A. Rosenberg (2012). "MicroDrop: a program for estimating and correcting for allelic dropout in nonreplicated microsatellite genotypes version 1.01."

Warton, D. I., R. A. Duursma, D. S. Falster and S. Taskinen (2012). "smatr 3–an R package for estimation and inference about allometric lines." Methods in Ecology and Evolution **3**(2): 257-259.

Weber, N., S. Bearhop, S. X. Dall, R. Delahay, R. McDonald and S. Carter (2013). "Denning behaviour of the European badger (*Meles meles*) correlates with bovine tuberculosis infection status." Behavioral Ecology and Sociobiology **67**(3): 471-479.

Weber, N., S. P. Carter, S. R. Dall, R. J. Delahay, J. L. McDonald, S. Bearhop and R. A. McDonald (2013). "Badger social networks correlate with tuberculosis infection." Current Biology **23**(20): R915-R916.

Wedlock, D. N., M. Denis, H. M. Vordermeier, R. G. Hewinson and B. M. Buddle (2007). "Vaccination of cattle with Danish and Pasteur strains of *Mycobacterium bovis* BCG induce different levels of IFNgamma post-vaccination, but induce similar levels of protection against bovine tuberculosis." Vet Immunol Immunopathol **118**(1-2): 50-58.

Weir, B. and C. C. Cockerham (1973). "Mixed self and random mating at two loci." Genetical research **21**(03): 247-262.

Whiteman, N. K., K. D. Matson, J. L. Bollmer and P. G. Parker (2006). Disease ecology in the Galápagos Hawk (Buteo galapagoensis): host genetic diversity, parasite load and natural antibodies.

Whittaker, R. H. (1960). "Vegetation of the Siskiyou mountains, Oregon and California." Ecological monographs **30**(3): 279-338.

Williams, E. S., M. W. Miller, T. J. Kreeger, R. H. Kahn and E. T. Thorne (2002). "Chronic Wasting Disease of Deer and Elk: A Review with Recommendations for Management." The Journal of Wildlife Management **66**(3): 551-563.

Winters, J. B. and P. M. Waser (2003). "Gene dispersal and outbreeding in a philopatric mammal." Molecular Ecology **12**(8): 2251-2259.

Woodroffe, R. (1993). "Alloparental behaviour in the European badger." Animal Behaviour **46**(2): 413-415.

Woodroffe, R., C. A. Donnelly, D. R. Cox, F. J. Bourne, C. L. Cheeseman, R. J. Delahay, G. Gettinby, J. P. McInerney and W. I. Morrison (2006). "Effects of culling on badger *Meles meles* spatial organization: implications for the control of bovine tuberculosis." Journal of Applied Ecology **43**(1): 1-10.

Woodroffe, R., C. A. Donnelly, C. Ham, S. Y. B. Jackson, K. Moyes, K. Chapman, N. G. Stratton and S. J. Cartwright (2016). "Badgers prefer cattle pasture but avoid cattle: implications for bovine tuberculosis control." Ecology Letters **19**(10): 1201-1208.

Woodroffe, R., C. A. Donnelly, W. T. Johnston, F. J. Bourne, C. L. Cheeseman, R. S. Clifton-Hadley, D. R. Cox, G. Gettinby, R. G. Hewinson, A. M. Le Fevre, J. P. McInerney and W. I. Morrison (2005). "Spatial association of *Mycobacterium bovis* infection in cattle and badgers *Meles meles*." Journal of Applied Ecology **42**(5): 852-862.

Woodroffe, R., C. A. Donnelly, G. Wei, D. Cox, F. J. Bourne, T. Burke, R. K. Butlin, C. Cheeseman, G. Gettinby and P. Gilks (2009). "Social group size affects *Mycobacterium bovis* infection in European badgers (*Meles meles*)." Journal of Animal Ecology **78**(4): 818-827.

Woodroffe, R., D. Macdonald and J. Silva (1995). "Dispersal and philopatry in the European badger, *Meles meles*." Journal of Zoology **237**(2): 227-239.

Woodroffe, R., K. C. Prager, L. Munson, P. A. Conrad, E. J. Dubovi and J. A. K. Mazet (2012). "Contact with Domestic Dogs Increases Pathogen Exposure in Endangered African Wild Dogs (*Lycaon pictus*)." PloS one **7**(1): e30099.

Wright, B., C. E. Willet, R. Hamede, M. Jones, K. Belov and C. M. Wade (2017). "Variants in the host genome may inhibit tumour growth in devil facial tumours: evidence from genome-wide association." Scientific Reports **7**(1): 423.

Wylie, J. L., T. Cabral and A. M. Jolly (2005). "Identification of Networks of Sexually Transmitted Infection: A Molecular, Geographic, and Social Network Analysis." Journal of Infectious Diseases **191**(6): 899-906.

Xiao, L. H. and U. M. Ryan (2004). "Cryptosporidiosis: an update in molecular epidemiology." Current Opinion in Infectious Diseases **17**(5): 483-490.

Yang, Z. (1994). "Maximum likelihood phylogenetic estimation from DNA sequences with variable rates over sites: approximate methods." Journal of Molecular evolution **39**(3): 306-314.

Yoshimura, A., M. Koketsu, H. Bando, E. Saiki, M. Suzuki, Y. Watanabe, H. Kanuka and S. Fukumoto (2014). "Phylogenetic Comparison of Avian Haemosporidian Parasites from Resident and Migratory Birds in Northern Japan." J Wildl Dis.