

Testing the climatic controls on plague (*Yersinia pestis*) using palaeobiogeographical modelling and experimental microbiology

Henry Gillies Fell

Thesis submitted to the University of Nottingham for the degree of
Doctor of Philosophy

March 2022

Abstract

Yersinia pestis is the bacterial agent of plague, a disease which has caused three trans-continental pandemics through the Common Era. *Y. pestis* is currently present in reservoir populations in Asia, Africa, North and South America and has recently been identified as a re-emergent disease primarily due to increasing recorded cases across Africa. Plague is a zoonotic disease, the bacterial agent of which (*Y. pestis*) is transmitted and maintained primarily through rodent reservoir species and the ectoparasites which feed on them. The *Y. pestis* zoonotic disease system is therefore highly complex as the dynamics of the bacterium, vector and reservoir species can all impact transmission and dissemination.

In this thesis I focus on the niche of *Y. pestis* at two opposing spatial scales, macroecological and microbiological, and address three gaps in the literature across these scales. Firstly the degree to which reservoir (mammalian host) species locations dictate the distribution of *Y. pestis* compared to the environmental niche of *Y. pestis* per se is debated and the transferability of such niches between regions is rarely tested. I assess the niche differences between the *Y. pestis* at a continental scale across native and invaded ranges, finding that biotic factors, in this case reservoir species distribution, drive such differences.

Secondly *Y. pestis* has commonly been demonstrated to be a highly climatically dynamic disease and subsequently correlations are regularly observed between climatic variables and plague infection among human populations. The mediating factor between these elements of the plague system is often assumed to be a response of reservoir species to climate however the exact mechanisms driving this correlation are poorly resolved. I therefore assess the role that reservoir species play in mediating changes in human plague cases driven by climate perturbations, as has regularly been hypothesised. I find that the response of reservoir species to climatic conditions is highly variable and that even within a similar climatic regime the response of reservoir species to climatic variation should not be assumed to be homogeneous.

At the opposing end of the spatial scale much of the literature assumes that unlike its most recent ancestors *Y. pestis* is incapable of persisting independently of host and vector species within the environment. Selected recent works have challenged this assumption by suggesting that *Y. pestis* may be capable of surviving in soils. However, little work has experimentally tested the survival of *Y. pestis* in “real world” unsterilised soils or attempted to determine the impact of variables within the soil such as moisture and temperature on *Y. pestis* survival. Through experimental microbiology, I test the niche of *Y. pestis* within soil environments independent of reservoir and vector species. Survival of *Y. pestis* within soils is an often suggested mechanism to explain prolonged periods of quiescence but has yet to be fully integrated into models of the complex plague transmission and maintenance system. My work suggests that soil moisture is a key variable in enabling the persistence of *Y. pestis* in soil environments.

Integrating findings from such disparate scales into a cohesive model of a complex zoonotic disease system is a significant challenge which my work only begins to address. This work is however necessary to avoid the erroneous transfer of assumptions between scales and contexts, which is particularly relevant for plague given the global breadth of distribution in varying environments. My work will hopefully aid in contributing to a more holistic multiscale view of the *Y. pestis* disease system.

Acknowledgements

Firstly, I would like to thank my supervisors Dr Adam Algar, Professor Matt Jones and Dr Steve Atkinson. All of whom have been incredibly patient with me and guided me through the learning process from the very start. This PhD represented my introduction to multiple new disciplines, and I am incredibly grateful that my supervisors gave their time and effort to help me get to grips with these new disciplines.

The ecology lab group in all of its various iterations has been a very supportive group throughout my PhD and feedback from all of the various members has helped shaped this PhD and the way I think about ecology. Dr Simon Tarr provided me with a huge amount of help getting to grips with R and ENMs which represents a large proportion of my work, so thank you Simon for answering all my stupid questions. Similarly working with Dr Owen Osborne has improved my understanding of null modelling and his help was invaluable in Chapter 3. Without the training, advice, and support of Dr Vanina Garcia I would not have made it into the CL3 lab so all the microbiological work in this thesis is a product of the help Vanina provided.

The School of Geography community has been a constant source of reassurance support and joy and I thank everyone who has made me smile in the past 4 years. Thanks to my parents for telling me to follow what I enjoy, still the best advice I have ever received. Finally thank you Helen.

Table of Contents

1	Introduction & literature review.....	10
1.1	Thesis rationale	10
1.2	Plague in a modern and historical context	12
1.3	Plague Evolution and Epidemiology.....	14
1.4	Plague Ecology	17
1.4.1	Climatic constraints on <i>Yersinia pestis</i>	17
1.4.1.1	Bacterial climatic tolerance.....	21
1.4.1.2	Vector species climatic tolerance	21
1.4.1.3	Reservoir species climatic tolerance.....	22
1.4.2	System-specific plague ecology	23
1.4.3	Climatic drivers of plague ecology	23
1.4.4	Environmental niche modelling of <i>Y. pestis</i>	27
2	Biotic factors limit the invasion of plague's pathogen (<i>Yersinia pestis</i>) in novel geographical settings	30
2.1	Introduction	30
2.2	Materials and methods	33
2.2.1	Data acquisition	33
2.2.1.1	<i>Yersinia pestis</i> localities.....	33
2.2.1.2	Host species localities.	33
2.2.1.3	Climate data	34
2.2.2	Niche modelling	35
2.2.3	Hypothesis testing.....	36
2.2.3.1	Native and invaded niche similarity.....	36
2.2.3.2	Biotic influences on the invaded niche	38
2.2.3.3	Geographical null modelling	39
2.3	Results	39
2.3.1	Native and invaded niche similarity.....	39
2.3.2	Biotic influences on the invaded niche	42
2.3.3	Geographical null modelling	44
2.4	Discussion.....	45
2.5	Chapter 3 Appendix.....	51
2.5.1	Supplementary tables	51
2.5.2	Supplementary figures	54
2.5.3	Details of null modelling algorithm.....	70

3	The role of reservoir species in mediating plague's dynamic response to climate.....	72
3.1	Introduction	72
3.2	Materials and methods	77
3.2.1	Data acquisition	77
3.2.1.1	Host species localities	77
3.2.1.2	Historical climate data	78
3.2.2	Niche modelling	80
3.2.2.1	Variable selection.....	80
3.2.2.2	Background selection	81
3.2.2.3	Model tuning.....	82
3.2.2.4	Hindcasting across multiple climate scenarios	84
3.2.3	Hypothesis testing.....	84
3.2.3.1	Response curves.....	85
3.2.3.2	Niche time series calculation	86
3.2.3.3	Correlation analysis.....	86
3.2.3.4	Impact of wet and dry periods.....	87
3.3	Results	88
3.3.1	Niche modelling	88
3.3.1.1	Model performance	88
3.3.2	Hypothesis testing.....	90
3.3.2.1	Response curves.....	90
3.3.2.2	Niche time series correlation	93
3.3.2.3	Wet and dry period niche analysis.....	96
3.4	Discussion.....	99
3.4.1	Insights from modelling approach	99
3.4.2	Correlation analysis.....	100
3.4.3	Wet and dry period analysis	101
3.4.4	Methodological considerations	102
3.4.5	Limitation of modelling approach.....	104
3.5	Conclusion	105
3.6	Chapter 4 Appendix.....	106
3.6.1	Supplementary tables	106
3.6.2	Supplementary figures.....	113
3.6.3	Attempted multi-temporal calibration	118
4	Testing the niche of <i>Y. pestis</i> in soil through experimental microbiology.....	120

4.1	Introduction	120
4.2	Material and methods.....	125
4.2.1	Bacterial strains used in this study.....	125
4.2.2	Growth media and growth conditions	125
4.2.3	<i>C. elegans</i> biofilm infection assays.	127
4.2.4	Long-term maintenance and recovery of <i>Y. pseudotuberculosis</i> from an isolated sterile environment.	127
4.2.5	Soil sterilisation and validation	128
4.2.6	CL3 general safety principles	129
4.2.7	<i>Y. pestis</i> soil assay	129
4.3	Results	131
4.3.1	<i>C. elegans</i> biofilm assay	131
4.3.2	Long term maintenance and recovery of <i>Y. pseudotuberculosis</i> from an isolated sterile environment.	132
4.3.3	Soil sterilisation and validation	135
4.3.4	<i>Y. pestis</i> soil assay	137
4.3.4.1	Experiment 1 – Preliminary methods work-up.....	137
4.3.4.2	Experiment 2 – Month-long run	138
4.3.4.3	Experiment 3 – Month-long run with rehydration	139
4.4	Discussion.....	142
4.4.1	Discussion of experimental findings	142
4.4.1.1	Work-up experiments	142
4.4.1.2	<i>Y. pestis</i> soil assay	143
4.4.2	Discussion of methodology	148
4.4.3	Conclusions and future directions	150
5	Future directions and synthesis	151
5.1	Combined conclusions	151
5.2	Chapter-specific conclusions.....	153
5.2.1	Chapter 2 -Biotic factors limit the invasion of plague’s pathogen (<i>Yersinia pestis</i>) in novel geographical settings.	153
5.2.2	Chapter 3 - The role of reservoir species in mediating plague’s dynamic response to climate.	154
5.2.3	Chapter 4 - Testing the niche of <i>Y. pestis</i> in soil through experimental microbiology.....	156
5.2.4	Future directions and conclusion - Synthesised microbiological and macroecological studies.....	157
6	Bibliography	160

List of Figures

Figure 2.1 Comparison of <i>Yersinia pestis</i> ' native and invaded niche.....	41
Figure 2.2. Differing projected distribution of <i>Y. pestis</i> across North America	43
Figure 2.3. Niche overlap values of the observed distribution models.	44
Figure 2.4. Null and observed overlap (Schoener's D) between <i>Y. pestis</i> ' North American niche and the predicted niche	45
Figure 2.5 Projected distribution of all species used in the Multi-host models	54
Figure 2.6 Points removed based on previous literature from the species occurrence points.....	55
Figure 2.7 Niche overlap (Schoener's D) values between the projected Asian mod (Asia-proj) inclusive and exclusive of biotic factors and the North American model (N.Am-obs).....	56
Figure 2.8 Identity test results	57
Figure 2.9 Background test results	58
Figure 2.10 Multivariate Environmental Similarity surface plot.	59
Figure 2.11 Niche overlap values of the observed distribution models using the most selected optimization settings across all models	60
Figure 2.12 Null and observed overlap (Schoener's D) for models with the consistent LQ 0.5 settings.....	61
Figure 2.13 AUC values for the empirical and null models.	62
Figure 2.14 Univariate response curves from the Asia + Multi-host model.....	63
Figure 2.15 Null and observed overlap (Schoener's D) between <i>Y. pestis</i> ' North American niche and the further models with negatively correlated species removed	64
Figure 2.16 Differing projected distribution of <i>Y. pestis</i> across North America	65
Figure 2.17 Niche overlap values of all the observed distribution models.	66
Figure 2.18 Figure 3.1 replicated using all threshold methods.	67
Figure 2.19 Comparison of between-species KL divergence values	68
Figure 2.20 Four exemplar geographic null occurrences	69
Figure 3.1. Comparison of mean annual precipitation and temperature values between the two climate scenarios (GFDL & IPSL) between the two regions (North and South China).	79
Figure 3.2. Bioclimatic variables selected for use in ENM of all reservoir species.....	81
Figure 3.3. Varying feature class and regularisation multiplier across repeated optimizations	83
Figure 3.4 Distribution of human plague cases across China during the Third Pandemic	85
Figure 3.5. Identification of wet and dry periods across both climate scenarios.	88
Figure 3.6. Permutation importance and response curves for select species.	90
Figure 3.7 Variable selection through correlograms and cluster dendrograms.	113
Figure 3.8 Niche suitability maps across selected species projected into the time averaged period (1970-2000 C.E.) for each climate and clamping scenario (noted in each plot title).	114
Figure 4.1 Variation in humidity through time in containment box.....	134
Figure 4.2. Comparison of sterility check methods in LB plates.....	137
Figure 4.3. <i>Y. pestis</i> recovery experiment 2- Month-long run.	139
Figure 4.4. <i>Y. pestis</i> recovery Experiment 3 - with rehydration	141
Figure 4.5. All <i>Y. pestis</i> soil survival treatment comparison.	142

List of Tables

Table 2.1. Mechanisms through which climate can impact <i>Y. pestis</i> survival distribution and transmission, organized by mode of transmission and mechanism.	30
Table 2.1. Naming protocol and performance for models used to predict <i>Y. pestis</i> ' North American niche. Predictors included in each model are given below, along with the accuracy of each model, calculated as the area under the curve (AUC) of the receiver operating.....	37
Table 2.2 Feature Class and regularisation multiplier of each host species model.	51
Table 2.3 Feature Class and regularisation multiplier of each model used for overlap calculations	52
Table 2.4 Naming protocol and performance for supplementary models used to predict <i>Y. pestis</i> ' North American niche.	53
Table 3.1. Minimum niche overlap between each species with multiple best models selected through optimization	82
Table 3.2. ENM performances across differing extents.....	88
Table 3.3. Correlation of reservoir species niche suitability values through time (summed) and annual precipitation values.....	94
Table 4.5. Mean standardised niche suitability of reservoir species through wet and dry periods.	98
Table 3.5 Chinese reservoir species were selected from Mahmoudi et al., (2021) and Cui et al., (2013).	106
Table 3.6 Correlation of reservoir species niche suitability values through time (BMCP summed) and annual precipitation values.	109
Table 3.7 Correlation of reservoir species niche suitability values through time (summed) and lagged (1 year) annual precipitation values.	110
Table 3.8 Correlation of reservoir species niche suitability values through time (BMCP summed) and lagged (1 year) annual precipitation values.	111
Table 3.9 Mean AUC difference for all species.	112
Table 4.1 Full list of bacterial strains used throughout the experimental study.....	125
Table 4.2. Rate of moisture loss from soils in falcon tubes at two different incubation temperatures	131
Table 4.3 Impact of solution used to recover <i>Y. pseudotuberculosis</i> from cover slips.....	133

1 Introduction & literature review

1.1 Thesis rationale

Plague is a zoonotic disease caused by the gram-negative bacterial pathogen *Yersinia pestis*. It has driven three pandemics throughout the Common Era which caused millions of deaths and led to periods of societal reorganisation (Stenseth et al., 2008). Plague, however, is not only a disease of historical importance but is currently regarded as a re-emergent disease due to increasing human rates of infection, primarily across Africa (World Health Organization, 2015). Plague is a climatically dynamic disease capable of rapidly responding to climatic change across the broad range of scales over which it functions and is also sensitive to anthropogenic land use changes (Stenseth et al., 2006, McCauley et al., 2015, Cui et al., 2020). It is therefore vital to gain an understanding of how all elements of the plague disease system are affected by their environments, to mitigate and plan for future risk.

Humans are not the primary reservoir of *Y. pestis* but are instead infected during “spill over” events when infections within reservoir species, usually rodents, reach a threshold level past which the likelihood of human infection increases (Samia et al., 2011, Mahmoudi et al., 2020). Within reservoir populations, *Y. pestis* is transmitted between individual hosts by vector species, usually fleas, which transmit the bacterium through regurgitation of infected blood meals during feeding (Jarrett et al., 2004). Following transmission through global shipping routes during the Third Pandemic (1772 C.E – 1964 C.E.), permanent plague reservoirs were established across North America, South America, Asia, and Africa (Stenseth et al., 2008, Cliff et al., 2009, Xu et al., 2011). *Y. pestis* is now therefore distributed in a broad range of reservoir and vector species across a range of environments (Mahmoudi et al., 2020). Growing evidence also suggests that *Y. pestis* may be able to persist in the soils which may explain periods of quiescence (where *Y. pestis* is not sampled from reservoir species or vectors) observed in plague systems (Bertherat et al., 2007, Eisen et al., 2008, Andrianaivoarimanana et al., 2013, Markman et al., 2018).

Environmental factors can impact reservoir and vector species from a range of differing habitats in a broad and varied manner. Monitoring of plague in the wild is spatially heterogeneous and only a limited subset of reservoir species have been monitored consistently through time (Davis et al., 2004, Bevins et al., 2012). Many studies focus on bacteria-vector-reservoir species dynamics within a single region, which contributes significantly to the understanding of plague within that system but may have limited applicability in other regions (Kausrud et al., 2007, Savage et al., 2011, Xu et al., 2015). This can lead to assumptions being made across a range of reservoir species by using findings for one well-studied species to predict the response of reservoir species more generally and further *Y. pestis* infection in human populations (Kausrud et al., 2010, Xu et al., 2011). Studies often identify correlations between climate and human plague records and infer host and vector reservoir mechanisms which could mediate this correlation, however, these mechanisms may not be applicable across all scales of study (Ben-Ari et al., 2010, Ben-Ari et al., 2011, Schmid et al., 2015). Understanding plague dynamics at broader scales, across multiple regions, will require the integration of multiple reservoir species, with partially or non-overlapping distributions, into the plague system.

The survival of *Y. pestis* in soils represents a further knowledge gap. Until recently, *Y. pestis* has not been considered capable of surviving within the environment independently of reservoir species and vector species (Hinnebusch et al., 2016), suggesting that environment-plague relationships must be mediated by reservoir species and vectors. However, a growing number of studies are suggesting that *Y. pestis* may be capable of surviving within soil environments and potentially utilising soil micro biota as vectors (Ayyadurai et al., 2008, Eisen et al., 2008, Boegler et al., 2012, Markman et al., 2018), raising new possibilities for how environment influences plague outbreaks and dynamics.

In the thesis, I investigate the role that multiple reservoir species may play in mediating the effects of climate upon plague infection and further apply niche theory

to survival of *Y. pestis* in soil environments, by experimentally testing the fundamental environmental niche of *Y. pestis* in this environment. In Chapter 2, I compare the environmental niche of *Y. pestis* between native and invaded regions, and investigate the impact that biotic factors may have upon disease distribution following introduction to novel environments. In Chapter 3 I test the hypothesis that reservoir species niche dynamics mediated the heterogeneous response in human plague cases to precipitation in China during the Third Pandemic. In Chapter 4 I test the hypothesis that *Y. pestis* can survive in a viable state in soil environments for prolonged periods. Each chapter represents a discrete area of investigation and is supplemented with the necessary background for each study. A broad literature review (Chapter **Error! Reference source not found.**) provides an overview of previous research on plague ecology across the disparate scales investigated in this thesis.

1.2 Plague in a modern and historical context

Y. pestis is capable of persisting in at least 351 mammalian species globally, predominantly within arid to semi-arid regions (Mahmoudi et al., 2020). Of these mammalian species rodents are by far the most important to the maintenance and transmission of *Y. pestis* (Mahmoudi et al., 2020). *Y. pestis* is maintained in these species through fluctuating enzootic/epizootic cycles leading to varying levels of infection in reservoir species. I have used the term reservoir species to describe any non-insect species which may host *Y. pestis* through infection and hence represent a natural reservoir for the bacteria. I do not suggest that all species are permanent reservoirs, but are capable of hosting *Y. pestis*, even if only for brief periods during epizootics. Within reservoir species, *Y. pestis* is transmitted through exchange of infected blood meals by insect vector species, most commonly fleas. Humans are also susceptible to *Y. pestis* infection and plague is responsible for at least three pandemics within the Common Era, as well as regular epidemic outbreaks in geographic proximity to active host species populations (Stenseth et al., 2008, Vallès et al., 2020).

The historic impact of plague on Eurasian societies is incredibly broad. For example, the Second Pandemic, which includes the Black Death Period (1347–1351 C.E.), devastated Eurasian populations from the fourteenth century onwards and persisted in isolated areas into the nineteenth century (Bramanti et al., 2016, Green, 2020). The impact of plague is not purely historical, with the World Health Organization (WHO) classifying plague as a re-emergent disease due to increasing reported human infection, largely across Africa, through the twentieth and early twenty first centuries, coupled with persistence within host populations globally (Stenseth et al., 2008, Mahmoudi et al., 2020). Madagascar, a country with an annual reported plague infection that accounts for three quarters of the cases reported to the WHO annually, suffered a primarily pneumonic plague epidemic in 2017, which led to 2,414 clinically suspected cases from two urban foci around the cities of Antananarivo and Toamasina (Randremanana et al., 2019). This was a large and unexpected outbreak as most of the cases reported annually are bubonic and generally isolated to the rural highlands where plague is endemic within mammalian host populations (Randremanana et al., 2019). This highlights the immediate contemporary threat posed by plague and the need for further research into both the distribution and mechanisms of transmission to enable better prediction and risk mitigation.

Although Madagascar currently has the highest annual human plague cases of any country with 200-700 cases annually (Randremanana et al., 2019), many countries and regions maintain permanent enzootic plague reservoirs within wild mammal populations. From 2010-2015 there were 3,248 human cases reported worldwide to the WHO (World Health Organization, 2017), the majority of which were in Africa (Stenseth et al., 2008). Away from Africa, there have been limited human outbreaks reported across several South America countries since 2000 in Peru, Bolivia, Ecuador and Brazil (Rivière-Cinnamond et al., 2018). Enzootic host populations are also present in the USA (California and Colorado), where isolated human cases have been reported, as well as across much of the Asian Steppe environment (China, Kyrgyzstan, Kazakhstan, Russia, and Mongolia) where plague has been maintained through enzootic and epizootic cycles for thousands of years (Vallès et al., 2020).

1.3 Plague Evolution and Epidemiology

Y. pestis is the most infamous member of the *Yersinia* genus. *Yersinia* is comprised of fifteen species (Reuter et al., 2014) of rod-shaped, Gram-negative bacteria, which are collectively members of the Enterobacteriaceae family. Three species are pathogenic to humans, namely *Yersinia enterocolitica*, *Yersinia pseudotuberculosis* and *Y. pestis* (Wren, 2003). *Y. enterocolitica* is an enteropathogenic (gastrointestinal tract) bacteria found within soil environments and commonly infects a range of mammalian and avian species with particularly high prevalence within domestic pig populations (Fredriksson-Ahomaa et al., 2009). In humans, large outbreaks are driven by the consumption of infected foods (Bottone, 1999). *Y. pseudotuberculosis* is, similarly, a gastrointestinal tract pathogen also found widely within the environment (soil and water) where it can survive for long periods through adaptation mechanisms enabling survival in the soil under biotic and abiotic environmental conditions significantly different from those encountered during its within-host life cycle (Buzoleva and Somov, 2003). *Y. pseudotuberculosis* primarily causes gastrointestinal infection following ingestion into the digestive system through infected food stuffs.

The transmission pathway of *Y. pestis* is significantly removed from the other *Yersinia* species. It is primarily transmitted through the regurgitation of infected blood meals from vector species (commonly fleas) into reservoir species while attempting to feed (Jarrett et al., 2004). Rodent species show varying resistance and susceptibility to infection which in combination with the effective dose that vector species can provide determines the dynamics of *Y. pestis* transmission and infection within host-vector-bacteria systems (Mahmoudi et al., 2020). Human infection can occur when infected reservoir and/or vector species come into close geographical proximity with human populations and share infected vectors (Stenseth et al., 2008).

Infection through transmission of *Y. pestis* via a blood meal from a vector most commonly leads to bubonic infection in humans. This is where the bacterial load migrates through the cutaneous lymphatics to the lymph nodes and subsequent rapid multiplication of *Y. pestis* causes necrosis of the lymph nodes leading to the formation of the tular bubo. Along with the formation of buboes the symptoms include sudden

onset of fever, chills, and weakness. These symptoms typically take 2-8 days to develop following initial infection (Inglesby et al., 2000). If untreated this can develop into septicemic plague which is an infection of the blood, and in rare cases this can occur without infection of the lymphatic system following a flea bite (primary septicemic plague) (Roussos, 2002, Olson and Anderson, 2019). The final form of *Y. pestis* infection is pneumonic plague, which is an infection of the lungs. This can develop like septicemic plague as a secondary infection following initial bubonic infection and then can be rapidly transmitted through the production and transmission of aerosolised droplets containing *Y. pestis*. The symptoms of pneumonic plague are chest pain, dyspnea, cough, and hemoptysis, and this form of *Y. pestis* infection is the most deadly with a mortality rate of 90% when untreated compared to 40-60% for bubonic infection (Butler, 2013, Pechous et al., 2016).

Despite *Y. pseudotuberculosis* and *Y. enterocolitica* sharing similar infection pathways and areas of environmental survival, *Y. pestis* and *Y. pseudotuberculosis* are in fact much more closely related with approximately 97 % DNA sequence homology compared to approximately 50 % between *Y. pseudotuberculosis* and *Y. enterocolitica* (Wren, 2003, Fredriksson-Ahomaa et al., 2009). It therefore follows that *Y. pestis* and *Y. pseudotuberculosis* diverged much later than *Y. enterocolitica*, with *Y. enterocolitica* estimated to have diverged within the last 200 million years whereas the former two diverged between 5700-6000 years ago (Spyrou et al., 2018, Demeure et al., 2019). As the infection pathways and clinical symptoms suggest there are several key genetic differences between *Y. pestis* and *Y. pseudotuberculosis* which aid in understanding the functional differences between the two and their differing niche preference and lifestyle.

Both species are capable of constructing biofilms which in this context are bacterial colonies suspended within an extracellular matrix, synthesised by the *Yersinia* species (Hinnebusch and Erickson, 2008). The extracellular matrix which supports the bacteria within the biofilm is constructed of β -1,6-N-acetyl-D-glucosamine (GlcNAc). For *Y. pestis*, biofilm is key to the vector transmission pathway as the formation within vector species can lead to blockages of the vector midgut effectively leading

to starvation preceded by frantic repeated feeding and hence transmission to reservoir species (Hinnebusch et al., 1996). In *Y. pseudotuberculosis* the formation of biofilms has been demonstrated in the laboratory upon the nematode species *Caenorhabditis elegans*, this is strain dependent (Atkinson et al., 2011) and has been hypothesised to aid with environmental survival (Sun et al., 2009). This has yet to be reported in its natural environment. Nematodes are the most abundant animals on earth and found commonly throughout soil environments therefore could potentially play a key role in maintenance of *Y. pestis* in soil environments as could a range of soil dwelling amebae species (Benavides-Montaña & Vadyvaloo 2017, Markmen et al 2018).

All pathogenic *Yersinia* species possess an extrachromosomal 70-kb virulence plasmid which encodes all components of the Type three secretion system (T3SS). The T3SS is a bacterial nanomachine analogous to a hypodermic needle which is composed of structural elements, regulators and effectors known as *Yersinia* outer proteins (*Yops*). The T3SS is responsible for the delivery of *Yops* into the host cells through its needle like delivery structure which is made up of *Yersinia* secretion complex (*Ysc*) proteins. Once inside mammalian host cells the *Yops* inhibit the phagocytic response of these cells while also downregulating the inflammatory responses of macrophages and hence enable the bacteria to persist in an extracellular state within the host (Fahlgren et al., 2014).

The functional differences between *Y. pestis* and *Y. pseudotuberculosis* are due to gene gain and gene loss events, which caused the transition from infection of host species from the environment to infection *via* an infected blood meal from a vector species (commonly, however, not solely flea species) (Sun et al., 2014). Using ancient DNA (aDNA) and high-throughput DNA sequencing technologies the timing of these gene loss and gain events can be estimated. Divergence began between 5700 and 6000 years ago with the acquisition of two virulence-associated plasmids pFra/pMT1, which enables effective vector transmission and pPla/ pPCP1, which improves survival in reservoir species (Hinnebusch et al., 2016, Demeure et al., 2019). These acquisitions were further accompanied by the inactivation of the virulence-

associated genes phosphodiesterases (*pde*) 2 and 3, and *rcaA* which respectively enabled an increase in biofilm formation (Sun et al., 2014). The pFra/pMT1 plasmid encodes the *Yersinia* murine toxin *ymt*, a phospholipase D that protects *Y. pestis* from cytotoxic blood plasma digestion products inside the flea gut (Hinnebusch et al., 1996). pPst/ pPCP1 encodes the plasminogen-activating protease (*Pla*) gene which facilitates pulmonary infection in the host leading to haemorrhage and inhibits T-cell defence (Lathem et al., 2007, Smiley et al., 2019, Wang et al., 2020). Prior to these gene acquisition and loss events *Y. pseudotuberculosis* infection of a vector would predominantly lead to mortality and in the rare case of survival *Y. pseudotuberculosis* was restricted to the hindgut and therefore not transmissible through further blood meals (Hinnebusch et al., 2016). Following these genetic and hence functional changes *Y. pestis* could grow in higher numbers in the flea mid gut, forming a biofilm blockage and therefore transmitting high titres of *Y. pestis* into the host as the flea takes a blood meal.

1.4 Plague Ecology

1.4.1 Climatic constraints on *Yersinia pestis*

Both the distribution and transmission of *Y. pestis* are reliant on a broad range of biotic and abiotic factors (). The third and most recent plague pandemic (Third Pandemic) began in 1772 C.E. in Yunnan province China, following gradual initial transmission *Y. pestis* spread to the South China coast and reached Hong Kong by the start of the twentieth century, from which point plague was transmitted globally through international shipping routes (Cliff et al., 2009). Following the global dissemination of *Y. pestis* during the Third Pandemic *Y. pestis* established global reservoirs within the 20°C summer isotherms (Cliff et al., 2009). Plague reservoirs globally are more prevalent at high elevations, however, the elevation threshold varies between regions (Eads & Biggins, 2021; Eisen et al., 2012; Giorgi et al., 2016; Qian et al., 2014). *Y. pestis* distributions are driven by complex and varying systems in which *Y. pestis* is maintained and transmitted across a range of scales.

Table 1.1. Mechanisms through which climate can impact *Y. pestis* survival distribution and transmission, organized by mode of transmission and mechanism.

Mode of transmission	Mechanism	Example	Evidence of mechanism	Additional evidence required	Sources
Airborne (Pneumonic)	Bacterial survival	Aerosolized virulent <i>Y. pestis</i> may remain active for up to 5 days on certain surfaces, however, this is dependent on the solution <i>Y. pestis</i> is maintained in.	Experimental survival of <i>Y. pestis</i> suspended in PBS and BHI upon a range of surfaces.	Experimentally determine the impact temperature and humidity may have upon survival on surfaces	(Rose et al., 2003)
	Human defences	Malnourishment has a negative influence on the immunity of an individual, with poverty driven malnutrition discussed with particular reference to the 1910-1911 pneumonic epidemic in Manchuria (Gamsa, 2006) and recent prevalence in Madagascar (Andrianaivoarimanana et al., 2013). Initial outbreaks during the Black Death and the Third Pandemic also occurred within highly stressed populations likely to be suffering from food shortages. These food shortages may in part be driven by climatic perturbations at the time (Fell et al., 2020).	Primarily correlative however, mechanistic links between malnutrition and weakened immunity is well established.	Case studies highlighting the association of plague outbreaks with periods of climatic perturbation and subsequent potential food shortages	(Gamsa, 2006, Andrianaivoarimanana et al., 2013)
Telluric (Soil dwelling)	Bacterial prevalence in soil	<i>Y. pestis</i> biotype Orientalis can remain viable and fully virulent after 40 weeks in soil (Ayyadurai et al., 2008), however, transmission rate has experimentally been shown to be low (<1%) (Boegler et al., 2012).	Experimental long-term persistence of <i>Y. pestis</i> in sterilised soils. Experimental infection of scarified mice on nutrient rich soil inoculated with <i>Y. pestis</i> .	Environmental influence on <i>Y. pestis</i> in soil is unknown, determining climatic conditions conducive to survival of bacterium in soil medium may aid in explaining quiescent periods.	(Ayyadurai et al., 2008, Boegler et al., 2012)

Vector-borne (Bubonic)	Vector development	Seasonal temperatures influence the timing of vector development. For example <i>X. cheopis</i> has a developmental threshold of 12.36°C (Kreppel et al., 2016). The timing of these threshold temperatures influences annual development of vectors which further impacts <i>Y. pestis</i> .	Experimental investigation of development time under varying temperature and relative humidity.	Similar analysis determining developmental temperature and humidity thresholds for a range of vector species would provide invaluable data for mechanistic modelling of <i>Y. pestis</i> .	(Kreppel et al., 2016)
	Vector geographic distribution	The environmental niche as defined by climatic variables can be constructed through correlative analysis and projected into novel temporal and spatial regions to estimate past or future distributions	Species distribution models can be constructed using the occurrence of vector species and explanatory climate variables. This has been completed for 13 vector species across California, however, there is limited occurrence data available (Adjemian et al., 2006).	Further occurrence data is required across all vector species, presence of, <i>Y. pestis</i> in association with each vector would also be hugely beneficial.	(Adjemian et al., 2006)
	Bacterial transmission efficiency	<i>Y. pestis</i> has been found to be most efficiently transmitted via <i>X. cheopis</i> at temperatures of 23-30°C, with slight advantage at 23°C.	Transmission efficacy experimentally determined through exposing mice to infected <i>X. cheopis</i> at varying temperatures. (Schotthoefer et al., 2011)	Expansion to determining efficiency in differing host-vector systems	(Schotthoefer et al., 2011)

Risk of Host Human Contact	Low frequency extreme weather events may contribute to increased contact between commensal rodents and humans, as identified through a correlation in plague spread velocity and wet conditions in China during the Third Pandemic.	Analysis of plague occurrence data from the Third pandemic in China with climatic data	Inclusion of extreme weather events in larger temporal and spatial models (discussed in Metcalf et al 2017 with regard to disease modelling)	(Xu et al., 2014, Metcalf et al., 2017)
Host Species geographic distribution.	The environmental niche of host and vector species will likely track climatic changes, this will lead to changes in the geographic range of <i>Y. pestis</i> .	Environmental niche modelling of the plague reservoir from surveillance of California ground squirrel plague hosts (Holt et al., 2009).	Expansion of environmental niche modelling to further plague host and vector species. This correlative approach, however, lacks/assumes the mechanisms of transmission.	(Holt et al., 2009)
Population dynamics of reservoir species	Population dynamics of host species proximal to Tian Shan are suggested to have contributed to repeated reintroduction of plague during the Black Death Pandemic.	Correlation between European plague reintroduction events determined from 7711 georeferenced plague occurrences and climate-driven reintroductions from Asian source regions.	Further analysis of the suggested mechanisms contributing to the reintroduction events proposed by Schmid et al. (2015)	(Schmid et al., 2015)

1.4.1.1 Bacterial climatic tolerance

Y. pestis is a gram negative, non-motile, rod-shaped bacterium and, depending on the environment, it is either aerobic, facultatively anaerobic or a facultatively intracellular pathogen. Fully virulent *Y. pestis* has been shown to survive on sterile surfaces suspended in phosphate-buffered saline (PBS) solution, which matches the ion concentration of the human body, for up to 48 hours and in brain heart infusion (BHI), a nutritionally dense organic solution, for up to 120 hours (Rose et al., 2003). Limited experimental and observational evidence also suggests that *Y. pestis* can survive in a viable community in both sterile and non-sterile soils for extended periods (40 weeks and 3 weeks respectively) (Ayyadurai et al., 2008, Eisen et al., 2008). Such findings are counter to the previously prevailing hypothesis that, following *Y. pestis*' divergence from its most recent common ancestor (MRCA), the soil dwelling enterobacteriaceae *Yersinia pseudotuberculosis*, its mechanism for environmental survival became redundant given the new life cycle within hosts and vectors and hence *Y. pestis* is no longer capable of survival within the environment (Sun et al., 2014, Hinnebusch et al., 2016).

1.4.1.2 Vector species climatic tolerance

The temperature tolerance and hence thermal niche of *Y. pestis* is further determined by vector species tolerances as vectors are required to maintain and transmit *Y. pestis* within reservoir species populations. Vector species are insulated to a degree by their mammalian host, but they must remain viable in the external environment to reproduce and hence persist in a region. For example, *Xenopsylla cheopis* (Oriental rat flea), is one of the most common vectors of *Y. pestis*. This species has a lower temperature threshold of 12.4 °C, below which larval and pupal development are inhibited (Kreppel et al., 2016). Further the transmission efficiency of *Y. pestis* from vector species is influenced by temperature and hence dictates distribution. Experimental testing has determined that *X. cheopis* has a maximum transmission efficiency occurring at 23°C with no statistical decrease in efficiency up to 30°C (Schotthoefer et al., 2011). *X. cheopis* is one of an array of vectors capable of transmitting *Y. pestis* with multiple *Xenopsylla* species along with *Ctenocephalides*

felis (Cat Flea) and multiple *Synopsyllus* species, all of which have varying climate tolerances, transmission efficiency and preferred hosts (Krasnov et al., 2006). Determining all these vectors' thermal niches is a substantial task. Even using correlative methods (Environmental niche models, ENMs) to determine the realised climatic niche of these vector species is challenging as there are very limited spatial data on the occurrences of these species (Adjemian et al., 2006).

1.4.1.3 Reservoir species climatic tolerance

There is a similarly broad range of reservoir species involved in the maintenance of *Y. pestis* globally. The majority of *Y. pestis* reservoir species are rodents, however, a wide taxonomic range of mammalian species are susceptible to infection, including predators of infected rodent species, and domesticated animals (e.g. camels and goats) (Christie et al., 1980, Holt et al., 2009). *Y. pestis* infection and hence presence is predominantly maintained by key rodent species and the role of non-rodent species (predators and domestic animals) in maintaining this presence requires further research (Holt et al., 2009, Mahmoudi et al., 2020). *Y. pestis* persists in three states within a reservoir population: quiescent (undetectable), enzoonotic (low levels of detection, <2% of rodents captured infected), or epizootic (epidemic among rodents in the region) (Davis et al., 2004, Xu et al., 2015, Mahmoudi et al., 2020). During the first two states, the disease is isolated to only one key species, or may be persisting in the environment, whereas in the third stage multiple host species are infected (Mahmoudi et al., 2020). Investigations regularly focus on a single host species as the dynamics of this primary host and associated vectors are often key to the transition between each of the infection phases, however, interactions with further species are likely common and seldom completely understood (Schmid et al., 2012, Andrianaivoarimanana et al., 2013). Both host and vector densities are fundamental in predicting plague prevalence, with climatic factors often positively influencing both (Davis et al., 2004, Stenseth et al., 2006). Regional studies further demonstrate geomorphological features such as soil type and topography influencing direction of population expansion and hence plague movement (Wilschut et al., 2013). Climate impacts each component of *Y. pestis*' life cycle at varying temporal and spatial scales, with each vector and host providing the potential for varying

responses to climate. For this reason there are many possible mechanisms through which climate can influence plague () and care should be taken to not extrapolate expected responses to climate between different scales (Ben-Ari et al., 2011).

1.4.2 System-specific plague ecology

When investigating plague ecology in differing locations and at differing scales it is incredibly important to note that mechanisms linking climate to *Y. pestis* presence and abundance in vectors, hosts or humans are specific to the system under investigation. Commonalities should not be assumed but interrogated and projections of expected mechanisms from one system or scale to another challenged. For this reason, macroecological studies of plague are rarer than those focused on a single system and can be criticised for oversimplifying the system under investigation as they often extrapolate findings focused on a single species across a spatial extent far beyond the range of that species (Kausrud et al., 2010, Schmid et al., 2015). Here, I summarise research of specific systems in which climate has been observed to influence plague ecology and the differences and commonalities that can be drawn from them.

1.4.3 Climatic drivers of plague ecology

One of most regularly studied geographical areas in plague ecology is East Kazakhstan due to the availability of high temporal resolution reservoir and vector species data recorded from 1949 C.E. to 1994 C.E (Davis et al., 2004, Heier et al., 2015). This region to the southeast of Lake Balkhash (henceforth Pre-Balkhash) is inhabited by the plague reservoir species *Rhombomys opimus* (great gerbil). *Y. pestis* persists in *R. opimus* populations through vector transmission of fleas (mostly in the genus *Xenopsylla*) and an estimate of the host and vector abundance has been recorded throughout the sampling period. Further to the host and vector monitoring, bacteriological tests (1955-1970) and serological tests (post-1970) were used to identify *Y. pestis* antibodies in the host and vector populations (Davis et al., 2004, Davis et al., 2007). The data for this region show *Y. pestis* progressing through each state of infection with periods of enzootic and epizootic infection separated by periods of quiescence. One of the first studies using this data identified that epizootic

periods of transmission were triggered when host populations surpassed a certain threshold, with either a one or two year lag between the host population and epizootic transmission (Davis et al., 2004). Studies that followed identified key climatic variables, specifically spring temperatures and summer precipitation, as influencing gerbil and flea densities on an annual scale, suggesting that climate was the driver of these epizootic periods (Stenseth et al., 2006). There are two mechanisms by which climate is hypothesized to influence plague within this system which are not mutually exclusive. The first mechanism is a hypothesized bottom-up trophic cascade in which increased precipitation and/or warming during spring and summer contributes to high *R. opimus* densities and subsequent epizootic periods across the region due to increased primary productivity within their arid environment (Kausrud et al., 2007). The second hypothesized mechanism relates to the population dynamics of the vector population, which is predominantly influenced by temperature. Spring warmth (early seasonal warming) has a positive impact on the flea vector populations (predominantly *Xenopsylla* species) which are only active above $\sim 10^{\circ}\text{C}$ () and a longer period of population growth results in a larger flea density and thus a higher plague risk (Stenseth et al., 2006). Summer humidity further correlates with flea burden (flea to host ratio), increasing plague prevalence. Both mechanisms are influenced by the same climatic conditions compounding the prevalence of *Y. pestis* within these populations under such conditions (Samia et al., 2011).

A trophic cascade mechanism influencing plague outbreaks in human populations was first suggested by Parmenter et al. (1999) in work focusing on plague infection in New Mexico. They found that winter-spring precipitation functioned as a predictor of risk for human plague infection and hypothesized that these climatic conditions initiated a trophic cascade influencing both host and vector populations. Host populations are expected to increase following an increase in primary productivity. As *Y. pestis* persists in many host species within arid to semi-arid environments (Mahmoudi et al., 2020), it is therefore logical that a broad application of this trophic cascade mechanism, as observed in both New Mexico and Central Asia, may be applied to plague (Bramanti et al., 2016). Vector populations increase as larval

development and flea reproductive success are both influenced by increased soil moisture in preceding years (Eads et al., 2020). Increased soil moisture in regions of key host species (*Cynomys ludovicianus* & *Cynomys gunnisoni*) particularly at lower depths (up to a max of 24 inches) show positive associations with plague presence, however, soil moisture is influenced by soil taxonomy which varies both vertically and horizontally highlighting the need for a large range of abiotic data when attempting to understand potential heterogeneity in mechanisms between regional plague systems (Pauling et al., 2021).

The trophic cascade mechanism described above is not a true “trophic cascade” as first described in the literature, where a top-level predator indirectly impacts the primary producers through preying on the herbivores which consume these primary producers (Paine, 1980, Schmitz et al., 2000). This top-down trophic cascade has also been suggested for plague systems with *Y. pestis* effectively acting as a predator and influencing primary productivity through very high mortality rates in certain rodent species (e.g. *C. ludovicianus*) (Stapp, 2007). There has, however, been no evidence to support the top-down hypothesis. In contrast, evidence consistent with the bottom-up hypothesis to varying degrees has been found across a range of plague systems. In *C. ludovicianus* populations in Montana support has been found for the bottom-up trophic cascade mechanism with precipitation correlating positively with plague cases when an optimum temperature range (26.7–35°C) was also achieved, however, the same study also found no support for the trophic cascade mechanism in *C. ludovicianus* populations in Colorado, which the authors suggested was due to differences in landscape structure, climate, or characteristics of the data between the study sites (Collinge et al., 2005). Furthermore, flea-borne transmission within *C. ludovicianus* in Colorado does not drive epizootic infection and other mechanisms (e.g. influence of further host species or a soil reservoir system) are required to understand this system (Webb et al., 2006). A similar study was completed investigating the validity of the bottom-up trophic cascade, and the links between each taxonomic level, in Inner Mongolia (China) for two plague reservoir species, *Meriones unguiculatus* (Mongolian gerbil) and *Spermophilus dauricus* (Daurian ground squirrel) (Xu et al., 2014). For the *M. unguiculatus* system the hypothesis was

supported but with caveats, as not all associations at different trophic levels were positive, however, the link between climate and plague persisted (Xu et al., 2014). For the *S. dauricus* system general support was found for a hypothesis with an increase in resources mediated by climate leading to increased abundance of host and vectors, however, this was not shown to directly influence plague prevalence. This finding is consistent with Webb et al. (2006) as there appears to be a disconnect in transmission within the plague system which cannot be explained when investigating one species in isolation or without the influence of non-biotic (environmental) reservoirs.

Studies investigating climatic influences on plague from a regional to continental scale often focus on only one species in geographic isolation (eg. *Rhombomys opimus* – Kazakhstan (Davis et al., 2004), *Otospermophilus beecheyi* – California (Holt et al., 2009), *Marmota himalayana* – Tibetan Plateau (Qian et al., 2014)). However, this may represent an overly simplistic model for *Y. pestis* maintenance through enzootic and epizootic cycles (Webb et al., 2006, Xu et al., 2014). Such studies can be highly informative within the range of the specific species selected, for example, through identifying areas of high suitability for infected individuals of the selected species or through projecting areas of transmission risk via the selected species into future climate conditions (Holt et al., 2009, Qian et al., 2014). The applications of these studies are, however, limited to the range of the selected species as the results may not be representative of climate-reservoir-*Y. pestis* dynamics for other species across a broader geographic range. Even within the range of one host species the potential impact of other species should be considered. Salkeld et al. (2010) investigated interactions of two North American reservoir species, *C. ludovicianus* and *Onychomys leucogaster* and found that the presence of one species facilitated the transmission of *Y. pestis* to the other. *C. ludovicianus* populations are extremely sensitive to *Y. pestis* infection and during epizootics can suffer 100% mortality within a single coterie (family burrow). This high mortality limits the spread as infected vectors are isolated to a single burrow however *O. leucogaster* can act as an intermediate host and facilitate transmission between coterie leading to large epizootics in the *C. ludovicianus* populations (Salkeld et al., 2010). This example demonstrates the

importance of considering multiple hosts even within a geographical range of one key reservoir species.

1.4.4 Environmental niche modelling of *Y. pestis*

Environmental niche modelling (ENM) methods have been used regularly in the investigations of the niche of *Y. pestis* across the globe (Adjemian et al., 2006, Neerinckx et al., 2008, Holt et al., 2009, Maher et al., 2010, Ben-Ari et al., 2012, Qian et al., 2014, Walsh and Haseeb, 2015). Previously, ENM and Species Distribution Models (SDM) have often been used interchangeably, however, all (above) plague niche studies state that they use ENM methods not SDM methods and this distinction is important. The calculation of an ENM is a modelling exercise in which a selection of commonly abiotic variables, such as annual and/or seasonal temperature and precipitation variables are used to determine a species environmental niche within hyperdimensional climate space through the abiotic variable values at locations which the species occurs. This model is then projected into space providing a geographical representation of the environmental niche which may also reflect the species' geographical distribution. This potential distribution will encompass the abiotic conditions which the species, according to the occurrence data provided, may inhabit and is likely close to the species' realized niche (Peterson and Soberón, 2012). However the niche estimation produced by ENM methods may be an overestimation of this realized niche as there are several factors, competition, predation, and dispersal limitations such as dispersal ability and the presence of geographic barriers which inhibit a species from inhabiting all available niche space and are absent from ENM methods. An SDM on the other hand will include dispersal limitations and therefore predicts species presence not just where environmental conditions are suitable but further limits the niche prediction to areas which are accessible (Peterson and Soberón, 2012). Dispersal limitations have rarely been included in studies of plague's niche, hence these studies may produce overestimates of the realized niche (Mendes et al., 2020).

The Maxent algorithm (Phillips et al., 2006) is used consistently across several plague ecology papers (cited above), however, the aims of these studies and hence the

occurrence data used varies significantly. The realized niche of *Y. pestis* is influenced by a combination of host and vector species and the interactions between these species determine the eventual realized niche. Maher et al (2010) presented two hypotheses for *Y. pestis*' distribution (explored in Chapter 2), host niche hypothesis (HNH), in which *Y. pestis*' niche represents a mosaic of combined host (reservoir) species niches and plague niche hypothesis (PNH), in which *Y. pestis*' niche is dictated by ecological conditions conducive to transmission and maintenance of *Y. pestis* rather than purely the presence of a viable reservoir species. Their findings suggest support for the plague niche hypothesis as *Y. pestis* infected species are sampled within a subset of reservoir species localities which is common across reservoir species, and they suggest that this niche is driven by the ecological requirements of vector species. This highlights the three scales at which the niche operates (bacterium, vector, and reservoir species) and the need to consider each when attempting to estimate the overall niche.

Reservoir species occurrences, both independent of and co-occurring with *Y. pestis*, are regularly used within ENMs to estimate the niche of *Y. pestis*. In regions of elevated plague monitoring (selected high-risk areas) reservoir species are often sampled and tested for the presence of *Y. pestis*. In these regions occurrence data of infected reservoir species, which will likely represent a subset of the selected reservoir species occurrences, are available. *Y. pestis* occurrence in an individual reservoir species is determined either through serological testing or bacterial culture from an infected individual. The niche of an infected reservoir species can then be constructed, which will represent a subset of the reservoir species overall niche and can be informative of the risk of infection through transmission from an identified key species (Maher et al., 2010, Qian et al., 2014). This method provides estimates of the niche of *Y. pestis* within a particular species (e.g. *Marmota Himalayana*, Qian et al. (2014)) which can be used in regional risk management, but is reliant on the assumption that the primary reservoir species in the region has been identified and that transmission and maintenance of *Y. pestis* are not significantly influenced by other species (Salkeld et al., 2010). Constructing the niche of species not commonly considered to drive or maintain regional infection but still regularly sampled as

infected (sentinel species), such as carnivores, can also be informative in planning regional plague monitoring systems (Holt et al., 2009). To my knowledge only one study has used ENM methods to produce estimates of the niche of vector species due to the limited data available for such studies. This study highlighted environmental conditions conducive to the occurrence of plague vectors and further suggested areas of potential plague risk through the distribution of each vector's niche (Adjemian et al., 2006). ENM methods have also been used to estimate the niche of *Y. pestis* independently of both reservoir and vector species, by using human infection data with the aim of investigating the distribution of *Y. pestis* infection in humans (Ben-Ari et al., 2010).

Through this thesis I will build on the existing literature through using ENM methods to investigate the impact of a large range of reservoir species in the context of the invasion of a novel region (Chapter 2) and historical climate-plague dynamics (Chapter 3). At the opposite end of the spatial scale, I will investigate the niche of *Y. pestis* at a microbiological scale, through experimentally testing the longevity of *Y. pestis* within soil environments under differing moisture conditions (Chapter 4).

2 The role of biotic factors in limiting the invasion of plague's pathogen (*Yersinia pestis*) in novel geographical settings

(This chapter has been published in the Journal of *Global Ecology and Biogeography* – <https://doi.org/10.1111/geb.13453>

[As such this work was contributed to by a range of authors whose contributions were as follows.](#)

Henry G. Fell, Adam C. Algar, Matthew D. Jones and Steve Atkinson designed the research and drafted the manuscript. Henry G. Fell collected the data. Henry G. Fell all performed all initial analysis, during revision both Henry G. Fell and Owen G. Osborne performed further analysis. Henry G. Fell, Adam C. Algar, Matthew D. Jones, Steve Atkinson, Owen G. Osborne, Simon Tarr and Suzanne H. Keddie contributed to the discussion and initial manuscript revisions.)

2.1 Introduction

Yersinia pestis, the bacterial agent of plague, has caused three trans-continental pandemics, led to millions of human fatalities, and repeatedly catalysed societal re-organisation (Stenseth et al., 2008). However, its importance for human health is not only historical; a large epidemic occurred in Madagascar in late 2017 leading to nearly 2,500 reported cases (Majumder et al., 2018). Given that all of the World Health Organisation's top emergent disease – like plague – are zoonotic, or caused by spill-over from animals (World Health Organization, 2015), it is important to understand what limits the geographical spread of pathogens, reservoirs, and reservoir species outside periods of human epidemics. Ongoing climate change has led to multiple vector-borne diseases expanding their ranges or intensifying within their current distributions with further changes expected. For example, malaria has moved into the East African highlands (Caminade et al., 2014), *Borrelia burgdorferi*, the bacterial agent of Lyme disease, is expanding northward through Canada (Eisen et al., 2016), and plague is predicted to intensify in the Balkhash region of Kazakhstan (Kausrud et al., 2010).

While the spread of plague, and other zoonotic diseases, through human populations depends on a range of epidemiological factors, predicting the likelihood, or risk, of

these outbreaks requires understanding of the pathogen's geographical distribution. While the presence of a pathogen does not guarantee that animal to human transmission will occur, absence of the pathogen ensures that it will not (Messina et al., 2015). Like other zoonotic diseases, non-human plague maintenance is affected by climate across a range of spatial and temporal scales (Ben-Ari et al., 2011). However, its dynamics are not solely determined by direct climate effects, but also depend on the dynamics of the pathogen, vectors and reservoir species, all of which may be affected by changing climate (Kausrud et al., 2007, Ben-Ari et al., 2011). The geographical distribution of *Y. pestis* may reflect both abiotic (e.g. climate) and biotic (e.g. host and vector species) interactions. Disentangling these pathogen-host-vector-climate relationships should improve predictions of the geographical spread of plague under future environmental change and shed light on the role of abiotic and biotic factors in determining species' distributions, a long-standing and fundamental goal of biogeography (Hutchinson, 1957, Soberón, 2007, Blois et al., 2013).

The fact that *Y. pestis* relies on both enzootic (maintenance) and epizootic (amplifying) reservoir species for transmission and expansion has led to two competing hypotheses to describe *Y. pestis*' distribution (Maher et al., 2010). The host niche hypothesis (HNH) suggests that the distribution of *Y. pestis* is determined by the combined niches of host species, which define the limits of the host range. In contrast, the plague niche hypothesis (PNH), proposes that the niche is independent of host niches. Under this hypothesis, the distribution of *Y. pestis* will be limited by environmental conditions (e.g. climate) and will fill only a subset of host ranges. Through comparing the total range of host species to the plague-infected range of host species across North America, Maher et al. (2010) found evidence for the PNH over the HNH. However, recent work using environmental niche modelling has contradicted this finding, suggesting that the strongest predictor of *Y. pestis* distribution across North America is the distribution of *Peromyscus maniculatus* (deer mouse), consistent with the HNH (Walsh and Haseeb, 2015). Further recent work suggests that *P. maniculatus* is not even a maintenance species and may only be infected with *Y. pestis* as a result of an endemic period within other reservoir species

(Danforth et al., 2018). Thus, there is no consensus on whether the current distribution of *Y. pestis* is determined primarily by reservoir species or its own environmental niche. An unexplored possibility is that the hypotheses are not mutually exclusive, and both contribute to *Y. pestis*' distributional limits. Given the large range of reservoir and vector species that are susceptible to *Y. pestis* (Gage and Kosoy, 2005) and its near global distribution (Stenseth et al., 2008), it is also plausible that the relative importance of different factors may vary geographically and with spatial and temporal scale (Ben-Ari et al., 2011).

Y. pestis evolved in Asia from *Y. pseudotuberculosis* roughly 6,000 years before present (BP) (Rascovan et al., 2019). It is enzootic within several host populations across Asia and has caused plague epidemics in humans since at least 5000 BP (Spyrou et al., 2018). *Y. pestis* was introduced to the west coast of the United States at the beginning of the 20th century and rapidly spread eastwards (Adjemian et al., 2007). Several human epidemics were recorded up until 1925 C.E., while *Y. pestis* simultaneously established itself as a permanent presence within rodent reservoir species (Gage and Kosoy, 2005). Plague initially spread at rates of up to 87 km per year, but expansion halted at the 103rd Meridian in the 1950s and has remained static since (Adjemian et al., 2007).

In this chapter, I test whether *Y. pestis*' climatic niche differs between its native and invaded regions and if so whether this difference can be explained by biotic factors. I use information on *Y. pestis*' current climatic niche in its native (Asia) and invaded (North America) range to first test whether the native and invaded niche differ—which is common in non-disease cross-continental invasions (Early and Sax, 2014)—and then evaluate whether incorporating information on putative host distributions can improve predictions of the American range. I find evidence that host distributions limit plague's North American distribution, and that biotic factors are key both independently and with abiotic factors to explain *Y. pestis*' distributional limits in North America.

2.2 Materials and methods

2.2.1 Data acquisition

2.2.1.1 *Yersinia pestis* localities

Y. pestis occurrence records were extracted from two sources: Walsh and Haseeb (2015) for the North America data and Cui et al. (2013) for the Asian data. The sample sizes of the *Y. pestis* occurrence data across Asia and North America were 68 and 62 respectively. The American data includes all *Y. pestis* infection cases in animals, across the USA from 2000-2015, reported through the Program for Monitoring Emerging Diseases (ProMED), to a spatial resolution of 1km². It is unlikely that all plague surveillance data for this period will have been reported through ProMed and hence the data can only represent a subset of *Y. pestis* occurrences for this time. Further to this, the North American data also include occurrences from domestic animals which, although not responsible for plague maintenance in the wild, function as a proxy for local plague infection due to the much higher sampling effort expended upon domesticated animals. Most of the Central Asian data were from China between 1952 and 2006. Once again these data only represent a subset of plague occurrences during this time and were initially selected to represent genetic diversity of *Y. pestis* (Cui et al., 2013). All data were from reservoir or vector species, with human cases removed. These datasets represent all the *Y. pestis* occurrences that I could access, and I make the assumption that they are representative of *Y. pestis* niche. I consider the potential for bias in these data to affect the conclusions in the discussion.

2.2.1.2 Host species localities.

I selected fifteen North America rodent species based on previous evidence suggesting that they may influence *Y. pestis*' North American distribution (Walsh and Haseeb, 2015) or commonly co-occur with *Y. pestis* as a reservoir species (Maher et al., 2010, Mahmoudi et al., 2020). Six of the species (*Peromyscus maniculatus*, *Otospermophilus beecheyi*, *Cynomys gunnisoni*, *Neotamias townsendii*, *Otospermophilus variegatus* and *Cynomys ludovicianus*) were the most commonly sampled rodent species co-occurring with *Y. pestis* (Maher et al., 2010) and have each been hypothesised as species regionally or continentally important to *Y. pestis*

distribution (Pauli et al., 2006, Hubbart et al., 2011, Mahmoudi et al., 2020). The nine further species (*Cynomys leucurus*, *Cynomys parvidens*, *Sciurus aberti*, *Sciurus griseus*, *Sciurus niger*, *Neotamias amoenus*, *Neotamias dorsalis*, *Neotamias minimus* and *Neotamias umbrinus*) are all known reservoir species (Lowell et al., 2009, Mahmoudi et al., 2020). Modelled distributions of these species are shown in Chapter 3 Appendix Figure 2.5. These are all either enzootic (maintenance), epizootic (amplifying) or resistant (*P. maniculatus* only) reservoir species (Gage and Kosoy, 2005). I do not consider predators and scavengers as they are not considered a primary driver or maintainer of *Y. pestis* (Savage et al., 2011), though more research in this area is required (Salkeld and Stapp, 2006).

Species occurrence data for the fifteen host species were obtained from the Global Biodiversity Information Facility (GBIF) (GBIF, 2019) and initially cleaned by removing NA values and duplicates. Likely erroneous localities were identified and removed through visual examination and comparison to previous literature for the species *N. amoenus*, *S. griseus*, *O. beecheyi*, *N. umbrinus*, *C. gunnisoni*, *C. ludovicianus*, *C. leucurus*, *N. townsendii* and *N. minimus* (Pizzimenti and Hoffmann, 1973, Sutton, 1992, Carraway and Verts, 1994, Miller and Cully Jr, 2001, Verts and Carraway, 2001, Braun et al., 2011, Smith et al., 2016) (Chapter 3 Appendix, Figure 2.6). Due to computational constraints, I limited the number of occurrences of each individual species to 1500. Following Boria and Blois (2018), I randomly selected 1500 occurrences from across its range without replacement for subsequent modelling. Treating each species separately, I thinned occurrence data so that occurrences were separated by at least 10 km using the *spThin* package (Aiello-Lammens et al., 2015) to increase computation speed while also addressing potential spatial sampling biases.

2.2.1.3 Climate data

The 30 arc-second WorldClim dataset was aggregated to 10 km resolution (Hijmans et al., 2005) to quantify climatic variation across North America (longitude -134.5° - -4.0°, latitude 26.4° - 52.5°) and Asia (longitude 60.0° - 144.3°, latitude 7.1° - 54.3°). This resource was constructed from data from the period 1950-2000, and therefore

represents a time-averaged estimate of climate variables. Eight Worldclim climate variables were selected for use by identifying uncorrelated climatic variables across the research areas (Asia and North America). To avoid the inclusion of overly correlated variables I ran pairwise correlation analysis on all climate variables across both regions. Variables were retained if all pairwise correlation values were less than 0.8 in both the North American and Asian region. I also included precipitation of the driest quarter as previous research has suggested a strong link between this variable and *Y. pestis* maintenance in host species from varying localities (Neerinckx et al., 2008, Holt et al., 2009, Ben-Ari et al., 2010, Kausrud et al., 2010). I validated the variable selection through a literature review of biologically significant variables in previous studies of climatic influence on *Y. pestis* (Neerinckx et al., 2008, Holt et al., 2009, Ben-Ari et al., 2010, Kausrud et al., 2010, Walsh and Haseeb, 2015). The selected variables were: mean diurnal range (BIO2), isothermality (BIO3), mean temperature of wettest quarter (BIO8), mean temperature of driest quarter (BIO9), precipitation seasonality (BIO15), precipitation of driest quarter (BIO17), precipitation of warmest quarter (BIO18) and precipitation of coldest quarter (BIO19). All data cleaning and analysis was completed in the R statistical programming environment (R Core Team, 2013).

2.2.2 Niche modelling

To quantify *Y. pestis*' climatic niche in its native (Asia) and invaded (North America) range, I fit environmental niche models in each region using the maxent algorithm (Phillips et al., 2006). ENMeval (Muscarella et al., 2014) was used to tune the regularisation multiplier and feature classes, with the linear, quadratic, hinge, threshold and product feature classes included and the regularisation multiplier ranging from 0.5 to 4 by 0.5 increments. I selected the best model for each region by sorting on lowest omission rate, and selecting the model with the highest AUC sequentially (Galante et al., 2018). I used the spatially structured partitioning method checkerboard2 from the *ENMeval* package which splits the data into 4 bins based on a nested checkerboard structure, models are then run iteratively using k-1 bins for training with the remaining one for testing. I found that the best model varied when the optimization was repeated due to variation in the random selection and

subsequent partitioning of the background data, therefore this process was repeated 50 times and the most commonly selected best model was subsequently used. This method enabled me to ensure that the regularisation multiplier and feature class selected was suited to each species occurrence data and not a product of background data selection and partitioning. The same approach as that described above was used to model the potential distributions in the USA of the fifteen putative reservoir species of *Y. pestis* for subsequent hypothesis testing (model settings and projections shown in Chapter 3 Appendix Table 2.2 and Figure 2.5).

2.2.3 Hypothesis testing

2.2.3.1 Native and invaded niche similarity

The Asian Model (Table 2.1) was projected into North America and calculated the niche overlap between the projected Asian Model (Asian-proj; Table 2.1) and the North American model (N.am-obs; Table 2.1) using Schoener's D. I used the identity test and background test through *ENMtools* (Warren et al., 2008, Warren et al., 2010) to test the null hypotheses that the Asian and North American climatic niches were identical (identity test), and that niches do not differ more than expected by chance given environmental differences between regions (background test). I further completed multivariate environmental similarity surfaces (MESS) analysis to identify areas of non-analogous climate in North America in comparison to the Asian occurrences.

2.2.3.2 Biotic influences on the invaded niche

The extent to which potential host species distributions may be responsible for *Y. pestis*' climatic niche difference in North America (see Results) was tested by comparing the overlap between the N.am-obs model (Table 2.1) and a series of ENMs of *Y. pestis* in North America that used the modelled distributions of putative reservoir species and the projected Asian climatic niche as predictors (Table 2.1). If host distributions are responsible for the observed differing niche in *Y. pestis*, then I expect higher overlap between the N.Am-obs model and models using reservoir species as predictors than between the N.Am-obs model and an ENM that included

Table 2.1. Naming protocol and performance for models used to predict *Y. pestis*' North American niche. Predictors included in each model are given below, along with the accuracy of each model, calculated as the area under the curve (AUC) of the receiver operating

Model name	Predictors	Mean AUC	Standard Deviation
Asian Model	Asian Worldclim variables at 10km resolution (BIO2, BIO3, BIO8, BIO9, BIO15, BIO17, BIO18, BIO19)	0.709	--
N.Am-obs	North American Worldclim variables at 10km resolution (BIO2, BIO3, BIO8, BIO9, BIO15, BIO17, BIO18, BIO19)	0.914	0.00276
Asian-proj	Model fit in Asia projected into North America	0.798	0.00334
<i>P.man.</i> -only	<i>P. maniculatis</i> suitability in North America	0.831	0.00277
Multi-host	<i>P. maniculatus</i> , <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>N. townsendii</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>C. parvidens</i> , <i>S. aberti</i> , <i>S. griseus</i> , <i>S. niger</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.937	0.00220
Multi-host (limited selection)	<i>P. maniculatus</i> , <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>C. ludovicianus</i> , <i>N. townsendii</i> and <i>O. variegatus</i> distributions	0.931	0.00162
Asia+ <i>P.man.</i>	Projected Asian Model and the projected <i>P. maniculatis</i> distribution	0.897	0.00223
Asia+Multi-host	Projected Asian Model and the projected <i>P. maniculatus</i> , <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>N. townsendii</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>C. parvidens</i> , <i>S. aberti</i> , <i>S. griseus</i> , <i>S. niger</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.941	0.00307
Multi-host, no <i>P.man</i>	<i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>N. townsendii</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>C. parvidens</i> , <i>S. aberti</i> , <i>S. griseus</i> , <i>S. niger</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.939	0.00192
Asia+Multi-host, no <i>P.man</i>	Projected Asian Model and the projected <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>N. townsendii</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>C. parvidens</i> , <i>S. aberti</i> , <i>S. griseus</i> , <i>S. niger</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.943	0.00280

only the Asian-proj model as a predictor. Projecting the Asian niche into North America and using this as a predictor in an ENM of North American *Y. pestis* allowed me to evaluate the overlap between the North American and Asian niche in a manner that is consistent with how I measured overlap from models using host species distributions and also allowed the inclusion of both the Asian niche and host distributions within the same model. If host distributions are not responsible for the climatic niche difference in North America, then a greater overlap would not be expected between the N.Am-obs and models using host distributions than with models using only the Asian-niche projection. I fit models including all host species, as well as models including only *P. maniculatus* distribution based on findings by Maher et al. (2010). Models were also fitted that included fourteen reservoir species and excluding *P. maniculatus*. Details on the predictors included in each model are given in Table 2.1.

I fit ENMs of *Y. pestis* in North America using the projected Asian niche and host distributions as predictors following the approach outlined above to select the regularisation multiplier and feature classes. I also repeated the analysis using the most commonly selected regularisation multiplier and feature class across all models (Linear and Quadratic and regularisation multiplier of 0.5) to determine to what degree the results are a product of varying model settings. To train each model, spatially structured partitioning methods (checkerboard2) were used to split the occurrence and background data into training and test sets. This partitioning resulted in very slightly different model parameters (see Table 2.1) in each run due to the variation in background localities and partitioning. To ensure this did not influence subsequent analyses, I re-selected and partitioned the background data 100 times and carried out all downstream analyses on these 100 model sets. After model fitting, I computed the overlap of each Asian-niche and/or host distribution model (Table 2.1) with the N.Am-obs model using Schoener's D. I further computed the niche overlap of Asian-niche and/or host distribution models using thresholded projections, using the specificity-sensitivity threshold (the threshold at which model sensitivity and specificity are highest) applied to confirm that the results were not an artefact of the continuous Maxent projections (Chapter 3 Appendix Figure 2.7).

2.2.3.3 Geographical null modelling

Spatial autocorrelation of species' occurrences and environmental variables can create spurious associations in ENMs (Dormann et al., 2007). Thus, relationships between *Y. pestis* occurrences and host species' suitability maps could arise solely because both are spatially structured. It could be that any spatially structured predictor would have a significant effect on *Y. pestis* distribution and thus overlap, even in the absence of a functional relationship (e.g. Fourcade et al. (2018)). To guard against this, a null modelling procedure was used. Following Beale et al. (2008), Algar et al. (2013) and Nunes and Pearson (2017), I generated a set of null occurrences for each host species with the same number of points and the same degree of spatial clustering as the observed species using the '*fauxcurrence*' package (version 1.1.0) in R (Osborne et al., 2021) (see Chapter 3 Appendix 2.5.3 Details of null modelling algorithm). The algorithm also preserved spatial clustering of interspecies centroid distances to account for the fact that host species may be found in similar environments (Chapter 3 Appendix 2.5.3). I then modelled each host's null distribution using Maxent as described for real species and used these modelled distributions as predictors in models of *Y. pestis* occurrence. I then measured overlap with N.am-obs. This null overlap value describes the expected overlap if a series of spatially structured predictors, from anywhere in the study area, were used to predict *Y. pestis*' distribution. This process was repeated 1000 times to get a null distribution of Schoener's D values for each model and compared the observed Schoener's D to this null distribution, computing two-tailed P-values.

2.3 Results

2.3.1 Native and invaded niche similarity

The identity test (Chapter 3 Appendix) found that there was significantly less overlap between the native and invaded niche of *Y. pestis* than expected under the null hypothesis of identical niches (p-value = 0.01). The background test (Chapter 3 Appendix) suggested that the native and invaded niche were significantly less similar than expected given the geographic regions in which they reside (p-value = 0.01). In

combination these results suggest that the current native and invaded niche are non-identical and that the difference is not purely due to their differing geographic regions. I constructed and projected models in both regions to identify factors that contributed to this dissimilarity. Within Asia, the best fitting Maxent niche model for *Y. pestis* had an AUC of 0.709, while in North America (N.Am-obs, Table 2.1) the equivalent model had a mean AUC of 0.914 (individual species model and North American settings shown in Chapter 3 Appendix Table 2.2 & Table 2.3 respectively). The overlap between the native and invaded niche of *Y. pestis* had a Schoener's D value of 0.497. The Asian model AUC was close to the AUC value of 0.7 above which models are considered fair (Swets, 1988), while the model was also relatively simple due to the singular linear feature. The Asian data did not include any occurrences in south-eastern coastal China and this area was thus poorly represented in the projected species distribution. Despite no occurrences in Kyrgyzstan and Kazakhstan, regions of permanent plague reservoirs (Kausrud et al., 2007), the model predicted plague's presence in Kyrgyzstan and moderately in Kazakhstan but did not include the Balkhash reservoir region (Kausrud et al., 2007).

The N.am-obs model showed the niche of *Y. pestis* generally falling into two regions, California, where *Y. pestis* was introduced, and eastern Rocky Mountains up the length of the USA through New Mexico, Colorado, Wyoming and Montana. The model matches previous modelled distributions of plague in California as well as records of *Y. pestis* within human populations across a similar period of sampling (Holt et al., 2009, Kugeler et al., 2015). The differences between the native and invaded niche were highlighted through projecting both into the alternative region (Figure 2.1 a-d). The North American model poorly represented the Asian distribution, primarily predicting *Y. pestis* presence adjacent to the Quighai-Tibet Plateaux which is likely due to the altitude of North American occurrence points within the Rockies. Meanwhile the Asian model substantially overpredicted *Y. pestis* distribution in North America with the largest overprediction into midwestern grasslands (Figure 2.1 c & e). The Asian model performs better when predicting distribution in California, however, consistently overpredicts elsewhere. MESS analysis suggests that non-analogous climate in North America (with regards to Asian

occurrences) is focused on the north western coast while California and the central belt of North America where the Asian model predicts distribution (Figure 2.1 c) is an area of more analogous climate (Chapter 3 Appendix Figure 2.10).

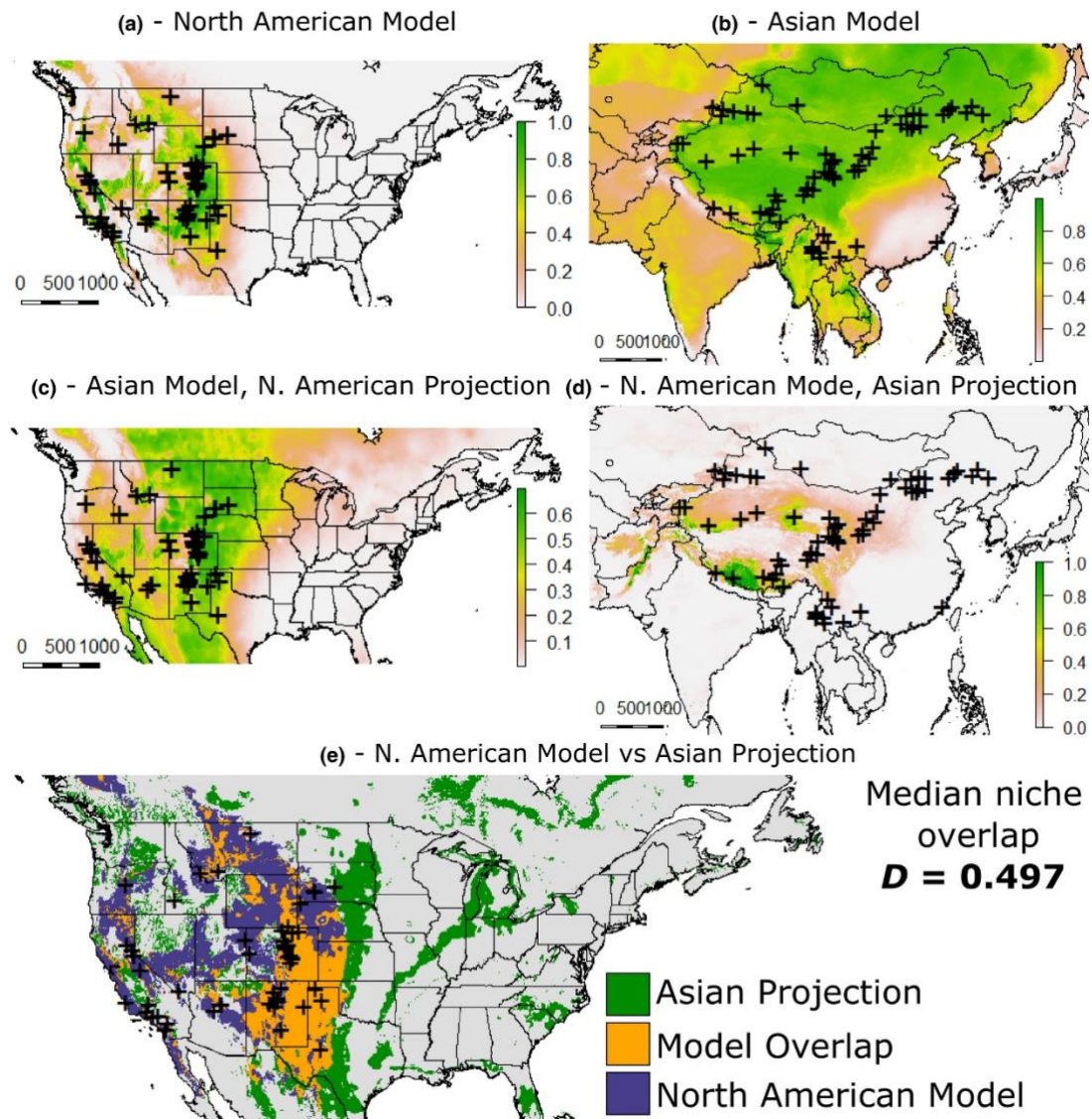


Figure 2.1 Comparison of *Yersinia pestis*' native and invaded niche. *Y. pestis* distribution models built within; (a) the invaded North American region, (b) the native Asian region and (c-d) both models projected into the alternate region. Differing projected distribution of *Y. pestis* across N. America using a model fitted in North America, and a model fit in Asia and projected to North America. Colour scale shows model predicted niche suitability where 1 is complete niche suitability. The Asian projection ("Asia-proj" in Table 2.1) over-predicts *Y. pestis* to the east of the N. American model. For visualisation purposes, the figure uses the spec_sens threshold, which is the threshold at which sensitivity and specificity are highest. The same pattern was observed using other threshold methods, however spec_sens is one of the more conservative visualisations (Chapter 3 Appendix Figure 2.18). Mean niche overlap values calculated using Schoener's D (Schoener, 1968) were derived from 100 replicates of the models with a random selection of background points leading to slight variation in predicted distribution. A scale bar (km) is provided (a-d).

2.3.2 Biotic influences on the invaded niche

To determine whether host distributions can explain *Y. pestis*' niche difference in North America, I first projected the Asian-fitted model into North American climate space and used the resulting suitability map as the only predictor to fit a Maxent model to North American *Y. pestis* occurrences. This model (*Asian-Proj*, Table 2.1) had an AUC of 0.798 which is above the 0.7 threshold which indicates a “good” model. These two models also showed a low niche overlap (Figure 2.1 e) in geographic space (Schoener's $D=0.497$). Next, to test whether putative host species can explain *Y. pestis*' North American distribution, I used the suitability maps of rodent species as predictors in a series of Maxent models' fitted to *Y. pestis* North American occurrences (Figure 2.2 & Figure 2.3, Table 2.1, Chapter 3 Appendix Table 2.3 & Table 2.4). All models including hosts achieved a greater niche overlap than the Asian projection alone (Figure 2.1 e and Figure 2.2). The model containing only *P. maniculatus* performed moderately better (Schoener's $D = 0.592$) than the Asian projection alone (Figure 2.2 a), while the model containing all putative host species performed better again and produced the greatest niche overlap of any of the models (Schoener's $D = 0.722$) (Figure 2.2 b). The second highest niche overlap was achieved by the model containing the Asian projection and the *P. maniculatus* model (Figure 2.2 c), however, I urge caution in interpreting these results by suggesting that *P. maniculatus* is a key host species as it is possible its higher modelled suitability in the western US, where the model tends to over-predict (Figure 2.2 a), reflects sampling density as *P. maniculatus* has a broad North American range. In contrast, all other models achieve greater niche overlap and show less overprediction to the west (Figure 2.2 b-f). The model containing all host species, which showed the greatest niche overlap values (Figure 2.2 b), performed well in the Californian region and the southern Rockies, with the largest area of underprediction in the northern Rockies (Montana and Wyoming). The models that contained both the Asian projection and one or more host performed marginally worse than the equivalent models with only reservoir species included. When I held model settings consistent across all models (Chapter 3 Appendix Figure 2.11 & Figure 2.12), I found that the key result, that including host distributions improved niche overlap remained. However, there were

minor differences in terms of which host-model performed best. The Asia + *P. maniculatus* model showed the greatest niche overlap followed by the multi-host with a limited selection of host species. The optimized models are most likely to show the most reliable results, however the major findings are robust to details in model selection.

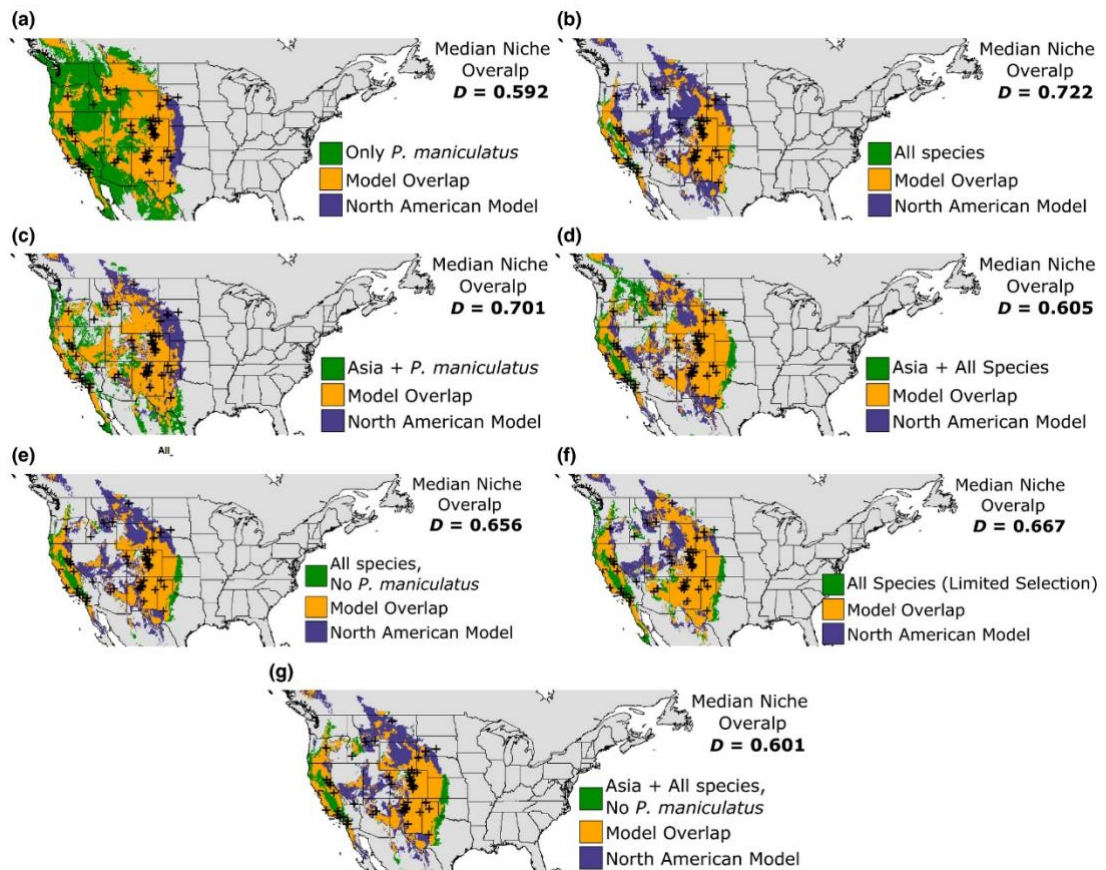


Figure 2.2. Differing projected distribution of *Y. pestis* across North America using a model fitted in North America and: (a) a model with only *P. maniculatus* distribution included (“P.man only”), (b) a model with fifteen rodent species (“Multi-host”), (c) a model fit in Asia and projected to North America with *P. maniculatus* distribution included (“Asia+P.man”), (d) a model fit in Asia and projected to North America with fifteen rodent species (“Asia+Multi-host”), (e) a model built with a limited selection of only six species based on co-occurrences with *Y. pestis* (f) a model without *P. maiculatus* but including the fourteen further selected host rodent species and (“Multi-host, no P.man”) and, (g) a model fit in Asia and projected to North America without *P. maniculatus* but including the fourteen further host rodent species (“Asia+Multi-host, no P.man”).

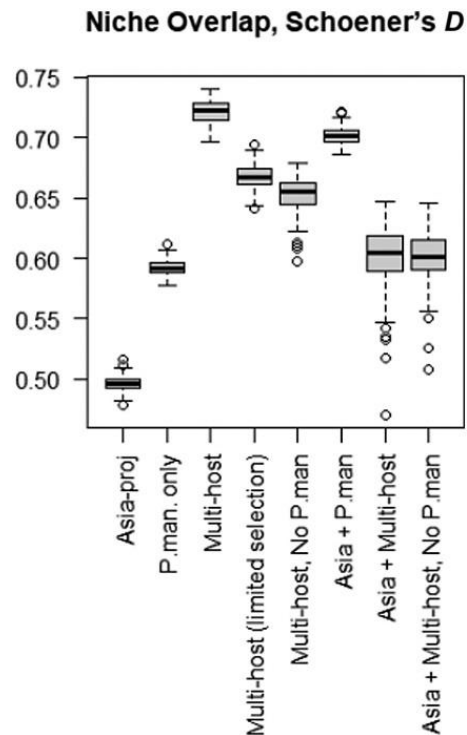


Figure 2.3. Niche overlap values of the observed distribution models. Boxplots of the niche overlap values (Schoener's D) produced through the 100 replicates of each niche model. Model descriptions are provided in Table 2.1. Higher values of Schoener's D indicate greater niche overlap.

2.3.3 Geographical null modelling

Niche models based on null occurrences had lower AUC values than models based on real data (Chapter 3 Appendix Figure 2.13), suggesting that the model fits are not solely an artefact of spatially clustered occurrence data. Moreover, most of the models had significantly greater overlap with the N.am-obs model than expected from the null expectation (Figure 2.4). The two exceptions were the Asia + Multi-host model ($P = 0.993$, Figure 2.4 c), and the Asia + Multi-host model excluding *P. maniculatus* ($P = 0.683$, Figure 2.4 f). These results suggest that even randomly placed sets of occurrence points can interact with the Asian-proj to produce relatively high apparent overlap. However, the overlap (Schoener's D) of these models was lower than that when multi-hosts was used alone, a result which was significantly different from the null expectation, indicating that the increase in overlap due to host distributions (either from the complete or limited-set) is not an artefact of spatially autocorrelated occurrences. Using consistent model settings I saw fewer results retain significance (Chapter 3 Appendix Figure 2.12) which likely reflects the poorer

performance of these sub-optimized models, but highlights the importance of considering how model selection and parameterization influence the outcome of environmental niche models.

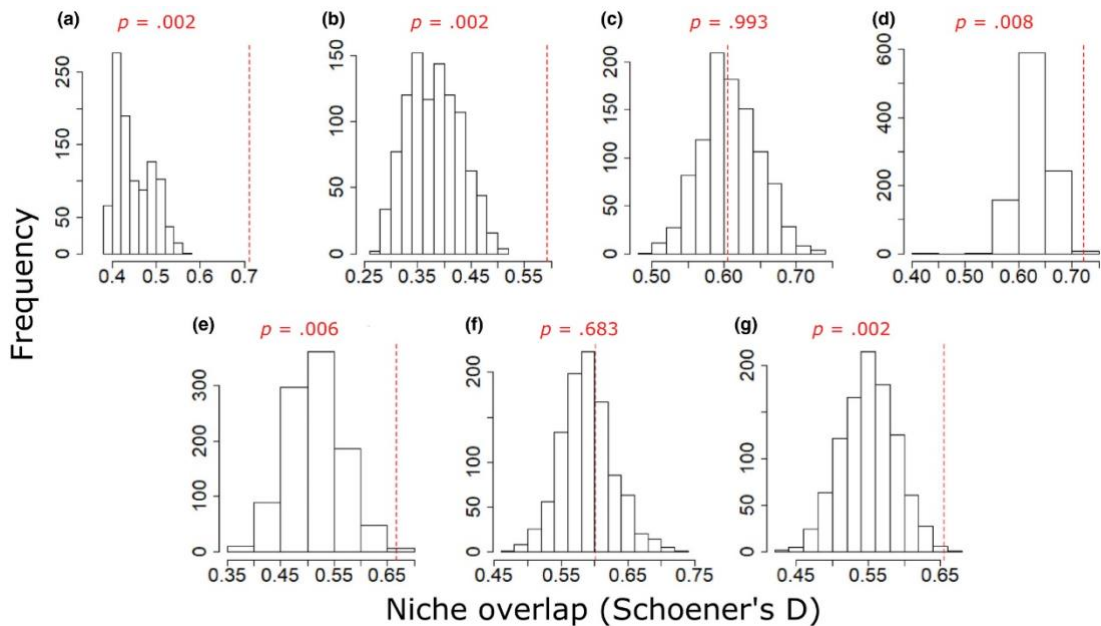


Figure 2.4. Null and observed overlap (Schoener's D) between *Y. pestis*' North American niche and the predicted niche based on; (a) *P. maniculatus* and the Asian niche (Asia + *P. man*), (b) just a single host species distribution (*P.man*-only), (c) multiple hosts and the Asian niche (Asia + Multi-host), (d) only multiple host species (Multi-host), (e) Limited selection of multiple hosts (Multi-host limited) (f) multiple hosts with the Asian niche but without *P. maniculatus* (Asia + Multi-host, no *P.man*) and (g) multiple host species without *P. maniculatus* (Multi-host, no *P.man*). Histograms show the distribution from 500 null simulations and red lines show observed values.

2.4 Discussion

I found that the current environmental niche of *Y. pestis* in Central Asia differs from the environmental niche in North America. Despite the presence of suitable climate space, *Y. pestis* has failed to fill its potential range in central North America. Incorporating information on mammalian host distributions greatly improved concordance between the observed and predicted range edge in North America. This indicates a role for biotic factors in limiting *Y. pestis*' North American distribution, at least at broad continental scales; though the same may not be true at more regional scales (Ben-Ari et al., 2011, Danforth et al., 2018). Including information on plague's native climatic niche had mixed impacts on the models, improving one of the model's niche overlap while reducing the rest, suggesting that although both climatic and biotic factors may interact to limit the spread of plague's pathogen in North America biotic factors appear to have the larger impact. Further effects of host distributions

may be revealed by conducting targeted regional studies, while consideration of vectors and their distributional limits as well as the impact of further biotic factors such as soil and vegetation variables will provide a more complete picture of the North American *Y. pestis* invasion (Adjemian et al., 2006).

The models allowed me to determine the impact that a subset of putative reservoir species distributions had upon *Y. pestis* distribution and evaluate competing hypotheses (plague niche, PNH, and host-niche, HNH) for the dominant factors limiting *Y. pestis*' niche in North America. Models using only host distributions (Figure 2.2 a & b) outperformed models based on the Asian-climatic niche, suggesting biotic factors are responsible for the discrepancy between the Asian and North American niche, as proposed by the HNH. Response curves from maxent models show positive relationships linking host and *Y. pestis* occurrences for twelve of the fifteen selected species (Chapter 3 Appendix Figure 2.14). This suggests that the three species that showed negative response curves, *Neotamias townsendii*, *Cynomys parvidens* and *Sciurus niger*, may not represent host species which drive *Y. pestis* infection at a continental scale. A further species, *Sciurus griseus*, had a permutation importance of 0 in the Asia + multi-host model. Each of these species may still be important regionally or across a subset of their range. Removing these species had mixed impact on empirical and null models suggesting further work is required to determine their impact on *Y. pestis* distribution at a continental scale (Chapter 3 Appendix , Figure 2.16, Figure 2.17 and Table 2.4). Overall, I found that the combination of host species that best explained *Y. pestis* distribution was sensitive to the model settings used (compare Figure 2.3 and Chapter 3 Appendix Figure 2.11), highlighting the importance of careful calibration of environmental niche models.

The role of *P. maniculatus* as a host species of *Y. pestis* is debated, with recent work suggesting that it may represent a spill-over host and not drive transmission, at least within California (Danforth et al., 2018). The null modelling results (Figure 2.4) showed that models with *P. maniculatus* in isolation, or as one of two variables, showed significantly greater niche overlap than the spatial nulls. Thus, the results, like those of Walsh and Haseeb (2015), cannot eliminate *P. maniculatus* as an

important North American host species, but nor can it confirm it drives transmission. Given regional evidence suggesting *P. maniculatus* may only represent an overflow host in some regions (Danforth et al., 2018) I urge caution in the use of *P. maniculatus* as a driver of *Y. pestis* distribution, especially as with the optimized models, including information on additional reservoir species provided improved overlap with *Y. pestis*' North American distribution.

The constraining effects of host distributions on *Y. pestis*' niche in North America means that it is in disequilibrium with climate (*sensu* Araújo et al. (2005)) in its invaded range. Given that host species limit *Y. pestis*' niche in North America, it seems likely that reservoir species also constrain *Y. pestis*' climatic niche within Asia, meaning it is probably in disequilibrium with climate in both its invaded and native range, a situation that may be common for invasive species more broadly (Early and Sax, 2014, Guisan et al., 2014). Conceptually, the differences in *Y. pestis*' current climatic niches in Asia and North America, therefore, reflects the relaxation of biotic constraints as the pathogen is transmitted from its native region, only to be replaced by novel biotic constraints in its invaded range. However, the constraints on the native range have occurred in a different part of climate space than in the invaded range, leading to an altered realized niche, driven by the change in host species from its native to invaded region. Another possibility is that the Asian climatic niche has expanded, either due to evolutionary change or relaxation of biotic constraints, since *Y. pestis*' introduction to the Americas, creating the appearance of reduction in niche width in the Americas. While this possibility cannot be discounted, to explain the results it would have had to occur in such a way as to coincidentally align with the distributions of known host species, and thus I suggest limitation by biotic factors in North America a more likely explanation. As both the native and invaded niches incorporate biotic factors, the degree of niche underfilling (Strubbe et al., 2013, Guisan et al., 2014) in North America is likely greater than detected here as the native niche model likely failed to capture fundamental niche limits.

Could differences observed between the current native and invaded niches reflect limiting factors other than host distributions and climate? Two alternative

possibilities are a limited time for dispersal and founder effects. As recent invaders often fail to fill their potential niche simply because they have had insufficient time to do so (Strubbe et al., 2013), *Y. pestis* could thus be limited to its current range in North America due to a lack of time to expand into viable climatic niche space further east of its current extent. However, this explanation is unlikely as *Y. pestis* reached its maximum eastern extent approximately 70 years ago and has remained static since (Adjemian et al., 2007), suggesting an alternative - such as biotic factors - has halted its eastward expansion.

Founder effects can induce fundamental niche changes as only a limited proportion of genetic variation is retained in the invaded population (Pearman et al., 2008), which can lead to invaders filling only a subset of the species' native niche (dos Santos Ribas et al., 2018). Given that only one strain of *Y. pestis* (1.ORI1) was introduced to North America (Morelli et al., 2010) I cannot discount this possibility, as I lack sufficient data to quantify the climatic niche of the 1.ORI1 strain in its native range. While I cannot discount founder effects, I suggest that it is unlikely that they would manifest in such a way that they produce spurious correlations with host distributions. Additionally, while I have focused on the pathogen and its host here, the former's distribution will also reflect influences of the vector and thus a fuller understanding of the effect of biotic factors on the pathogen's distribution will require data not just on reservoir species, but also the vector's niche and distribution.

The results suggest a role for host distributions in limiting the expansion of *Y. pestis* in North America, but these findings must be interpreted cautiously due to the limited data available for modelling. The predicted distributions are based on Asian and North American occurrence data that only partially overlap in time (2000-2006), with the Asian data collected over a longer time period. I therefore assume that the Asian climatic niche has not undergone substantial change during this period. As time-averaged climatic data was used (Hijmans et al., 2005), I also assume that recent climate changes have not drastically altered the measurement of the climatic niche.

The North American data represent only a subset of plague cases recorded in animals as these data were originally sourced through ProMED which may not represent all plague surveillance data and may represent cases only investigated following human infection (Cowen et al., 2006). Previous studies have utilised alternative data sources, but these were inaccessible to us. There are also limitations of the Asian data, which are biased towards China and were originally selected to represent the genetic diversity of *Y. pestis* in China (Cui et al., 2013). While more complete data would provide more certainty in the estimates of *Y. pestis* niche, there are two reasons why a more extensive dataset is unlikely to change the major conclusion. Firstly, additional data on *Y. pestis*' occurrence throughout Asia would expand, not contract, the breadth of the Asian niche and increase the range of environment in which *Y. pestis* could occur in North America, further reducing niche overlap. Similarly, including data on *Y. pestis*' distribution in other regions of the world, outside of China, in the estimate of the non-Americas climatic niche would only expand the measured breadth of the climatic niche and expand the potential distribution in the Americas. Underestimation of the North American distribution of *Y. pestis*, possibly from a reduced range of sampling years relative to the Asian data, could result in reductions in observed overlap and thus it is possible that I have over-estimated the niche difference in North America. However, the major area of under-filling is at the eastern edge of *Y. pestis*' distribution and the expansion and recent stasis of this eastern edge of the distribution is well described (Adjemian et al., 2007), suggesting that the observed niche contraction is unlikely to be solely due to a sampling artefact. Nonetheless, improved accessibility of existing *Y. pestis* distribution data, increased surveillance of non-human associated cases and, just as importantly, increased information on vector distribution and niches will all contribute to improved understanding of the dynamics of *Y. pestis* invasions.

Predicting the geography of zoonotic diseases, past, present and future, depends on understanding the factors limiting the geographic distribution of the pathogen. Here I have shown that, as for many ecological invasions (Strubbe et al., 2013, Early and Sax, 2014, Guisan et al., 2014), the climatic niche of *Y. pestis* has shifted after introduction to a novel region, suggesting that environmental niche models based

solely on climate may not provide reliable predictions of zoonotic disease distributions under rapid environmental change, particularly as they invade new regions. This work suggests that reliable information on host distributions is more important in predicting pathogen distributions than information on the climatic niche of the pathogen itself. Thus, identifying potential reservoir species, and what influences their distributions, is key to understanding the future geography of emerging and established zoonotic pathogens.

2.5 Chapter 3 Appendix

2.5.1 Supplementary tables

Table 2.2 Feature Class and regularisation multiplier of each host species model.

Model name	Feature Class	Regularisation Multiplier
<i>C. gunnisoni</i>	Linear, Quadratic, Hinge, Product	1.0
<i>C. leucurus</i>	Linear, Quadratic, Hinge, Product	4.0
<i>C. ludovicianus</i>	Linear, Quadratic, Hinge, Product	3.0
<i>C. parvidens</i>	Linear, Quadratic, Hinge, Product	2.0
<i>N. amoenus</i>	Linear	0.5
<i>N. dorsalis</i>	Linear, Quadratic, Hinge, Product	1.0
<i>N. minimus</i>	Hinge	1.5
<i>N. townsendii</i>	Linear, Quadratic, Hinge, Product, Threshold	2.0
<i>N. umbrinus</i>	Linear, Quadratic, Hinge	4.0
<i>O. beecheyi</i>	Linear, Quadratic	3.5
<i>O. variegatus</i>	Linear, Quadratic, Hinge, Product, Threshold	2.0
<i>P. maniculatus</i>	Linear	0.5
<i>S. aberti</i>	Linear, Quadratic, Hinge, Product	1.0
<i>S. griseus</i>	Linear, Quadratic	0.5
<i>S. niger</i>	Linear, Quadratic, Hinge, Product	0.5

Table 2.3 Feature Class and regularisation multiplier of each model used for overlap calculations

Model name	Feature Class	Regularisation Multiplier
Asian Model	Linear	3.0
N.Am-obs	Linear, Quadratic	1.0
Asian-proj	Linear, Quadratic, Hinge, Product, Threshold	0.5
<i>P.man.</i> -only	Linear, Quadratic	1.5
Multi-host	Linear, Quadratic, Hinge	4.0
Multi-host (Limited)	Linear, Quadratic	0.5
Asia+ <i>P.man.</i>	Linear, Quadratic, Hinge	3.0
Asia+Multi-hosts	Linear, Quadratic	0.5
Multi-host, no <i>P.man</i>	Linear, Quadratic	2.5
Asia+Multi-host, no <i>P.man</i>	Linear, Quadratic	0.5
Multi-host, no neg cor	Linear, Quadratic	0.5
Asia+Multi-host, no neg cor	Linear, Quadratic	2.5
Multi-host, no neg cor, no <i>P.man</i>	Linear	0.5
Asia+Multi-host, no neg cor, no <i>P.man</i>	Linear, Quadratic, Hinge	3.0

Table 2.4 Naming protocol and performance for supplementary models used to predict *Y. pestis*' North American niche. Predictors included in each model are given below, along with the accuracy of each model, calculated as the area under the curve (AUC) of the receiver operating characteristic, using known occurrences of *Y. pestis* in North America. The mean AUC and standard deviation are calculated from 100 replicates of each model.

Model name	Predictors	Mean AUC	Standard Deviation (3 sig fig)
Multi-host, no neg cor	<i>P. maniculatus</i> , <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>S. aberti</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.942	0.00254
Asia+Multi-host, no neg cor	Projected Asian Model and the projected <i>P. maniculatus</i> , <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>S. aberti</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.937	0.00192
Multi-host, no neg cor, no <i>P.man</i>	<i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>S. aberti</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.941	0.00168
Asia+Multi-host, no neg cor, no <i>P.man</i>	Projected Asian Model and the projected <i>O. beecheyi</i> , <i>C. gunnisoni</i> , <i>O. variegatus</i> , <i>C. ludovicianus</i> , <i>C. leucurus</i> , <i>S. aberti</i> , <i>N. amoenus</i> , <i>N. dorsalis</i> , <i>N. minimus</i> and <i>N. umbrinus</i> distributions	0.939	0.00201

2.5.2 Supplementary figures

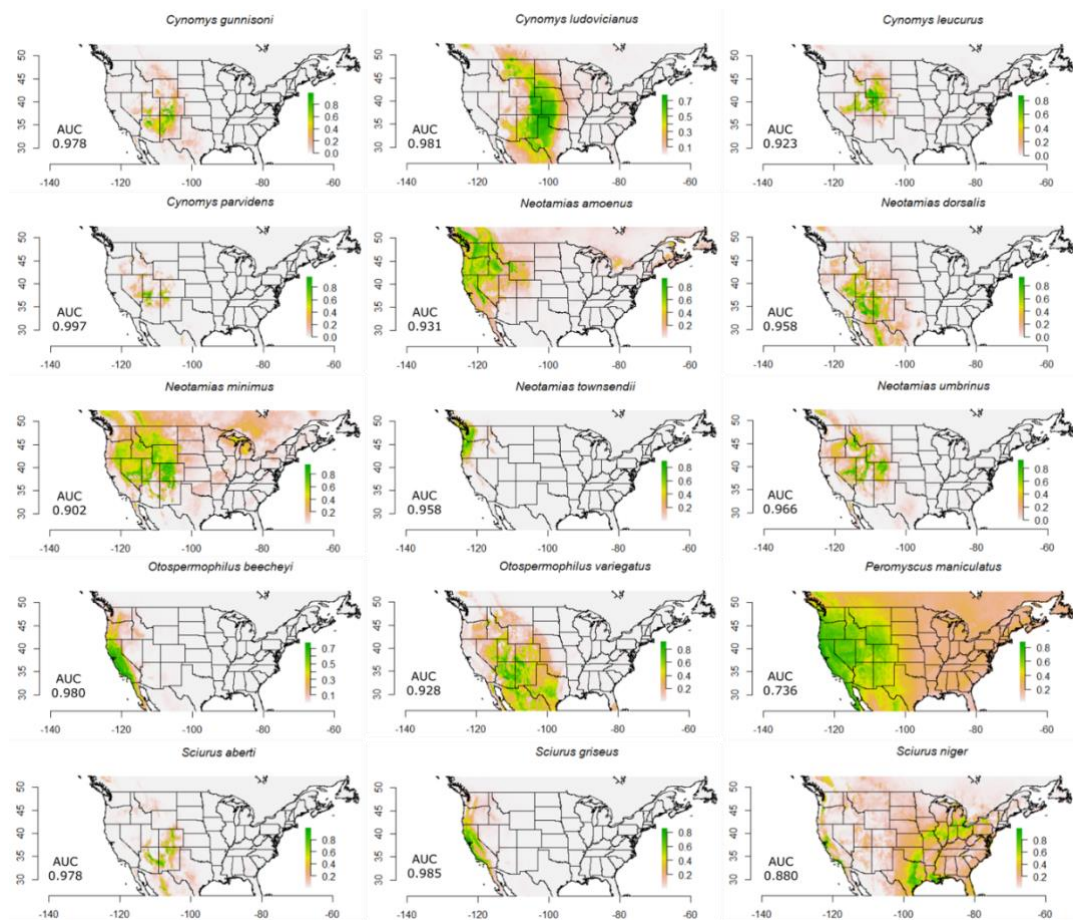


Figure 2.5 Projected distribution of all species used in the Multi-host models (Figure 2.2 a-f), *C. gunnisoni* (Gunnison's prairie dog), *C. ludovicianus* (Black-tailed prairie dog), *C. leucurus* (White-tailed prairie dog), *C. parvidens* (Utah prairie dog), *N. amoenus* (Yellow-pine chipmunk), *N. dorsalis* (Cliff chipmunk), *N. minimus* (Least chipmunk), *N. townsendii* (Townsend's chipmunk), *N. umbrinus* (Uinta chipmunk), *O. beecheyi* (California ground squirrel), *O. variegatus* (rock squirrel), *P. maniculatus* (North American Deer Mouse), *S. aberti* (Abert's squirrel), *S. griseus* (Western gray squirrel) and *S. niger* (Fox squirrel). Each model was built using 8 Worldclim variables (BIO2, BIO3, BIO8, BIO9, BIO15, BIO17, BIO18 and BIO19) and the colour scale represents predicted niche suitability.

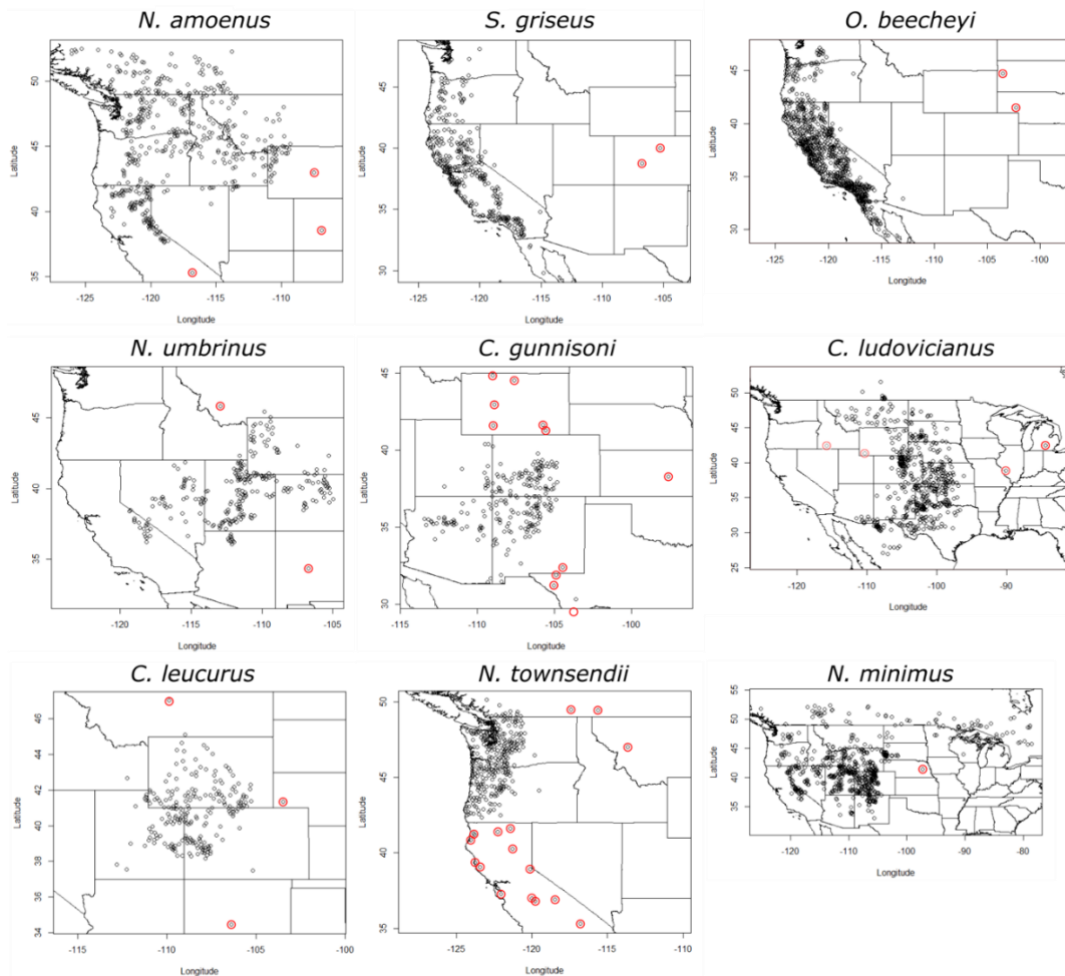


Figure 2.6 Points removed based on previous literature from the species occurrence points sourced from GBIF for *N. amoenus*, *S. griseus*, *O. beecheyi*, *N. umbrinus*, *C. gunnisoni*, *C. ludovicianus*, *C. leucurus*, *N. townsendii* and *N. minimus*. All removed points are shown in red circle.

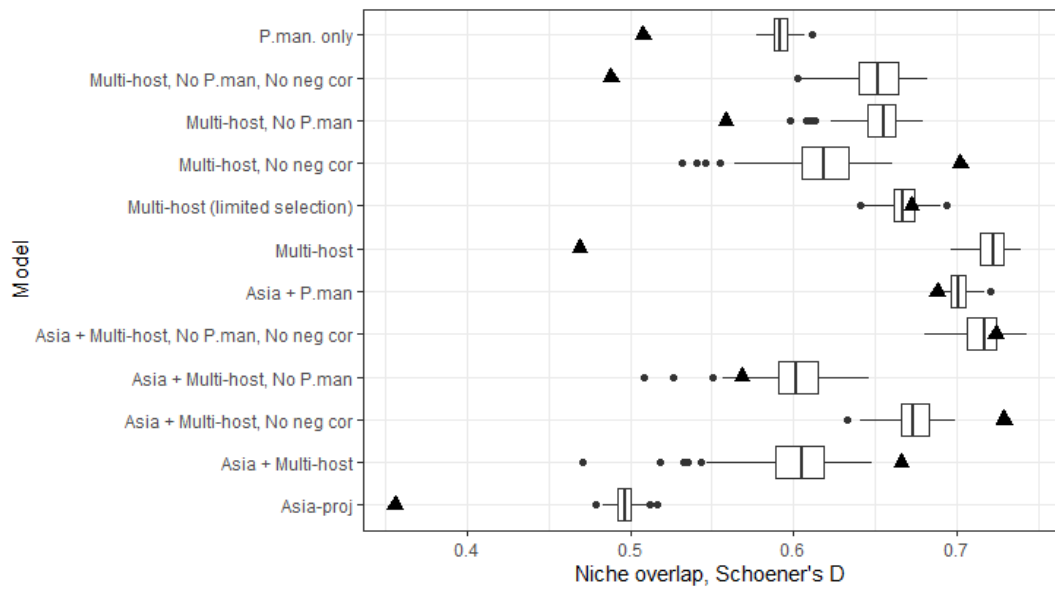


Figure 2.7 Niche overlap (Schoener's D) values between the projected Asian mod (Asia-proj) inclusive and exclusive of biotic factors and the North American model (N.Am-obs). Triangles are overlap values of threshold models (using the spec_sens threshold at which the sum of the sensitivity (true positive rate) and specificity (true negative rate) are highest). Boxplots represent the niche overlap values (Schoener's D) produced through the 100 replicates of each niche model. Visualisations of further threshold projections are shown in Figure 2.16.

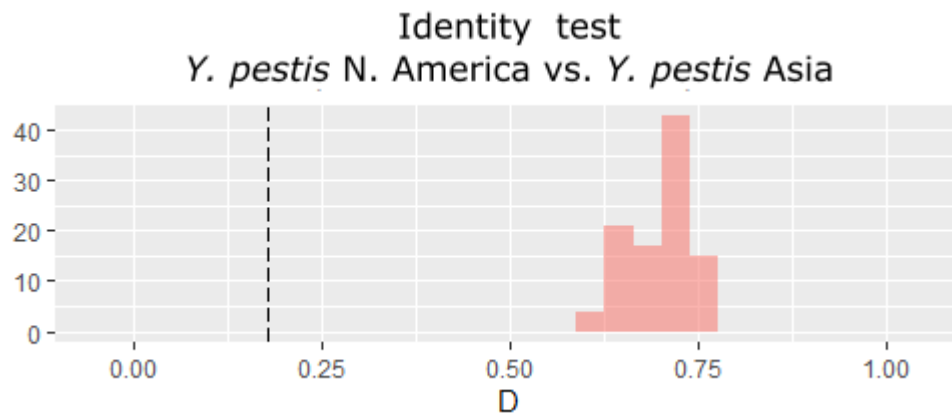


Figure 2.8 Identity test results. The black dashed line represents the actual niche overlap score of the niche comparisons between the native and invaded niche while the red shaded histograms represent 100 randomised model runs, where occurrence points are randomly selected from the pooled native and invaded occurrences. D is the niche overlap metric Schoener's D. The niche overlap value is significantly smaller (p -value = 0.01) than the null values indicating that the niches are non-identical. The niche overlap values recorded in this test and the background test are not the same as the calculated niche overlap as they were constructed using the combined Asian and North American regions which is a requirement for each test.

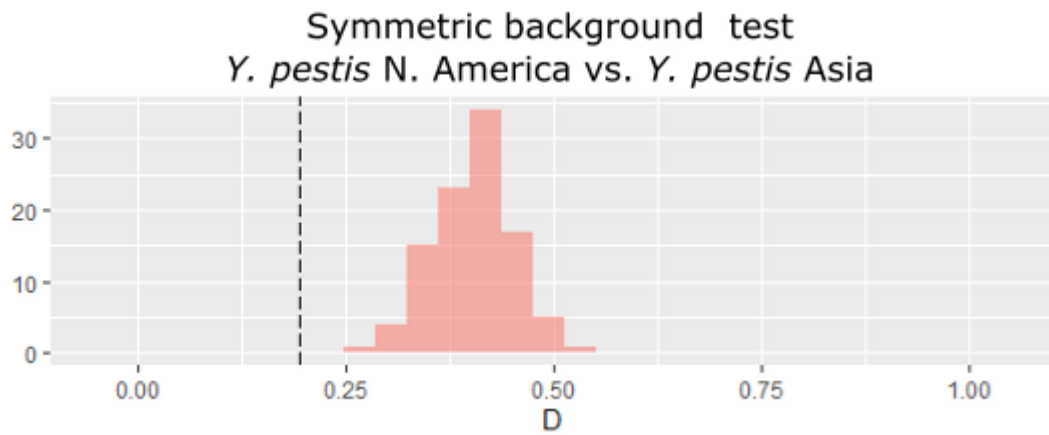


Figure 2.9 Background test results. The black dashed line represents the actual niche overlap score of the niche comparisons between the native and invaded niche while the red shaded histograms represent 100 randomised model runs, where occurrence points are randomly selected from background region. The niche overlap value is significantly smaller (p -value = 0.01 in both cases) than the null values indicating that the niches are less similar than expected given the difference in geographic regions in which they reside.

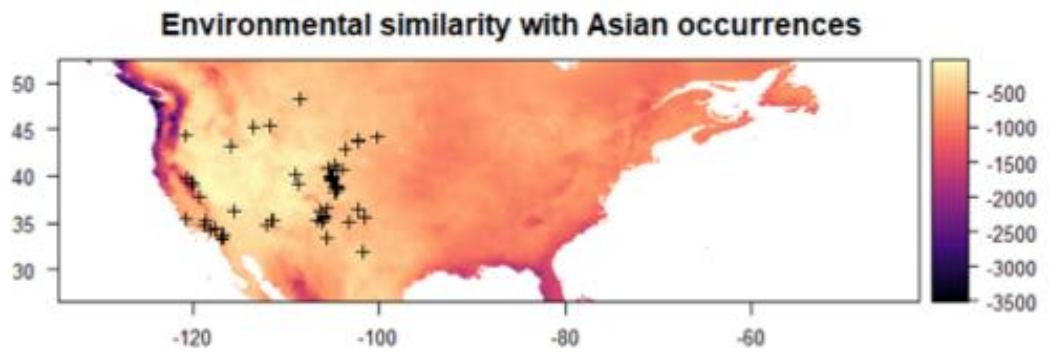


Figure 2.10 Multivariate Environmental Similarity surface plot. Increasing negative values indicate dissimilarity from the reference data, which was the climate data extracted from the Asian occurrence points. Crosses indicate North American occurrence points. Generally the North American occurrences are clustered in the areas of most analogous climate.

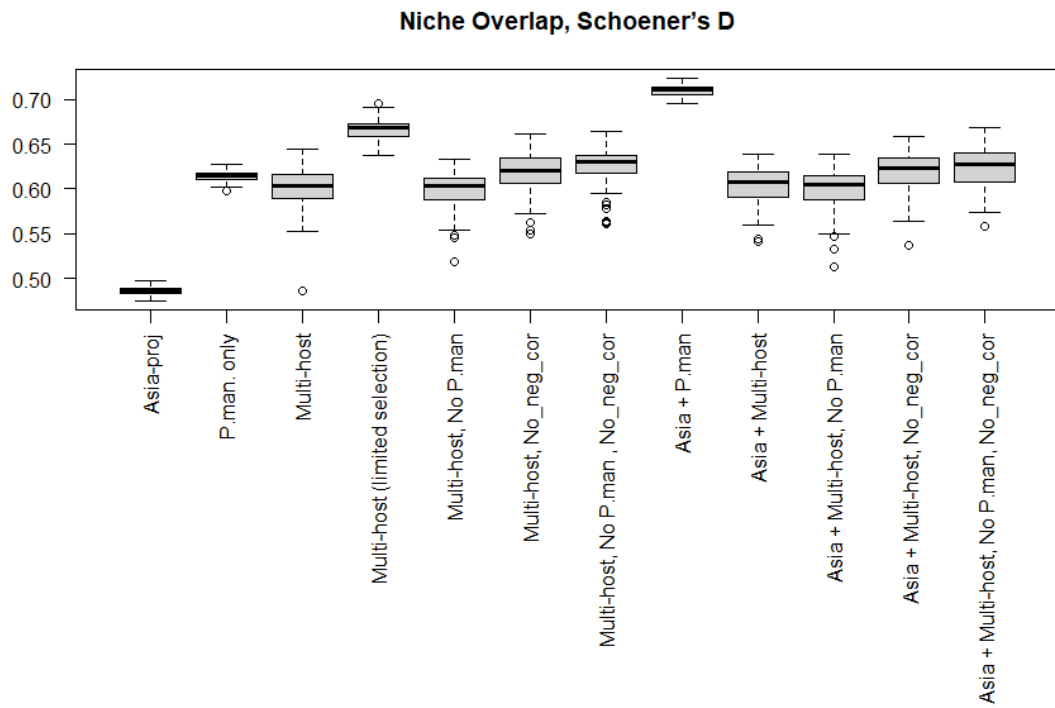


Figure 2.11 Niche overlap values of the observed distribution models using the most selected optimization settings across all models (Feature class = linear & quadratic, Regularisation multiplier = 0.5). Boxplots of the niche overlap values (Schoener's D) produced through the 100 replicates of each niche model.

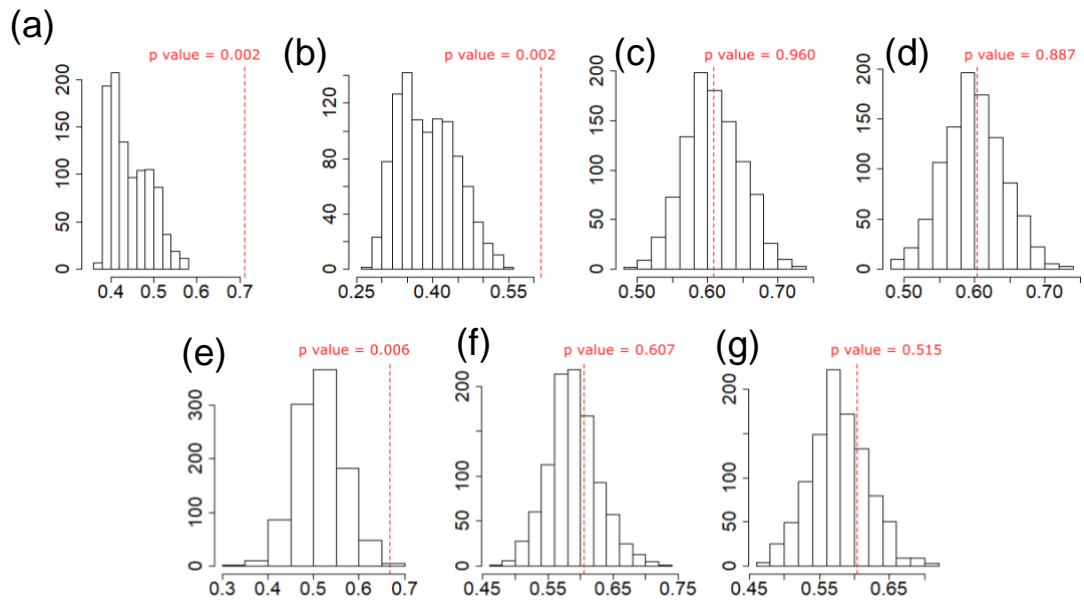


Figure 2.12 Null and observed overlap (Schoener's D) for models with the consistent LQ 0.5 settings. The overlap measured is between *Y. pestis*' North American niche and the predicted niche based on; (a) *P. maniculatus* and the Asian niche (Asia + *P. man*), (b) just a single host species distribution (*P. man*-only), (c) multiple hosts and the Asian niche (Asia + Multi-host), (d) only multiple host species (Multi-host), (e) Limited selection of multiple hosts (Multi-host limited) (f) multiple hosts with the Asian niche but without *P. maniculatus* (Asia + Multi-host, no *P. man*) and (g) multiple host species without *P. maniculatus* (Multi-host, no *P. man*). Histograms show the distribution from 1000 null simulations and red lines show observed values.

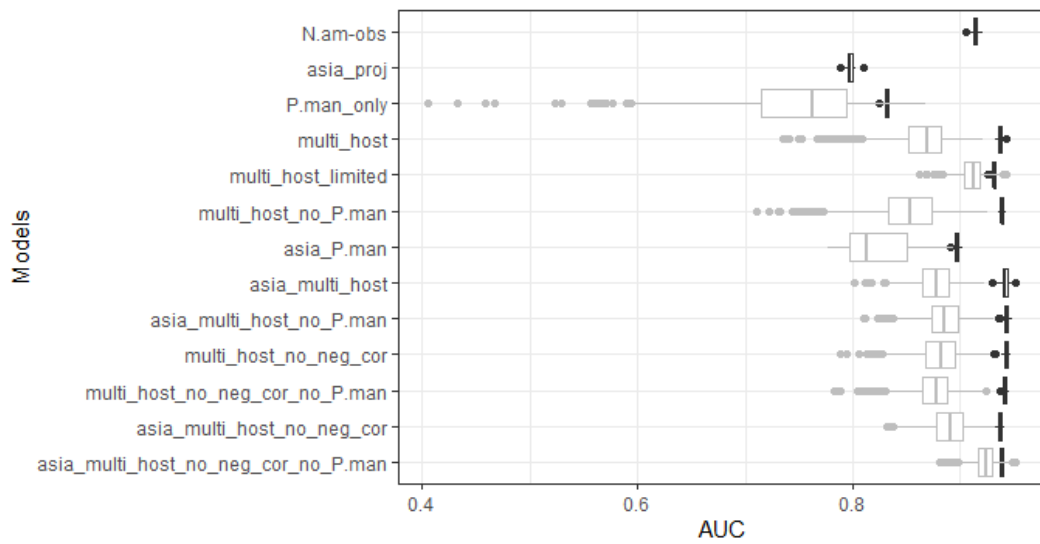


Figure 2.13 AUC values for the empirical and null models. Null model boxplots are in grey and the empirical models are in black. There is only an empirical AUC for the North American model (N.am-obs) and the Asian projection (asia_proj) as these were not simulated through the nulls.

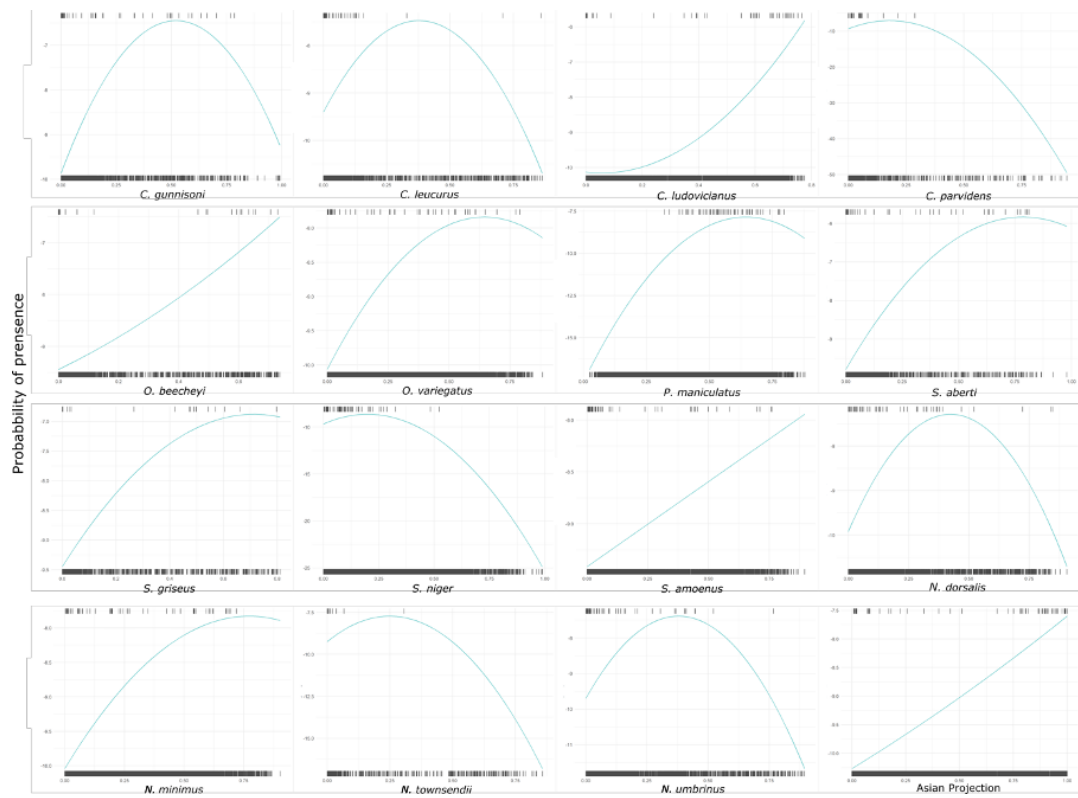


Figure 2.14 Univariate response curves from the Asia + Multi-host model. All variables show a predominantly positive trend with the probability of plague presence apart from *N. townsendii* and *C. parvidens* and *S. niger* which were removed in further supplementary models. *S. griseus* was further removed from these supplementary models as the permutation importance of this species was zero. Each of these species was retained in the primary “All species” models and was only removed in the supplementary models identified as “No negative correlation species” (no neg cor).

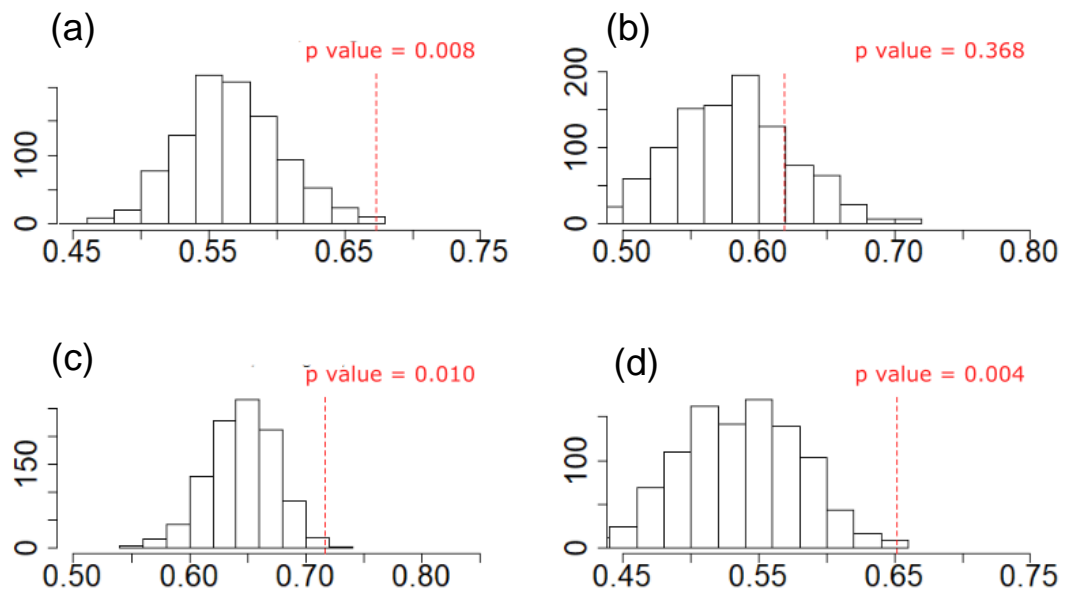


Figure 2.15 Null and observed overlap (Schoener's D) between *Y. pestis*' North American niche and the further models with negatively correlated species removed, *P. maniculatus* removed and both removed. (a) Multiple hosts and the Asian niche with negatively correlated species removed (Asia + Multi-host, no neg cor), (b) Multiple hosts with negatively correlated species removed (Multi-host, no neg cor), (c) Multiple hosts with the Asian niche, without *P. maniculatus* and the negatively correlated species (Asia + Multi-host, no neg cor, no *P.man*) and (d) Multiple hosts, without *P. maniculatus* and the negatively correlated species (Asia + Multi-host, no neg cor, no *P.man*)

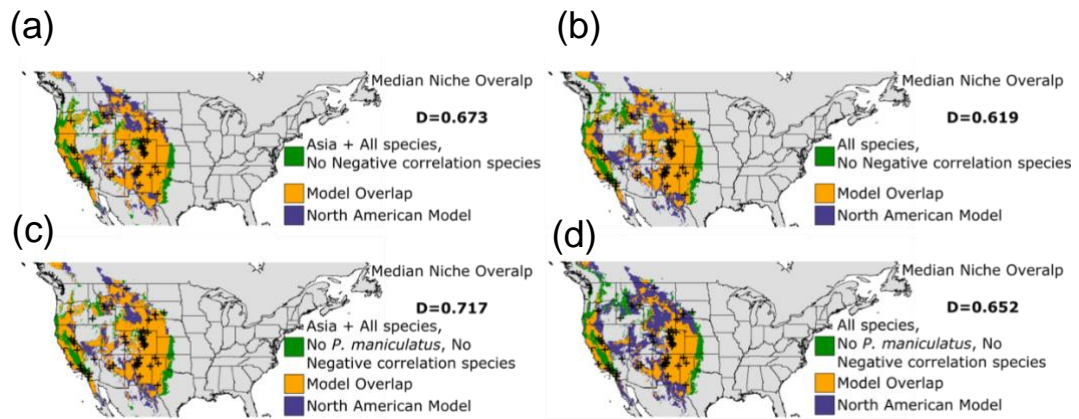


Figure 2.16 Differing projected distribution of *Y. pestis* across North America using a model fitted in North America and: (a) a model fit in Asia and projected to North America with only four positively correlated rodent species, (Asia + Multi-host, no neg cor in Table 2.4), (b) a model with only four positively correlated rodent species, (Multi-host, no neg cor in Table 2.4), (c) a model fit in Asia and projected to North America without *P. maniculatus* distribution included and only the three remaining positively correlated rodent species, (Asia+Multi-host, no neg cor, no *P.man* in Table 2.4) and (d) a model without *P. maiculatus* distribution included and only the three remaining positively correlated rodent species, (Multi-host, no neg cor, no *P.man* in Table 2.4).

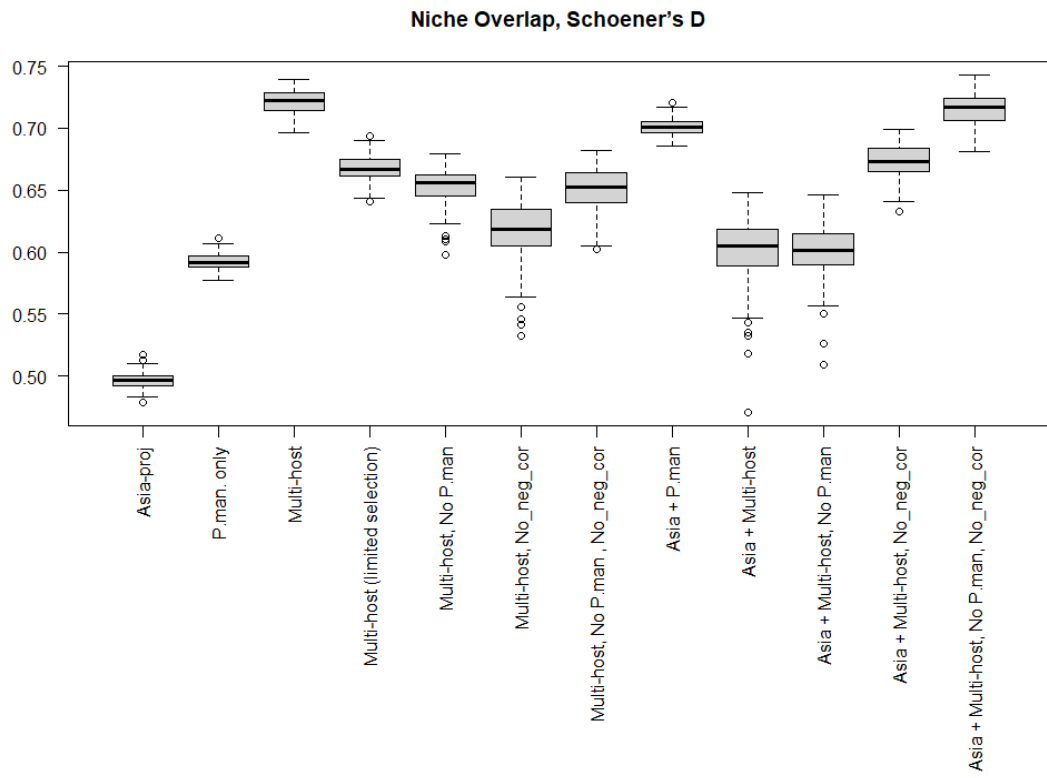


Figure 2.17 Niche overlap values of all the observed distribution models. Boxplots of the niche overlap values (Schoener's D) produced through the 100 replicates of each niche model.

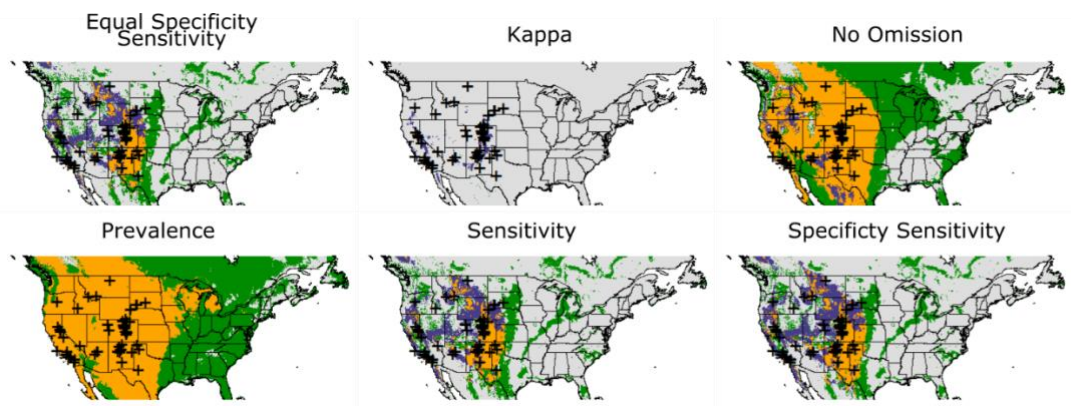


Figure 2.18 Figure 2.1 replicated using all threshold methods. The same trend of eastward over-projection is visible in using all threshold methods, apart from Kappa. The spec_sens threshold was selected as a conservative estimate.

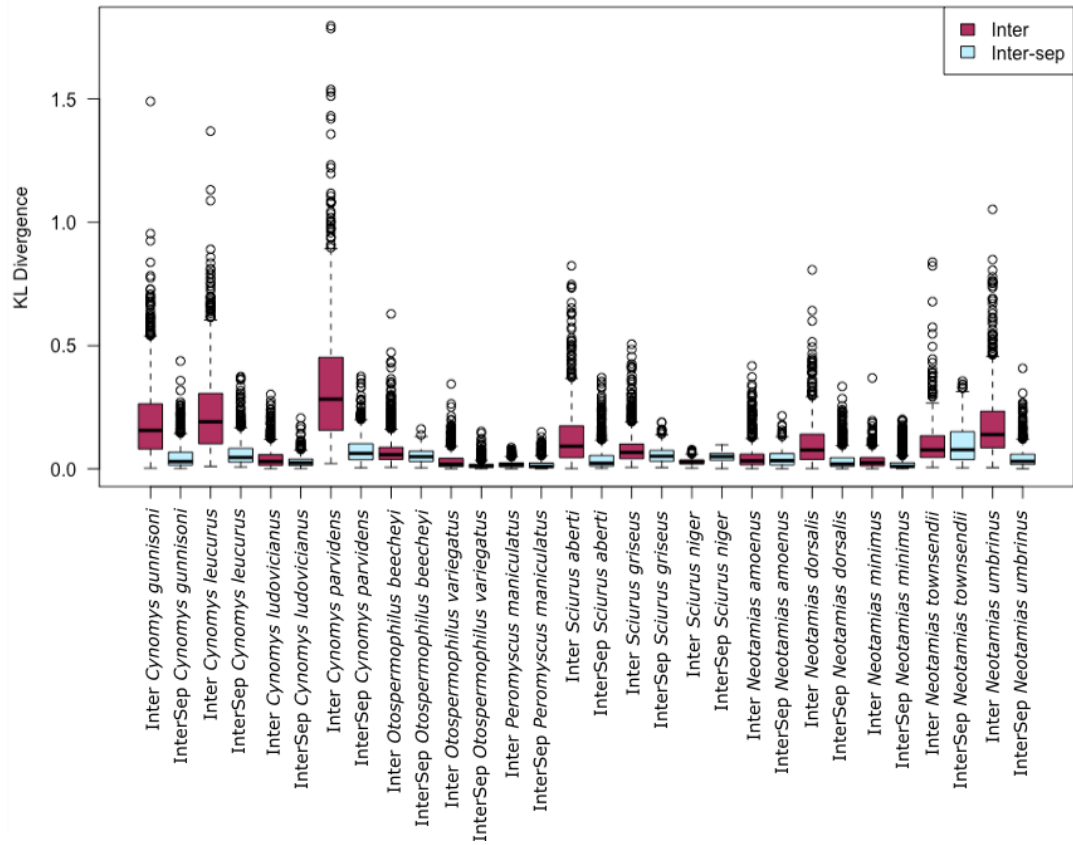


Figure 2.19 Comparison of between-species KL divergence values for the inter (maroon) and inter-sep (light blue) models in fauxcurrence. Each boxplot summarizes the KL divergence between the null and observed between-species DPD across 1000 replicates for a single species and model.

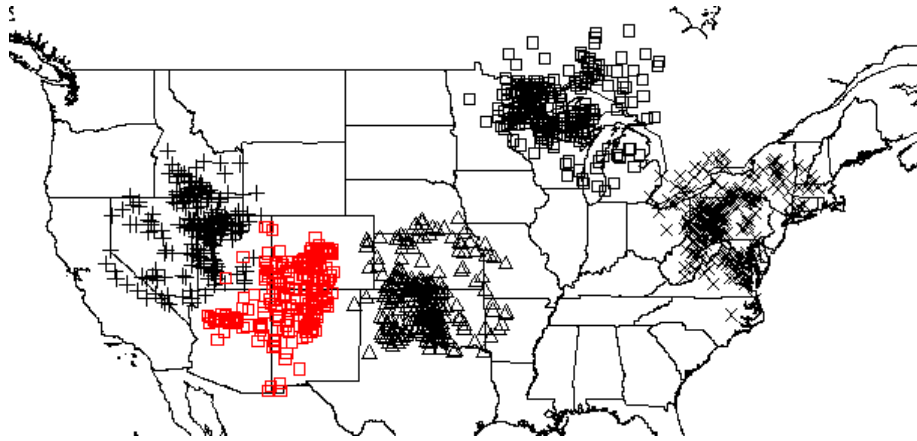


Figure 2.20 Four exemplar geographic null occurrences (black) empirical occurrences for *Cynomys gunnisoni* (Red). Each symbol has varying occurrence localities but a near identical distribution of pairwise distances (DPD) among occurrences.

2.5.3 Details of null modelling algorithm

The *'fauxcurrence'* package (version 1.1) was used to construct null occurrences in R (Osborne et al., 2021). The approach builds on a previous family of methods (Beale, Lennon, & Gimona, 2008; Algar, Mahler, Glor, & Losos, 2013) which iteratively improve the similarity in spatial structure of an initially randomized set of null points to that of a set of observed occurrences. Specifically, the method minimises the Kullback-Liebler (KL) divergence between the distribution of pairwise distances (DPD) among observed occurrences and the DPD among null occurrences.

When multiple host species are considered, spatial structuring can occur at two different scales. Occurrences for a single species will be spatially clustered. However, the range centroids of a set of host species may also be spatially clustered if the reservoir species are found in similar environments. *fauxcurrence* allows both of these levels of spatial structure (i.e. both within- and between-species distances) to be accounted for. For each replicate, the *"stg.1.only=TRUE"* option was first used to produce one randomized initial point for each species separated by between-species distances which were within the observed ranges. Null points were then generated for each species separately, using the point generated in the first step as a fixed seed point (using the option: *fix.seed.pts*). The fixed seed point cannot be replaced during the iterative procedure, so the approach guarantees that between-species distances in the null model are similar to those in the observed dataset. KL-divergence was considered to be sufficiently minimised when no improvement had occurred for 2000 iterations (option: *div.n.flat = 2000*). The above procedure was repeated to generate 1000 sets of host species occurrences.

When producing the initial point for each species, the package can consider either one between-species DPD for each pair of species (options: *inter.spp=TRUE, sep.inter.spp=TRUE*; referred to as the *inter-sep* model) or one between-species DPD for each individual species (in this case, the DPD describes the distances between all points of that species to all heterospecific points; options: *inter.spp=TRUE, sep.inter.spp=FALSE*; referred to as the *inter-spp* model). 1000 replicates were produced using each of these methods and compared the KL-divergence of the final

nulls. This showed that the *inter-sep* model produced substantially smaller KL-divergence values (Figure 2.19), so the *inter-sep* replicates were used as the final null models. An example of the null distributions produced for the species *Cynomys gunnisoni* can be seen in Figure 2.20.

To ensure the null model generation procedure was repeatable, the randomization seed was fixed in the randomization seed in *R* before each iteration. This was achieved by first generating a random integer (using the code: `my.seed <- sample.int(.Machine$integer.max,1)` in *R*) which was recorded and used to set the seed for the *R* random number generator (using the code: `set.seed(my.seed)`). This was done before producing the seed points for each null replicate, and then again before producing each species' full set of null points for each replicate. The random seeds used are available in figshare (<https://doi.org/10.6084/m9.figshare.17025230.v1>). By setting these seeds before running *fauxcurrence* with the input data and options, the results can be reproduced exactly.

3 The role of reservoir species in mediating plague's dynamic response to climate.

3.1 Introduction

The distribution and dissemination of the plague bacterium *Yersinia pestis* is inextricably linked to climate dynamics (Stenseth et al., 2006). Studies from across plague's near global distribution have observed climate can influence plague distribution and dynamics at a range of scales (Kausrud et al., 2007, Ben-Ari et al., 2010, Schmid et al., 2015, Kreppel et al., 2016). The observation that plague systems respond to climate is consistent across all reservoir systems, however, the response is far from homogenous (Ben-Ari et al., 2011, Xu et al., 2011). Heterogeneous responses to climate are to be expected given the huge range of reservoir species, vector species and environments in which *Y. pestis* persists (Stenseth et al., 2008, Mahmoudi et al., 2020). The dynamics of reservoir species, vectors, and the bacterium are all key elements in understanding the mechanisms which drive epidemic and enzootic cycles (Ben-Ari et al., 2011, Xu et al., 2015, Cui et al., 2020). Assumptions are often made about the ecological mechanisms that link climate to the transmission of plague within human populations, based on the response of selected reservoir species which may not be representative of the complex multi-reservoir species system under investigation (Kausrud et al., 2010). In this work I attempt to address one such assumption by using niche modelling methods to test the hypothesis presented by Xu et al. (2011), that plague's heterogeneous response to climate (precipitation) across China during the Third Pandemic was driven by differential responses of reservoir species to changing precipitation.

The complexity of plague systems increases as the spatial and temporal scale of the system under investigation increases, therefore although mechanistic pathways linking climate and *Y. pestis*' responses are calculable at a narrow scale (e.g. one reservoir species in one region) this becomes more difficult at larger scales (Ben-Ari et al., 2011). When moving between local-scale or regional studies to much larger continental or global scales the mechanisms observed at one scale seldom hold true when applied to the other; this is known as the transmutation issue and is an area of

continued investigation across macroecology (O'Neill, 1977, McGill, 2019). The range of differing environments in which *Y. pestis* persists further confounds this issue. Although generally *Y. pestis* is mostly found in reservoir species within arid to semi-arid environments, these are not the only environments where *Y. pestis* persists. For example within China *Y. pestis* reservoirs are found in both arid northern environments as well as much wetter southern environments (Xu et al., 2011).

Y. pestis was incredibly active in the northern and southern reservoir regions during the Third Pandemic in China. The Third Pandemic began in Yunnan province in 1772 C.E., but the initial spread was gradual, with cases outside of Yunnan only reported \approx 100 years later (Benedict, 1996). Cases across Yunnan and eventually the rest of southern China began to increase from 1850 C.E. *Y. pestis* entered Hong Kong in 1894 C.E. and was subsequently transmitted globally through international shipping routes and the peak in cases across China was reported shortly after that, at the turn of the twentieth century (Ben-Ari et al., 2012). In northeast China the peak in plague mortality was recorded from 1910-1920 C.E. and the Third Pandemic persisted across China until the second half of the twentieth century with a final very large peak in cases recorded immediately prior to the start of the plague control program around 1950 C.E. (Benedict, 1996, Ben-Ari et al., 2012).

During the Third Pandemic, the reservoir species that maintained *Y. pestis* inhabited very different environments in North and South China. North China had a predominantly arid to semi-arid environment, which is most common to *Y. pestis* reservoir species, while South China is within the influence of the South Asian Monsoon and thus experiences high annual precipitation (Ning and Qian, 2009). Human plague records have revealed that plague intensity varied heterogeneously across China during the Third Pandemic (Xu et al., 2011). Generally, across North China plague intensity increased as wetness increased, whereas the opposite was observed in South China with plague intensity correlating with dryness. In both regions extreme precipitation events had the opposing impact, with extreme wetness and dryness in North China correlating with decreased plague intensity and extreme wetness and dryness in South China correlating with increased plague intensity (Xu

et al., 2011). When comparing unlagged (henceforth contemporary) climate data to plague intensity, Xu et al. (2011) found a hump-shaped relationship between plague intensity and a dryness/wetness index in the North suggesting that extreme wetness and dryness are detrimental to plague intensity. Plague intensity in the North was observed to correlate with periods of heavy rainfall in the prior year. The opposite response in contemporary and lagged data was observed across Southern China with a U-shaped contemporary response and a correlation between high plague intensity and dryness (Xu et al., 2011). In the North, extreme precipitation events may have a detrimental impact on species through flooding of burrows, while in the south extreme wetness may drive behavioural change in commensal rodents pushing them into closer contact with humans to avoid flooding. This heterogeneous response to extreme wetness may also be impacted by human population density differences as larger human populations correlate with larger commensal rodent populations, which could compound the impact extreme wetness may have in densely populated areas such as Southwest China (Sun et al., 2019).

Xu et al. (2011) hypothesised that these results were driven by the response of reservoir species to precipitation in each region. In the North, this is consistent with a bottom-up trophic cascade, posited as a driver of plague epizootic cycles in some regions (Collinge et al., 2005). This hypothesis proposes that reservoir species populations, as well as the population of vector species, are positively impacted by increased precipitation via increased primary productivity, leading to greater food resources for reservoir species within these precipitation-limited environments. Xu et al. (2011) further suggested that reservoir populations become decoupled from this hypothesis in extreme precipitation scenarios, where the reservoir populations may decrease due to increased mortality through the flooding of burrows. In contrast, the trophic cascade mechanism is not consistent with the findings in South China where dryness is observed to correlate with plague intensity. The exact mechanism driving this correlation is still unknown. *Rattus spp.* which are more common in the South, are human-commensal rodents whose populations have been observed to correlate negatively with precipitation in this region (Chen, 1996). Additional species, such as *Microtus fortis*, live on river and lake banks and are

therefore highly sensitive to precipitation due to changing water levels, which may partially contribute to the observed plague-dryness relationship (Zhang et al., 2010, Xu et al., 2011). Further research is required on host and vector systems in this region to determine the response of reservoir species to precipitation and the exact mechanisms involved.

In the North, the correlation between precipitation and plague intensity observed by Xu et al. (2011) is consistent with the prevailing trophic cascade hypothesis which is regularly applied to plague systems, whereas in the South the mechanisms are much harder to identify and are inferred from the findings of studies of individual species across this region. Ideally the role of reservoir species in mediating the impact of precipitation on plague intensity could be tested using temporal records of reservoir species population sizes. However, such records are only available for the Pre-Balkhash region of Kazakhstan and not across China therefore alternate methods must be devised (Davis et al., 2004). In the absence of time-series of reservoir population sizes, an alternative approach is to use data on climatic suitability of reservoir species estimated from environmental niche models (ENMs). ENMs can be used as a reasonable proxy for abundance, however, results should be interpreted with caution (Weber et al., 2017, A. Lee-Yaw et al., 2021). These models can then be hindcast to estimate responses of species to past climatic change. Hindcasting is the projection of a model into past conditions, as opposed to forecasting where a model is projected into estimated future conditions. In the context of niche modelling, hindcasting is the projection of ENMs into historical climate space to estimate how niche suitability may have varied under historic climate conditions. Hindcasting has been used to investigate past species distributions and gain insight into extinctions (Martínez-Meyer et al., 2004), areas of refugia (Svenning et al., 2008), and migration pathways (Waltari and Guralnick, 2009). Species past distributions are incredibly important in understanding both current distributions (Svenning et al., 2008, Maiorano et al., 2013) and how species may react to changing climates. Literature in this area generally focusses upon reconstructing species distributions many thousands of years in the past with much research focussed on changes in distributions which occurred prior to and following the Last Glacial Maximum

(Nogués-Bravo, 2009). Hindcasting within recent history (e.g. the past 100-300 years) may be used to test specific hypotheses rather than purely to construct the past distribution of a species (Dobrowski et al., 2011). Here, I use ENMs and hindcasting to test the hypothesis put forward by Xu et al. (2011) that plague's heterogeneous response to precipitation across China is driven by consistent differences in how host species respond to precipitation.

To test the hypothesis presented by Xu et al. (2011) the response of reservoir species to precipitation must be tested. I apply an ENM approach which is well suited to testing the hypothesis as I can assess the impact that precipitation may have upon a broad range of species from across China using occurrence data, with contemporary and historic climate data. Based on the hypothesis presented by Xu et al. (2011), I made three predictions. Firstly, to test if precipitation drove the variation in plague intensity response through impacting reservoir species, I predicted that precipitation variables would be important in determining the species niche suitability with a positive relationship between niche suitability and precipitation variables in the North and the opposite in the South. Secondly, I predicted that the species niche suitability would correlate with precipitation in differing directions for northern and southern species through time. I predicted a positive correlation with precipitation in northern species and a negative correlation in southern species. Thirdly, I predicted that the niche suitability of northern species would increase during wet periods while for southern species the niche suitability would increase during dry periods. All these predictions were tested using ENM and hindcasting methods in combination with historical climate data (ISIMIP) to estimate niche suitability of the selected reservoir species during the Third Pandemic.

3.2 Materials and methods

3.2.1 Data acquisition

3.2.1.1 Host species localities

A broad range of reservoir species were selected across China with distributions ranging from one Chinese province to global coverage. The selection was based on Mahmoudi et al. (2020) who identified plague reservoir species throughout the world. To be selected, reservoir species had to be present in China or Tibet (30 species). This list of species was further supplemented by species in which *Y. pestis* had been isolated and sequenced across China (Cui et al., 2013) (16 species). In total 38 species were initially selected as there was overlap in the species suggested through the two sources; five of these species were subsequently excluded from the analysis due to limited occurrence records (ten or fewer). Reservoir species were predominantly from the order Rodentia, but also included Carnivora, Artiodactyla, and Lagomorpha (Chapter 4 Appendix Table 3.5). The maintenance of *Y. pestis* in the wild is predominantly driven by rodent species, however, *Y. pestis* has been identified as present in a range of further species particularly predators, such as mustelids (*Meles leucurus*, *Mustela eversmanii*, *Mustela nivalis*) (Salkeld and Stapp, 2006). The exact role of predator species in the maintenance of plague is unknown and warrants further investigation (Salkeld and Stapp, 2006). I therefore retained these species as well as Artiodactyla, and Lagomorpha species to assess the potential impacts of species across a broad range of niches.

Species occurrences were sourced from the Global Biodiversity Information Facility (GBIF, full list of GBIF DIOs Chapter 4 Appendix Table 3.5) using the *rgbif* package (Chamberlain and Boettiger, 2017). Occurrences were clipped to a broad Eurasian extent (-10 E, 150 W, 0 S, 80 N). Many of the species have a distribution much broader than China (eg. *Mus musculus*, *Rattus norvegicus*) therefore, this broad Eurasian extent was selected to avoid imposing bias upon the models by limiting the available climate space to only the selected region of investigation (Feng et al., 2019). All species occurrences were then examined to identify potentially erroneous occurrences which I attempted to validate and if unsuccessful removed (Chapter 4 Appendix Table 3.5). All species occurrence data were thinned such that there was

never more than one occurrence per cell of climate data (0.5° resolution, see below) using the `elimCellDups` function in the *enmSdm* package (Morelli et al., 2020, Sillero and Barbosa, 2021). Species with ten or fewer occurrence points were then excluded (*Microtus juldaschi*, *Alexandromys fortis*, *Neodon fuscus*, *Eothenomys eleusis* and *Spermophilus alashanicus*) which decreased the number of species to 33.

3.2.1.2 Historical climate data

I used the bias-corrected atmospheric climate simulation data produced through the Inter-Sectoral Impact Model Intercomparison Project (ISIMIP) as it covered the period from the start of the Third Pandemic (1772 C.E.) until relatively recently (2005 C.E.) and includes multiple climate scenarios for comparison. The ISIMIP project aims to produce data which can be used to investigate the impacts of climate change across a range of disciplines and scales and has produced several iterations of global climate models (Warszawski et al., 2014). I used data from the ISIMIP2b project, specifically two of the bias-corrected ISIMIP data sets based on differing global climate models (GCMs), the National Oceanic and Atmospheric Administration (NOAA) geophysical fluid dynamics laboratory GCM, henceforth GFDL and the Institut Pierre-Simon Laplace GCM, henceforth IPSL. The resulting ISIMIP GCMs represent two different climate scenarios and hence help capture the uncertainty in past climate variation and reconstruction. The climatic variables included mean, max and min daily temperature as well as total daily precipitation. Comparing the two scenarios across China (split into northern and southern regions, discussed below), I observed (**Error! Reference source not found.**) broader range in annual precipitation and marginally higher median annual mean precipitation across the North in the GFDL scenario. In the South there is very little variation in annual precipitation between the scenarios. In both regions the mean near-surface air temperature is roughly 1 °C warmer in the GFDL scenario.

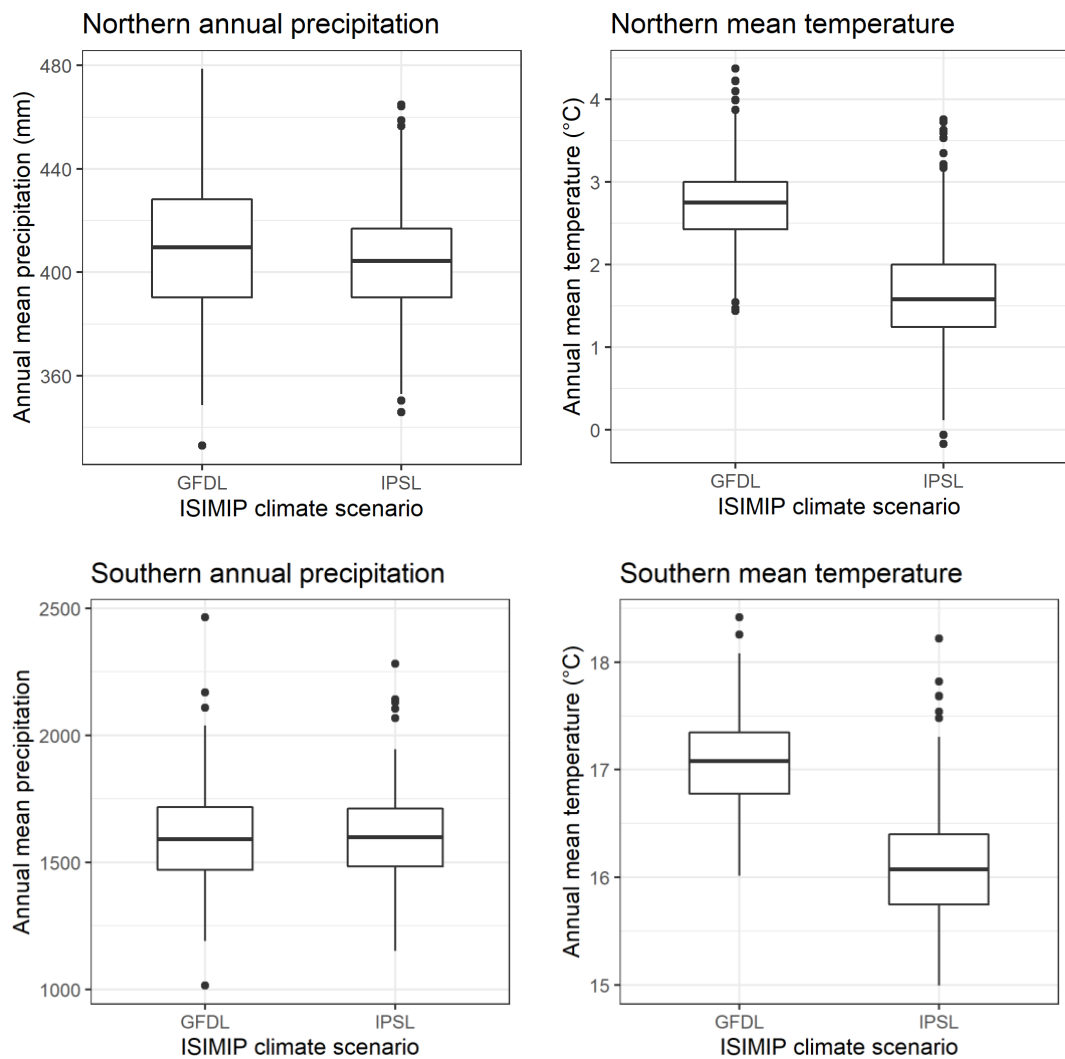


Figure 3.1. Comparison of mean annual precipitation and temperature values between the two climate scenarios (GFDL & IPSL) between the two regions (North and South China). Boxplots are constructed from the mean annual data for across all years in the ISIMIP data (1661-2005 C.E.).

For use in ENMs, climate data were converted into the 19 bioclimatic variables (biovars) produced in the WorldClim data which are used near ubiquitously across niche modelling (Fick and Hijmans, 2017). This required monthly minimum and maximum temperature and total precipitation for each pixel, which was calculated using the biovar function of the *dismo* package (Hijmans et al., 2017). Precipitation in ISIMIP is in the form of precipitation flux ($\text{kg m}^{-2} \text{s}^{-1}$) which requires conversion to precipitation per day (mm/day) prior to conversion to monthly data for the biovar calculation. To convert from precipitation flux to precipitation per day, the density of water was assumed to be 1 kg L^{-1} , precipitation flux was subsequently multiplied by 86400 to convert to precipitation (mm) per day. As a check of this calculation the

annual mean precipitation values across the combined North and South regions (**Error! Reference source not found.**) were checked against current and past precipitation records for China (Guo et al., 2020, Li et al., 2021). All variables were then transformed from daily to monthly resolution and then used to calculate 19 annual biovars for each year in the data (1661-2005 C.E.).

3.2.2 Niche modelling

I constructed ENMs for each of the selected reservoir species using contemporary climate data and then projected each of these models into historical climate space to estimate the variation in each species niche suitability through time. The niche models for the 33 selected reservoir species were fitted using the Maxent algorithm (Phillips et al., 2006), implemented and tuned using the *ENMeval* package (Muscarella et al., 2014). The block spatial partitioning method was used throughout, which splits the occurrence data into four geographical bins of equal numbers (k), the models are then tested using a k-1 holdout method with each of the bins. The block method is recommended when a model is transferred across space or time (Wenger and Olden, 2012, Muscarella et al., 2014). To construct a contemporary ENM for each species I constructed a time averaged climate data set from the annual climate data. I followed the WorldClim2 dataset and built a time averaged raster of each of the 19 bioclimatic variables across the 1970-2000 C.E. period (Fick and Hijmans, 2017). All the niche modelling analysis, apart from the variable selection and the model optimisation were completed in parallel across both climate scenarios. The variable selection and model optimization for each species were completed using only the GFDL scenario to limit the variation between scenarios attributable to model settings.

3.2.2.1 Variable selection

I selected a subset of the least correlated climate variables for constructing the species ENMs (Sillero and Barbosa, 2021). As the hypothesis I am testing focuses on the impact of precipitation, I ensured the inclusion of precipitation variables by calculating the correlation of all precipitation and temperature values independently across the selected China plague extent (CPE, Latitude 71.2 E, 135.1 E, Longitude 15.4 N, 54.6 N). The least correlated variables were identified for each variable type

through pairwise correlation analysis and visualised through cluster dendrograms and correlograms (Chapter 4 Appendix Figure 3.7). Four precipitation and four temperature variables were selected to use consistently across all species and I validated this selection against the variables selected for each species individually using the same methods but without separating temperature and precipitation variables. The eight most selected variables coincided with the independently selected variables validating their use across all ENMs (Figure 3.2).

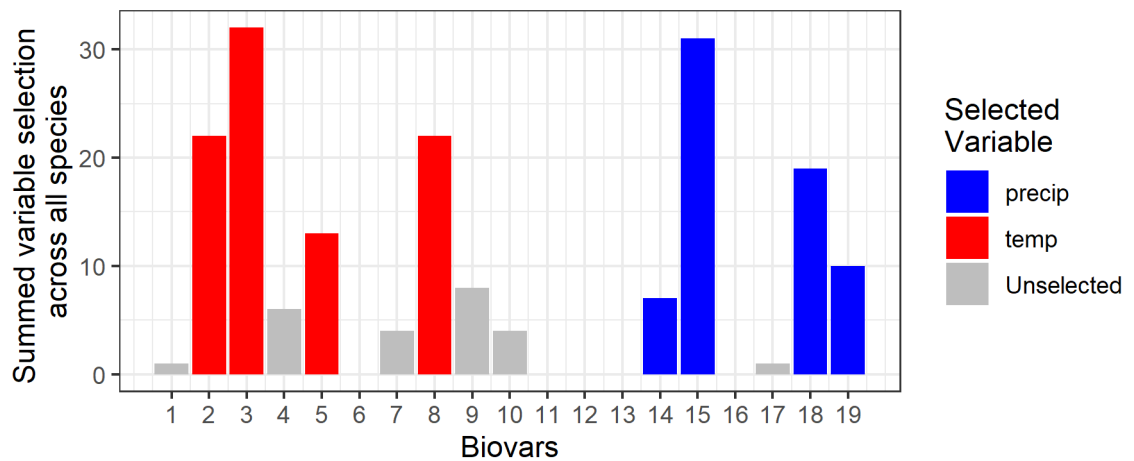


Figure 3.2. Bioclimatic variables selected for use in ENM of all reservoir species. The colours indicate the variables selected when temperature and precipitation variables were selected in isolation. The height of the bars is the summed total that each variable was selected when precipitation and temperature were not selected for in isolation.

3.2.2.2 Background selection

Background occurrences were selected from a rectangular buffer of 4 degrees around each occurrence. Ten thousand background points were selected across each species which given the limited extent of some of the species means that all pixels for some species may have been sampled as background points, but no resampling was completed.

3.2.2.3 Model tuning

Table 3.1. Minimum niche overlap between each species with multiple best models selected through optimization

Species	Minimum Niche overlap (D)
<i>A. agrarius</i>	0.98
<i>C. erythraeus</i>	0.97
<i>M. eversmannii</i>	0.92
<i>M. leucurus</i>	0.51
<i>M. musculus</i>	0.98
<i>M. nivalis</i>	0.93
<i>M. unguiculatus</i>	0.68
<i>R. norvegicus</i>	0.77
<i>R. rattus</i>	0.96
<i>V. vulpes</i>	0.96

I used the ENMevaluate function in the *ENMeval* package to tune the regularisation multiplier and feature class of each of the models to determine the best model settings for each species (Muscarella et al., 2014). The included feature classes were linear (L), quadratic (Q), hinge (H), threshold (T) and product (P) feature classes (in the following combinations L, LQ, LQH, H, LQHT, LQHTP) and the regularization multiplier ranged from 0.5 to 4 in increments of 0.5. The best model for each species was selected by sorting on the lowest average omission rate and highest AUC sequentially (Galante et al., 2018). Based on previous findings (Fell et al. (2022), Chapter 2) that the selection of the best model settings changed when optimization was repeated due to variation in the random selection and subsequent partitioning of the background data and marginal variation inherent in models produced using Maxent, I replicated optimization 50 times (Sillero and Barbosa, 2021). A variation in best model selection was observed for 10 of the species, notably those with large Eurasian extents, as the background data in more range-restricted species will not have varied due to the complete sampling of the available area (Figure 3.3). For each of the 10 species, the niche overlap of all possible best settings was calculated to

assess the variation in model prediction (). Niche overlap (Schoener's D) was generally very high across the different models for each species (> 0.9), apart from for two species (*M. leucurus* and *M. unguiculatus*). Hence the modal settings were used for all further analysis.

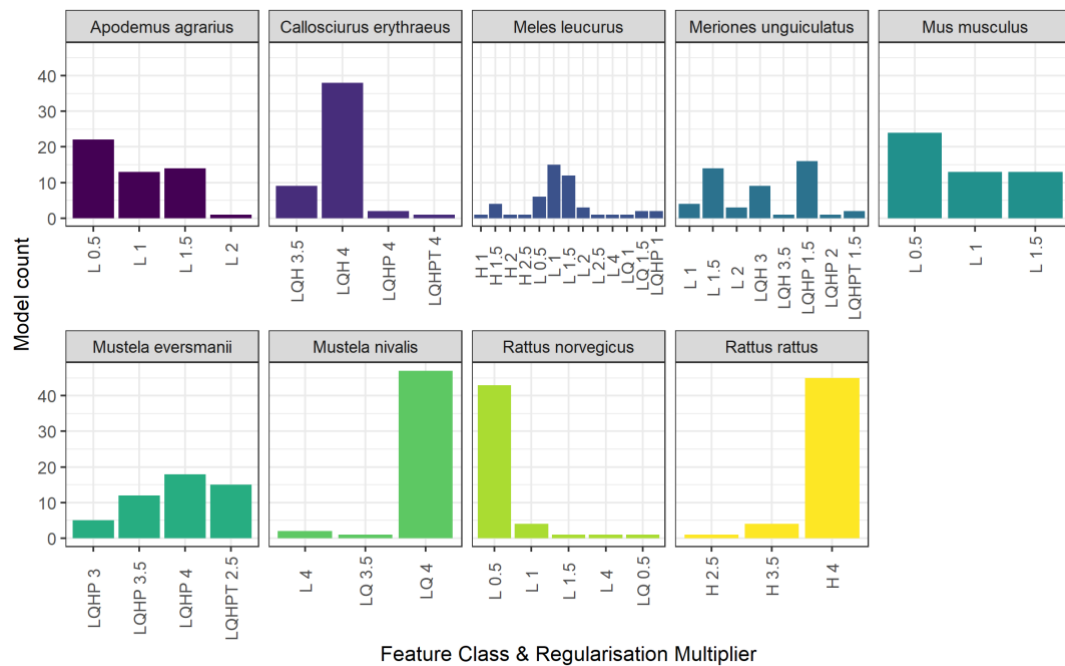


Figure 3.3. Varying feature class and regularisation multiplier across repeated optimizations for reservoir species where consistent settings were not selected through model optimization.

A final best model was then produced using the optimized settings and the results, permutation importance and response curves for each variable were extracted for further analysis. The response curves in combination with the permutation importance of each variable provided the first opportunity to test the hypothesis that host species' response to precipitation drove the heterogeneous response of plague. The AUC values of these final models were calculated across the full Eurasian range of each species and over the CPE which was used for hypothesis testing. The AUC of each species was calculated across the CPE using the env.evaluate function within the ENMtools package (Table 3.2) (Warren et al., 2021).

3.2.2.4 Hindcasting across multiple climate scenarios

The niche models constructed for the contemporary climate period were then hindcast into the annual historic climate space. There are four differing scenarios over which I construct hindcast niche projections. All the niche modelling analysis, apart from the variable selection and the model optimisation were completed in parallel across both climate scenarios (GFDL & IPSL). The variable selection and model optimization were completed using the GFDL scenario to limit the variation in the models attributable to model settings. Within both scenarios I hindcast clamped and unclamped versions of each model. Model clamping is the process which constrains model features to remain within the range of the training data, which confines the projection of an ENM to analogous climate space, limiting the uncertainty in model extrapolation (Elith et al., 2010).

3.2.3 Hypothesis testing

To test the hypothesis, I compared the trend in species environmental niche through time to the precipitation in the northern and southern regions. In line with Xu et al. (2011), human plague data across China were used to define the northern and southern regions (Figure 3.4). The North-South divide is located at 31°N with each region encompassing the natural plague reservoirs in the North and South and the divide placed in the central region where there are no natural plague reservoirs due to extensive land cultivation and crop plantations (Xu et al., 2011).

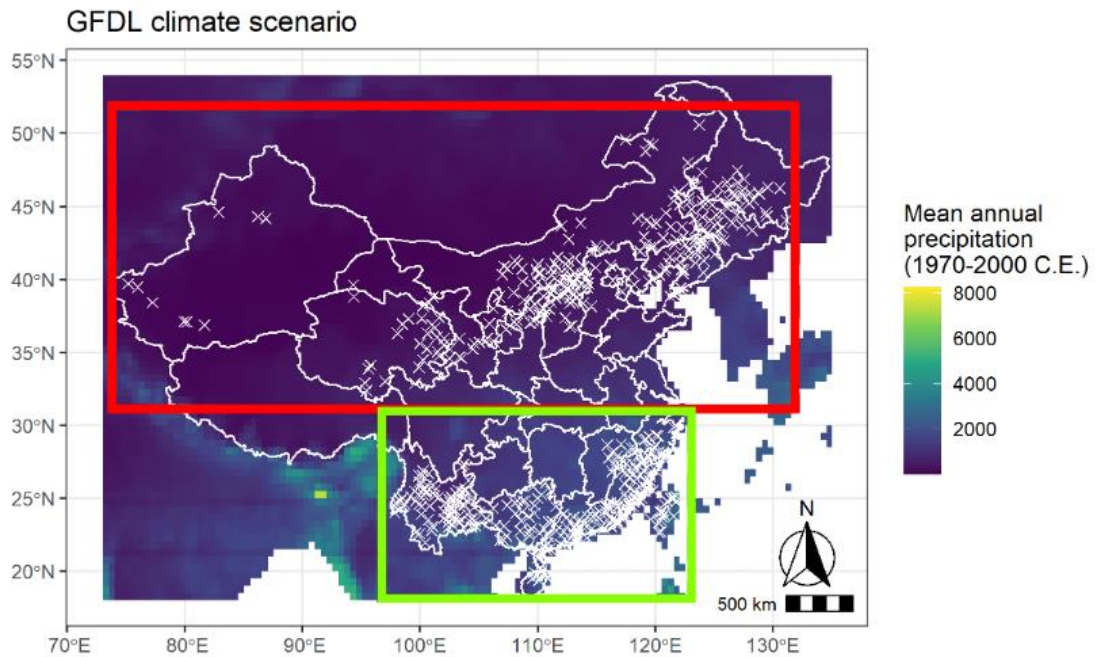


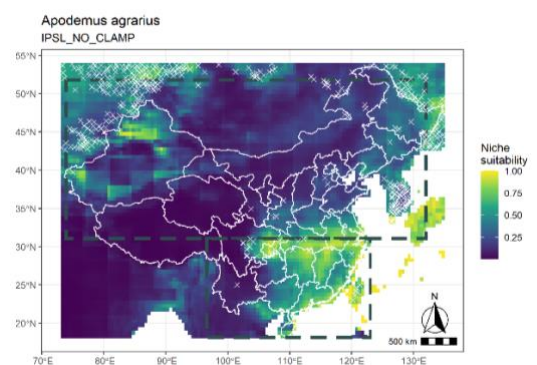
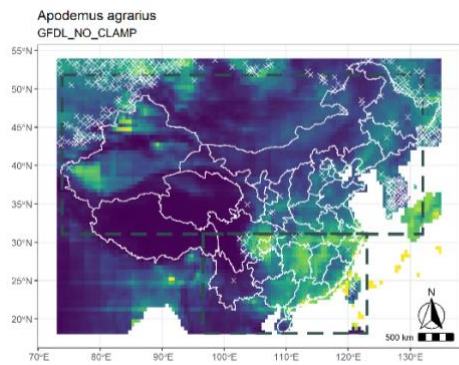
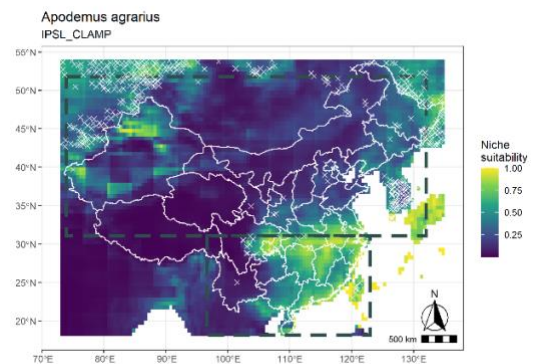
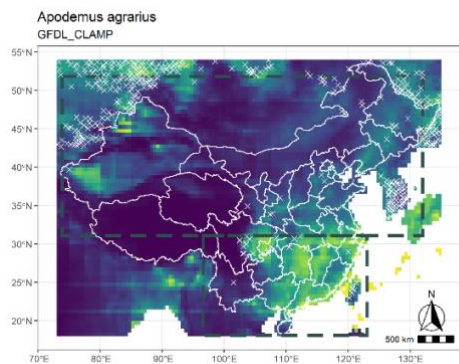
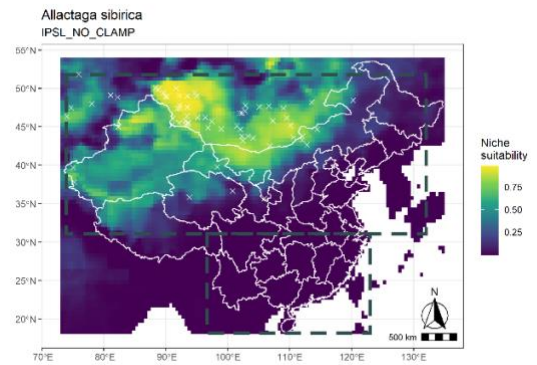
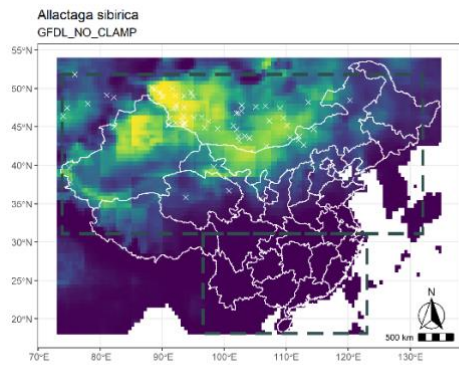
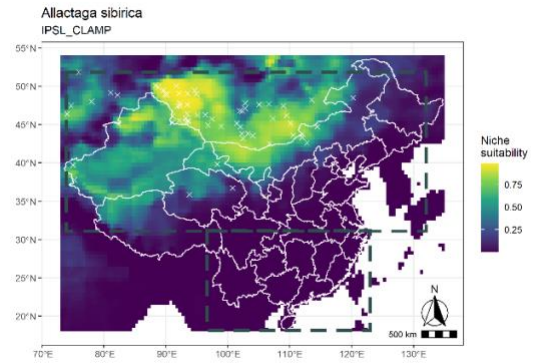
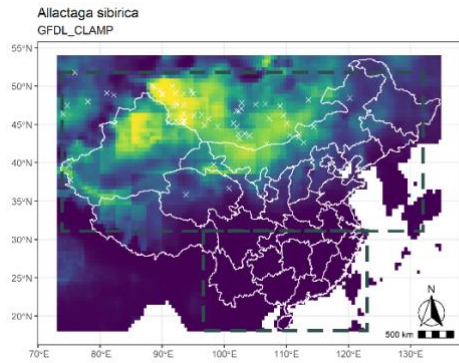
Figure 3.4 Distribution of human plague cases across China during the Third Pandemic (white crosses) with mean annual precipitation between 1970-2000 C.E and Chinese regional borders (white lines) The northern region is bounded by a red box and the southern region is bounded by a yellow box.

3.2.3.1 Response curves

The response curves and permutation importance of the reservoir species ENMs provided the first opportunity to test the hypothesis that reservoir species mediated the response of plague intensity to precipitation. I established criteria to determine if the response curves and permutation importance were consistent with this hypothesis. For northern species the hypothesis predicts that niche suitability increases during periods of elevated precipitation, with the opposite expected for southern species. I therefore expected that modelled niche suitability would show a positive trend with precipitation variables within the variable response curves and that the precipitation variables would be important variables in the reservoir species ENMs in the North with the opposite expected in the South. To be consistent with the hypothesis both of the following criteria had to be met. Firstly at least one of the two highest permutation importance variables must be a precipitation variable, and secondly the high permutation importance precipitation variable must show the expected response (positive in the North, negative in the South) dependant on which region the species is from.

3.2.3.2 Niche time series calculation

To transform the annual niche model projections into a time series comparable with annual precipitation data the sum of the environmental niche suitability across all pixels was calculated for each time step. Further, due to the CPE over which environmental niches were predicted, several species with limited ranges were predicted to occur with a high probability in regions that are assumed inaccessible, either due to abiotic factors such as geographic distance and topographic barriers or biotic factors such as predation or competition (*Marmota baibacina*). I therefore imposed distance constraints on the range of ENM projections using a buffered minimum convex polygon (BMCP) calculated from species occurrences. The BMCP method was developed by (Mendes et al., 2020) and is an adaptation of minimum convex polygon (MCP) method (Kremen et al., 2008). The MCP approach excludes suitable pixels not encompassed by a minimum convex polygon, with interior angles smaller than 180° , enclosing all occurrences of a species; the BMCP is the same as the MCP but includes a buffer around the MCP area based on the maximum distance of minimal pairwise distances between occurrences (Mendes et al., 2020). The BMCP calculations were completed using the `MSDM_Posteriori` function in the *MSDM* package (Mendes et al., 2020) (a visualisation of each method for several species can be found in Chapter 4 Appendix



).

3.2.3.3 Correlation analysis

Distribution time series for each species scenario and sum method were compared to the mean annual precipitation variable calculated across each region (as well as the regions combined). Each time series was standardized through z-score transformation to enable the comparison of respective changes in niche distribution between species. The correlation was then calculated between the standardised summed annual suitability of each species and the mean annual precipitation across each scenario region and method, with mean correlation values across each species and region further calculated to attempt to identify the hypothesised trends (Table 3.3 and Chapter 4 Appendix Table 3.6, Table 3.7 & Table 3.8). The correlation statistics were calculated with no lag and a 1-year lag period following Xu et al. (2011).

3.2.3.4 Impact of wet and dry periods

To investigate the specific impact that wet and dry periods may have had upon reservoir species across either region I identified such periods within annual mean precipitation in the North and South regions across both climate scenarios (Figure 3.5). Across each region I calculated a 15-year running average in annual mean precipitation and identified the upper and lower quartiles of this data as wet and dry periods respectively. Periods of less than three consecutive years were excluded due to the inflated impact these periods could have when calculating summary statistics across each period. The mean standardised suitability was then calculated across all scenarios enabling trends during these periods to be identified. This data was combined across both sum methods within each climate scenario to provide the mean niche suitability across wet and dry periods.

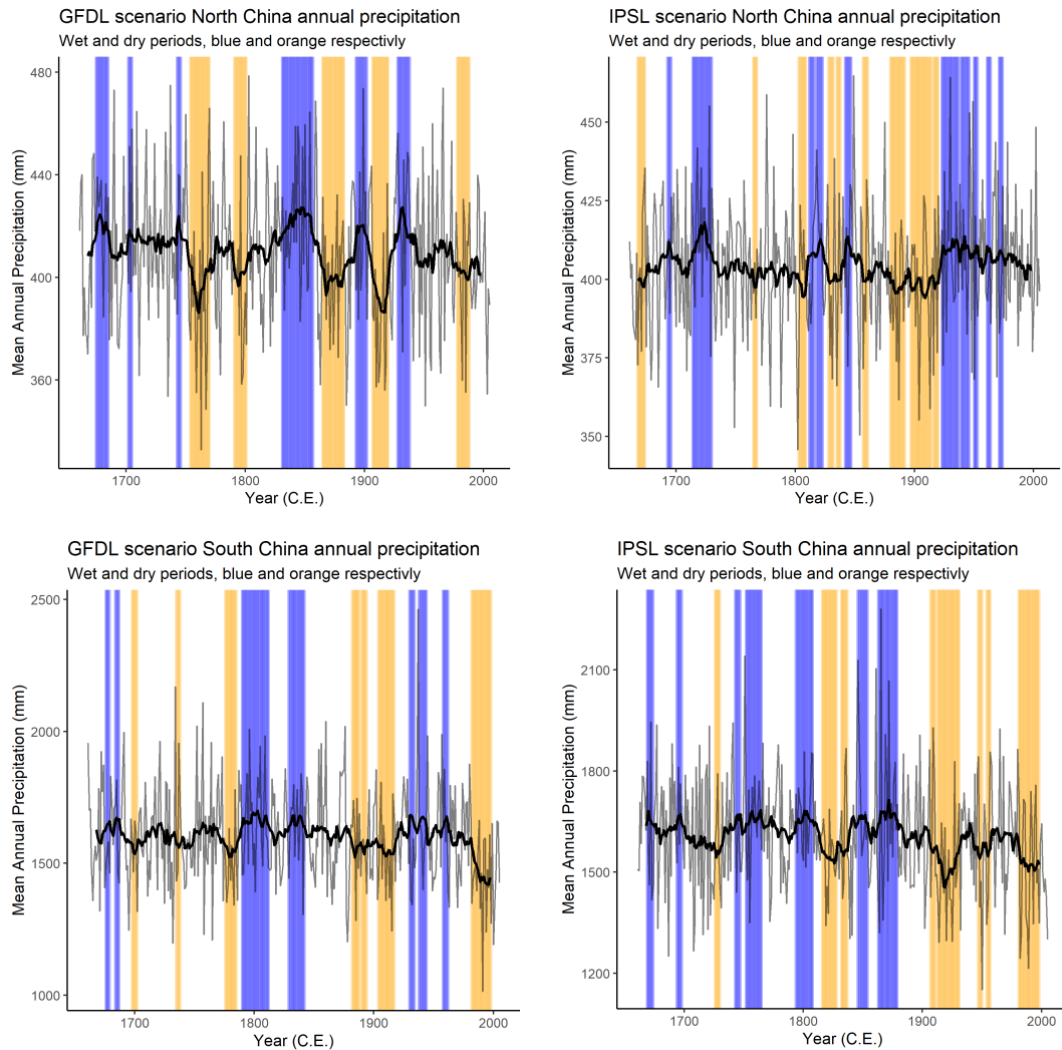


Figure 3.5. Identification of wet and dry periods across both climate scenarios. Blue indicates the wettest quartile and orange indicates the driest quartile. Both annual (thin black line) and 15 year running average (thick black line) mean annual precipitation shown.

3.3 Results

3.3.1 Niche modelling

3.3.1.1 Model performance

Species ENM's generally performed well over the extent determined by their respective buffered Eurasian range (species extent) (Table 3.2). The mean area under the curve (AUC) of the receiver operating characteristic across the species extent models and all climate scenarios (GFDL clamped, GFDL unclamped, IPSL clamped, IPSL unclamped) was 0.811 well above the 0.7 threshold above which models are considered fair (Swets, 1988). Two species fell below this threshold, *Niviventer*

andersoni and *Rhombomys opimus*. There was also little variation across the climate scenarios with a maximum standard deviation in AUC for any species of 0.055 (*Marmota baibacina*), this shows a consistency of model performance despite varying climate scenarios. I separately assessed the performance of each of the models across the CPE study region. Greater variation across species' mean AUC was observed in the geographically limited region from very good predictions of 0.996 - 0.999 (*Allactaga sibirica*, *Marmota sibirica*, *Meriones meridianus* and *Ochotona dauurica*) to poor predictions of 0.6 or less. Species with AUC values below 0.7 were retained during hypothesis testing, however, subsequent results from these species should be treated with caution. As AUC values do not provide direct information regarding overfitting (Warren and Seifert, 2011), I further present the AUC difference (difference between test and training AUC values) for all species across the species extent (Chapter 4 Appendix Table 3.9). Across all the species 70% have a very low AUC difference of less than 0.1 suggesting little overfitting for most of the species. The highest AUC difference is for the species *R. opimus* (0.2), which although still low may be suggestive of some overfitting for this species model.

Table 3.2. ENM performances across differing extents, region of each reservoir species and support for prediction based upon response curves. AUC values provide an estimate of model performance across the species defined extent (Species extent) and the China plague extent (CPE). A subset of the response curves are presented in .

Species	Mean AUC (Species extent)	SD	Mean AUC (CPE)	SD	Region	Response curves conform to predictions
<i>Allactaga sibirica</i>	0.805	0.007	0.999	0.001	N	No
<i>Apodemus agrarius</i>	0.824	0.010	0.789	0.034	N + S (Eurasian)	-
<i>Apodemus chevrieri</i>	0.866	0.016	0.908	0.028	S	No
<i>Callosciurus erythraeus</i>	0.952	0.001	0.824	0.057	S	No
<i>Cricetulus barabensis</i>	0.760	0.011	0.991	0.001	N	No
<i>Dipus sagitta</i>	0.851	0.002	0.992	0.002	N	No
<i>Eothenomys melanogaster</i>	0.821	0.018	0.762	0.020	S	No
<i>Marmota baibacina</i>	0.724	0.056	0.539	0.031	N	Yes
<i>Marmota caudata</i>	0.893	0.007	0.945	0.044	N	Yes (GFDL) / No (IPSL)

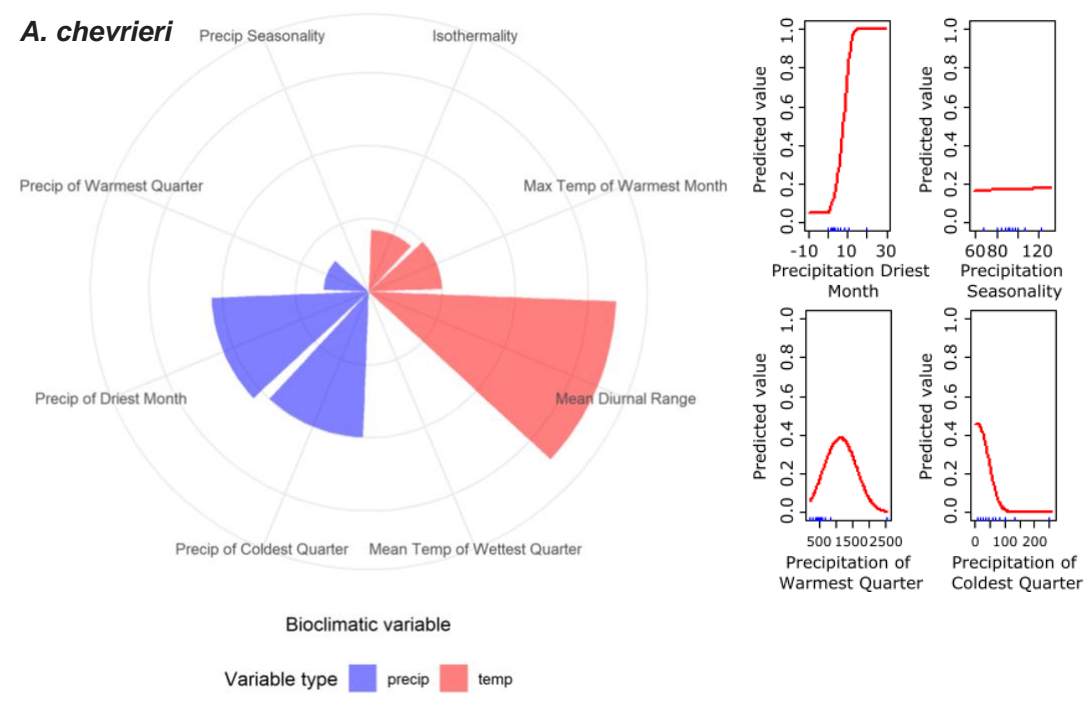
<i>Marmota himalayana</i>	0.805	0.000	0.481	0.028	N + S (Tibetan Plateau)	-
<i>Marmota sibirica</i>	0.763	0.010	0.998	0.001	N	No
<i>Meles leucurus</i>	0.730	0.035	0.554	0.204	N	Yes
<i>Meriones meridianus</i>	0.776	0.000	0.996	0.001	N	No
<i>Meriones unguiculatus</i>	0.893	0.017	0.974	0.012	N	No
<i>Micromys minutus</i>	0.908	0.007	0.468	0.024	N + S	-
<i>Microtus fortis</i>	0.771	0.016	0.786	0.065	N + S	-
<i>Mus musculus</i>	0.766	0.020	0.482	0.041	N + S (Eurasian)	-
<i>Mustela eversmanii</i>	0.744	0.020	0.957	0.011	N	No
<i>Mustela nivalis</i>	0.871	0.015	0.485	0.074	N	No
<i>Niviventer andersoni</i>	0.621	0.018	0.930	0.034	S	Yes
<i>Ochotona dauurica</i>	0.858	0.019	0.979	0.010	N	No
<i>Pseudois nayaur</i>	0.901	0.007	0.751	0.008	N + S (Tibetan Plateau)	-
<i>Rattus losea</i>	0.856	0.004	0.583	0.199	S	No
<i>Rattus nitidus</i>	0.906	0.006	0.461	0.038	S	No
<i>Rattus norvegicus</i>	0.816	0.031	0.625	0.020	N + S (Eurasian)	-
<i>Rattus rattus</i>	0.893	0.002	0.706	0.035	S (Mostly)	No
<i>Rattus tanezumi</i>	0.859	0.006	0.734	0.066	S	No
<i>Rhombomys opimus</i>	0.674	0.005	0.948	0.044	N	No
<i>Spermophilus dauricus</i>	0.791	0.019	0.980	0.019	N	No
<i>Spermophilus erythrogenys</i>	0.739	0.019	0.718	0.028	N	No
<i>Tscherskia triton</i>	0.863	0.030	0.870	0.017	N	Yes
<i>Urocitellus undulatus</i>	0.814	0.015	0.878	0.108	N	No
<i>Vulpes vulpes</i>	0.864	0.002	0.622	0.158	N + S (Eurasian)	-

3.3.2 Hypothesis testing

3.3.2.1 Response curves

According to my criteria (3.2.3.1) only four species' response curves were consistent with the predictions in showing the expected response to precipitation in either

region with a sufficient (1st or 2nd highest) permutation importance (Table 3.2 &). These species were *M. baibacina*, *M. leucurus*, *N. andersoni* and *T. triton*, all of which apart from *N. andersoni* inhabited the Northern region. Two of these species (*N. andersoni* & *M. baibacina*) had poor model performance (<0.7 AUC) over either the species or CPE. One species (*M. caudata*) met the criteria in the GFDL model but not in the IPSL model. Across the remaining species, the direction of response to the precipitation variables was counter to the predicted response (see *A. chevrieri*, *R. nitidus* and *R. opimus*,). Over 33 species and the two climate scenarios, GFDL and IPSL, no variation was observed between clamped and unclamped results. Overall only 9 of the 66 models (14%) fulfilled my criteria.



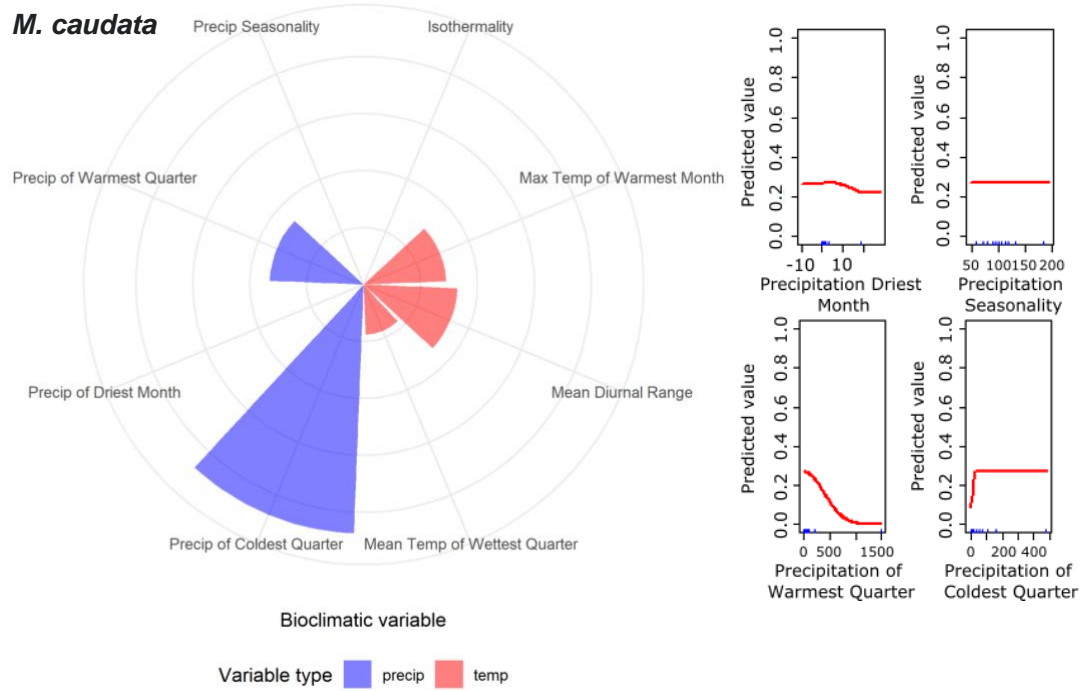
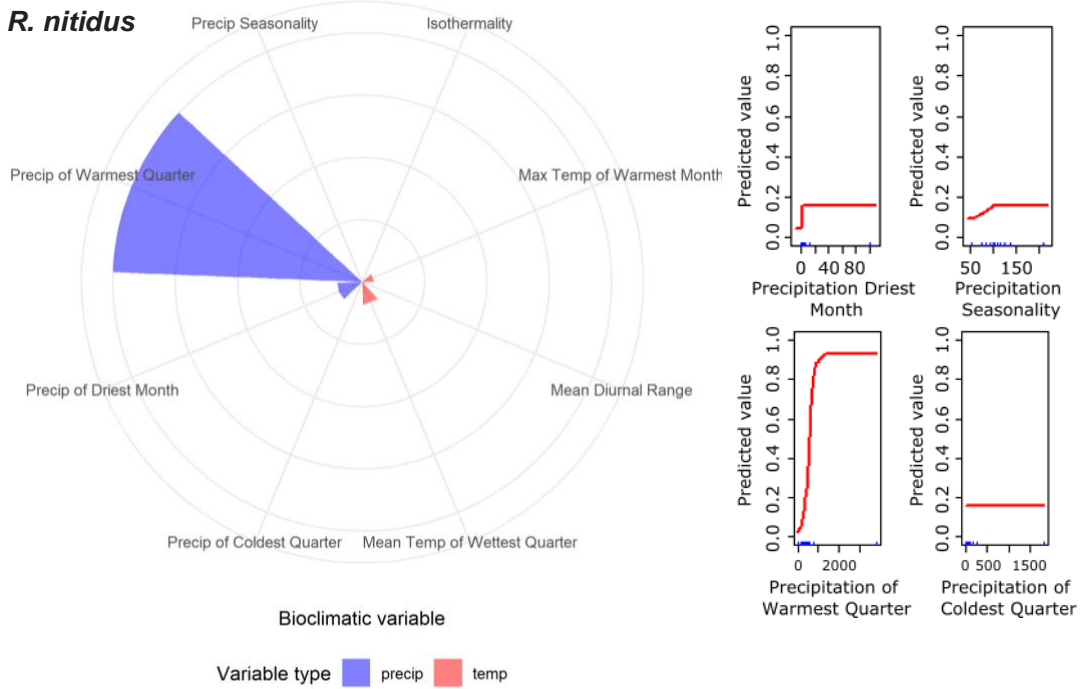


Figure 3.6 Permutation importance and response curves for select species. The polar bar chart represents the permutation importance of the precipitation (blue) and temperature variables. In the variable response curves for the precipitation variables (red line) are on the right with the blue ticks showing the variable values at the occurrence points. (Figure continued on subsequent pages)

R. nitidus



R. opimus

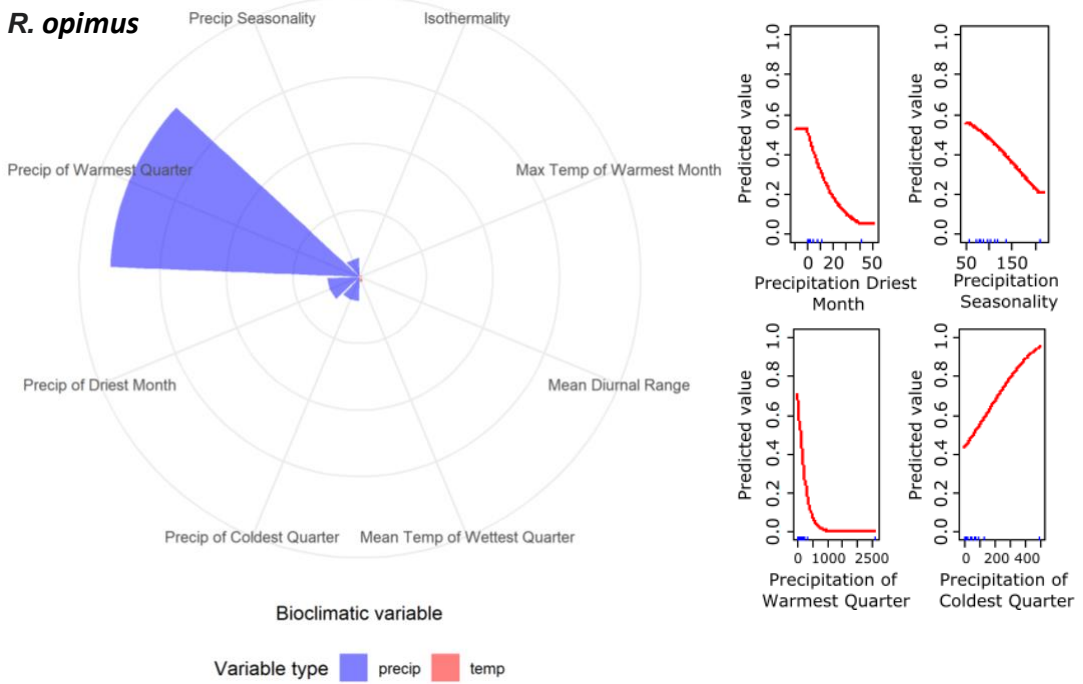


Figure 3.6 continued

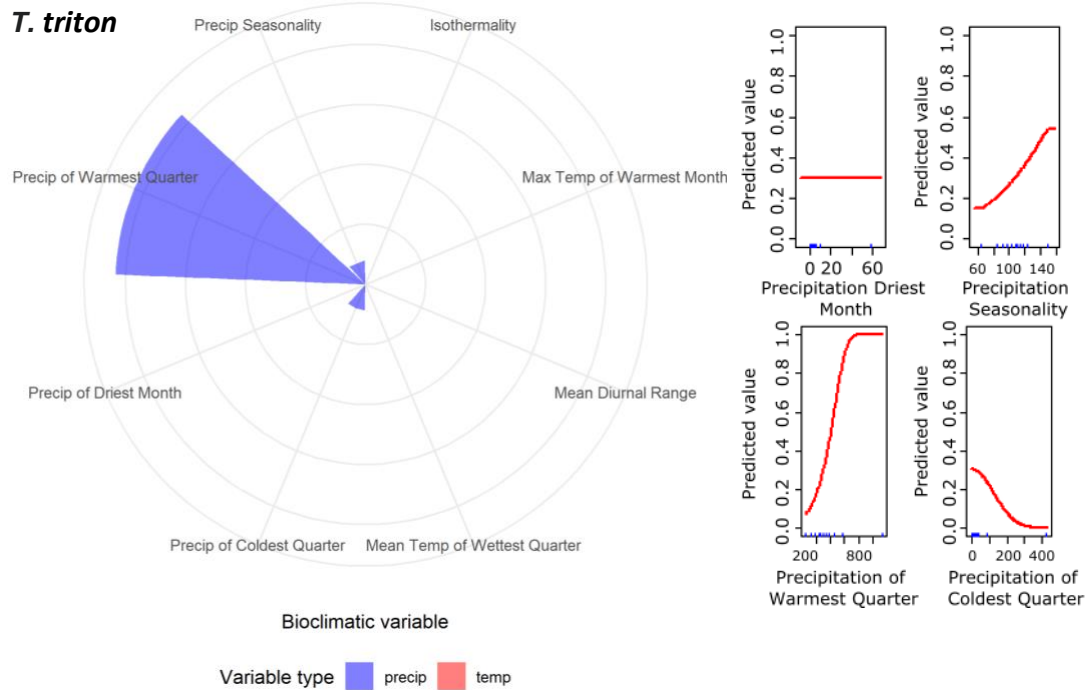


Figure 3.6 continued

3.3.2.2 Niche time series correlation

I did not find the expected correlation when species niche suitability results were combined across a region when comparing against both contemporary and 1-year lagged annual precipitation. For the summed contemporary niche suitability time series in the North, 9 of the 17 species (53%) showed the predicted positive mean correlation with annual precipitation across all climate scenarios (Table 3.3). Of these 9 species only one (*M. caudata*) had a standard deviation larger than the mean across the climate scenarios suggesting the direction of correlation varied across the scenarios; this is consistent with the response curve results which varied for *M. caudata* depending on the climate scenario. The northern species with response curves which conformed to the predictions showed consistency between methods through positive correlations with precipitation (*M. baibacina*, *M. leucurus* and *T. triton*). The remaining northern species showed a negative correlation with precipitation with two species (*R. opimus* and *M. meridianus*) showing a moderate negative correlation (<-0.4). The split in positive and negative correlation across the region meant there was near zero mean correlation across all Northern species ($r = -0.017$) and a large standard deviation (0.296) highlighting the varied correlation.

In the South only two of the species had the predicted negative correlation with annual precipitation (*N. andersoni* and *A. chevrieri*). *N. andersoni* was the only Southern species consistent with the response curve results. The remainder of the Southern species showed a moderate to strong mean correlation with annual precipitation ($r = 0.371-0.629$). The standard deviations of the mean values for these species were all much lower than the mean correlation values suggesting a consistent trend across all climate scenarios. The strong and consistent positive correlations led to an overall mean correlation across all species of $r = 0.313$ in the South, counter to the hypothesis' predictions. The standard deviation was larger than this mean value (0.361) due to the positive and negative correlations.

Species classified as both northern and southern, either due to a broad Eurasian extent or distribution in central China, all showed a positive correlation with precipitation. Two of the species with the broadest range (*M. musculus* and *R. norvegicus*) had standard deviation values larger than their mean suggesting variation across the scenarios. The combined North and South region showed a small positive mean correlation across all scenarios ($r = 0.192$). Across all species the high agreement in correlation direction across the scenarios and low standard deviation across most reservoir species suggests a good agreement between the two climate scenarios.

The results for the spatially constrained (BMCP) niche models are consistent with the unconstrained models (Chapter 4 Appendix Table 3.6). Generally, there were marginally weaker correlations across all species when compared to the unconstrained models (Table 3.3). For three species (*A. chevrieri*, *M. caudata* and *M. himalayana*) there was a change in correlation direction. In *M. caudata* and *M. himalayana* the mean correlation is less than the standard deviation across the scenarios. The change in correlation direction for *A. chevrieri*, limits the negatively correlated (conforming to the hypothesis) species in the South to only one (*N.*

Table 3.3. Correlation of reservoir species niche suitability values through time (summed) and annual precipitation values. Correlation was compared across two climate scenarios each with clamped and unclamped model

versions. Red-green colour scale was applied with red representing the lowest (most negative values) and green representing the highest (most positive values) correlations. Reservoir species are grouped by region.

Region	Species	Sum correlation with annual precipitation						Mean by Region	SD
		GFDL CLAMP	GFDL NO CLAMP	IPSL CLAMP	IPSL NO CLAMP	Mean	SD		
North	<i>Allactaga sibirica</i>	-0.207	-0.207	-0.096	-0.096	-0.152	0.064	-0.017	0.296
	<i>Cricetulus barabensis</i>	-0.049	-0.049	-0.054	-0.054	-0.051	0.002		
	<i>Dipus sagitta</i>	-0.379	-0.379	-0.308	-0.308	-0.343	0.041		
	<i>Marmota baibacina</i>	0.527	0.527	0.369	0.369	0.448	0.091		
	<i>Marmota caudata</i>	-0.032	-0.032	0.123	0.123	0.045	0.089		
	<i>Marmota sibirica</i>	-0.215	-0.215	-0.170	-0.170	-0.192	0.026		
	<i>Meles leucurus</i>	0.234	0.224	0.072	0.090	0.155	0.086		
	<i>Meriones meridianus</i>	-0.562	-0.562	-0.410	-0.410	-0.486	0.088		
	<i>Meriones unguiculatus</i>	0.207	0.217	0.014	0.039	0.120	0.108		
	<i>Mustela eversmanii</i>	0.429	0.429	0.181	0.186	0.306	0.141		
	<i>Mustela nivalis</i>	0.071	0.079	0.044	0.044	0.060	0.018		
	<i>Ochotona dauurica</i>	0.089	0.089	0.019	0.114	0.078	0.041		
	<i>Rhombomys opimus</i>	-0.613	-0.613	-0.360	-0.360	-0.486	0.147		
	<i>Spermophilus dauricus</i>	-0.373	-0.373	-0.074	-0.074	-0.223	0.172		
	<i>Spermophilus erythrogenys</i>	-0.183	-0.183	-0.184	-0.184	-0.183	0.001		
<i>Tscherskia triton</i>	0.465	0.003	0.599	0.599	0.417	0.283			
<i>Urocitellus undulatus</i>	0.162	0.465	0.073	0.073	0.193	0.186			
South	<i>Apodemus chevrieri</i>	-0.061	-0.061	-0.206	-0.206	-0.133	0.084	0.313	0.361
	<i>Callosciurus erythraeus</i>	0.400	0.379	0.503	0.470	0.438	0.058		
	<i>Eothenomys melanogaster</i>	0.514	0.514	0.417	0.417	0.465	0.056		
	<i>Niviventer andersoni</i>	-0.475	-0.475	-0.322	-0.322	-0.399	0.088		
	<i>Rattus losea</i>	0.541	0.541	0.717	0.717	0.629	0.101		
	<i>Rattus nitidus</i>	0.591	0.591	0.650	0.650	0.621	0.034		
	<i>Rattus rattus</i>	0.384	0.362	0.374	0.365	0.371	0.010		
<i>Rattus tanezumi</i>	0.570	0.570	0.451	0.451	0.510	0.069			
North & South	<i>Apodemus agrarius</i>	0.164	0.132	0.300	0.284	0.220	0.084	0.192	0.150
	<i>Marmota himalayana</i>	0.070	0.070	0.195	0.195	0.133	0.072		
	<i>Micromys minutus</i>	0.195	0.179	0.515	0.508	0.349	0.188		
	<i>Microtus fortis</i>	0.258	0.258	0.165	0.165	0.211	0.053		
	<i>Mus musculus</i>	-0.032	-0.046	0.214	0.218	0.089	0.147		
	<i>Pseudois nayaur</i>	0.411	0.411	0.117	0.117	0.264	0.170		
	<i>Rattus norvegicus</i>	-0.074	-0.069	0.241	0.238	0.084	0.180		
<i>Vulpes vulpes</i>	0.084	0.024	0.332	0.315	0.189	0.158			

andersoni). The mean correlations across the regions show little change, the North was almost neutral ($r = -0.014$) but was outweighed by the standard deviation (0.245), the South was slightly more positive ($r = 0.347$) due to the change in correlation direction of *A. chevrieri* and the combined region was still weakly positive ($r = 0.145$) but is now outweighed by the standard deviation (0.159). Across the two modelling methods (spatially constrained and non-spatially constrained) the variation in correlation between clamped and unclamped is negligible for most species, however,

two species, *T. triton* and *U. undulatus* in the GFDL scenario show more variation (Sum, SD = 0.327 & 0.214, BMCP Sum, SD = 0.090 & 0.074 respectively) suggesting that in the GFDL scenario, these species are both projected into non-analogous climate space.

Following Xu et al. (2011), I completed the correlation analysis between the niche time series and climate data from the year prior (Chapter 4 Appendix Table 3.7 & Table 3.8). Generally, all correlations are weaker in the lagged version, the direction of correlation is maintained for the majority of species. In the few cases where a change in correlation direction occurs (*O. daurica*, *M. leucurus*, *U. undulatus*, *A. agrarius*, *M. fortis* & *M. hamalyana*), the correlation is consistently weak ($|r| < 0.1$).

3.3.2.3 Wet and dry period niche analysis

Across the Northern wet periods in both climate scenarios, 44% of the species showed increased mean niche suitability, however, only 15% of the species had an increased mean niche suitability greater than 0.1 (Table 3.4), suggesting a weak response that thus supplies weak evidence consistent with the prediction of increased niche suitability during wet periods in the North and dry periods in the South. As the mean niche suitability values were standardised, this represents an increase of mean niche suitability of 0.1 standard deviations, highlighting the generally limited variation in mean niche suitability in many species through periods of high and low precipitation. The mean niche suitability increased for more species in northern dry periods where 59% of species had an increased mean niche suitability and for 35% of the species this increase was greater than 0.1. The species with the greatest standardised niche probability increase during the wet periods were *M. baibacina* (GFDL climate scenario), *M. eversmanii* (GFDL climate scenario), *U. undulatus* (GFDL climate scenario) and *M. nivalis* (Both scenarios). Each of these species also showed consistency across investigation method as they showed a positive correlation with annual precipitation in the North as expected given the initial prediction (Table 3.3).

In the southern region 31% of the species showed an increased mean niche suitability during the identified dry periods, as predicted, but only one species (*N. andersoni*) had an increase of greater than 0.1. This species conformed to the hypothesis predictions through each method. As in the North under the alternate precipitation scenario I find a much larger proportion of species with increased mean niche suitability, with 87.5% showing an increase while 56% of species had an increase in mean niche suitability greater than 0.1.

There was generally high agreement in directional change (increase or decrease above the mean) across the two climate scenarios with 83% of species agreeing across each scenario (81% when the species found in both the northern and southern regions are included). A further small percentage (9%) of the species under all scenarios had a standard deviation greater than the mean niche suitability, effectively suggesting that the niche suitability during wet or dry periods overlapped with the mean variation in climate scenario predictions. This was due to the difference in clamped and unclamped niche projections.

Table 3.4. Mean standardised niche suitability of reservoir species through wet and dry periods. Mean and standard deviation values were calculated across periods identified as wet or dry and a red-green colour scale is applied with red representing the lowest (most negative values) and green representing the highest (most positive values). Reservoir species are grouped by region.

		Mean niche suitability							
		Wet				Dry			
		GFDL		IPSL		GFDL		IPSL	
Region	Species	Mean	SD	Mean	SD	Mean	SD	Mean	SD
North	<i>Allactaga sibirica</i>	-0.154	0.007	-0.119	0.000	0.216	0.012	0.104	0.003
	<i>Cricetulus barabensis</i>	-0.094	0.004	-0.061	0.029	0.038	0.000	0.071	0.020
	<i>Dipus sagitta</i>	-0.132	0.005	-0.052	0.011	0.242	0.006	-0.146	0.001
	<i>Marmota baibacina</i>	0.317	0.000	0.051	0.022	-0.274	0.090	0.006	0.162
	<i>Marmota caudata</i>	0.043	0.036	0.011	0.026	0.236	0.031	0.066	0.056
	<i>Marmota sibirica</i>	-0.139	0.001	-0.113	0.039	0.196	0.032	0.088	0.057
	<i>Meles leucurus</i>	0.044	0.004	-0.149	0.007	0.004	0.006	-0.267	0.002
	<i>Meriones meridianus</i>	-0.351	0.010	-0.098	0.016	0.261	0.063	-0.059	0.007
	<i>Meriones unguiculatus</i>	0.050	0.009	-0.013	0.004	-0.091	0.009	-0.068	0.007
	<i>Mustela eversmanii</i>	0.288	0.002	-0.066	0.001	-0.134	0.001	-0.132	0.003
	<i>Mustela nivalis</i>	0.138	0.012	0.353	0.013	0.162	0.009	-0.011	0.008
	<i>Ochotona dauurica</i>	0.040	0.002	-0.075	0.039	0.075	0.012	0.101	0.025
	<i>Rhombomys opimus</i>	-0.271	0.041	-0.054	0.046	0.326	0.091	-0.136	0.016
	<i>Spermophilus dauricus</i>	-0.166	0.001	-0.063	0.017	0.346	0.009	0.018	0.028
	<i>Spermophilus erythrogegens</i>	-0.072	0.013	0.023	0.021	0.238	0.006	0.135	0.031
	<i>Tscherskia triton</i>	0.046	0.049	0.097	0.025	-0.105	0.186	-0.258	0.107
<i>Urocitellus undulatus</i>	0.101	0.023	0.020	0.019	-0.232	0.116	0.073	0.015	
South	<i>Apodemus chevrieri</i>	0.185	0.046	0.132	0.025	-0.116	0.010	0.014	0.049
	<i>Callosciurus erythraeus</i>	0.142	0.003	0.239	0.013	0.001	0.013	-0.130	0.006
	<i>Eothenomys melanogaster</i>	0.030	0.001	0.356	0.001	-0.172	0.000	-0.063	0.004
	<i>Niviventer andersoni</i>	0.115	0.001	-0.054	0.042	0.075	0.001	0.149	0.055
	<i>Rattus losea</i>	0.092	0.004	0.291	0.024	-0.282	0.020	-0.123	0.007
	<i>Rattus nitidus</i>	-0.054	0.029	0.078	0.016	-0.158	0.004	-0.286	0.012
	<i>Rattus rattus</i>	0.046	0.004	0.269	0.003	0.039	0.009	-0.234	0.003
	<i>Rattus tanezumi</i>	0.118	0.000	0.242	0.000	-0.101	0.000	-0.062	0.000
North & South	<i>Apodemus agrarius</i>	-0.033	0.008	0.174	0.003	0.279	0.015	-0.077	0.002
	<i>Marmota himalayana</i>	0.182	0.003	0.046	0.072	-0.229	0.033	-0.124	0.129
	<i>Micromys minutus</i>	0.141	0.003	0.262	0.001	0.037	0.007	-0.259	0.005
	<i>Microtus fortis</i>	-0.156	0.011	0.158	0.008	0.094	0.001	0.044	0.031
	<i>Mus musculus</i>	-0.031	0.001	0.030	0.003	0.187	0.007	-0.152	0.002
	<i>Pseudois nayaur</i>	0.153	0.067	0.022	0.034	-0.138	0.134	-0.003	0.121
	<i>Rattus norvegicus</i>	-0.017	0.001	0.084	0.001	0.171	0.001	-0.174	0.004
	<i>Vulpes vulpes</i>	0.148	0.077	0.171	0.012	-0.045	0.236	-0.254	0.014

3.4 Discussion

Using three differing methods based upon ENMs, I have found minimal evidence consistent with the hypothesis that in China during the Third Pandemic the heterogeneous response of plague intensity to precipitation was mediated by reservoir species. In contrast to predictions, I found no consistent positive correlation between reservoir species niche suitability and precipitation in North China. In South China, I predominantly found positive, instead of negative correlations between reservoir species niche suitability, counter to the prediction. Species niche suitability was found to be no higher during periods of increased precipitation in the North or decreased precipitation in the South as predicted and in fact the opposite was observed with generally niche suitability increasing in the North in dry periods and decreasing in the South during wet periods.

Certain species did perform as predicted under all investigation methods, specifically *M. baibacina* in the North and *N. andersoni* in the South. Both species have been identified as reservoir species (Mahmoudi et al., 2020), *M. baibacina* is a key host in the Tian Shen region (He et al., 2021) and there is little further literature on the role of *N. andersoni* in plague maintenance in Yunnan. However, there is no further evidence to suggest that these species were key to infection and transmission during the Third Pandemic or that they may have had a disproportionately large impact on plague transmission compared to other species. The results therefore suggest that further elements of the plague system aside from reservoir species, such as human population movement (Xu et al., 2014), or an element of reservoir species ecology not captured in the models (eg. Behavioural or biotic factors) may have mediated the observed heterogeneous response of plague intensity to precipitation across China.

3.4.1 Insights from modelling approach

The modelling approach used provided several insights prior to the use of the model predictions for hypothesis testing. Variable selection, prior to the selection of equal precipitation and climate variables showed a limited selection of precipitation variables (Biovars 12-19) comparable to temperature variables (Biovars 1-11) when strongly correlated variables are removed (Figure 3.2). Only one precipitation

variable was selected for four species (*A. sibirica*, *A. agrarius*, *O daurica* and *T. triton*) suggesting that across the range of these species precipitation variables were unimportant to defining the niche or co-varied with temperature variables. With many of the remaining species, only two precipitation variables were selected, suggesting a similar lack of importance or co-variation with temperature variables. Therefore, even when precipitation variables are maintained in the model, co-variation between variables may limit the interpretation of certain variables as biologically relevant to the species niche. The inclusion of potentially co-varying variables may also explain the permutation importance of several variables in species models having no contribution (See . *T. triton* and *R. nitidus*).

3.4.2 Correlation analysis

Just over half of the Northern species showed the predicted positive correlation between niche suitability and precipitation. This is consistent with the bottom-up trophic cascade mechanism regularly applied to plague systems (Collinge et al., 2005, Stenseth et al., 2006, Kausrud et al., 2007, Xu et al., 2015). However while it suggests a link between precipitation and reservoir species dynamics, the subsequent increase in *Y. pestis* transmission and hence plague intensity has to be assumed and therefore is an area in need of further investigation. The remaining northern species showed a negative correlation between precipitation and niche suitability suggesting that these species are not influenced by a bottom-up trophic cascade. However, alternate mechanisms may mediate the impact of precipitation on plague intensity through these species. For example, a population decline mediated by climate may increase the vector load per individual host (Schmid et al. 2015). This would facilitate increased transmission within the reservoir as well as drive transmission to alternate reservoir species (Reijniers et al., 2014). Therefore, although my findings do not support the trophic cascade mechanism as a driving transmission for all species in North China, alternate mechanisms across the diverse range of species may have differing impacts on reservoir dynamics yet still facilitate an increase in plague intensity.

In the South only two species conformed to predictions, with the majority of species showing a positive correlation between precipitation and niche suitability. While this does not match my predictions, it is potentially still consistent with the “U” shaped response to precipitation observed in South China (Xu et al., 2011) where very wet periods correlate with increased plague intensity. However, this “U” shaped response is difficult to falsify as the opposite response in plague intensity is expected under wet and very wet conditions and the boundaries between these conditions are likely spatially heterogeneous. As in the North, my findings suggests that there is not a homogenous response from reservoir species to changes in precipitation, and thus the mechanism driving plague transmission through such species is unlikely to be consistent across a broad range of species. Moving forward from these findings, a key goal should be to determine which species are most important to the transmission of *Y. pestis*, particularly at the initiation of epidemics. This should draw together the large-scale correlative methods utilised here and mechanistic methods applied to individual species of interest (Kausrud et al., 2007, Xu et al., 2015)

3.4.3 Wet and dry period analysis

Very few species conformed to the initial predictions of high niche suitability during wet periods in the North and high niche suitability during dry periods in the South. In fact, the greatest niche suitability values were observed under the opposite precipitation regimes than predicted. Were the results representative of extreme precipitation conditions (see below), to conform to the hump and “U” shaped response of plague intensity to precipitation observed by Xu et al. (2011), I would expect relatively symmetric results across the wet and dry periods, which I do not observe. The positive niche suitability observed across the majority of southern species during wet periods shows some support for the “U” shaped southern response as plague intensity was high during southern wet periods. However, my findings are counter to the mechanism Xu et al. (2011) suggest for the high plague intensity during wet to very wet periods as they suggest that this increase is primarily due to a behavioural response whereas my results suggest increased niche suitability during these periods. As many of the southern species, particularly the *Rattus spp.* are commensal rodents closely associated with human habitation and crops their

response to climatic variation may be partially independent of their climatic niche and instead relate to land use factors such as the distribution of agricultural land and the response of such land to climatic variables e.g. crop failure (McCauley et al., 2015).

Wet and dry period analysis likely fails to explicitly capture the extreme wet and dry conditions which Xu et al. (2011) identify as having opposing responses in plague intensity to more moderate wet and dry conditions. Although across China the ISIMIP data has been found to simulate extreme continuous multiday rainfall well (Yuan et al., 2017), the conversion to annual climate variables will likely partially mute the impact of extreme weather within the data. I therefore suggest that they identified wet and dry periods are representative of moderate not extreme precipitation and subsequently the climate data over which the niche models are hindcast will not represent extreme conditions. This suggests that the results are valid in testing the general trends observed by Xu et al. (2011), Northern wet-plague correlation and Southern dry-plague correlation, but not the hump and “U” shaped correlations observed when extreme precipitation conditions are included.

The results suggest either that, the reservoir mechanisms proposed to mediate the connection between precipitation and plague intensity are not correct in either region or, alternatively, the species I observe conforming to the hypothesis have a disproportionate impact on plague intensity. A third and likely alternative is that the environmental conditions impacting each species are not consistent across all reservoir species but vary and that the scale of impact each species may have upon plague transmission is also not consistent and may fluctuate spatio-temporally. A further fourth alternative which has not been explored in this work is that human movement, transmission and population dynamics may contribute to the heterogeneous response to precipitation (Xu et al., 2014).

3.4.4 Methodological considerations

The hindcasting methodology used in this work followed many of the recommended practices suggested to improve the reliability of hindcasting predictions however

there are further practices which I was not capable of addressing. The reliability of hindcasting ENM projections can be influenced by a range of factors such as the niche stability of the species or the variation in climate conditions between the time period the model was constructed in and projected into. The BMCP method limited species predicted suitability based upon the current distribution of each species assuming that the dispersal ability of each species has not changed throughout the period of projection. This assumption is unlikely to be valid for all species, particularly species which may be impacted by changes in land use (eg. Commensal rodents) which can strongly impact plague reservoir dynamics and distribution (Miao et al., 2013, McCauley et al., 2015). To assess the impact of non-analogous climate upon my models all analyses were completed in duplicate using clamped and unclamped models. This enabled me to identify variation in model predictions due to non-analogous climate. I found very little variation due to clamping across most niche models, however, variation was observed in two species results (*T. triton* and *U. undulatus*) which is attributable to the difference between clamped and unclamped niche projections. Therefore it is unlikely that projection into non-analogous climate space has affected the overall conclusions, drawn from the remainder of the species.

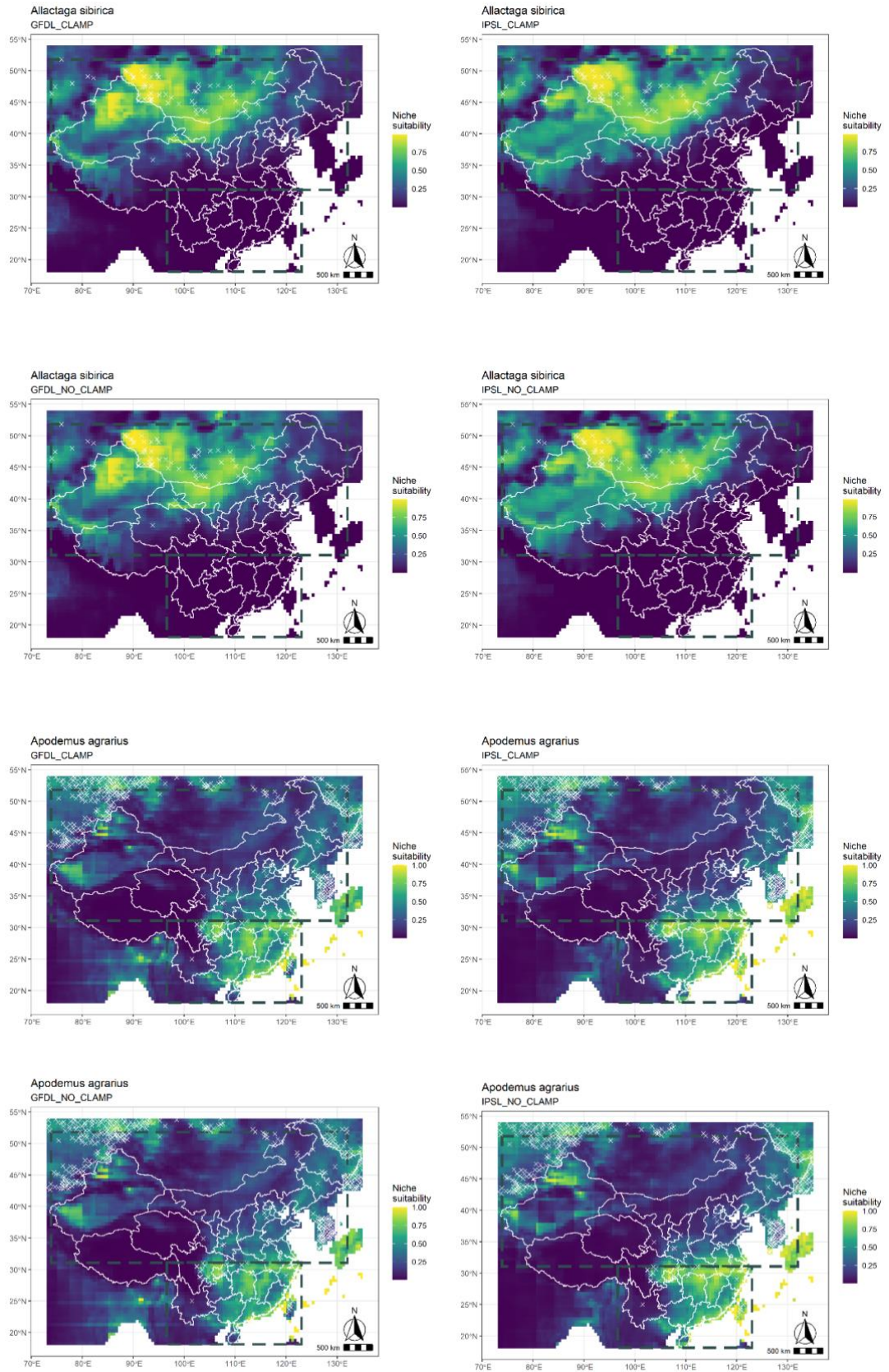
It is recommended to use multiple climate scenarios when hindcasting ENMs to confirm that results are not an artefact of the scenarios used (Nogués-Bravo, 2009). Agreement in results between the climate scenarios was high across the correlation and wet-dry period analysis, however, variation was observed in the response curves with permutation importance of bioclimatic variables varying between the climate scenarios. The limited variation in results between climate scenarios is consistent with the limited time period of hindcasting, in comparison to the majority of hindcasting studies (Nogués-Bravo, 2009).

The niche stability of all species has been assumed throughout this work. Given the relatively short time-span over which the models were projected, in comparison to much longer time periods commonly covered in hindcasting studies (Nogués-Bravo, 2009, Maiorano et al., 2013), the need to address niche stability in this work is relatively unknown. An initial attempt to address the potential instability of species

niches' was made through multi-temporal calibration of the models however the results were inconclusive and hence the method was not used (Chapter 4 Appendix 3.6.3 Attempted multi-temporal calibration). If the rapid recent morphological adaptation of commensal rodent species (over the past 100 years) (Pergams and Lawler, 2009) has driven changes in these species niche (Van Valen, 1965) then the assumption of niche stability for species key to the transmission of *Y. pestis* may not be appropriate. Future work should focus on investigating the stability of the niche of key reservoir species as, if these species have highly dynamic and rapidly fluctuating niches this should be factored in to both hindcasting and forecasting of zoonotic disease distributions. Inclusion of eco-physiological data would provide further validation of the niche models and hindcast projections. Further, disease specific data such as species susceptibility to infection could improve the estimates of disease distribution calculable using reservoir species data.

3.4.5 Limitation of modelling approach

Poor model performance for some species across both the species specified extent and the CPE suggest that the climate variables used throughout the ENMs may not be the most important in determining the niche of several of the reservoir species. Poor model performance over the species extent could be indicative of a missing variable which informs the species niche. As many of the selected species are burrowing rodents, a key aspect of the species niche may be dictated by soil variables which is not accounted for in models; this may partially explain the poor performance of the *R. opimus* model (Wilschut et al., 2015). Across the CPE, poor model performance could be attributed to two main factors: limited dispersal ability and a geographical bias in occurrence points towards Europe. *M. baibacina*, a species with a limited range, predicting a high probability of distribution across a broad range of China which is inaccessible to the species due to a combination of abiotic and biotic factors not captured in the models, such as geographic distance or competition. Several of the species have a broad Eurasian range with most occurrences located within Europe compared to a limited number across China (*M. minutus*, *M. musculus*, *M. nivalis*, *R. nitidus* and *V. vulpes* Chapter 4 Appendix



). For these species the Eurasian extent models perform well, however, the limited proportion of test and training data in China inhibited their performance across this

area. The alternative method of selecting only species occurrences from across the investigation region (CPE) would produce “better” (higher AUC) models across this region but would be less representative of the species niche given the bias sub-setting of the data.

3.5 Conclusion

My findings suggest that reservoir species dynamics do not mediate the heterogeneous response of plague intensity across China to precipitation as hypothesised (Xu et al., 2011). The results suggests that assuming consistent host dynamics in a region is an oversimplification of the highly complex plague transmission system and that mechanisms should not be assumed based on studies of individual species. Although I found two species which conform to the hypothesis predictions (*M. baibacina* and *N. andersoni*) I do not suggest that these species had an inflated impact upon plague intensity but instead suggest that diverse and interrelated responses of reservoir species may drive the observed response. The findings suggest that the impact of further factors such as behavioural response to climatic change as well as vector and human population dynamics should be investigated to understand the drivers of the observed heterogeneous response. The methods represent a novel and simple (comparative to mechanistic approaches) means of investigating a complex zoonotic system and counter to regional and species-specific mechanistic models, patterns can be easily assessed over large spatial and temporal extents and a broad range of reservoir species, however, the relative simplicity of this approach has limitations. Such disease systems therefore represent an ideal scenario to test hybrid mechanistic-correlative models where physiological data relating to reservoir species as well as vectors could be combined with correlative methods and a balance between data requirements and the simplification of such systems could be struck (c.f. Kearney et al. (2010), Redding et al. (2016)).

3.6 Chapter 4 Appendix

3.6.1 Supplementary tables

Table 3.5 Chinese reservoir species sourced from Mahmoudi et al., (2021) and Cui et al., (2013). An asterisk next to the species indicates ten or less occurrence points were available through GBIF and the species was subsequently removed from the analysis.

Species	Order	Source	GBIF DOI	Location of removed erroneous data points (Bounding box x,x,y,y)
<i>Alexandromys fortis*</i>	Rodentia	Mahmoudi et al. (2020)	(Not found within GBIF)	-
<i>Allactaga sibirica</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.xj3ajx	40,50, 40,60 (Central Asia), 100,130, 20,30 (South China)
<i>Apodemus agrarius</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.fq6acu	-10,10, 0,20 (West Africa)
<i>Apodemus chevrieri</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.44fckb	-10,0, 0,20 (West Africa)
<i>Callosciurus erythraeus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.y4958a	60,90, 15,25 (India)
<i>Cricetulus barabensis</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.wn2yjf	0,75, 50,80 (West Russia), 20,30, 100,150 (South China)
<i>Dipus sagitta</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.zsxjc4	0,50, 20,40 (East Africa)
<i>Eothenomys eleusis*</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.ajgygj	-
<i>Eothenomys melanogaster</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.pk7s7b	-10,0, 0,20 (West Africa)
<i>Marmota baibacina</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.qgdww5	-
<i>Marmota caudata</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.edyx9b	-
<i>Marmota himalayana</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.w99q4r	-
<i>Marmota sibirica</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.thtbau	125,150, 20,40 (Japan)

<i>Meles leucurus</i>	Carnivora	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.h862ak	-
<i>Meriones meridianus</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.zhkuku	-
<i>Meriones unguiculatus</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.9m9hbs	-
<i>Micromys minutus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.xesm4r	-4,4, 0,4 (South Atlantic)
<i>Microtus fortis</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.qujxzu	-
<i>Microtus juldaschi*</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.x98mfq	-
<i>Mus musculus</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.82gk7w	-4,4, 0,4 (South Atlantic)
<i>Mustela eversmanii</i>	Carnivora	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.6r6dan	-
<i>Mustela nivalis</i>	Carnivora	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.x3v5nz	-10,10, 0,10 (South Atlantic)
<i>Neodon fuscus*</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.xhrxvj	-
<i>Niviventer andersoni</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.qm7w4p	-
<i>Ochotona dauurica</i>	Lagomopha	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.anzd7b	95,130, 20,30 (South China)
<i>Pseudois nayaur</i>	Artiodactyla	(Cui et al., 2013)	https://doi.org/10.15468/dl.unykv5	-
<i>Rattus losea</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.qad8vj	-
<i>Rattus nitidus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.6ubg65	-
<i>Rattus norvegicus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.rm8wv5	-1,1 0,5 (South Atlantic)
<i>Rattus rattus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.acx9hv	-
<i>Rattus tanezumi</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.h25ynz	-

<i>Rhombomys opimus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.w3yvey	-
<i>Spermophilus alashanicus*</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.eepgmt	-
<i>Spermophilus dauricus</i>	Rodentia	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.ted62t	-
<i>Spermophilus erythrogenys</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.kd4c49	-
<i>Tscherskia triton</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.fkhdk	-
<i>Urocitellus undulatus</i>	Rodentia	Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.rur2wc	-
<i>Vulpes vulpes</i>	Carnivora	Cui et al. (2013), Mahmoudi et al. (2020)	https://doi.org/10.15468/dl.sff7b6	-10,10, 0,10 (South Atlantic)

Table 3.6 Correlation of reservoir species niche suitability values through time (BMCP summed) and annual precipitation values. Correlation was compared across two climate scenarios each with clamped and unclamped model versions. Red-green colour scale was applied with red representing the lowest (most negative values) and green representing the highest (most positive values) correlations. Reservoir species are grouped by region.

Region	Species	BMCP Sum correlation with annual precipitation						Mean by Region	SD
		GFDL CLAMP	GFDL NO CLAMP	IPSL CLAMP	IPSL NO CLAMP	Mean	SD		
North	<i>Allactaga sibirica</i>	-0.187	-0.187	-0.084	-0.084	-0.135	0.060	-0.014	0.245
	<i>Cricetulus barabensis</i>	-0.048	-0.048	-0.043	-0.043	-0.046	0.003		
	<i>Dipus sagitta</i>	-0.361	-0.361	-0.296	-0.296	-0.328	0.038		
	<i>Marmota baibacina</i>	0.376	0.376	0.285	0.285	0.330	0.052		
	<i>Marmota caudata</i>	-0.058	-0.058	0.055	0.055	-0.002	0.065		
	<i>Marmota sibirica</i>	-0.111	-0.111	-0.098	-0.098	-0.104	0.007		
	<i>Meles leucurus</i>	0.233	0.223	0.070	0.088	0.153	0.086		
	<i>Meriones meridianus</i>	-0.464	-0.464	-0.308	-0.308	-0.386	0.090		
	<i>Meriones unguiculatus</i>	0.204	0.213	0.012	0.037	0.116	0.107		
	<i>Mustela eversmanii</i>	0.430	0.430	0.183	0.187	0.307	0.142		
	<i>Mustela nivalis</i>	0.068	0.074	0.022	0.023	0.047	0.028		
	<i>Ochotona dauurica</i>	0.097	0.097	0.024	0.122	0.085	0.043		
	<i>Rhombomys opimus</i>	-0.446	-0.446	-0.290	-0.290	-0.368	0.090		
	<i>Spermophilus dauricus</i>	-0.355	-0.355	-0.120	-0.120	-0.238	0.136		
	<i>Spermophilus erythrogenys</i>	-0.224	-0.224	-0.190	-0.190	-0.207	0.019		
<i>Tscherskia triton</i>	0.322	0.195	0.375	0.375	0.317	0.085			
<i>Urocitellus undulatus</i>	0.272	0.377	0.120	0.120	0.222	0.126			
South	<i>Apodemus chevrieri</i>	0.113	0.113	0.017	0.017	0.065	0.056	0.347	0.305
	<i>Callosciurus erythraeus</i>	0.400	0.379	0.503	0.471	0.438	0.058		
	<i>Eothenomys melanogaster</i>	0.523	0.523	0.412	0.412	0.467	0.064		
	<i>Niviventer andersoni</i>	-0.363	-0.363	-0.244	-0.244	-0.303	0.069		
	<i>Rattus losea</i>	0.530	0.530	0.708	0.708	0.619	0.103		
	<i>Rattus nitidus</i>	0.577	0.577	0.643	0.643	0.610	0.038		
	<i>Rattus rattus</i>	0.384	0.362	0.374	0.365	0.371	0.010		
	<i>Rattus tanezumi</i>	0.570	0.570	0.451	0.451	0.510	0.069		
North & South	<i>Apodemus agrarius</i>	0.164	0.132	0.300	0.284	0.220	0.084	0.145	0.159
	<i>Marmota himalayana</i>	0.016	0.016	-0.082	-0.082	-0.033	0.056		
	<i>Micromys minutus</i>	0.195	0.179	0.515	0.508	0.349	0.188		
	<i>Microtus fortis</i>	0.223	0.223	0.158	0.158	0.191	0.038		
	<i>Mus musculus</i>	-0.032	-0.046	0.214	0.218	0.089	0.147		
	<i>Pseudois nayaur</i>	0.152	0.152	-0.015	-0.015	0.068	0.096		
	<i>Rattus norvegicus</i>	-0.074	-0.069	0.241	0.238	0.084	0.180		
<i>Vulpes vulpes</i>	0.084	0.024	0.332	0.315	0.189	0.158			

Table 3.7 Correlation of reservoir species niche suitability values through time (summed) and lagged (1 year) annual precipitation values. Correlation was compared across two climate scenarios each with clamped and unclamped model versions. Red-green colour scale was applied with red representing the lowest (most negative values) and green representing the highest (most positive values) correlations. Reservoir species are grouped by region.

Region	Species	Sum correlation with annual precipitation						Mean by Region	SD
		GFDL CLAMP	GFDL NO CLAMP	IPSL CLAMP	IPSL NO CLAMP	Mean	SD		
North	<i>Allactaga sibirica</i>	-0.157	-0.157	-0.029	-0.029	-0.093	0.074	-0.029	0.086
	<i>Cricetulus barabensis</i>	-0.106	-0.106	-0.080	-0.080	-0.093	0.015		
	<i>Dipus sagitta</i>	-0.097	-0.097	-0.069	-0.069	-0.083	0.016		
	<i>Marmota baibacina</i>	0.122	0.122	-0.008	-0.008	0.057	0.075		
	<i>Marmota caudata</i>	0.089	0.089	0.015	0.015	0.052	0.043		
	<i>Marmota sibirica</i>	-0.149	-0.149	-0.095	-0.095	-0.122	0.031		
	<i>Meles leucurus</i>	-0.112	-0.111	0.007	0.008	-0.052	0.069		
	<i>Meriones meridianus</i>	-0.215	-0.215	-0.024	-0.024	-0.119	0.110		
	<i>Meriones unguiculatus</i>	0.007	0.014	-0.013	-0.014	-0.001	0.014		
	<i>Mustela eversmannii</i>	0.109	0.107	0.071	0.070	0.089	0.022		
	<i>Mustela nivalis</i>	0.117	0.116	0.066	0.068	0.092	0.028		
	<i>Ochotona dauurica</i>	-0.115	-0.115	-0.060	-0.040	-0.083	0.039		
	<i>Rhombomys opimus</i>	-0.152	-0.152	0.014	0.014	-0.069	0.096		
	<i>Spermophilus dauricus</i>	-0.140	-0.140	-0.017	-0.017	-0.078	0.071		
	<i>Spermophilus erythrogenys</i>	-0.022	-0.022	0.019	0.019	-0.002	0.024		
<i>Tscherskia triton</i>	0.040	-0.085	0.065	0.065	0.021	0.072			
<i>Urocitellus undulatus</i>	0.028	0.040	-0.042	-0.042	-0.004	0.044			
South	<i>Apodemus chevrieri</i>	-0.026	-0.026	-0.024	-0.024	-0.025	0.001	0.034	0.080
	<i>Callosciurus erythraeus</i>	0.021	0.024	0.012	0.010	0.017	0.007		
	<i>Eothenomys melanogaster</i>	0.059	0.059	0.081	0.081	0.070	0.013		
	<i>Niviventer andersoni</i>	-0.142	-0.142	-0.027	-0.027	-0.085	0.067		
	<i>Rattus losea</i>	0.159	0.159	0.021	0.021	0.090	0.080		
	<i>Rattus nitidus</i>	-0.011	-0.011	0.055	0.055	0.022	0.038		
	<i>Rattus rattus</i>	0.194	0.190	0.117	0.116	0.154	0.043		
<i>Rattus tanezumi</i>	0.071	0.071	-0.015	-0.015	0.028	0.050			
North & South	<i>Apodemus agrarius</i>	-0.056	-0.051	-0.058	-0.061	-0.056	0.004	0.085	0.109
	<i>Marmota himalayana</i>	0.227	0.227	0.078	0.078	0.153	0.087		
	<i>Micromys minutus</i>	0.225	0.227	0.086	0.083	0.155	0.082		
	<i>Microtus fortis</i>	-0.002	-0.002	-0.121	-0.121	-0.061	0.069		
	<i>Mus musculus</i>	0.152	0.152	0.085	0.081	0.117	0.040		
	<i>Pseudois nayaur</i>	0.255	0.255	0.108	0.108	0.182	0.085		
	<i>Rattus norvegicus</i>	0.153	0.153	0.092	0.092	0.123	0.035		
<i>Vulpes vulpes</i>	0.205	-0.021	0.035	0.043	0.065	0.097			

Table 3.8 Correlation of reservoir species niche suitability values through time (BMCP summed) and lagged (1 year) annual precipitation values. Correlation was compared across two climate scenarios each with clamped and unclamped model versions. Red-green colour scale was applied with red representing the lowest (most negative values) and green representing the highest (most positive values) correlations. Reservoir species are grouped by region.

Region	Species	Sum correlation with annual precipitation						Mean by Region	SD
		GFDL CLAMP	GFDL NO CLAMP	IPSL CLAMP	IPSL NO CLAMP	Mean	SD		
North	<i>Allactaga sibirica</i>	-0.145	-0.145	-0.032	-0.032	-0.088	0.065	-0.022	0.085
	<i>Cricetulus barabensis</i>	-0.096	-0.096	-0.071	-0.071	-0.083	0.015		
	<i>Dipus sagitta</i>	-0.094	-0.094	-0.072	-0.072	-0.083	0.013		
	<i>Marmota baibacina</i>	0.095	0.095	0.049	0.049	0.072	0.027		
	<i>Marmota caudata</i>	0.111	0.111	0.053	0.053	0.082	0.033		
	<i>Marmota sibirica</i>	-0.155	-0.155	-0.098	-0.098	-0.127	0.033		
	<i>Meles leucurus</i>	-0.113	-0.112	0.007	0.008	-0.052	0.069		
	<i>Meriones meridianus</i>	-0.197	-0.197	-0.010	-0.010	-0.103	0.108		
	<i>Meriones unguiculatus</i>	0.006	0.012	-0.015	-0.016	-0.003	0.014		
	<i>Mustela eversmanii</i>	0.109	0.107	0.070	0.069	0.089	0.022		
	<i>Mustela nivalis</i>	0.103	0.101	0.083	0.083	0.093	0.011		
	<i>Ochotona dauurica</i>	-0.110	-0.110	-0.060	-0.040	-0.080	0.035		
	<i>Rhombomys opimus</i>	-0.154	-0.154	0.024	0.024	-0.065	0.103		
	<i>Spermophilus dauricus</i>	-0.105	-0.105	-0.038	-0.038	-0.071	0.039		
	<i>Spermophilus erythrogegens</i>	-0.035	-0.035	0.012	0.012	-0.011	0.027		
<i>Tscherskia triton</i>	0.017	0.035	0.028	0.028	0.027	0.007			
<i>Urocitellus undulatus</i>	0.056	0.114	-0.033	-0.033	0.026	0.072			
South	<i>Apodemus chevrieri</i>	0.075	-0.051	0.003	0.003	0.008	0.052	0.051	0.085
	<i>Callosciurus erythraeus</i>	0.021	0.126	0.012	0.011	0.042	0.056		
	<i>Eothenomys melanogaster</i>	0.057	0.227	0.084	0.084	0.113	0.077		
	<i>Niviventer andersoni</i>	-0.175	-0.005	-0.037	-0.037	-0.063	0.076		
	<i>Rattus losea</i>	0.152	0.152	0.033	0.033	0.092	0.069		
	<i>Rattus nitidus</i>	-0.015	0.173	0.059	0.059	0.069	0.078		
	<i>Rattus rattus</i>	0.194	0.153	0.117	0.116	0.145	0.037		
<i>Rattus tanezumi</i>	0.071	-0.021	-0.015	-0.015	0.005	0.044			
North & South	<i>Apodemus agrarius</i>	-0.056	0.075	-0.058	-0.061	-0.025	0.067	0.046	0.100
	<i>Marmota himalayana</i>	0.126	0.024	-0.027	-0.027	0.024	0.072		
	<i>Micromys minutus</i>	0.225	0.057	0.086	0.083	0.113	0.076		
	<i>Microtus fortis</i>	-0.005	-0.175	-0.099	-0.099	-0.094	0.070		
	<i>Mus musculus</i>	0.152	0.152	0.085	0.081	0.117	0.040		
	<i>Pseudois nayaur</i>	0.173	-0.015	-0.055	-0.055	0.012	0.109		
	<i>Rattus norvegicus</i>	0.153	0.190	0.092	0.092	0.132	0.048		
<i>Vulpes vulpes</i>	0.205	0.071	0.035	0.043	0.088	0.079			

Table 3.9 Mean AUC difference for all species. Mean and SD are calculated across the species extent for all climate scenarios

Species	Mean AUC Difference	SD
<i>Allactaga sibirica</i>	0.041	0.039
<i>Apodemus agrarius</i>	0.082	0.060
<i>Apodemus chevrieri</i>	0.095	0.094
<i>Callosciurus erythraeus</i>	0.016	0.010
<i>Cricetulus barabensis</i>	0.066	0.039
<i>Dipus sagitta</i>	0.103	0.025
<i>Eothenomys melanogaster</i>	0.168	0.132
<i>Marmota baibacina</i>	0.151	0.124
<i>Marmota caudata</i>	0.048	0.025
<i>Marmota himalayana</i>	0.099	0.084
<i>Marmota sibirica</i>	0.095	0.081
<i>Meles leucurus</i>	0.178	0.155
<i>Meriones meridianus</i>	0.095	0.060
<i>Meriones unguiculatus</i>	0.074	0.078
<i>Micromys minutus</i>	0.039	0.038
<i>Microtus fortis</i>	0.154	0.067
<i>Mus musculus</i>	0.150	0.092
<i>Mustela eversmanii</i>	0.057	0.063
<i>Mustela nivalis</i>	0.058	0.038
<i>Niviventer andersoni</i>	0.100	0.064
<i>Ochotona dauurica</i>	0.052	0.026
<i>Pseudois nayaur</i>	0.030	0.023
<i>Rattus losea</i>	0.092	0.055
<i>Rattus nitidus</i>	0.050	0.042
<i>Rattus norvegicus</i>	0.125	0.050
<i>Rattus rattus</i>	0.072	0.028
<i>Rattus tanezumi</i>	0.076	0.054
<i>Rhombomys opimus</i>	0.200	0.186
<i>Spermophilus dauricus</i>	0.117	0.079
<i>Spermophilus erythrogenys</i>	0.097	0.092
<i>Tscherskia triton</i>	0.098	0.072
<i>Urocitellus undulatus</i>	0.087	0.094
<i>Vulpes vulpes</i>	0.065	0.065

3.6.2 Supplementary figures

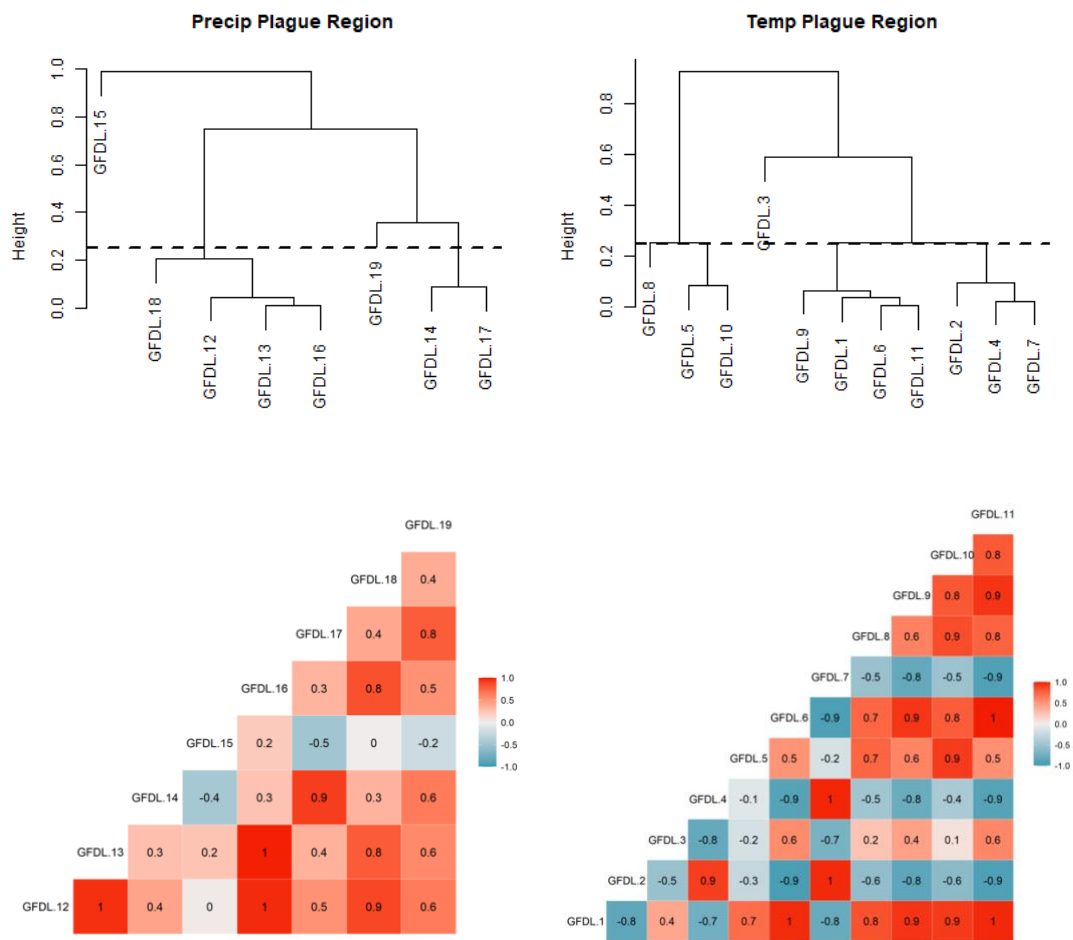


Figure 3.7 Variable selection through correlograms and cluster dendrograms. The top two plots are cluster dendrograms, the variables above the dashed line are the least correlated. As few variables were suggested as uncorrelated through the cluster dendrograms, correlograms (lower plots) were also used to identify the 4 least correlated precipitation (left) and temperature variables.

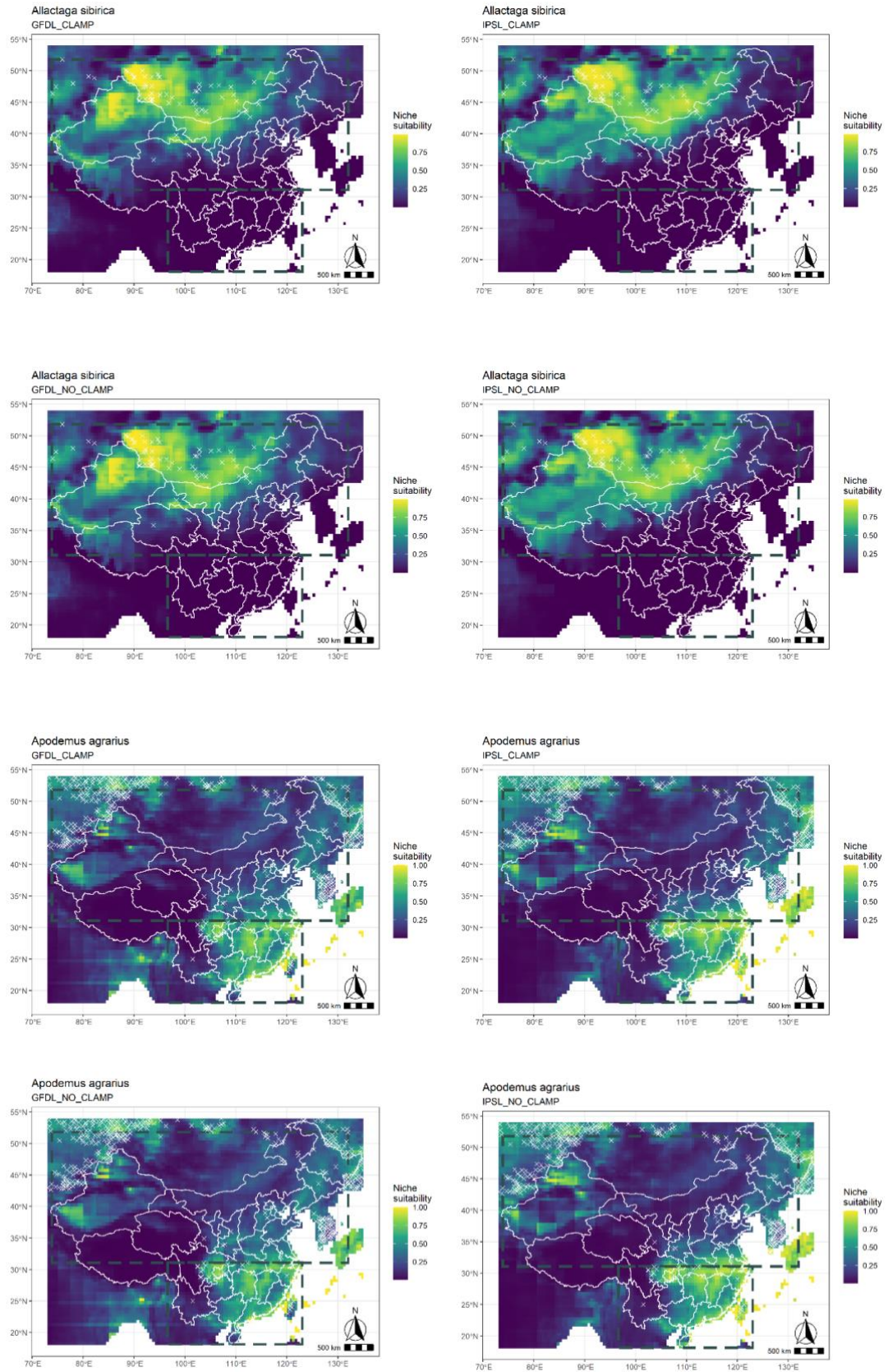


Figure 3.8 Niche suitability maps across selected species projected into the time averaged period (1970-2000 C.E.) for each climate and clamping scenario (noted in each plot title). Chinese regional boundaries (white line) and the Northern and Southern China regions (grey dashed lines) are shown. (Figure continued on following pages)

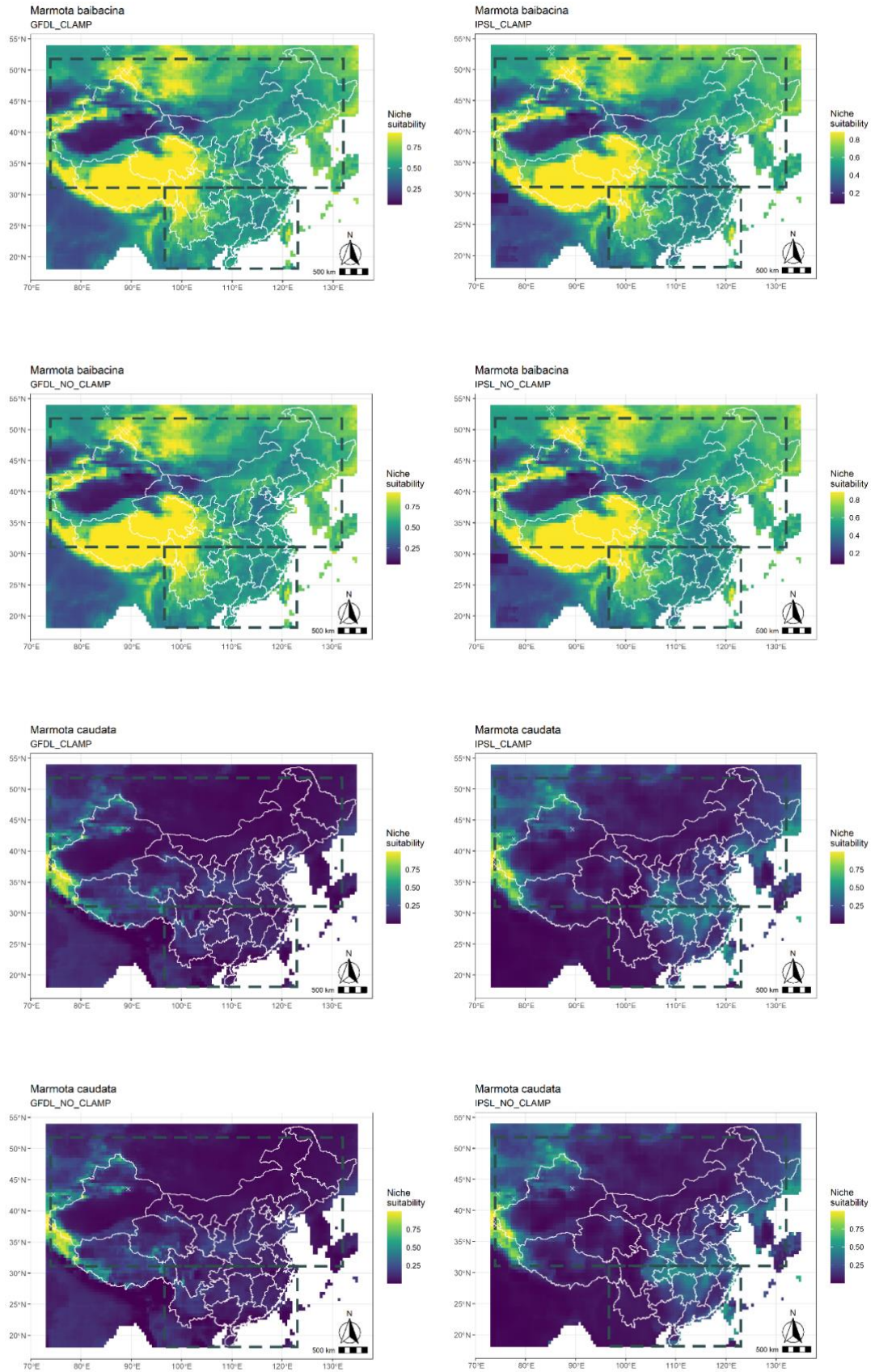


Figure 3.8 continued

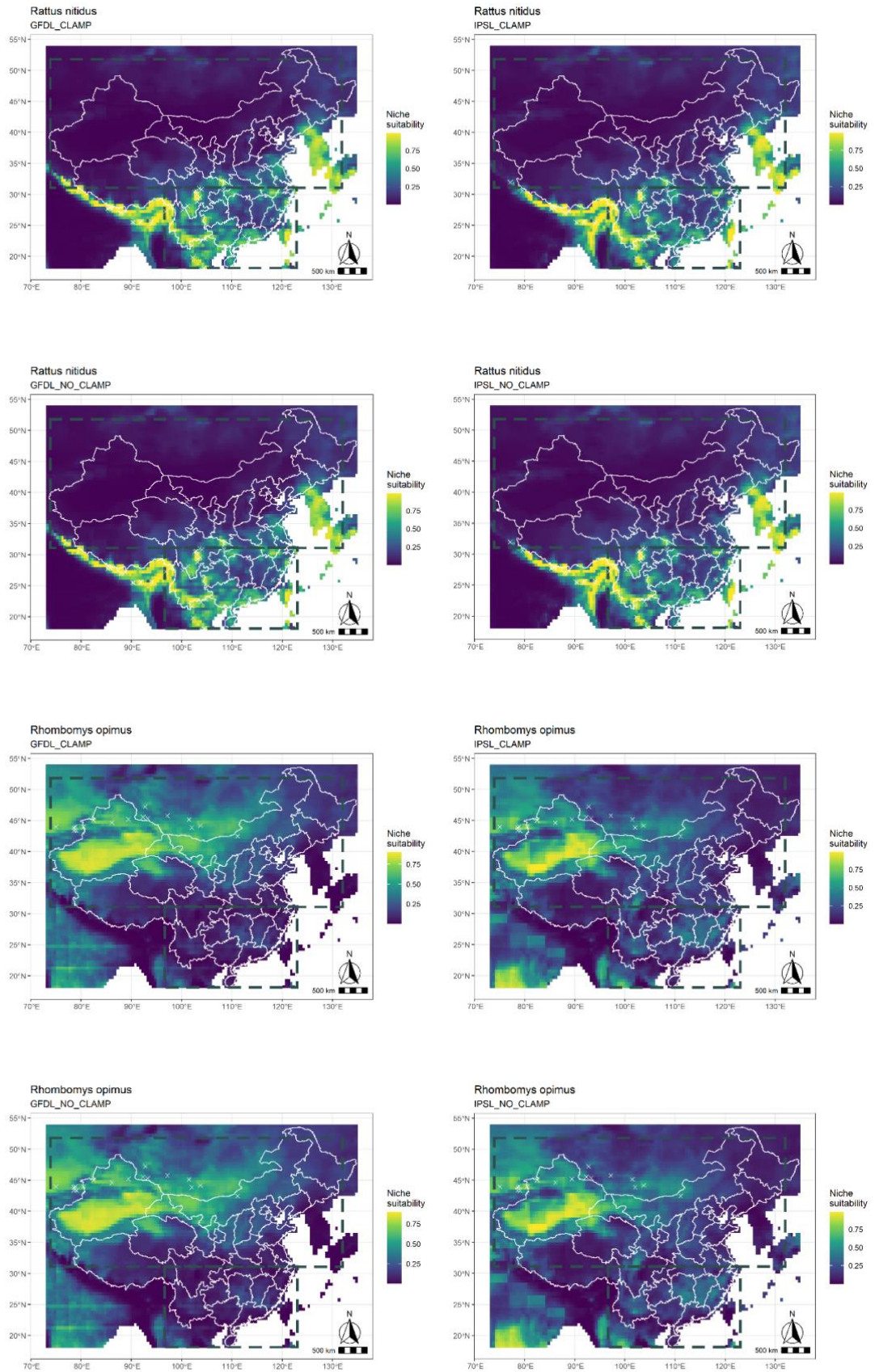


Figure 3.8 continued

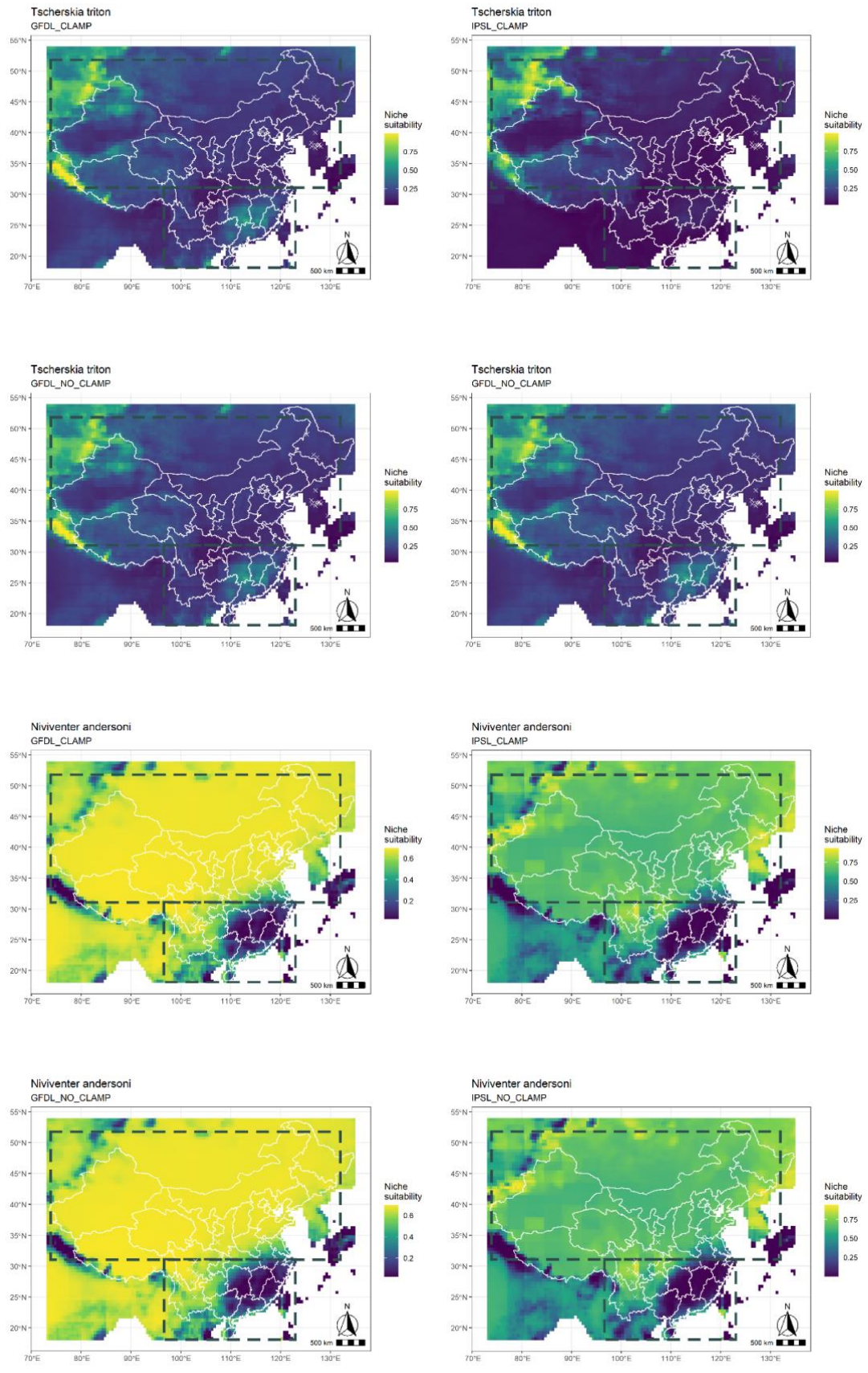


Figure 3.8 continued

3.6.3 Attempted multi-temporal calibration

Hindcasting ENM's into past climate space requires the assumption of temporal niche stability, where I assumed that the niche of a species was the same in both the period the model was built and projected (Nogués-Bravo, 2009). This assumption was critical when predicting any species over a several thousands of years and was further relevant for many species over hundreds of years for example highly climatically dynamic species such as pathogens or insects, both of which are key plague dynamics (Pearman et al., 2008). Further rapid recent morphological changes in rodents are suggested to be strongly associated with human population density as well as changing climate trends, which may suggest a lack of stability across some common commensal host species, which warrants further investigation (Pergams and Lawler, 2009).

This assumption should be considered whenever hindcasting as well as the assumption that species are at equilibrium with climatic conditions, which must be considered even when projecting a species within a single time period. Nogués-Bravo (2009) suggests a range of practices which should be implemented to limit the potential impact which these assumptions may have upon results. Several of these practices may be applicable to this study and I explore one particular method (multi-temporal calibration) explicitly.

Initial hindcasting focused upon the niche of *Y. pestis* instead of the niches of host species. Although this was not subsequently utilised for the hypothesis testing within this chapter it provided an example of a species likely to show niche plasticity over the study period due to rapid phylogenetic response to climatic change (Cui et al., 2020).

The suggested method which I attempted to include in hindcasting analysis is the use of multi-temporal calibration. Multi-temporal calibrations, where climatic and occurrence data from multiple time steps are combined into a single model or projection, are suggested to better represent the range of climatic conditions that a species utilises through time and therefore may lead to a greater transferability of

ENMs through time (Nogués-Bravo, 2009). This methodology is not isolated to hindcasting but is also applied to contemporary distributions where the species under investigation are highly climatically dynamic (Fancourt et al., 2015) or when extreme weather events, not visible in time averaged data, may have a large impact on species distributions (Bateman et al., 2012). I suggest that such methods may be applicable to the distribution of *Y. pestis* across Asia. Further the period of investigation covering the Third Pandemic is a highly volatile climatic period, particularly across Yunnan province. The period from the first human fatalities to the international pandemic transmission represents a particularly climatically unstable period with the eruptions of Tambora (1815 C.E.) and a large unknown eruption in 1809 C.E. which combined had a significant impact on Yunnan and the climate of South China (Cao et al., 2012). This combination of rapid response to climatic conditions by *Y. pestis* and the volatility of climatic conditions across the period of investigation lead to the hypothesis that; *Y. pestis* distribution prior to and during the early stages of the Third Pandemic would be best represented using models built with multi-temporal calibrations rather than time averaged bioclimatic data.

This analysis did not validate this hypothesis as there was negligible difference between a multi-temporal calibration methods and a time averaged methods both when testing the models against all occurrences and when testing against a temporal subset of the data. This method was therefore not developed any further with the integration of reservoir species.

4 Testing the niche of *Y. pestis* in soil through experimental microbiology

4.1 Introduction

The previous chapters (2 & 3) have focused on the impacts that reservoir species can have upon the transmission and distribution of *Y. pestis*. However, *Y. pestis* infection has regularly been recorded to re-emerge following prolonged periods of quiescence, during which it has not been sampled from reservoir populations (Ramalingaswami, 1995, Bertherat et al., 2007, Andrianaivoarimanana et al., 2013, Cabanel et al., 2013). One possible hypothesis which may explain these re-emergences is that *Y. pestis* can potentially persist in soils (Drancourt et al., 2006). This is counter to the prevailing hypothesis that, following the adaptive diversification which enabled vector transmission in *Y. pestis*, the potential for *Y. pestis* to survive in the soil may have been gradually lost (Hinnebusch et al., 2016). Several studies have suggested *Y. pestis* is capable of surviving in soil both experimentally and observationally (Karimi, 1963, Mollaret, 1963, Ayyadurai et al., 2008, Eisen et al., 2008, Malek et al., 2017) and I aim to further this literature by testing the impact that soil abiotic variables may have upon *Y. pestis*' potential for survival.

The prevailing hypothesis that *Y. pestis* is incapable of surviving in the environment is based on the assumption that the adaptive diversification which *Y. pestis* underwent led to an ecological change from generalist to specialist and hence following the transition to a host-vector life cycle, mutations in genes with functions influencing environmental survival may accumulate leading to host dependency (Hinnebusch et al 2016). However this hypothesis has been challenged with recent work suggesting *Y. pestis* can persist within soil environments, both independently and through infection of soil dwelling biota (Benavides-Montaño and Vadyvaloo, 2017, Markman et al., 2018). This alternate hypothesis has been forwarded through observational (Eisen et al., 2008) and experimental (Ayyadurai et al., 2008, Markman et al., 2018) studies, however, much further work is required regarding the mechanisms and environmental tolerances of *Y. pestis*' survival within soil.

Discussions of *Y. pestis* surviving in the soil are not an entirely new hypothesis with one of the earliest reports of *Y. pestis* by Alexandre Yersin, the co-discover of *Y. pestis* and after whom it was named, who cultured *Y. pestis* from the soil floor of a hut whose inhabitants had died of plague (Yersin, 1894, Ayyadurai et al., 2008). Further early evidence of *Y. pestis* surviving in soil can be found from the burrow of infected host species (Karimi, 1963, Ayyadurai et al., 2008) as well as early experimental soil survival experiments (Mollaret, 1963, Ayyadurai et al., 2008). There also appears to be a wealth of Russian studies investigating the survival of *Y. pestis* in soil for up to 8 years in experimental studies and 3 years in observational studies as well as interaction with soil bacteria and amebae, however, much of this work, including relatively recent papers, are only accessible to non-Russian speakers through rare review papers (Bukharin et al., 2005, Karimova et al., 2010).

Recent studies have furthered the evidence for *Y. pestis* survival in soil. One such paper focuses on a case from 2007 where a wildlife biologist in Arizona contracted plague in Grand Canyon National Park and subsequently died after contact with a recently deceased mountain lion (Eisen et al., 2008). The animal died as a result of the *Y. pestis* infection and post-mortem, a large pool of blood had formed in the soil under the animal's mouth and nose. The soil in which the blood had pooled was sampled and fully virulent *Y. pestis* was recovered (Eisen et al., 2008). This evidence appears to suggest that *Y. pestis* can survive for at least 3 weeks in soil under natural conditions, however, there are several caveats to these findings. The area where plague was sampled was heavily saturated in blood which will have formed an incredibly nutrient enriched environment for bacterial growth which is unlikely to be consistent throughout the environment. The area was also observed to be shaded and was in late October which led to limited UV exposure and an ambient temperature, these conditions are not consistent across environments where *Y. pestis* is found, however, the author notes that these conditions are analogous to conditions found in epizootic host burrows (Eisen et al., 2008). Much further work is required to gain greater understanding of the mechanisms of both environmental survival and subsequent transmission to susceptible reservoir species.

A recent publication found evidence of *Y. pestis* (Orientalis strain) recovery from a high salinity soil specimen from the edge of a salted lake “Chott” in Algeria (Malek et al., 2017). This work further suggested a geographical correlation between plague reservoirs in North Africa (and further the northern Hemisphere in general) and proximity (>3km) to salt sources, either the Mediterranean Sea or a Chott lake suggesting an adaptive tolerance to potential high salinity soils. The authors also tested the survival of strain Algeria1 in three natural Algerian soils, sterilised through autoclaving, and supplemented NaCl to mimic low and high salinity soils. After 5 weeks there was no difference in the survival of *Y. pestis* between these two treatments and both had only shown a minor decrease in recovered colonies through time. A further study has demonstrated the capability of *Y. pestis* to survive in soil for up to 40 weeks independent of other organisms (Ayyadurai et al., 2008). The soil used in the experiment was only identified as “Natural sand collected in the Marseille area”, this soil sample was sterilised to eliminate all other bacterial growth and subsequently inoculated with fully virulent *Y. pestis*. Samples were maintained at room temperature (range 18.7°C to 24.0°C) for the full 40 week sampling period. At each of the 26 sampling points across the 40 weeks mouse infection experiments were performed with *Y. pestis* recovered from the sand which was demonstrably shown to be fully virulent and caused similar symptoms as expected in mouse models of primary plague pneumonia (Lathem et al., 2007).

One mechanism for soil survival often discussed is the possibility of adherence to or infection of soil dwelling invertebrates. *Y. pestis* has been observed to both survive and replicate in several species of soil dwelling amoebae and nematodes. Markman et al. (2018) found that of 5 amoebae species found within prairie dog burrows, one species, *Dictyostelium discoideum* was observed to phagocytose *Y. pestis* which would then replicate intracellularly. The authors further suggested that through maintenance in soil amoebae within host burrows *Y. pestis* could re-enter the bloodstream of reservoir species by entering wounds either caused by burrowing or antagonistic host-to-host interactions (Markman et al., 2018). Such amoebae also form spores (other species under investigation in the study form cysts) and can survive dormant in the environment for months even under relatively extreme conditions

(Lawal et al., 2020), which Markman et al. (2018) suggested may explain *Y. pestis* 'cryptic' inter-epizootic re-emergence after 2-5 years. Benavides-Montaña and Vadyvaloo (2017) completed similar experiments with *Acanthamoeba castellanii* with results suggesting that multiple strains of *Y. pestis* can survive phagocytosis by *A. castellanii* and survive for at least 5 days at temperatures representative of winter host burrow conditions (4°C). *Y. pestis* also localises in *A. castellanii* trophozoites and it was further speculated that *Y. pestis* may persist in cysts for prolonged periods which may help explain interepizootic re-emergence observed in many global reservoirs (Barbieri et al., 2020). Benavides-Montaña and Vadyvaloo (2017) further investigated the mechanisms which contribute to survival of amoebae phagocytosis. The *PhoP* regulator, which is required for intracellular survival within host cells was found necessary for survival within *A. castellanii*, while expression of the Type 3 secretion system (T3SS), which inhibits phagocytosis by host cells was, as expected found to also block phagocytosis by the amoebae. However, the conditions previously found to induce expression of the T3SS (low Ca^{2+} and 37°C (Chen and Anderson, 2011) are unlikely to occur in the soil environment. This therefore suggests that within soil environments the T3SS is also active at low temperatures. *Y. pestis* has been observed to produce biofilm which can envelope the "head" and inhibit the feeding of the common soil nematode species *Caenorhabditis elegans* (Darby et al., 2002).

A recent investigation into the impact of climate, in the Tien Shan Mountains in north-western China, on the prevalence of *Y. pestis* strains with varying rates of biofilm production (*rpoZ* variants) has tentatively suggested that during cold and dry periods there is positive selective pressure towards the *rpoZ* variants (Cui et al., 2020). If a bottom-up trophic cascade relationship exists within the host-vector-bacteria system in this region, as is observed in several other plague reservoir systems (Collinge et al., 2005, Stenseth et al., 2006, Kausrud et al., 2007, Xu et al., 2015), then this may suggest greater biofilm production is promoted as a response to decreasing vector and host populations and densities during colder and drier periods. The authors state that the selection of the increased biofilm producing strains may alternatively be related to factors influencing the survival of *Y. pestis* in soil, however, the survival mechanisms require much further research to support this claim.

The survival of the bacterial species in soil is determined by a combination of biotic factors, such as co-habitation, predation and competition by other soil dwelling organisms, and abiotic factors, such as organic carbon content, nutrient content, pH, water tension and temperature (van Veen et al., 1997). Both Ayyadurai et al. (2008) and Malek et al. (2017) eliminated all biotic factors through autoclaving as it effectively sterilises the soil but also alters the soil structure causing the release of ammonium-N and amino acids, and may further decrease the surface area of clays (Trevors, 1996, Buchan et al., 2012). Inoculating the soils with chemicals such as methyl bromide or formaldehyde is another alternative sterilisation technique, however, they lead to contamination of the soil with toxic residues (Buchan et al., 2012). An alternate method which does not impact soil structure and has a much lesser impact on organic soil chemistry but may cause depolymerisation of carbohydrates is Gamma-Irradiation (γ -Irradiation) (Trevors, 1996). Any sterilization method eliminates biotic factors which may inhibit the survival of a given species in 'real world' conditions limiting the application of such results in epidemiological or ecological studies. The alternative use of non-sterile soils for inoculation experiments with *Y. pestis* is absent from modern literature despite the observational studies of *Y. pestis* surviving in 'real world' all be it niche circumstances (Eisen et al., 2008). As recent studies have worked to include further elements into the plague's hypothesised soil reservoir such as interactions with soil biota I must further look at *Y. pestis* in non-sterile soils to determine the viability of this reservoir when not controlling for biotic factors in the environment (Benavides-Montaña and Vadyvaloo, 2017, Markman et al., 2018). Generally, non-soil adapted bacteria, when introduced to natural soils will show a decline in growth of populations and long-term persistence is rare, therefore if any survival of *Y. pestis* can be demonstrated in soils this may further demonstrate the role of soil reservoirs in the long term persistence and evolution of plague (van Veen et al., 1997, Drancourt et al., 2006).

I therefore developed an experimental framework for testing the survival of *Y. pestis* in soils while varying specific soil variables in an attempt to determine the niche which *Y. pestis* may inhabit within the soils. Given strict protocols when working with fully virulent *Y. pestis*, significant prior training with repeated observation and testing was

required to complete experimental work within the Containment Level 3 (CL3) Laboratory where all work with *Y. pestis* must take place. I therefore completed several experiments with *Y. pseudotuberculosis* under the required Containment Level 2 (CL2) conditions to improve my aseptic technique while testing protocols for long term survival experiments. These initial work-up experiments included, *Y. pseudotuberculosis* biofilm assays with the soil nematode species *Caenorhabditis elegans*, long term maintenance and recovery of *Y. pseudotuberculosis* from an isolated sterile environment and soil sterilisation. All of which are presented below and provided the opportunity to improve my aseptic technique and develop novel methodology for use on the final *Y. pestis* soil survival experiment.

4.2 Material and methods

4.2.1 Growth media and growth conditions

Y. pseudotuberculosis strains (Table 4.1) were cultured in Luria broth Lennox 5 g/L NaCl, 5 gL⁻¹ peptone (Difco-Bacto™), 5 gL⁻¹ yeast extract (YLB) (Luria and Delbrück, 1943) at 30°C and *E. coli* OP50 were grown at 37°C in Luria broth (LB) (10 g/L NaCl, 5 gL⁻¹ peptone, 5 gL⁻¹ yeast extract) (Luria and Burrous, 1957) with aeration and shaking at 250 rpm. When necessary, 15 g/L agar was added to the media for plate assays. *Y. pestis* strains were grown at 30°C shaking (200 rpm) unless specified otherwise in Brain-heart-infusion (BHI) broth (Oxoid™) using 37 gL⁻¹ BHI powder.

For biotic biofilm assays, *Caenorhabditis elegans* were maintained at 22°C on modified nematode Growth Media (NGM) containing 1 g peptone, 1.2 g NaCl, 6.8 g Agar in 400 ml sterile distilled water, with further additions following autoclaving of 0.4 ml cholesterol (5 mg/ml in EtOH), 0.4 ml 1M CaCl₂, 0.4 ml Mg SO₄ and 10.0 ml 1M KH₂ PO₄ pH6 seeded with *E. coli* OP50 (Epstein and Shakes, 1995).

Throughout all *Y. pestis* soil survival assays the soils were maintained at 20°C. *Y. pestis* strains, following soil inoculation and recovery, were cultured in *Yersinia* selective agar YSA (Sigma- Aldrich) treated with the fungicide cycloheximide (cycloheximide 18079-10X10MI-F, 0.1% sol, Sigma- Aldrich) (40 ml/L of medium) and the antibiotic erythromycin (Erm) at 25 µg/ml. YSA plates seeded with aliquots from the soil assays were cultured at 30 °C.

Table 4.1 Full list of bacterial strains used throughout the experimental study

Bacterial Strain	Description	Source
<i>Escherichia coli</i>		
OP50	A uracil auxotroph nutrient source for <i>C. elegans</i> / Tc ^R	(Epstein and Shakes, 1995)
<i>Yersinia pseudotuberculosis</i>		
YpIII Parent	Parent strain of <i>Y. pseudotuberculosis</i> harbouring the virulence plasmid pYV Serotype O:3/ NaI ^R	(Rosqvist et al., 1988)
YpIII Δ nagC	YpIII lacking the repressor of the GlcNAc-operon NagC (Km ^R)	(Wiechmann, 2016)
<i>Yersinia pestis</i>		
CO92 parent	Parent strain, isolated from a patient who died Colorado, pCD1, pFra, pPst (Erm ^R)	(Doll et al., 1994)

4.2.2 *C. elegans* biofilm infection assays.

C. elegans (wild-type N2 Bristol, obtained from the *Caenorhabditis* Genetics Centre, University of Minnesota, USA) were maintained on NGM plates with 400 µl of *E. coli* OP50 spread on the surface and incubated at 22°C. Biofilm infection assays were performed as described by Atkinson et al. (2011). Bacterial strains (Table 4.1) were cultured overnight and subsequently diluted to an OD₆₀₀ of 1.0 in 2 ml. 400 µl of each culture was spread on fresh NGM plates and approximately 30 *C. elegans* were picked from the stock plates onto the plate seeded with *Y. pseudotuberculosis* and incubated at 22°C. Following 24h of growth a biofilm severity index was then assigned to each worm by rating the size and coverage of the biofilm as follows:

Biofilm severity index (%)

$$= \left\{ \left[\frac{\sum \{ \textit{severity} \times \textit{number of worms in this level} \}}{\textit{highest severity} \times \textit{total number of worms}} \right] \right\} \times 100$$

0 = no biofilm, 1 = small biofilm (often) around the buccal cavity, 2 = large, mono-focal biofilm, 3 = very large, multi-focal biofilm (Atkinson et al., 2011).

4.2.3 Long-term maintenance and recovery of *Y. pseudotuberculosis* from an isolated sterile environment.

13mm glass coverslips were placed in a petri dish with 2 strips of masking tape affixed to the base, all of which had been previously UV sterilised through use of UV sterilisation cabinet. The tape was used to provide stability for the cover slips and facilitate the easy picking of cover slips using tweezers. *Y. pseudotuberculosis* YpIII was cultured overnight, after which the OD₆₀₀ was normalised to 0.1 in fresh YLB and returned to the incubator until the OD₆₀₀ values representative of the exponential (0.6-0.8) and stationary (1-1.2) growth phases were achieved. 2 ml of each normalised culture was added to 2 ml tubes and centrifuged for 5 minutes at 10,000 x g. The supernatant was then discarded and washed with PBS, this process was repeated twice and following the second wash the pellet was re-suspended in 1 ml of PBS. 50 µl aliquots of the sample were pipetted onto the sterilised 13 mm cover

slips within the petri dishes. Each petri dish was further placed within a sealed plastic container (which had also been UV sterilised) with a digital hygrometer taped to the lid to enable the recording of humidity at each sampling point. All results were recorded within a sealed container with a 2 L internal volume maintained at a constant temperature of 30°C. To sample viable bacteria cover slips were carefully removed taking care to keep the coverslip horizontal as at high humidity the sample had often not dried to the cover slip. Each coverslip was then placed in a small tightly sealed petri dish with 0.5 ml of PBS. Using an orbital shaker each sample was agitated for three fifteen second periods to re-suspend the attached cells. Serial dilutions in PBS by factors of 10 were completed up to a dilution of 1/10000 and aliquots (25 µl, 50 µl and 100 µl) were spread onto YLB plates and incubated overnight at 30°C. Undiluted samples were centrifuged to concentrate into a pellet and re-suspended in 20 µl of PBS, which was further plated on to YLB. Four further plates (25 µl, 50 µl and 100 µl and remainder) were stored in the 4°C and were only checked following overnight growth at 30°C if the initial samples showed no growth to confirm that no viable cells survived. Following the overnight of the initial samples all plates were photographed as record of cell survival and to allow survival to be later quantified.

4.2.4 Soil sterilisation and validation

Model soils were prepared using a combination of peat and sand homogenised to a consistent grain size of $355 \mu\text{m} < x < 710 \mu\text{m}$ and a respective ratio of 0.75 and 0.25 to ensure that the availability of organic matter (assuming that the organic matter is bioavailable) would not be a factor in limiting bacterial growth (Fontaine and Barot, 2005). Homogenising the grain size was an attempt to normalise the pore space as if the grain size is consistent throughout the samples the pore space should remain similarly consistent. Following the mixing of the constituent parts the soil samples were sterilised by autoclaving three times in autoclave bags. Soil sterility was assessed by culturing autoclaved samples on LB agar, using a loop to mix the soil then streak on to LB agar.

4.2.5 CL3 general safety principles

For all work completed in the Containment Level 3 (CL3) laboratory detailed safety procedures and protocols were implemented to prevent exposure. This included ongoing training and monitoring to show continued competence of workers and was further accompanied by quarterly emergency training exercises. Whenever the lab was in use another CL3 trained member of the team was on site to provide support in the case of an emergency. CL3 lab work was always completed within a Class I Microbial Safety Cabinet (MSC) which maintains a negative air pressure gradient from outside the lab through the lobby and laboratory and out through the MSC to ensure the worker is protected at all times. *Y. pestis* cultures were double contained whenever they were removed from the MSC with custom carry cases suitable for agar plates or liquid cultures. The maximum liquid culture volume was limited to 20ml and racks were always placed perpendicular to the front of the MSC to limit the chance of a spillage leaving the MSC. All surfaces were cleaned with either 2 or 10 % Distel with a 3 minute contact time (Star Lab Group) and weekly cleaning of the lobby and lab was mandatory. All liquid cultures of *Y. pestis* were treated with 10% Distel prior to removal and autoclaving.

Appropriate personal protective equipment (PPE) was worn at all times and waste from the MSC was placed in sweetie jars (Sigma Aldrich) prior to removal from the MSC and combined with all other waste produced in the CL3 suite and was autoclaved prior to removal where it was autoclaved a second time before final disposal.

4.2.6 *Y. pestis* soil assay

The experimental work up was a multi-day process with sampling and recording results also taking multiple days to allow the longest possible period to observe bacterial growth. The methodology described was in a process of iterative improvement as restrictions due to the CL3 operating procedures required all changes to be approved by the University of Nottingham Safety Office prior to actioning these changes in the lab. Improvements were being approved and preparations made for a final multi-month experiment as the COVID-19 pandemic

forced lab closures, after which work was not able to restart. The methods described below are therefore in their final form, however, some of the methods had not yet been used within the lab due to the premature end to the work.

Y. pestis (Table 4.1) was grown overnight at 30 °C with shaking in 50 ml vented falcon tubes containing 10 mL BHI broth. Following centrifugation (5000 x g, 15 min) the pellet was re-suspended in x (determined by number of soil samples) ml dH₂O, the OD₆₀₀ noted and 1 mL aliquots were used to seed 5 g samples of pre-warmed sterile and non-sterile soils in 50 mL vented tubes along with up to 0.5 ml of sterile dH₂O to ensure that each soil sample was completely saturated so that no dry areas could be observed in the tube but without any pooling on the surface of the sample. Alterations were being made to improve the standardisation for future iterations of this experiment and will be discussed in Discussion of methodology.

Following inoculation and subsequently at each sampling period the tube and soil sample was weighed and CFUs were then determined. 1.5 mL of dH₂O was added to the soils to be sampled and then mixed by gentle tapping. The samples were then left for 5 minutes to allow the soil particles to settle and to reduce the likelihood of aerosolised *Y. pestis*. 1 mL of the supernatant was then transferred to 1.5 ml screw cap tubes and serially diluted to 10⁻⁵ and 50 ul of each dilution was transferred onto YSA plates (segmented ¼ plates used), treated with cycloheximide fungicide and incubated at 30°C for up to 5 days. Colony counts were then completed at both sampling periods.

For the final experimental run the moisture of the soil was maintained to avoid desiccation. The rate of moisture loss from the 5 g soil samples had previously been calculated at 22°C and 37°C (personal communication Vanina Garcia) and to maintain constant moisture throughout the volume of dH₂O to be added prior to each sampling point maintain a moisture content of ≈ 1.5 ml was calculated (to match the starting moisture) throughout the 28 day sampling period.

Table 4.2. Rate of moisture loss from soils in falcon tubes at two different incubation temperatures

Temperature	Day	Percentage H2O	Percentage evaporated H2O	Average percentage evaporated H2O/day
37°C	0	100	0	
	4	72.7	27.3	6.8%/day
	6	60.7	39.3	6.5%/day
22°C	0	100	0	
	1	96.9	3.1	3.1%/day
	4	90.5	9.5	2.4%/day

4.3 Results

4.3.1 *C. elegans* biofilm assay

The aim of early experimental work was to gain practical and theoretical knowledge working with *Yersinia* spp. Following previously published methods (Atkinson et al., 2011, Wiechmann, 2016, Elton, 2018), my early experiments focussed on improving proficiency using laboratory equipment and developing essential aseptic techniques. This covered a range of activities including correct preparation of liquid and solid growth media, maintenance of *C. elegans* and culturing of *Y. pseudotuberculosis* bacterial populations.

C. elegans, a free-living nematode species, is a commonly used model organism which has been used in infection assays alongside *Y. pseudotuberculosis* to develop an understanding of genetic pathways of *Yersinia* spp. which mediate biofilm formation (Atkinson et al., 2011). To gain experience working with the *C. elegans* biofilm infection model, assays were performed with the *Y. pseudotuberculosis* parent which produces large biofilm masses on the surface on *C. elegans*, and a genetically modified parent strain in which a regulatory protein NagC was inactivated through mutation. The NagC protein is responsible for metabolism of GlcNAc and is necessary for biofilm formation (Wiechmann, 2016) and the NagC mutant lacks the NagC

regulator protein and produces much reduced or no biofilm masses on the infected *C. elegans*.

As expected, infection assays revealed that the *Y. pseudotuberculosis* parent (*YP111 wt*) formed large biofilms around the anterior end of *C. elegans* (BSI, 37%), while the NagC mutant was incapable of forming any observable biofilm (BSI, 0%).

4.3.2 Long term maintenance and recovery of *Y. pseudotuberculosis* from an isolated sterile environment.

The primary aim of this work package was to investigate long term survival and recovery of *Y. pseudotuberculosis* in a nutrient depleted environment and develop experience and associated methodologies required for long term survival experiments for later application to *Y. pestis* soil survival experiments. The methods for this experiment were developed in an iterative process, with incremental changes made to the evolving methodology as repetitions of the experiment were completed. This enabled a range of factors which may impact bacterial survival to be investigated during method development.

The initial experiments were primarily to develop the laboratory techniques and construct the standard operating procedures (SOPs) for the novel methodology developed for this experiment, while the latter experiments aimed to provide experimental results at a publishable standard. The work covered a range of techniques; optical density normalisation, UV sterilisation and serial dilutions and ensured that conditions and protocols would be appropriate for work with *Y. pestis* under CL3 conditions.

An initial short-term (7 day) experiment was designed to compare the recovery of *Y. pseudotuberculosis* following incubation on sterile cover slips using Phosphate-Buffered Saline (PBS) solution or dH₂O to recover the bacteria from the glass slides at the end of the experiment (Table 4.3). PBS is used ubiquitously for cell recovery and washing and contains NaCl, KCl, Na₂HPO₄, KH₂PO₄ at concentrations intended to match the human body. However, for both the *Y. pseudotuberculosis* and *Y. pestis*

recovery experiments the focus was on recovery from the environment and not in association with a human host. There is evidence of *Y. pestis* surviving in soils with high salt content (Malek et al., 2017) which is marginally analogous to PBS solution. The presence of a solution with a comparable mineral content to PBS within the environment is unlikely, whereas water is ubiquitous within the environment. Table 4.3 shows the colony counts against the dilutions for the initial comparison of the PBS when compared to dH₂O and suggest that PBS promotes greater recovered colony growth than dH₂O.

Table 4.3 Impact of solution used to recover *Y. pseudotuberculosis* from cover slips. TNTC stands for too numerous to count.

Treatment	Dilution	Colony Count
dH ₂ O	1	Confluent
	1/100	TNTC
	1/1000	407
	1/10000	22
PBS	1	Confluent
	1/100	TNTC
	1/1000	TNTC
	1/10000	TNTC

Although humidity was not stabilised it was monitored over time and the recorded humidity from each experiment was combined (Figure 4.1) and suggest that humidity within the sealed container is initially very high, near 100%, and gradually declines until around day 20 when there is an increased rate of humidity decrease. The highest rate of humidity decrease is from day 20 to 40 after which the gradient decreases until sampling ends at day 60.

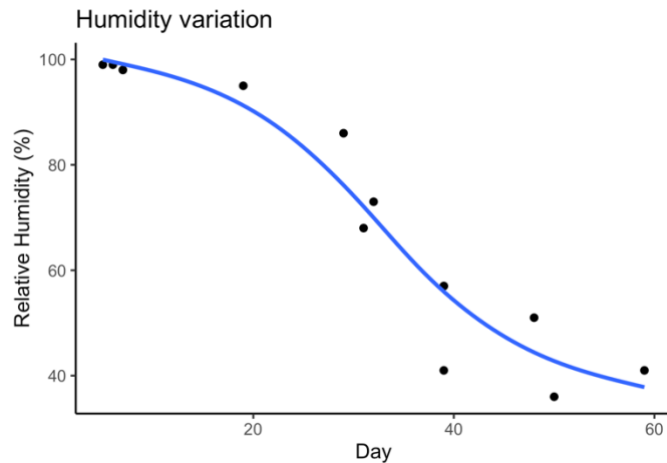


Figure 4.1 Variation in humidity through time in containment box

Optical density (OD) is routinely used to monitor bacterial growth by measuring the amount of light scatter (600 nm wavelength) through a liquid culture. This value acts as a proxy for the number of bacteria in a culture when combined with serial dilutions and colony counts on agar plates with the resulting growth curve showing the different growth phases. Based on previous estimates of when *Y. pseudotuberculosis* entered late exponential phase (OD₆₀₀ values of 0.8-1.0 - personal communication Vanina Garcia) cultures were incubated at 30°C and OD₆₀₀ monitored regularly and usually reached the required OD₆₀₀ between 2 h-3.5 h.

A follow-up experiment monitored growth until day 39 (RH = 41%) with data recorded as binary growth or no growth followed by an experiment to quantify the survival of *Y. pseudotuberculosis* at each sampling point by photographing each plate to allow later counting. However, survival was only observed to day 20, with no survival observed at the next sampling point, day 32 (RH = 72%).

In the final experiment I tested the hypothesis that *Y. pseudotuberculosis* survival was a function of the bacterial growth phase of *Y. pseudotuberculosis* at which the cover slip had been inoculated. This hypothesis is based upon the idea that Gram-negative bacterial cells in a stationary growth phase can transition to resistance cells capable of surviving in a sterile environment for long periods (Navarro Llorens et al., 2010). To determine if survival of *Y. pseudotuberculosis* in a sterile environment was a function of the growth phase at which the cover slips had been inoculated, cultures were standardised to an OD₆₀₀ of 0.6-0.8 (exponential growth phase) and 1-1.2

(stationary phase). The use of two separate treatments provided an analogue for later *Y. pestis* soil experiments as at each time step results needed to be recorded for both soil treatments and a suitable number of replicates. Samples were taken at 19, 29, 39, 50 and 59 days starting at day 19 as previous experiments had consistently demonstrated survival up to day 20. Samples were recovered and serially diluted to 10^{-4} . However, from day 50 onwards the colony counts at 10^{-4} exceeded the estimated 30-300 colonies and all serial dilution treatments was not varying by the expected factor, either suggesting a systematic sampling error or that greater serial dilution factors were required. The results therefore only confirm the survival of *Y. pseudotuberculosis* but unfortunately cannot be used to accurately quantify survival. The comparison of the two differing treatments is also inconclusive as the colony counts were not consistently greater for either treatment across all sampling periods. Independent of the treatments the general trend in the colony counts is an initial decrease from day 19-39 followed by an increase and stabilisation for the remaining time steps. The humidity was recorded as 57% at Day 39 and 36% at Day 50 after falling 95% at the initial sampling point (day 19) and eventually reaching 41% at the penultimate sampling point (day 59).

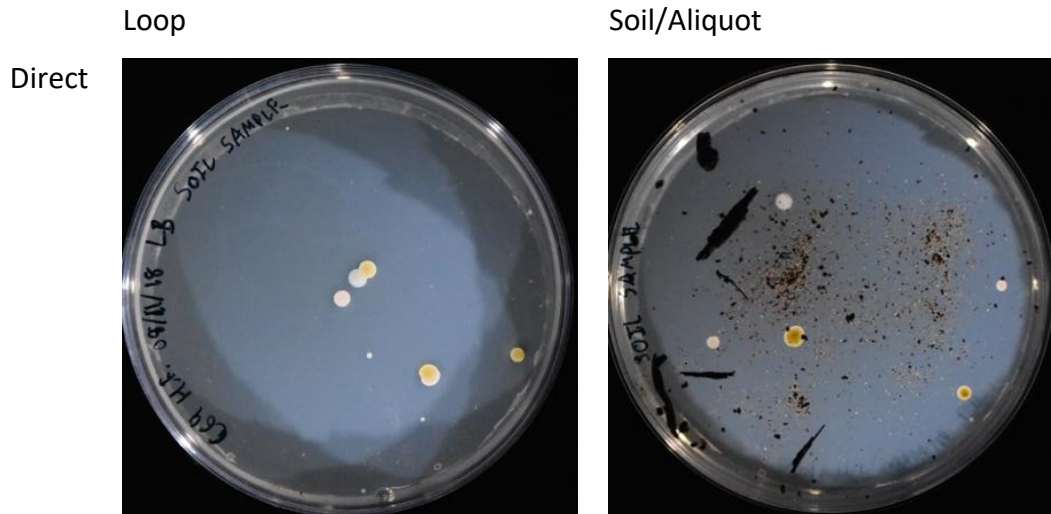
4.3.3 Soil sterilisation and validation

Previous experimental work exploring the survival of *Y. pestis* in soil was completed using sand sampled from the Marseille area (Ayyadurai et al., 2008). The aim of this methodology was to build on this previous work to formulate a standardised soil more representative of those where *Y. pestis* has been recovered previously (Eisen et al., 2008). The decline of introduced bacterial species within soil environments is generally dependent on protective pore space availability, nutrient availability, and the presence or absence of antagonistic microorganisms (van Veen et al., 1997). An attempt was made to normalise the pore space and nutrient availability while investigating the impact that potentially antagonistic microorganisms may have upon the survival of *Y. pestis*. The pore space was normalised through homogenising the grain size of the combined sand and peat to around that of medium grained sand ($355 \mu\text{m} < x < 710 \mu\text{m}$). The aim of the initial work was to confirm survival in sterilised

and un-sterilised homogenised soils with the potential of also varying environmental conditions such as temperature and humidity.

Soils were successfully sterilised through three autoclave runs using limited volumes to avoid clumping which can inhibit the penetration of steam (autoclave run: normal temperature 128.6 °C, pressure 1.4 bar, run cycle 22 mins). An issue with this methodology is that autoclaving alters the soil structure while releasing ammonium-N and amino acids (Trevors, 1996).

As the soil was not sterile after one autoclave run, this was an ideal juncture to test how well differing sterilisation methods performed (Figure 4.2). Soil samples were added directly to LB agar which was effective in identifying bacterial colonies as was using a loop and then streaking this onto an LB plate. In addition, small lumps of soil were added to LB broth, incubated at 30°C overnight and then the supernatant streaked onto LB agar which returned higher colony counts and was used for all subsequent experiments.



LB
Broth

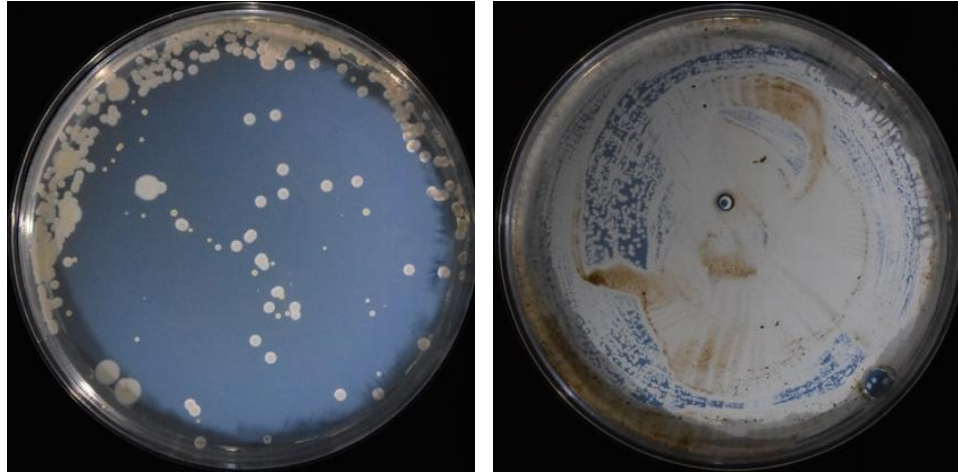


Figure 4.2. Comparison of sterility check methods in LB plates

4.3.4 *Y. pestis* soil assay

Three soil assay experiments with recorded results were completed prior to laboratory closures due to the COVID-19 pandemic, with each experiment intended to test and improve the methodology prior to the final three-month long survival experiment.

4.3.4.1 Experiment 1 – Preliminary methods work-up

Initially, survival of *Y. pestis* was tested in sterilised and non-sterilised soil under CL3 laboratory conditions over 12 days with colony counts completed on day 2 and 6 on *Yersinia* selective agar (YSA). Growth was observed in all four samples with much more growth observed in the sterilised soils. Colony counts could not be completed for the autoclaved soils as the colonies were too numerous to count (TNTC) at a serial dilution of 1×10^{-2} , however, no colonies were observed at 10^{-3} . Colonies were countable in the non-sterilised soils up to 10^{-2} .

As there was only one sampling period for the 12 day experiment there was limited opportunities to observe potential short falls in the sampling methodology to enable improvements. However, one area of potential improvement was the serial dilution process. The serial dilution was completed using 1.5ml screw cap Eppendorf tubes mounted in a rack which required 16 tubes for 4 treatments and 4 levels of serial dilution. Each serial dilution was then plated onto an individual plate and incubated.

As all samples in the CL3 laboratory require double containment and the transfer boxes can store a maximum of 8 plates this required 2 transfer boxes. Two improvements were made to this process. A colony count was completed for the non-sterile soil but the sterile soils were TNTC. A larger and flexible range of serial dilutions was needed, as suggested from the earlier long term *Y. pseudotuberculosis* experiments. To achieve this, instead of increasing the number of 1.5 mL tubes, serial dilutions were completed using a 96 well plate, greatly reducing the amount of equipment required. Space limitations meant that it was not feasible to use two transfer boxes for each sampling period, so to address this issue four compartment petri dishes were used which enabled a greater range of 10-fold serial dilutions to 10^{-7} , while reducing the number of plates. This enabled all samples from one sampling period to be stored in one transfer box.

4.3.4.2 Experiment 2 – Month-long run

The second recorded experiment extended the sampling period to 23 days (days 6, 12, 21 and 23) with soil samples stored in vented tubes and double contained in transfer boxes when outside the MSC. Ideally each sample would have been in triplicate but due to space issues this was not feasible during the early experiments. Instead at each sampling period one soil sample of each treatment was sampled twice (pseudo-replicate), which must be addressed when interpreting results.

Figure 4.3 shows a decrease in *Y. pestis* CFUs with no survival in the non-sterile sample following 12 days, whereas *Y. pestis* survived in the sterile soil sample until the end of the sampling period showing an exponential decline from inoculation. Although all other samples suggested an exponential decrease in recovered CFUs this cannot be confirmed for the non-sterile sample in this experiment due to only 2 non-zero sampling periods. One hypothesised mechanism which may facilitate *Y. pestis* survival in soil environments could be the formation of small colony variants (SCV) as a response to adverse conditions as observed in *Staphylococcus aureus* and several other pathogenic species (Kahl, 2014). Although small colony variants have not been observed in published literature for the *Yersinia spp*, it is plausible that this may be a survival mechanism in adverse conditions, such as *Y. pestis* in soil (Personal

communication Vanina Garcia). Potential small colony variants were observed at the day 12 sampling period in both non-autoclaved samples, however, only at lower dilutions with few large colonies observed. These potential small colonies were not included in the CFU calculation. Genetic sequencing would be required to confirm any small colony variants in future.

In order to recover samples from soil 1.5ml of dH₂O was added to each tube and mixed through gentle tapping. The soil sample was then left to settle for 5 minutes to reduce aerosols and then 1 ml of the soil water was removed with a 1 ml pipette. However, on occasion the pipette tips became blocked, thus limiting the amount of recoverable sample. This was particularly evident on day 21 and therefore the CFU value for this time point may be an underestimate. This issue would be partially resolved if the soils were maintained at a constant moisture increasing the recoverable liquid and avoiding sediment clogging the filter tips in future.

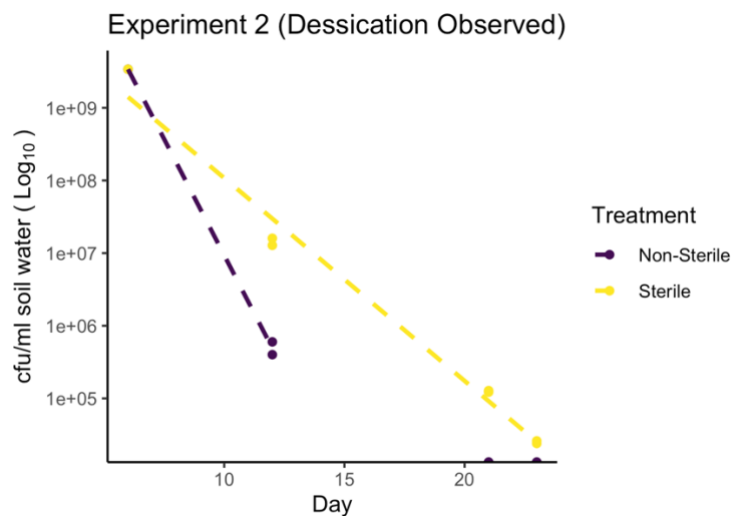


Figure 4.3. *Y. pestis* recovery experiment 2- Month-long run. Treatment methods are shown in differing colours (points and dashed line).

4.3.4.3 Experiment 3 – Month-long run with rehydration

In experiment 2, from day 21 onwards, cracks were observed in the surface of the soils of both treatments suggesting early stages of desiccation. To avoid this the moisture of the soils was maintained at a constant level equal to that at the beginning of the experiment, therefore limiting the impact of desiccation. This addressed the

soil moisture issue found in experiment 2 by attempting to maintain a constant moisture level in each of the soil samples throughout the experiment. When soil moisture was kept constant colony survival increased (Figure 4.4 compared to Figure 4.3) suggesting that moisture had been a limiting factor in experiment 2. At each sampling point the soil did not visually appear over-saturated, however, due to the large gap in sampling due to scheduling between day 5 and 17, 0.5 ml of dH₂O was added on day 7 leading to a large amount of surface pooling in both treatments, which appears counter to the aim of maintaining a constant moisture. A further issue is the lack of moisture validation throughout the sampling, as the moisture content throughout was calculated from estimates from an earlier experiment (Table 4.2). To avoid this fluctuation and provide ongoing validation in future, each sample would be weighed initially with the weight regularly checked and maintained through the addition of dH₂O at regular intervals.

As observed in experiment 2 the decrease in *Y. pestis* colonies follows an exponential decrease (Figure 4.3) while in this experiment the starting inoculum was also recorded allowing an opportunity to assess surviving colonies as a percentage of the initial inoculum if necessary. Experiment 3 confirms the earlier observations (Figure 4.4) that there is greater colony survival in sterile soil. Survival was recorded in non-sterile soils up to day 17 whereas sterile soil showed survival until the final sampling period (day 28) and although still decreasing in number the exponential decrease suggests that the colonies may be able to persist in low numbers well beyond the month-long sampling period (Figure 4.4).

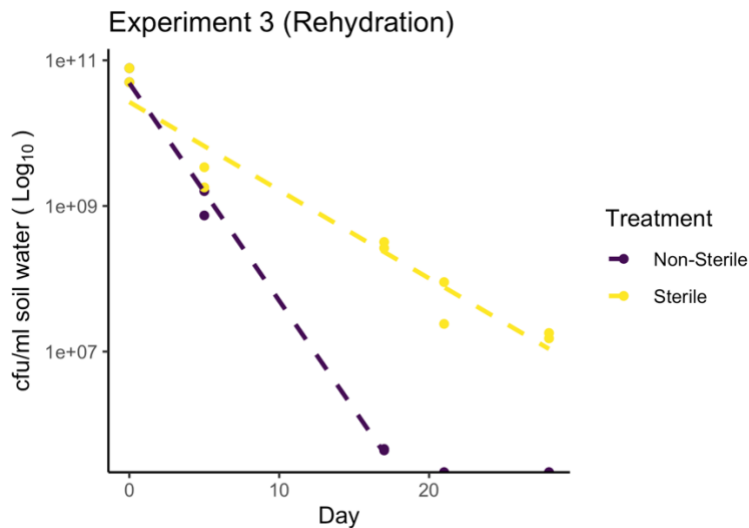


Figure 4.4. *Y. pestis* recovery Experiment 3 - with rehydration Treatment methods are shown in differing colours (points and dashed line).

When the data from Figure 4.3 and Figure 4.4 are combined all samples show an exponential decrease in CFUs (Figure 4.5), however, the lack of inoculum data for experiment 2 and particularly the lack of data for the non-sterile non-hydrated sample mean further work would be needed to confirm the consistency of the exponential decline. The sterile and rehydrated samples show the lowest rate of CFU decrease and hence greatest survival at the end of the sampling period whereas the non-sterile rehydrated sample and sterile non-hydrated sample show a similar rate of decrease, however, the sterile sample showed recoverable colonies for 23 days (the full sampling period of that experiment) whereas the non-sterile only showed recoverable colonies for 17 days. Therefore *Y. pestis* survived the complete sampling period in both sterile soil treatments, confirming validity of a longer sampling period in future work. The poorest performance of all treatments was the non-sterile non-hydrated sample which showed an extremely rapid decline and no recoverable colonies beyond day 12 (Figure 4.5).

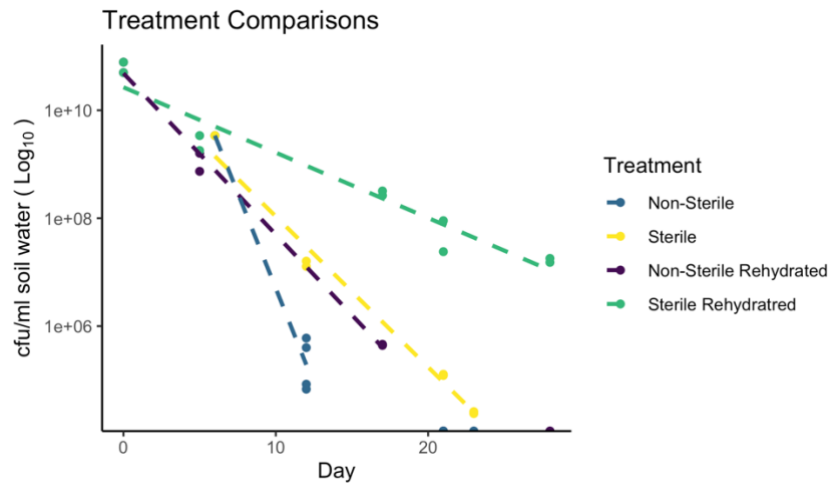


Figure 4.5. All *Y. pestis* soil survival treatment comparison. Treatment methods are shown in differing colours (points and dashed line).

4.4 Discussion

4.4.1 Discussion of experimental findings

4.4.1.1 Work-up experiments

The aims of the initial experiments, working with *Y. pseudotuberculosis*, were primarily to learn the necessary microbiological skills and techniques but also generated some novel data relating to *Y. pseudotuberculosis* survival. *Y. pseudotuberculosis* and *Y. enterocolitica* both survive in soil but little is known about the survival of these species on abiotic surfaces. *Y. enterocolitica* has been shown to survive for more than 21 days on the outside of milk containers (Stanfield et al., 1985) and *Y. pestis* can survive 48 h on sterile abiotic surfaces and up to 5 days on nutrient enriched abiotic surfaces (Rose et al., 2003). There is little literature on the survival of *Y. pseudotuberculosis* on sterile environments therefore the findings of this study suggesting the survival up to 72 days on a sterile abiotic surface and under varying humidity represent novel findings. The trend observed in the final experiment where colony counts under both treatments increased and plateaued after the humidity dropped below 57% at Day 39 potentially suggests a response relationship between humidity and *Y. pseudotuberculosis* survival. However, this is far from conclusive work and would need repetition with higher resolution sampling periods, a greater range of serial dilutions and a longer total sampling period to provide clarity for any

trends observed. The primary purpose of the initial run of *Y. pseudotuberculosis* experiments, however, was to improve aseptic technique and to gain a greater understanding of methods required to set up and sustain long term survival experiments, while training was ongoing to enable solo work within the CL3 suite.

4.4.1.2 *Y. pestis* soil assay

Although the final planned long term soil survival experiment could not be completed due to the impact of the COVID-19 pandemic lockdown, the data that was generated represents significant advances in the knowledge of *Y. pestis* survival in soil environments. It is clear that *Y. pestis* can survive in both sterile and non-sterile soil environments, with the most recoverable cells observed in hydrated sterile soil and the least in non-hydrated non-sterile soils. An exponential decrease in CFUs was observed in all experiments suggesting a rapid initial decrease in recoverable cells but with the potential for stabilisation in cell populations rather than falling to zero. Unfortunately the month-long experiments were not sufficient to confirm this hypothesis. It is therefore possible that *Y. pestis* may not survive indefinitely, however, both early and modern research suggests a much longer survival period than was demonstrated through this study (Ayyadurai et al., 2008; Mollaret, 1963; Karimi, 1963).

In the 12 day experiment small colonies were observed in both the sterile and non-sterile soils, however, due to the size of these colonies confirming the identity of the colony by visual inspection of morphology was not possible therefore they may represent some form of contamination instead. To completely remove the possibility of observations being influenced by fungal contamination as observed in the first experiment the YSA was treated with the fungicide cycloheximide in all subsequent experiments.

The observation that SCVs were present is noteworthy but further work is required to investigate the phenotypic and genotypic differences between these colonies and their normal colony variant counterparts and the environmental drivers which dictate the change in morphology. It is likely that the development of SCVs are a response to

the immediate environmental conditions that *Y. pestis* experiences within the soil and may, in part, be a stress response. Environmental stress can trigger the production of small colony variants in several bacterial species (*Staphylococcus aureus*, *Enterococcus faecalis*, and *Pseudomonas aeruginosa*) and is likely driven by the presence of antibiotics (von Eiff, 2008, Chandra and Kumar, 2017). Microbial production of antibiotics in the soil could potentially trigger this response (Chandra and Kumar, 2017). Alternatively the introduction of *Y. pestis* into a soil environment may represent a similar form of stress (Dekic et al., 2020) requiring an adaptive response which could include decreased growth rate and a reduced synthesis of adenosine triphosphate (ATP) (Johns et al., 2015), both of which could provide significant advantage to survival in nutrient depleted (at least compared to within reservoir species) environments over longer time scales. The decreased growth rate could enable *Y. pestis* to survive for longer periods and therefore increase the likelihood of colonies encountering a viable amebae or nematode species (Atkinson et al., 2011, Markman et al., 2018) or directly with burrow-dwelling reservoir species (Boegler et al., 2012). *Y. pestis* has been cultured from nutrient-enriched soil under the corpse of an infected animal 3 weeks following death (Eisen et al., 2008), however, survival in lower nutrient soils would be required for *Y. pestis* to overwinter and re-infect host species recolonising borrows of a deceased host population the following spring. SCVs may also exhibit increased adhesion and biofilm formation (Kahl, 2014), which would be beneficial within the soil environment (von Eiff, 2008, Johns et al., 2015) since natural survival is partially dependent on access to and colonisation of protected pore space. Increased biofilm formation may provide further protection to bacterial species once within soil pores and micro-pores (Vos et al., 2013). In addition, once stable within the soil the increased adhesion and biofilm production may provide an advantage to colonising soil dwelling amebae or nematodes as well as maintaining the cohesion of the *Y. pestis* colony within the environment.

Throughout the experiments soil moisture was shown to have an impact upon *Y. pestis* survival and recovery. However, it cannot be assumed that this result is purely due to moist soils better suiting a *Y. pestis* soil specific niche, as the soil moisture may

have impacted either the initial colonisation, or the bacterial recovery. All the soils were inoculated with *Y. pestis* and a maximum of 0.5 ml of sterile dH₂O, this was to ensure that each soil sample appeared fully saturated, e.g. moisture could be seen reaching to the bottom of the falcon tube. The logic behind this was to provide the bacteria with all available pore space within the sample (Vos et al 2013). In practice the full 0.5 ml was added in most cases but this was not recorded, and it therefore represents an unknown which should be removed in future work and in addition, the moisture content of the soils was not recorded prior to inoculation which is a further potential source of error. The moisture of the soils prior to inoculation is important as soil moisture can impact the long-term survival of bacteria in a soil sample (van Veen et al., 1997, Burmølle et al., 2011). This is due to the comparative accessibility of pore space dependent on moisture, in dryer soils a greater proportion of pore space is effectively vacant and can therefore be inoculated by the incoming bacterial cells whereas in moist soils much of the pore space is filled already and the incoming inoculum may have a more limited distribution (Postma et al., 1989). Sterile soils appeared to be moister than the non-sterile soils and this was likely due to the use of autoclaving as the steam undoubtedly increases the moisture content of the sterile soil. The *Y. pestis* inoculum may therefore have had less access to pore space within the soil samples, however, as results suggested higher recoverable populations from the sterilised soil the impact of access to pore space may be negligible when compared to the bacterial competition within the non-sterile soils.

Assuming that the results regarding moisture are not an artefact of either inoculation or recovery the results therefore suggest that soil moisture content is an environmental factor which influences the survival of *Y. pestis*. Previous research has suggested interactions between soil moisture and *Y. pestis* prevalence through both microbiological and macroecological investigation, (Ayyadurai et al., 2008, Pauling et al., 2021). *Y. pestis* is not universally thought to survive as a free-living organism due to adaptive specialisation following evolution from the much more generalist *Y. pseudotuberculosis* (Sun et al 2014). As such there have been limited studies on the potential niche of *Y. pestis* external of host and vector species. *Y. pestis* has been shown to survive desiccation when inoculated onto sterile surfaces for 24 and 48

hours (Rose et al., 2003), *Y. pestis* survival under varying moisture conditions has not been well documented. The results do not suggest that *Y. pestis* can survive desiccation in soil but instead indicate that increased soil moisture content correlates with recoverable CFUs. Attempting to determine a mechanistic response between soil moisture and survival of *Y. pestis* is difficult as the exact survival mechanisms of even *Y. pseudotuberculosis* within soil, remain poorly resolved (Santos-Montañez et al., 2015). As the longevity of *Y. pestis* in both sterile and non-sterile soils was reduced in non-hydrated soils it is possible that the absence of water was a limiting factor.

These findings have some broader implications and may suggest that *Y. pestis* can survive independently within the soil environment and moisture can dictate the survival period and number of surviving bacteria. This is supportive of previous work suggesting that *Y. pestis* may survive in rodent burrows and could be implicated in inter-annual transmission, as well as longer periods of quiescence (Drancourt et al., 2006, Boegler et al., 2012). The limited temporal scale of this study cannot directly support either of these hypotheses, however, the findings around moisture may in future be used to inform further work into this phenomenon. High soil moisture can be expected in the burrow systems of host species and as such has been found to correlate with *Y. pestis* infection in *C. ludovicianus* coterries due to the impacts this has upon flea populations through benefitting larval survival and development rate (Savage et al., 2011). These findings suggest that the high soil moisture content within the borrows may also represent a suitable environment for *Y. pestis* independent of reservoir species and vectors. Burrow systems represent the most likely environment for *Y. pestis* to be introduced to the soil and if soil dwelling amoebae or nematodes are to be implicated in the transmission cycle, prolonging survival following introduction to the soil increases the chance of encountering either of these common soil dwelling species (Markman et al., 2018). Away from burrow systems our results may suggest that preservation of *Y. pestis* in soils may be more likely in areas of generally moist climate. The chance of a reservoir or vector species coming into contact with *Y. pestis* infected soils away from a burrow environment are much lower than in a burrow. However, if the regional moisture regime does impact survival of *Y. pestis* in soils this

may play a role in explaining varying responses of *Y. pestis* to precipitation observed across China (Chapter 3) and therefore warrants further investigation.

The sterility of the soils is the second factor which the results suggest impact the survival of *Y. pestis* within soil environments. This is an expected finding as bacterial competition is a key biotic factor that limits the longevity of an inoculum in a soil environment (van Veen et al 1998). This is both driven by limited availability of protected pore spaces, competition for resources and predation by soil biota (Acea et al., 1988). Previous findings suggest that *Y. pestis* can survive for at least 3 weeks in non-sterile soils in natural conditions, however, this may be due to the scale of inoculation and the nutrient content of the soils which can both be expected to be very large from the blood of a deceased mountain lion (Eisen et al., 2008). This study only showed recoverable colonies from non-sterile soils for 17 days, however, bacterial competition may not be the only cause of the 11 day discrepancy between the sterile and non-sterile soils (hydrated). This may in part be due to the sterilisation process (autoclaving) which impacts both soil structure and chemistry (Buchan et al., 2012). Prior to autoclaving grain size was normalised to 355:710 μm representative of medium grain sands. Assuming that the soils were well mixed and there was minimal clumping the homogenous grain size should lead to relatively consistent pore size between samples. A However autoclaving destroys the soil aggregates to some degree which will break up the grains and lead to an increase in finer (clay) grains (Berns et al., 2008). The heterogeneity in grain size following autoclaving and the clumping of soils caused by the increased moisture content through the penetration of steam will likely lead to differing grain and pore sizes between the sterile and non-sterile soils. The increase in clay fraction can increase the soil surface area and the amount of micro-pores, particularly beneficial for bacterial colonisation (Berns et al., 2008), and survival (van Veen et al., 1997), however, previous studies have found results counter to this suggesting autoclaving may also lead to a reduction in surface area through the collapse of small pores (Lotrario et al., 1995). Further soil analysis of the sterile and non-sterile soils are therefore required, as results vary across soils, to determine the exact structural change following autoclaving and the potential impact this may have upon bacterial growth. Autoclaving also leads to a

significant release of dissolved organic carbon when compared to non-sterilised soils through the release of trapped organic carbon from within soil particles and through detaching organic carbon from particle surfaces (Berns et al., 2008). The impact of increased organic carbon content on the growth and replication of *Y. pestis* should therefore be tested.

4.4.2 Discussion of methodology

A large proportion of the laboratory work completed was in the development of the novel methods required for soil assays as well as soil sterilisation and validation. Discussion of the limitations and potential improvements which could be made to these methods are vital moving forward.

As discussed above the use of autoclaving to sterilise the soils had several unintended impacts on soil chemistry and structure. Although all possible soil sterilisation methods will to some degree impact the soils, the use of autoclaving has notably more impact than other feasible methods such as gamma radiation (λ -radiation) (Trevors, 1996). Importantly λ -radiation has a lesser impact on grain size throughout the sample and has a significantly reduced impact on organic soil chemistry when compared to autoclaving (Berns et al., 2008). Access to the necessary equipment to enable sterilisation by λ -radiation was not feasible during the project, however, future work should prioritise access to such equipment.

Further to the sterilisation of the soils prior to inoculation the methods of sterility checks although adequate could also benefit from refinement as the long-term sterility of the soils was not monitored. Soils were sterilised three times and validated by mixing the soils with a loop which was then streaked onto an agar plate. This method checked for viable cells having survived sterilisation initially, however, slow growing cells may take longer (weeks to months) to form colonies if present (Trevors, 1996). To provide long-term sterility checks a control soil sample which is not inoculated with *Y. pestis* should be maintained at the same temperature as the soil assay. This control should also receive the same moisture treatment as the soil assay. The control can then be repeatedly validated as sterile through suspension of soil

samples in LB broth followed by plating out and inspecting (Berns et al., 2008). This enables the potential competitive impact of bacterial colonies to be flagged during prolonged soil assay experiments.

The sampling method used through the soil assay experiment can also be assessed critically. As stated in the methodology at each sampling point 1.5 mL of dH₂O was added to the falcon tube containing the soil sample and this was then mixed by “gentle tapping” with the aim of partially homogenising the sample allowing a representative sample to be extracted from dH₂O pooled on the surface. This method of mixing, however, was far from sufficient and likely did little to homogenize or dislodge *Y. pestis* colonies from within the soil. I therefore expect that the recorded CFU values were likely conservative. Vortexing the sample would be more effective and provide a more representative sample of the bacterial community within the soil environment but would require a set period to enable any particulates aerosolised within the tube to settle.

Finally, the construction of the soil samples used was somewhat arbitrary and is therefore an area of potential experimental expansion. For all the soil assays a mix 3:1 of peat and sand respectively was used. Previous *Y. pestis* soil assay research had used only “natural sand” collected in the Marseille area as the “soil” (Ayyadurai et al., 2008) therefore my study aimed to build on the previous work while formalising the soil selection. The selection of the peat:sand ratio was intended to provide a more “real world” example of soil than the previous research, however, in future work I would recommend using a range of “end member soils” to determine the impact that various soil variables may have upon *Y. pestis* survival. Future work should use high and low organic content soils, for example peat and sand (separately) homogenised to the same grain size, further each of these soil types should be run in duplicate with a coarse and very fine (clay) version of each to determine the potential impact that fine grains (clays) may have on *Y. pestis* survival through the colonisation of micro-pore space (England et al., 1993). A control of each of these treatments should also be maintained to allow a thorough soil analysis to be completed on each to determine a range of soil variables such as moisture content, permeability as well as chemical

analysis to attempt to understand the mechanisms which facilitate the survival of *Y. pestis* within the soil.

4.4.3 Conclusions and future directions

The results of this study represent an expansion of knowledge regarding the potential survival of *Y. pestis* in soil environments and the impacts that this finding has upon broader plague ecology and supports previous findings in suggesting that *Y. pestis* can survive, for brief periods, in non-sterilised soils (Eisen et al., 2008). Moisture may also limit bacterial survival in this study suggesting that *Y. pestis* has the greatest chance of survival in soils held at a constant high moisture content such as within the burrows of a range of host species (Kreppel et al., 2016). This stable high moisture environment is also key to the development of vector species and hence research in understanding these systems is already pertinent to understanding the complex mechanistic pathways within the plague system. Monitoring of climate conditions within burrow systems, as completed by Kreppel et al. (2016), and subsequently recreating these conditions in the laboratory to test the survival of *Y. pestis* within soils both independent of vectors and in association with flea vectors as well as potential soil biota vectors would aid in shedding light on this under investigated area (Markman et al., 2018).

5 Future directions and synthesis

5.1 Combined conclusions

In this thesis I have investigated a portion of the highly complex system which mediates the transmission of *Y. pestis*. This work has focused at two ends of the scalar spectrum, from macroecological investigation around the interaction between the *Y. pestis* bacterium, reservoir species and climate to microbiological investigation of the potential survival of *Y. pestis* completely independent of vector and reservoir species within the environment. I am interested in the environmental niche of plague at both ends of this spectrum and suggest that an increased understanding of a zoonotic disease's niche at a broad range of scales is key to understanding the past, current and future risks of such diseases. My work highlights the importance of host reservoir niches particularly in invaded regions and the potential to test historical hypotheses through hindcasting reservoir niches, while also trying to establish conditions which dictate the niche of *Y. pestis* within soil environments. I am, however, cognisant that I only focus on two disparate elements in the plague system which at this stage are difficult to combine into a single cohesive methodology. I therefore suggest that work focuses on establishing plague's niche conditions across a range of scales and taxonomies. In particular, focus should be placed on vector species, which I was unable to integrate into this thesis at either the micro or macro scale due to methodological and data constraints. To further integrate the human element of the plague cycle and inform societal risk, a methodological step needs to be taken which progresses from purely correlative ecological models to integrating mechanistic epidemiological methods (Redding et al., 2016).

In Chapter 2, I compared the environmental niche of *Y. pestis* between its native region and one of the recently invaded regions. I found the niche was not consistent between the regions and that reservoir species may partially explain this difference. The variation of reservoir species between regions will impact the niche space available to *Y. pestis* and vector species, and while the presence of suitable reservoir species does not guarantee the presence of *Y. pestis*, it is a key factor that must be considered when investigating the invasion of any zoonotic disease into a novel

region (Messina et al., 2015). In Chapter 3, I tested a proposed hypothesis suggesting that reservoir species may drive the observed heterogeneous response of plague intensity to precipitation during the Third Pandemic in China (Xu et al., 2011). I found little evidence to support this hypothesis with results showing a diverse response of reservoir species to precipitation which demonstrates the complexity of the plague system even when only considering reservoir species in one region. This work highlights the need to better integrate mechanisms impacting plague transmission within one species with large scale correlative studies which may identify patterns across very large scales yet provide little evidence of the mechanisms behind them. The final area of research within the thesis (Chapter 4) sits at the opposite end of the spatial scale and focused upon the potential survival of *Y. pestis* within soil environments. The initial findings from this work demonstrated the survival of *Y. pestis* in both sterile and non-sterile soils for 28 days (the complete experimental run period) and 17 days respectively. I found that the drying of soils had a negative impact on the survival of *Y. pestis* and thus maintaining a constant moisture content in the soils enabled a greater period of bacterial recovery. Unfortunately, the proposed final long-term experiment was abandoned due to laboratory closures caused by the COVID-19 pandemic, however, this work still provided initial novel findings regarding the survival of *Y. pestis* under different soil sterility and moisture conditions and the development of a methodology for recovering *Y. pestis* from soils which can be built on in future work.

These disparate scales of study only capture limited elements of the plague system. A range of elements have not yet been integrated, including the impact of vector species and the role which humans play directly or indirectly in transmission. However, my work hopefully represents steps towards a more holistic view of the plague system. I intend to build from this work by integrating a broader range of biotic and abiotic variables and constructing model systems across multiple scales (both spatial and temporal) while attempting to avoid the assumption that mechanisms visible at one scale are present at another (Ben-Ari et al., 2011). Here I present potential future directions and areas of synthesis within and between each chapter.

5.2 Chapter-specific conclusions

5.2.1 Chapter 2 - invasion of plague's pathogen (*Yersinia pestis*) in *novel geographical settings*.

My findings suggest that plague reservoir species play an important role in defining the niche of plague in a particular region, suggesting that the reservoir species (biotic factors) are a source of niche variation between regions. Biotic factors are a key element of a bacterial pathogen's niche as the pathogen is reliant on biotic factors for transmission and maintenance (Escobar and Craft, 2016). This is counter to the previous research suggesting that across North America host species are less important to the niche of *Y. pestis* than climatic conditions (Maher et al., 2010). My findings emphasise the importance of biotic factors, particularly reservoir species dynamics and suggest that when variation in the niche of *Y. pestis* is observed between regions, biotic factors may explain these differences (Fell et al., 2022). At large coarse scales, such as the continental scale of my study, biotic factors have been suggested to play a limited role in defining the niche of a species and are instead only important when studying disease dynamics within a population (Peterson, 2011). I suggest that in the case of plague, large scale biotic factors (reservoir species niche) are important in defining the niche of *Y. pestis* particularly when projecting the niche to novel spatial or temporal environments.

At the scale of my study, the impact of the biotic and abiotic factors used implies correlation but should not imply mechanisms dictating the limits of the observed niche. The niche of a complex zoonotic disease may be investigated across a range of scales and at smaller scales inferences regarding mechanisms may be made with confidence. For example, within an infected body *Y. pestis* is confined to one or more of the; lungs, circulatory system or lymphatic system, which represent the in-host niche (Rollins et al., 2003). At this scale the niche is linked directly to the mechanism by which the body was infected as there is a differing infection pathway which leads to *Y. pestis* inhabiting each "environment". I therefore suggest that future correlative work should consider a range of "nested niches" at different scales (eg. bacterial

niche, vector niche, soil niche, reservoir niche). The impact of a range of abiotic variables across these scales could then be tested and, through the responses, reveal a deeper understanding of the scales at which different mechanisms are applicable and avoid the transference of assumptions to inappropriate scales (Escobar and Craft, 2016). Alternatively the construction of hybrid correlative-mechanistic models could enable integration of narrower scale mechanistic models, for example of vector lifecycle or bacterial survival in soil environments, into broad scale macro-ecological models (Kearney et al., 2010). The microclimatic elements of such models could be tested in the laboratory through the methodology developed in Chapter 4. Defining the abiotic niche of *Y. pestis* in the soil environment can then be integrated into hybrid models as micro-climatic systems and could further be tested in the field through extensive soil sampling in plague reservoir regions.

5.2.2 Chapter 3 - The role of reservoir species in mediating plague's dynamic response to climate.

This investigation focused on the niche suitability of reservoir species and showed that it could not explain the observed response in plague intensity to precipitation in China during the Third Pandemic. This suggests that an element not captured in the models may mediate the impact of precipitation upon plague intensity as well as challenging the assumption that reservoir species responses' to climate are homogenous even within a single climatic regime (Xu et al., 2015). I have thus far focused only on ecological elements of the plague system, which limits the degree to which findings can be applied to human plague data, which is by far the most common plague occurrence data and the data in which the response (plague intensity) is observed. The ecological elements of the plague cycle are key to understanding the transmission of *Y. pestis* within host reservoirs which will cover the vast majority of transmissions and is still an area where much further research is required. However, during a human epidemic or pandemic I do not expect plague ecology to explain observed transmission, although re-introduction from reservoir species may play a significant part in plague re-introduction particularly in historical periods (Schmid et al., 2015). More holistic understanding of the full plague system will come from integrating ecological and epidemiological theory, data, and models.

Environmental-mechanistic models represent a potential method which could fill the gap between an ENM and human disease data and I suggest that such models should be tested with the plague system (Redding et al., 2016).

The use of ENMs for investigating other historical epidemics, pandemics or introductions using hindcasting methodology is an area of huge potential. However, the above criticisms (limitations of purely ecological methods in epidemiological systems) must be remembered. An area where hindcasting ENMs may be particularly applicable is to understand the introduction of a bacterial pathogen to a novel region. During these periods, in the case of plague, there is often an initial peak in human cases, followed by the establishment of a plague reservoir in novel species (Adjemian et al., 2007, Andrianaivoarimanana et al., 2013). The establishment of this reservoir can then dictate, through biotic factors, the dynamics of plague in this region, which may be key to predicting future risk (Fell et al., 2022). Hindcasting niche models could therefore be highly informative in the introduction of invasive zoonotic diseases as long as dispersal and biotic variables are adequately considered and suggested hindcasting methodology is adhered to (Nogués-Bravo, 2009). Such methods can be used to test proposed hypotheses, as in Chapter 3, but only if these hypotheses make clear falsifiable predictions. For example, in Madagascar (a possible area of research focus given recent epidemics), hindcasting methods could be used to test if landuse change since the introduction of *Y. pestis* may have impacted the niche of reservoir and vector species and potentially the soil niche of *Y. pestis*.

There are two main areas of future study prompted by this Chapter: 1) Exploring which elements of the plague system, if not host species dynamics, as my findings suggests, drove the heterogeneous response of plague intensity to precipitation across China during the Third Pandemic. And 2) Using ENMs to test further hypotheses relevant to earlier historic pandemics or recent introductions to novel environments such as Madagascar or North America at the turn of the 20th Century.

5.2.3 Chapter 4 - Testing the niche of *Y. pestis* in soil through experimental microbiology

As the final planned experiment was cut short due to the COVID-19 laboratory closures, the first step beyond this research would be to complete the planned three-month-long experiment. The methodology developed throughout the pilot experiments would be further supplemented with the use of newly approved equipment for use in the CL3 laboratory (scales and shaking plate), which improve the accuracy and ease of maintaining a constant moisture content, and the homogenisation of the inoculated soil which would likely increase bacterial recovery. This three-month period was suggested as a logical step up from the month-long experiments regarding available space for samples while maintaining a minimum sampling every two weeks to produce moderate temporal resolution data. The subsequent dynamics of *Y. pestis* survival over this time period would then dictate whether increased sampling resolution would be required or a longer overall time period. The selection of three months is also a biologically meaningful period with regards to annual plague transmission as, in most reservoirs, *Y. pestis* transmission is greatly reduced during winter months due to limited vector numbers, and hibernation or decreased activity in reservoir species (Duplantier et al., 2005, Levick et al., 2015).

The *Y. pestis* genome is capable of rapid adaptation in response to climatic conditions (Cui et al., 2020), raising the question of whether a similar genetic response follows introduction to a soil environment? This genetic response could be investigated using high-throughput sequencing techniques to sequence colonies following recovery from soil environments. Observed up or down regulation of various genes may provide a valuable insight into mechanisms enabling prolonged soil survival and would require integration with current hypotheses of *Y. pestis* evolution which predominantly suggest the development of vector borne transmission inhibited the potential for *Y. pestis* to survive in the environment (Hinnebusch et al., 2016). Biofilm formation within vectors is a key factor in the transmission of *Y. pestis* and is an area of consistent ongoing research (Cui et al., 2020, Hinnebusch et al., 2021), however, the potential impact that biofilm formation may have within the soil environment

requires much further investigation (Burmølle et al., 2011). Formation of highly adhesive biofilms such as observed in the SCV of several bacterial species could potentially be beneficial within this environment. The impact of biofilms within soils could be tested using biofilm deficient mutant strains and help to elucidate the mechanisms which may facilitate the survival of *Y. pestis* within soil environments.

Following the completion of the initial planned long-term experiment there are three key areas I would like to focus on. Firstly, I would use high-throughput DNA sequencing to search for genetic variation (e.g. SCVs, or upregulated biofilm production) during maintenance within soils that could be indicative of a rapid genetic response to introduction to such environments. As part of this work, I would also investigate the survival of *Y. pestis* mutants, such as NagC which has a greatly inhibited biofilm production, to test which potential mechanisms *Y. pestis* may utilise in soil environments. Secondly, I would like to vary the soil variables (e.g. grain size, porosity, mineral content and organic content) and abiotic conditions (e.g. temperature and moisture) within the soil to determine which conditions are most conducive to survival and replication. Constructing an experimental soil niche for *Y. pestis* would enable integration with the macro ecological methods through testing this experimental niche in real world scenarios and the use of globally gridded soil data (eg. SoilGrids Hengl et al. (2014)) as well as data collection in the field from plague reservoir environments (Kreppel et al., 2016).

5.2.4 Future directions and conclusion - Synthesised microbiological and macroecological studies

Several areas of potential synthesis have been discussed above and all attempt to deal with the issue of scale and the difficulty in extrapolating findings from one scale to an environment of a different scale (i.e. the transmutation problem) (Ben-Ari et al., 2011, McGill, 2019). There is also a limit to the modelling systems used, for example although ENMs can be used to assess the niche of a zoonotic disease, integrating human disease data requires integration of mechanistic epidemiological methods (Escobar and Craft, 2016, Redding et al., 2016). This is not a requirement for

studies of zoonotic diseases using ENM methods, as demonstrated in Chapters 2 & 3 through the use of such methods to test ecological hypotheses.

Integrating experimental microbiological findings and macro-scale ecological or epidemiological data represents a novel challenge through combining experimental and observational data. Response relationships observed under laboratory conditions, for example the impact of soil moisture on *Y. pestis* survival, are difficult to apply directly to a “real world” ecological context. This could be addressed through working to better align the laboratory conditions and materials (soils) used with expected conditions in a plague reservoir region through use of soil samples from these regions. Further fine scale observational studies where micro-climatic variables specific to reservoir burrow locations are collected could inform the abiotic conditions to be replicated within the laboratory. As well as informing laboratory conditions, micro-scale climate data can be used to test micro-climate models constructed to bridge across scales of investigation and integrate mechanistic elements into regional and continental scale correlative investigations (Kearney et al., 2014). Integrating micro-scale high temporal resolution data into such models may lead to much more accurate estimates of the dynamics of *Y. pestis* across the range of environments it operates within, all of which are climatically dynamic (Ben-Ari et al., 2011, Xu et al., 2015, Cui et al., 2020). At the opposite end of the scale, an integration of ecological and epidemiological modelling methods could enable the meaningful integration of human plague records, which could then aid in directly informing risks under changing abiotic conditions (Redding et al 2016).

The above suggestions represent a highly complex modelling system, however, this is representative of the plague transmission system which is a highly complex system operating across a broad range of scales. The macroecological methods I have used in chapters 2 & 3 can be hugely informative when testing specific ecological hypothesis at the scale of the investigation, however, the transferability of results to different scales is limited. Therefore, such models should be used with extreme care, particularly when informing local risk of climatically dynamic zoonotic diseases (Ben-Ari et al., 2011, Escobar and Craft, 2016, Redding et al., 2016). Similarly, the

experimental microbiological methods used in Chapter 4 should not be extrapolated to larger scales without further investigation to bridge the gap in scale. My macroecological and microbiological work represents two vital end members of the plague system. To fully integrate the findings of such studies, work at intermediate scales is necessary to provide a more complete estimate of the system. Planned future work on the plague system will focus on constructing model systems across a range of scales, avoiding the transference of inappropriate assumptions between scales and lead to a deeper understanding of this highly complex and important disease in both a historical and contemporary context.

6 Bibliography

- A. LEE-YAW, J., L. MCCUNE, J., PIRONON, S. & N. SHETH, S. 2021. Species distribution models rarely predict the biology of real populations. *Ecography*.
- ACEA, M. J., MOORE, C. R. & ALEXANDER, M. 1988. Survival and growth of bacteria introduced into soil. *Soil Biology and Biochemistry*, 20, 509-515.
- ADJEMIAN, J. Z., FOLEY, P., GAGE, K. L. & FOLEY, J. E. 2007. Initiation and spread of traveling waves of plague, *Yersinia pestis*, in the western United States. *The American journal of tropical medicine and hygiene*, 76, 365-375.
- ADJEMIAN, J. Z., GIRVETZ, E. H., BECKETT, L. & FOLEY, J. E. 2006. Analysis of Genetic Algorithm for Rule-Set Production (GARP) modeling approach for predicting distributions of fleas implicated as vectors of plague, *Yersinia pestis*, in California. *Journal of medical entomology*, 43, 93-103.
- AIELLO-LAMMENS, M. E., BORJA, R. A., RADOSAVLJEVIC, A., VILELA, B. & ANDERSON, R. P. 2015. spThin: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38, 541-545.
- ALGAR, A. C., MAHLER, D. L., GLOR, R. E. & LOSOS, J. B. 2013. Niche incumbency, dispersal limitation and climate shape geographical distributions in a species-rich island adaptive radiation. *Global Ecology and Biogeography*, 22, 391-402.
- ANDRIANAIVOARIMANANA, V., KREPPPEL, K., ELISSA, N., DUPLANTIER, J.-M., CARNIEL, E., RAJERISON, M. & JAMBOU, R. 2013. Understanding the persistence of plague foci in Madagascar. *PLoS neglected tropical diseases*, 7, e2382.
- ARAÚJO, M. B., PEARSON, R. G. & RAHBEK, C. 2005. Equilibrium of species' distributions with climate. *Ecography*, 28, 693-695.
- ATKINSON, S., GOLDSTONE, R. J., JOSHUA, G. W., CHANG, C.-Y., PATRICK, H. L., CÁMARA, M., WREN, B. W. & WILLIAMS, P. 2011. Biofilm development on *Caenorhabditis elegans* by *Yersinia* is facilitated by quorum sensing-dependent repression of type III secretion. *PLoS pathogens*, 7, e1001250.
- AYYADURAI, S., HOUHAMDI, L., LEPIDI, H., NAPPEZ, C., RAOULT, D. & DRANCOURT, M. 2008. Long-term persistence of virulent *Yersinia pestis* in soil. *Microbiology*, 154, 2865-2871.
- BARBIERI, R., SIGNOLI, M., CHEVÉ, D., COSTEDOAT, C., TZORTZIS, S., ABOUDHARAM, G., RAOULT, D. & DRANCOURT, M. 2020. *Yersinia pestis*: the natural history of plague. *Clinical microbiology reviews*, 34, e00044-19.
- BATEMAN, B. L., VANDERWAL, J. & JOHNSON, C. N. 2012. Nice weather for bettongs: using weather events, not climate means, in species distribution models. *Ecography*, 35, 306-314.
- BEALE, C. M., LENNON, J. J. & GIMONA, A. 2008. Opening the climate envelope reveals no macroscale associations with climate in European birds. *Proceedings of the National Academy of Sciences*, 105, 14908-14912.
- BEN-ARI, T., GERSHUNOV, A., TRISTAN, R., CAZELLES, B., GAGE, K. & STENSETH, N. C. 2010. Interannual variability of human plague occurrence in the Western United States explained by tropical and North Pacific Ocean climate variability. *The American journal of tropical medicine and hygiene*, 83, 624-632.
- BEN-ARI, T., NEERINCKX, S., AGIER, L., CAZELLES, B., XU, L., ZHANG, Z., FANG, X., WANG, S., LIU, Q. & STENSETH, N. C. 2012. Identification of Chinese plague foci from long-term epidemiological data. *Proceedings of the National Academy of Sciences*, 109, 8196-8201.

- BEN-ARI, T., NEERINCKX, S., GAGE, K. L., KREPEL, K., LAUDISOIT, A., LEIRS, H. & STENSETH, N. C. 2011. Plague and climate: scales matter. *PLoS pathogens*, 7, e1002160-e1002160.
- BENAVIDES-MONTAÑO, J. A. & VADYVALOO, V. 2017. *Yersinia pestis* resists predation by *Acanthamoeba castellanii* and exhibits prolonged intracellular survival. *Applied and environmental microbiology*, 83, e00593-17.
- BENEDICT, C. A. 1996. *Bubonic plague in nineteenth-century China*, Stanford University Press.
- BERNS, A., PHILIPP, H., NARRES, H. D., BURAUDEL, P., VEREECKEN, H. & TAPPE, W. 2008. Effect of gamma-sterilization and autoclaving on soil organic matter structure as studied by solid state NMR, UV and fluorescence spectroscopy. *European Journal of Soil Science*, 59, 540-550.
- BERTHERAT, E., BEKHOUCHA, S., CHOUGRANI, S., RAZIK, F., DUCHEMIN, J. B., HOUTI, L., DEHARIB, L., FAYOLLE, C., MAKREROUGRASS, B. & DALI-YAHIA, R. 2007. Plague reappearance in Algeria after 50 years, 2003. *Emerging infectious diseases*, 13, 1459.
- BEVINS, S. N., BAROCH, J. A., NOLTE, D. L., ZHANG, M. & HE, H. 2012. *Yersinia pestis*: examining wildlife plague surveillance in China and the USA. *Integrative zoology*, 7, 99-109.
- BLOIS, J. L., ZARNETSKE, P. L., FITZPATRICK, M. C. & FINNEGAN, S. 2013. Climate change and the past, present, and future of biotic interactions. *Science*, 341, 499-504.
- BOEGLER, K. A., GRAHAM, C. B., MONTENIERI, J. A., MACMILLAN, K., HOLMES, J. L., PETERSEN, J. M., GAGE, K. L. & EISEN, R. J. 2012. Evaluation of the infectiousness to mice of soil contaminated with *Yersinia pestis*-infected blood. *Vector-Borne and Zoonotic Diseases*, 12, 948-952.
- BORIA, R. A. & BLOIS, J. L. 2018. The effect of large sample sizes on ecological niche models: Analysis using a North American rodent, *Peromyscus maniculatus*. *Ecological Modelling*, 386, 83-88.
- BOTTONE, E. J. 1999. *Yersinia enterocolitica*: overview and epidemiologic correlates. *Microbes and infection*, 1, 323-333.
- BRAMANTI, B., STENSETH, N. C., WALLØE, L. & LEI, X. 2016. Plague: A disease which changed the path of human civilization. *Yersinia pestis: retrospective and perspective*. Springer.
- BRAUN, J. K., JOHNSON, A. A. & MARES, M. A. 2011. *Tamias umbrinus* (Rodentia: Sciuridae). *Mammalian Species*, 43, 216-227.
- BUCHAN, D., MOESKOPS, B., AMELOOT, N., DE NEVE, S. & SLEUTEL, S. 2012. Selective sterilisation of undisturbed soil cores by gamma irradiation: Effects on free-living nematodes, microbial community and nitrogen dynamics. *Soil Biology and Biochemistry*, 47, 10-13.
- BUKHARIN, O., GINTSBURG, A., ROMANOVA, Y. M. & EL'-REGISTAN, G. 2005. Mekhanizmy vyzhivaniya bakterii. *Survival Mechanisms in Bacteria*, Moscow: Meditsina.
- BURMØLLE, M., KJØLLER, A. & SØRENSEN, S. J. 2011. Biofilms in soil. In: GLIŃSKI, J., HORABIK, J. & LIPIEC, J. (eds.) *Encyclopedia of agrophysics*. Springer Science & Business Media.
- BUTLER, T. 2013. Plague gives surprises in the first decade of the 21st century in the United States and worldwide. *The American journal of tropical medicine and hygiene*, 89, 788.
- BUZOLEVA, L. & SOMOV, G. 2003. Adaptation variability of *Yersinia pseudotuberculosis* during long-term persistence in soil. *Bulletin of experimental biology and medicine*, 135, 456-459.
- CABANEL, N., LECLERCQ, A., CHENAL-FRANCISQUE, V., ANNAJAR, B., RAJERISON, M., BEKHOUCHA, S., BERTHERAT, E. & CARNIEL, E.

2013. Plague outbreak in Libya, 2009, unrelated to plague in Algeria. *Emerging Infectious Diseases*, 19, 230.
- CAMINADE, C., KOVATS, S., ROCKLOV, J., TOMPKINS, A. M., MORSE, A. P., COLÓN-GONZÁLEZ, F. J., STENLUND, H., MARTENS, P. & LLOYD, S. J. 2014. Impact of climate change on global malaria distribution. *Proceedings of the National Academy of Sciences*, 111, 3286-3291.
- CAO, S., LI, Y. & YANG, B. 2012. Mt. tambora, climatic changes, and China's decline in the nineteenth century. *Journal of World History*, 587-607.
- CARRAWAY, L. N. & VERTS, B. 1994. *Sciurus griseus*. *Mammalian Species*, 1-7.
- CHAMBERLAIN, S. A. & BOETTIGER, C. 2017. R Python, and Ruby clients for GBIF species occurrence data. PeerJ Preprints.
- CHANDRA, N. & KUMAR, S. 2017. Antibiotics producing soil microorganisms. *Antibiotics and antibiotics resistance genes in soils*. Springer.
- CHEN, A. 1996. The ecological characteristics and integrative management of agricultural rodent pests in southern China. *Theory and Practice of Rodent Pest Management*, pp 247–312.
- CHEN, Y. & ANDERSON, D. M. 2011. Expression hierarchy in the Yersinia type III secretion system established through YopD recognition of RNA. *Molecular microbiology*, 80, 966-980.
- CHRISTIE, A., CHEN, T. & ELBERG, S. S. 1980. Plague in camels and goats: their role in human epidemics. *Journal of Infectious Diseases*, 141, 724-726.
- CLIFF, A., SMALLMAN-RAYNOR, M., HAGGETT, P., STROUP, D. & THACKER, S. 2009. *Emergence and re-emergence. Infectious diseases: a geographical analysis*.
- COLLINGE, S. K., JOHNSON, W. C., RAY, C., MATCHETT, R., GRENSTEN, J., CULLY, J. F., GAGE, K. L., KOSOY, M. Y., LOYE, J. E. & MARTIN, A. P. 2005. Testing the generality of a trophic-cascade model for plague. *EcoHealth*, 2, 102-112.
- COWEN, P., GARLAND, T., HUGH-JONES, M. E., SHIMSHONY, A., HANDYSIDES, S., KAYE, D., MADOFF, L. C., POLLACK, M. P. & WOODALL, J. 2006. Evaluation of ProMED-mail as an electronic early warning system for emerging animal diseases: 1996 to 2004. *Journal of the American Veterinary Medical Association*, 229, 1090-1099.
- CUI, Y., SCHMID, B. V., CAO, H., DAI, X., DU, Z., RYAN EASTERDAY, W., FANG, H., GUO, C., HUANG, S. & LIU, W. 2020. Evolutionary selection of biofilm-mediated extended phenotypes in Yersinia pestis in response to a fluctuating environment. *Nature communications*, 11, 1-8.
- CUI, Y., YU, C., YAN, Y., LI, D., LI, Y., JOMBART, T., WEINERT, L. A., WANG, Z., GUO, Z. & XU, L. 2013. Historical variations in mutation rate in an epidemic pathogen, Yersinia pestis. *Proceedings of the National Academy of Sciences*, 110, 577-582.
- DANFORTH, M., TUCKER, J. & NOVAK, M. 2018. The deer mouse (Peromyscus maniculatus) as an enzootic reservoir of plague in California. *EcoHealth*, 15, 566-576.
- DARBY, C., HSU, J. W., GHORI, N. & FALKOW, S. 2002. Plague bacteria biofilm blocks food intake. *Nature*, 417, 243-244.
- DAVIS, S., BEGON, M., DE BRUYN, L., AGEYEV, V. S., KLASSOVSKIY, N. L., POLE, S. B., VILJUGREIN, H., STENSETH, N. C. & LEIRS, H. 2004. Predictive thresholds for plague in Kazakhstan. *Science*, 304, 736-738.
- DAVIS, S., LEIRS, H., VILJUGREIN, H., STENSETH, N. C., DE BRUYN, L., KLASSOVSKIY, N., AGEYEV, V. & BEGON, M. 2007. Empirical assessment of a threshold model for sylvatic plague. *Journal of the Royal Society Interface*, 4, 649-657.

- DEKIC, S., HRENOVIC, J., DURN, G. & VENTER, C. 2020. Survival of extensively- and pandrug-resistant isolates of *Acinetobacter baumannii* in soils. *Applied Soil Ecology*, 147, 103396.
- DEMEURE, C. E., DUSSURGET, O., FIOL, G. M., LE GUERN, A.-S., SAVIN, C. & PIZARRO-CERDÁ, J. 2019. *Yersinia pestis* and plague: an updated view on evolution, virulence determinants, immune subversion, vaccination, and diagnostics. *Genes & Immunity*, 20, 357-370.
- DOBROWSKI, S. Z., THORNE, J. H., GREENBERG, J. A., SAFFORD, H. D., MYNSBERGE, A. R., CRIMMINS, S. M. & SWANSON, A. K. 2011. Modeling plant ranges over 75 years of climate change in California, USA: temporal transferability and species traits. *Ecological Monographs*, 81, 241-257.
- DOLL, J. M., ZEITZ, P. S., ETTESTAD, P., BUCHOLTZ, A. L., DAVIS, T. & GAGE, K. 1994. Cat-transmitted fatal pneumonic plague in a person who traveled from Colorado to Arizona. *The American journal of tropical medicine and hygiene*, 51, 109-114.
- DORMANN, C. F., MCPHERSON, J. M., ARAÚJO, M. B., BIVAND, R., BOLLIGER, J., CARL, G., DAVIES, R. G., HIRZEL, A., JETZ, W. & KISSLING, W. D. 2007. Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30, 609-628.
- DOS SANTOS RIBAS, L. G., DE CÁSSIA-SILVA, C., PETSCH, D. K., SILVEIRA, M. J. & LIMA-RIBEIRO, M. S. 2018. The potential invasiveness of an aquatic macrophyte reflects founder effects from native niche. *Biological invasions*, 20, 3347-3355.
- DRANCOURT, M., HOUHAMDI, L. & RAOULT, D. 2006. *Yersinia pestis* as a telluric, human ectoparasite-borne organism. *The Lancet infectious diseases*, 6, 234-241.
- DUPLANTIER, J.-M., DUCHEMIN, J.-B., CHANTEAU, S. & CARNIEL, E. 2005. From the recent lessons of the Malagasy foci towards a global understanding of the factors involved in plague reemergence. *Veterinary research*, 36, 437-453.
- EADS, D. A., ABBOTT, R. C., BIGGINS, D. E. & ROCKE, T. E. 2020. Flea parasitism and host survival in a plague-relevant system: theoretical and conservation implications. *The Journal of Wildlife Diseases*, 56, 378-387.
- EARLY, R. & SAX, D. F. 2014. Climatic niche shifts between species' native and naturalized ranges raise concern for ecological forecasts during invasions and climate change. *Global ecology and biogeography*, 23, 1356-1365.
- EISEN, R. J., EISEN, L., OGDEN, N. H. & BEARD, C. B. 2016. Linkages of weather and climate with *Ixodes scapularis* and *Ixodes pacificus* (Acari: Ixodidae), enzootic transmission of *Borrelia burgdorferi*, and Lyme disease in North America. *Journal of medical entomology*, 53, 250-261.
- EISEN, R. J., PETERSEN, J. M., HIGGINS, C. L., WONG, D., LEVY, C. E., MEAD, P. S., SCHRIEFER, M. E., GRIFFITH, K. S., GAGE, K. L. & BEARD, C. B. 2008. Persistence of *Yersinia pestis* in soil under natural conditions. *Emerging infectious diseases*, 14, 941.
- ELITH, J., KEARNEY, M. & PHILLIPS, S. 2010. The art of modelling range-shifting species. *Methods in ecology and evolution*, 1, 330-342.
- ELTON, L. 2018. *The Role of NagC in Yersinia Pestis and Yersinia Pseudotuberculosis Biofilm Development and Insect Transmission*. University of Nottingham.
- ENGLAND, L. S., LEE, H. & TREVORS, J. T. 1993. Bacterial survival in soil: effect of clays and protozoa. *Soil Biology and Biochemistry*, 25, 525-531.
- EPSTEIN, H. F. & SHAKES, D. C. 1995. *Caenorhabditis Elegans: Modern Biological Analysis of an Organism*, Academic Press.

- ESCOBAR, L. E. & CRAFT, M. E. 2016. Advances and limitations of disease biogeography using ecological niche modeling. *Frontiers in microbiology*, 7, 1174.
- FAHLGREN, A., AVICAN, K., WESTERMARK, L., NORDFELTH, R. & FÄLLMAN, M. 2014. Colonization of cecum is important for development of persistent infection by *Yersinia pseudotuberculosis*. *Infection and immunity*, 82, 3471-3482.
- FANCOURT, B. A., BATEMAN, B. L., VANDERWAL, J., NICOL, S. C., HAWKINS, C. E., JONES, M. E. & JOHNSON, C. N. 2015. Testing the role of climate change in species decline: is the eastern quoll a victim of a change in the weather? *PLoS One*, 10, e0129420.
- FELL, H. G., BALDINI, J. U., DODDS, B. & SHARPLES, G. J. 2020. Volcanism and global plague pandemics: Towards an interdisciplinary synthesis. *Journal of Historical Geography*, 70, 36-46.
- FELL, H. G., OSBORNE, O. G., JONES, M. D., ATKINSON, S., TARR, S., KEDDIE, S. H. & ALGAR, A. C. 2022. Biotic factors limit the invasion of the plague pathogen (*Yersinia pestis*) in novel geographical settings. *Global Ecology and Biogeography*, 31, 672-684.
- FENG, X., PARK, D. S., WALKER, C., PETERSON, A. T., MEROW, C. & PAPEŞ, M. 2019. A checklist for maximizing reproducibility of ecological niche models. *Nature Ecology & Evolution*, 3, 1382-1395.
- FICK, S. E. & HIJMANS, R. J. 2017. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International journal of climatology*, 37, 4302-4315.
- FONTAINE, S. & BAROT, S. 2005. Size and functional diversity of microbe populations control plant persistence and long-term soil carbon accumulation. *Ecology Letters*, 8, 1075-1087.
- FOURCADE, Y., BESNARD, A. G. & SECONDI, J. 2018. Paintings predict the distribution of species, or the challenge of selecting environmental predictors and evaluation statistics. *Global Ecology and Biogeography*, 27, 245-256.
- FREDRIKSSON-AHOMAA, M., LINDSTRÖM, M. & KORKEALA, H. 2009. *Yersinia enterocolitica* and *Yersinia pseudotuberculosis*. *Pathogens and toxins in foods: challenges and interventions*, 164-180.
- GAGE, K. L. & KOSOY, M. Y. 2005. Natural history of plague: perspectives from more than a century of research. *Annu. Rev. Entomol.*, 50, 505-528.
- GALANTE, P. J., ALADE, B., MUSCARELLA, R., JANSA, S. A., GOODMAN, S. M. & ANDERSON, R. P. 2018. The challenge of modeling niches and distributions for data-poor species: a comprehensive approach to model complexity. *Ecography*, 41, 726-736.
- GAMSA, M. 2006. The epidemic of pneumonic plague in Manchuria 1910–1911. *Past & present*, 190, 147-183.
- GBIF 2019. GBIF Home Page. Available from: <https://www.gbif.org> [13 August 2018].
- GREEN, M. H. 2020. The four black deaths. *The American Historical Review*, 125, 1601-1631.
- GUISAN, A., PETITPIERRE, B., BROENNIMANN, O., DAEHLER, C. & KUEFFER, C. 2014. Unifying niche shift studies: insights from biological invasions. *Trends in ecology & evolution*, 29, 260-269.
- GUO, B., ZHANG, J., MENG, X., XU, T. & SONG, Y. 2020. Long-term spatio-temporal precipitation variations in China with precipitation surface interpolated by ANUSPLIN. *Scientific reports*, 10, 1-17.
- HE, Z., WEI, B., ZHANG, Y., LIU, J., XI, J., CIREN, D., QI, T., LIANG, J., DUAN, R. & QIN, S. 2021. Distribution and Characteristics of Human Plague Cases and *Yersinia pestis* Isolates from 4 Marmota Plague Foci, China, 1950–2019. *Emerging Infectious Diseases*, 27, 2544.

- HEIER, L., VILJUGREIN, H. & STORVIK, G. O. 2015. Persistence of plague outbreaks among great gerbils in Kazakhstan: effects of host population dynamics. *Population ecology*, 57, 473-484.
- HENGL, T., DE JESUS, J. M., MACMILLAN, R. A., BATJES, N. H., HEUVELINK, G. B., RIBEIRO, E., SAMUEL-ROSA, A., KEMPEN, B., LEENAARS, J. G. & WALSH, M. G. 2014. SoilGrids1km—global soil information based on automated mapping. *PloS one*, 9, e105992.
- HIJMANS, R. J., CAMERON, S. E., PARRA, J. L., JONES, P. G. & JARVIS, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, 25, 1965-1978.
- HIJMANS, R. J., PHILLIPS, S., LEATHWICK, J., ELITH, J. & HIJMANS, M. R. J. 2017. Package 'dismo'. *Circles*, 9, 1-68.
- HINNEBUSCH, B. J., CHOUIKHA, I. & SUN, Y.-C. 2016. Ecological opportunity, evolution, and the emergence of flea-borne plague. *Infection and immunity*, 84, 1932-1940.
- HINNEBUSCH, B. J. & ERICKSON, D. 2008. Yersinia pestis biofilm in the flea vector and its role in the transmission of plague. *Bacterial biofilms*, 229-248.
- HINNEBUSCH, B. J., JARRETT, C. O. & BLAND, D. M. 2021. Molecular and Genetic Mechanisms That Mediate Transmission of Yersinia pestis by Fleas. *Biomolecules*, 11, 210.
- HINNEBUSCH, B. J., PERRY, R. D. & SCHWAN, T. G. 1996. Role of the Yersinia pestis hemin storage (hms) locus in the transmission of plague by fleas. *Science*, 273, 367-370.
- HOLT, A. C., FRITZ, C. L., SALKELD, D. J., TUCKER, J. R. & GONG, P. 2009. Spatial analysis of plague in California: niche modeling predictions of the current distribution and potential response to climate change. *International Journal of Health Geographics*, 8, 38.
- HUBBART, J. A., JACHOWSKI, D. S. & EADS, D. A. 2011. Seasonal and among-site variation in the occurrence and abundance of fleas on California ground squirrels (*Otospermophilus beecheyi*). *Journal of Vector Ecology*, 36, 117-123.
- HUTCHINSON, G. E. 1957. Cold spring harbor symposium on quantitative biology. *Concluding remarks*, 22, 415-427.
- INGLESBY, T. V., DENNIS, D. T., HENDERSON, D. A., BARTLETT, J. G., ASCHER, M. S., EITZEN, E., FINE, A. D., FRIEDLANDER, A. M., HAUER, J. & KOERNER, J. F. 2000. Plague as a biological weapon: medical and public health management. *Jama*, 283, 2281-2290.
- JARRETT, C. O., DEAK, E., ISHERWOOD, K. E., OYSTON, P. C., FISCHER, E. R., WHITNEY, A. R., KOBAYASHI, S. D., DELEO, F. R. & HINNEBUSCH, B. J. 2004. Transmission of Yersinia pestis from an infectious biofilm in the flea vector. *Journal of Infectious Diseases*, 190, 782-792.
- JOHNS, B. E., PURDY, K. J., TUCKER, N. P. & MADDOCKS, S. E. 2015. Phenotypic and genotypic characteristics of small colony variants and their role in chronic infection. *Microbiology insights*, 8, MBI. S25800.
- KAHL, B. C. 2014. Small colony variants (SCVs) of Staphylococcus aureus—a bacterial survival strategy. *Infection, Genetics and Evolution*, 21, 515-522.
- KARIMI, Y. 1963. Natural preservation of plague in soil. *Bulletin de la Societe de pathologie exotique et de ses filiales*, 56, 1183-1186.
- KARIMOVA, T. Y., NERONOV, V. & POPOV, V. 2010. Development of views on natural focality of plague. *Biology bulletin*, 37, 725-732.
- KAUSRUD, K. L., BEGON, M., ARI, T. B., VILJUGREIN, H., ESPER, J., BÜNTGEN, U., LEIRS, H., JUNGE, C., YANG, B. & YANG, M. 2010. Modeling the epidemiological history of plague in Central Asia: palaeoclimatic forcing on a disease system over the past millennium. *Bmc Biology*, 8, 112.

- KAUSRUD, K. L., VILJUGREIN, H., FRIGESSI, A., BEGON, M., DAVIS, S., LEIRS, H., DUBYANSKIY, V. & STENSETH, N. C. 2007. Climatically driven synchrony of gerbil populations allows large-scale plague outbreaks. *Proceedings of the Royal Society B: Biological Sciences*, 274, 1963-1969.
- KEARNEY, M. R., SHAMAKHY, A., TINGLEY, R., KAROLY, D. J., HOFFMANN, A. A., BRIGGS, P. R. & PORTER, W. P. 2014. Microclimate modelling at macro scales: a test of a general microclimate model integrated with gridded continental-scale soil and weather data. *Methods in Ecology and Evolution*, 5, 273-286.
- KEARNEY, M. R., WINTLE, B. A. & PORTER, W. P. 2010. Correlative and mechanistic models of species distribution provide congruent forecasts under climate change. *Conservation letters*, 3, 203-213.
- KRASNOV, B. R., SHENBROT, G. I., MOUILLOT, D., KHOKHLOVA, I. S. & POULIN, R. 2006. Ecological characteristics of flea species relate to their suitability as plague vectors. *Oecologia*, 149, 474-481.
- KREMEN, C., CAMERON, A., MOILANEN, A., PHILLIPS, S., THOMAS, C., BEENTJE, H., DRANSFIELD, J., FISHER, B., GLAW, F. & GOOD, T. 2008. Aligning conservation priorities across taxa in Madagascar with high-resolution planning tools. *Science*, 320, 222-226.
- KREPPPEL, K. S., TELFER, S., RAJERISON, M., MORSE, A. & BAYLIS, M. 2016. Effect of temperature and relative humidity on the development times and survival of *Synopsyllus fonquerniei* and *Xenopsylla cheopis*, the flea vectors of plague in Madagascar. *Parasites & vectors*, 9, 1-10.
- KUGELER, K. J., STAPLES, J. E., HINCKLEY, A. F., GAGE, K. L. & MEAD, P. S. 2015. Epidemiology of human plague in the United States, 1900–2012. *Emerging infectious diseases*, 21, 16.
- LATHEM, W. W., PRICE, P. A., MILLER, V. L. & GOLDMAN, W. E. 2007. A plasminogen-activating protease specifically controls the development of primary pneumonic plague. *Science*, 315, 509-513.
- LAWAL, H. M., SCHILDE, C., KIN, K., BROWN, M. W., JAMES, J., PRESCOTT, A. R. & SCHAAP, P. 2020. Cold climate adaptation is a plausible cause for evolution of multicellular sporulation in *Dictyostelia*. *Scientific reports*, 10, 1-9.
- LEVICK, B., LAUDISOIT, A., WILSCHUT, L., ADDINK, E., AGEYEV, V., YESZHANOV, A., SAPOZHNIKOV, V., BELAYEV, A., DAVYDOVA, T. & EAGLE, S. 2015. The perfect burrow, but for what? Identifying local habitat conditions promoting the presence of the host and vector species in the Kazakh plague system. *PLoS One*, 10, e0136962.
- LI, W., ZHAO, S., CHEN, Y., WANG, Q. & AI, W. 2021. State of China's Climate in 2020. *Atmospheric and Oceanic Science Letters*, 14, 100048.
- LOTRARIO, J., STUART, B., LAM, T., ARANDS, R., O'CONNOR, O. & KOSSON, D. 1995. Effects of sterilization methods on the physical characteristics of soil: implications for sorption isotherm analyses. *Bulletin of environmental contamination and toxicology*, 54, 668-675.
- LOWELL, J. L., EISEN, R. J., SCHOTTHOEFER, A. M., XIAOCHENG, L., MONTENIERI, J. A., TANDA, D., PAPE, J., SCHRIEFER, M. E., ANTOLIN, M. F. & GAGE, K. L. 2009. Colorado animal-based plague surveillance systems: relationships between targeted animal species and prediction efficacy of areas at risk for humans. *Journal of Vector Ecology*, 34, 22-31.
- LURIA, S. & BURROUS, J. W. 1957. Hybridization between *Escherichia coli* and *Shigella*. *Journal of bacteriology*, 74, 461-476.
- LURIA, S. E. & DELBRÜCK, M. 1943. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics*, 28, 491.

- MAHER, S. P., ELLIS, C., GAGE, K. L., ENSCORE, R. E. & PETERSON, A. T. 2010. Range-wide determinants of plague distribution in North America. *The American journal of tropical medicine and hygiene*, 83, 736-742.
- MAHMOUDI, A., KRYŠTUFEK, B., SLUDSKY, A., SCHMID, B. V., DE ALMEIDA, A. M., LEI, X., RAMASINDRAZANA, B., BERTHERAT, E., YESZHANOV, A. & STENSETH, N. C. 2020. Plague reservoir species throughout the world. *Integrative Zoology*.
- MAIORANO, L., CHEDDADI, R., ZIMMERMANN, N., PELLISSIER, L., PETITPIERRE, B., POTTIER, J., LABORDE, H., HURDU, B., PEARMAN, P. & PSOMAS, A. 2013. Building the niche through time: using 13,000 years of data to predict the effects of climate change on three tree species in Europe. *Global Ecology and Biogeography*, 22, 302-317.
- MAJUMDER, M. S., COHN, E. L., SANTILLANA, M. & BROWNSTEIN, J. S. 2018. Estimation of Pneumonic Plague Transmission in Madagascar, August–November 2017. *PLoS currents*, 10.
- MALEK, M. A., BITAM, I., LEVASSEUR, A., TERRAS, J., GAUDART, J., AZZA, S., FLAUDROPS, C., ROBERT, C., RAOULT, D. & DRANCOURT, M. 2017. *Yersinia pestis* halotolerance illuminates plague reservoirs. *Scientific reports*, 7, 1-10.
- MARKMAN, D. W., ANTOLIN, M. F., BOWEN, R. A., WHEAT, W. H., WOODS, M., GONZALEZ-JUARRERO, M. & JACKSON, M. 2018. *Yersinia pestis* survival and replication in potential ameba reservoir. *Emerging infectious diseases*, 24, 294.
- MARTÍNEZ-MEYER, E., TOWNSEND PETERSON, A. & HARGROVE, W. W. 2004. Ecological niches as stable distributional constraints on mammal species, with implications for Pleistocene extinctions and climate change projections for biodiversity. *Global Ecology and Biogeography*, 13, 305-314.
- MCCAULEY, D. J., SALKELD, D. J., YOUNG, H. S., MAKUNDI, R., DIRZO, R., ECKERLIN, R. P., LAMBIN, E. F., GAFFIKIN, L., BARRY, M. & HELGEN, K. M. 2015. Effects of land use on plague (*Yersinia pestis*) activity in rodents in Tanzania. *The American journal of tropical medicine and hygiene*, 92, 776.
- MCGILL, B. J. 2019. The what, how and why of doing macroecology. *Global Ecology and Biogeography*, 28, 6-17.
- MENDES, P., VELAZCO, S. J. E., DE ANDRADE, A. F. A. & JÚNIOR, P. D. M. 2020. Dealing with overprediction in species distribution models: How adding distance constraints can improve model accuracy. *Ecological Modelling*, 431, 109180.
- MESSINA, J. P., BRADY, O. J., PIGOTT, D. M., GOLDING, N., KRAEMER, M. U., SCOTT, T. W., WINT, G. W., SMITH, D. L. & HAY, S. I. 2015. The many projected futures of dengue. *Nature Reviews Microbiology*, 13, 230-239.
- METCALF, C. J. E., WALTER, K. S., WESOLOWSKI, A., BUCKEE, C. O., SHEVLIAKOVA, E., TATEM, A. J., BOOS, W. R., WEINBERGER, D. M. & PITZER, V. E. 2017. Identifying climate drivers of infectious disease dynamics: recent advances and challenges ahead. *Proceedings of the Royal Society B: Biological Sciences*, 284, 20170901.
- MIAO, L., ZHU, F., HE, B., FERRAT, M., LIU, Q., CAO, X. & CUI, X. 2013. Synthesis of China's land use in the past 300 years. *Global and Planetary Change*, 100, 224-233.
- MILLER, S. D. & CULLY JR, J. F. 2001. Conservation of black-tailed prairie dogs (*Cynomys ludovicianus*). *Journal of Mammalogy*, 82, 889-893.
- MOLLARET, HH 1963. Experimental preservation of plague in soil. *Bulletin de la Societe de Pathologie Exotique et de ses Filiales*, 56, 1168-1182.
- MORELLI, G., SONG, Y., MAZZONI, C. J., EPPINGER, M., ROUMAGNAC, P., WAGNER, D. M., FELDKAMP, M., KUSECEK, B., VOGLER, A. J. & LI, Y.

2010. *Yersinia pestis* genome sequencing identifies patterns of global phylogenetic diversity. *Nature genetics*, 42, 1140-1143.
- MORELLI, T. L., SMITH, A. B., MANCINI, A. N., BALKO, E. A., BORGERSON, C., DOLCH, R., FARRIS, Z., FEDERMAN, S., GOLDEN, C. D. & HOLMES, S. M. 2020. The fate of Madagascar's rainforest habitat. *Nature Climate Change*, 10, 89-96.
- MUSCARELLA, R., GALANTE, P. J., SOLEY-GUARDIA, M., BORIA, R. A., KASS, J. M., URIARTE, M. & ANDERSON, R. P. 2014. ENM eval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for Maxent ecological niche models. *Methods in Ecology and Evolution*, 5, 1198-1205.
- NAVARRO LLORENS, J. M., TORMO, A. & MARTÍNEZ-GARCÍA, E. 2010. Stationary phase in gram-negative bacteria. *FEMS microbiology reviews*, 34, 476-495.
- NEERINCKX, S. B., PETERSON, A. T., GULINCK, H., DECKERS, J. & LEIRS, H. 2008. Geographic distribution and ecological niche of plague in sub-Saharan Africa. *International Journal of Health Geographics*, 7, 54.
- NING, L. & QIAN, Y. 2009. Interdecadal change in extreme precipitation over South China and its mechanism. *Advances in Atmospheric Sciences*, 26, 109-118.
- NOGUÉS-BRAVO, D. 2009. Predicting the past distribution of species climatic niches. *Global Ecology and Biogeography*, 18, 521-531.
- NUNES, L. A. & PEARSON, R. G. 2017. A null biogeographical test for assessing ecological niche evolution. *Journal of Biogeography*, 44, 1331-1343.
- O'NEILL, R. V. 1977. Transmutations across hierarchical levels. Oak Ridge National Lab.
- OLSON, R. M. & ANDERSON, D. M. 2019. Shift from primary pneumonic to secondary septicemic plague by decreasing the volume of intranasal challenge with *Yersinia pestis* in the murine model. *Plos one*, 14, e0217440.
- OSBORNE, O. G., FELL, H. G., ATKINS, H., VAN TOL, J., PHILLIPS, D., HERRERA-ALSINA, L., MYNARD, P., BOCEDI, G., GUBRY-RANGIN, C. & LANCASTER, L. T. 2021. Fauxcurrence: simulating multi-species occurrences for null models in species distribution modelling and biogeography. *bioRxiv*.
- PAINE, R. T. 1980. Food webs: linkage, interaction strength and community infrastructure. *Journal of animal ecology*, 49, 667-685.
- PARMENTER, R. R., YADAV, E. P., PARMENTER, C. A., ETTESTAD, P. & GAGE, K. L. 1999. Incidence of plague associated with increased winter-spring precipitation in New Mexico. *The American journal of tropical medicine and hygiene*, 61, 814-821.
- PAULI, J. N., BUSKIRK, S. W., WILLIAMS, E. S. & EDWARDS, W. H. 2006. A plague epizootic in the black-tailed prairie dog (*Cynomys ludovicianus*). *Journal of Wildlife Diseases*, 42, 74-80.
- PAULING, C. D., FINKE, D. L. & ANDERSON, D. M. 2021. Interrelationship of soil moisture and temperature to sylvatic plague cycle among prairie dogs in the Western United States. *Integrative Zoology*, 16, 852-867.
- PEARMAN, P. B., GUISAN, A., BROENNIMANN, O. & RANDIN, C. F. 2008. Niche dynamics in space and time. *Trends in Ecology & Evolution*, 23, 149-158.
- PECHOUS, R. D., SIVARAMAN, V., STASULLI, N. M. & GOLDMAN, W. E. 2016. Pneumonic plague: the darker side of *Yersinia pestis*. *Trends in microbiology*, 24, 190-197.
- PERGAMS, O. R. & LAWLER, J. J. 2009. Recent and widespread rapid morphological change in rodents. *PloS one*, 4, e6452.
- PETERSON, A. T. 2011. Ecological niche conservatism: A time-structured review of evidence. *Journal of Biogeography*, 38, 817-827.

- PETERSON, A. T. & SOBERÓN, J. 2012. Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservação*, 10, 102-107.
- PHILLIPS, S. J., ANDERSON, R. P. & SCHAPIRE, R. E. 2006. Maximum entropy modeling of species geographic distributions. *Ecological modelling*, 190, 231-259.
- PIZZIMENTI, J. J. & HOFFMANN, R. S. 1973. *Cynomys gunnisoni*. *Mammalian Species*, 1-4.
- POSTMA, J., VAN VEEN, J. & WALTER, S. 1989. Influence of different initial soil moisture contents on the distribution and population dynamics of introduced *Rhizobium leguminosarum* biovar *trifolii*. *Soil Biology and Biochemistry*, 21, 437-442.
- QIAN, Q., ZHAO, J., FANG, L., ZHOU, H., ZHANG, W., WEI, L., YANG, H., YIN, W., CAO, W. & LI, Q. 2014. Mapping risk of plague in Qinghai-Tibetan plateau, China. *BMC infectious diseases*, 14, 382.
- R CORE TEAM 2013. R: A language and environment for statistical computing.
- RAMALINGASWAMI, V. 1995. Plague in India. *Nature medicine*, 1, 1237-1239.
- RANDREMANANA, R., ANDRIANAIVOARIMANANA, V., NIKOLAY, B., RAMASINDRAZANA, B., PAIREAU, J., TEN BOSCH, Q. A., RAKOTONDRA MANGA, J. M., RAHAJANDRAIBE, S., RAHELINIRINA, S. & RAKOTOMANANA, F. 2019. Epidemiological characteristics of an urban plague epidemic in Madagascar, August–November, 2017: an outbreak report. *The Lancet Infectious Diseases*, 19, 537-545.
- RASCOVAN, N., SJÖGREN, K.-G., KRISTIANSEN, K., NIELSEN, R., WILLERSLEV, E., DESNUES, C. & RASMUSSEN, S. 2019. Emergence and spread of basal lineages of *Yersinia pestis* during the Neolithic decline. *Cell*, 176, 295-305. e10.
- REDDING, D. W., MOSES, L. M., CUNNINGHAM, A. A., WOOD, J. & JONES, K. E. 2016. Environmental-mechanistic modelling of the impact of global change on human zoonotic disease emergence: a case study of Lassa fever. *Methods in Ecology and Evolution*, 7, 646-655.
- REIJNIERS, J., BEGON, M., AGEYEV, V. S. & LEIRS, H. 2014. Plague epizootic cycles in Central Asia. *Biology letters*, 10, 20140302.
- REUTER, S., CONNOR, T. R., BARQUIST, L., WALKER, D., FELTWELL, T., HARRIS, S. R., FOOKES, M., HALL, M. E., PETTY, N. K. & FUCHS, T. M. 2014. Parallel independent evolution of pathogenicity within the genus *Yersinia*. *Proceedings of the National Academy of Sciences*, 111, 6768-6773.
- RIVIÈRE-CINNAMOND, A., SANTANDREU, A., LUJÁN, A., MERTENS, F., ESPINOZA, J. O., CARPIO, Y., BRAVO, J. & GABASTOU, J.-M. 2018. Identifying the social and environmental determinants of plague endemicity in Peru: insights from a case study in Ascope, La Libertad. *BMC public health*, 18, 1-11.
- ROLLINS, S. E., ROLLINS, S. M. & RYAN, E. T. 2003. *Yersinia pestis* and the plague. *Pathology Patterns Reviews*, 119, S78-S85.
- ROSE, L. J., DONLAN, R., BANERJEE, S. N. & ARDUINO, M. J. 2003. Survival of *Yersinia pestis* on environmental surfaces. *Applied and Environmental Microbiology*, 69, 2166-2171.
- ROSQVIST, R., SKURNIK, M. & WOLF-WATZ, H. 1988. Increased virulence of *Yersinia pseudotuberculosis* by two independent mutations. *Nature*, 334, 522-525.
- ROUSSOS, D. 2002. Plague. *Primary care update for ob/gyns*, 9, 125-128.
- SALKELD & STAPP, P. 2006. Seroprevalence rates and transmission of plague (*Yersinia pestis*) in mammalian carnivores. *Vector-Borne & Zoonotic Diseases*, 6, 231-239.

- SALKELD, D. J., SALATHÉ, M., STAPP, P. & JONES, J. H. 2010. Plague outbreaks in prairie dog populations explained by percolation thresholds of alternate host abundance. *Proceedings of the National Academy of Sciences*, 107, 14247-14250.
- SAMIA, N. I., KAUSRUD, K. L., HEESTERBEEK, H., AGEYEV, V., BEGON, M., CHAN, K.-S. & STENSETH, N. C. 2011. Dynamics of the plague–wildlife–human system in Central Asia are controlled by two epidemiological thresholds. *Proceedings of the National Academy of Sciences*, 108, 14527-14532.
- SANTOS-MONTAÑEZ, J., BENAVIDES-MONTAÑO, J. A., HINZ, A. K. & VADYVALOO, V. 2015. *Yersinia pseudotuberculosis* IP32953 survives and replicates in trophozoites and persists in cysts of *Acanthamoeba castellanii*. *FEMS Microbiology Letters*, 362, fnv091.
- SAVAGE, L. T., REICH, R. M., HARTLEY, L. M., STAPP, P. & ANTOLIN, M. F. 2011. Climate, soils, and connectivity predict plague epizootics in black-tailed prairie dogs (*Cynomys ludovicianus*). *Ecological Applications*, 21, 2933-2943.
- SCHMID, B., BÜNTGEN, U., EASTERDAY, W., GINZLER, C., WALLØE, L., BRAMANTI, B. & STENSETH, N. 2015. Climate-driven introduction of the Black Death and successive plague reintroductions into Europe. *Proceedings of the National Academy of Sciences*, 112, 3020-3025.
- SCHMID, B., JESSE, M., WILSCHUT, L., VILJUGREIN, H. & HEESTERBEEK, J. 2012. Local persistence and extinction of plague in a metapopulation of great gerbil burrows, Kazakhstan. *Epidemics*, 4, 211-218.
- SCHMITZ, O. J., HAMBÄCK, P. A. & BECKERMAN, A. P. 2000. Trophic cascades in terrestrial systems: a review of the effects of carnivore removals on plants. *The American Naturalist*, 155, 141-153.
- SCHOTTHOEFER, A. M., BEARDEN, S. W., VETTER, S. M., HOLMES, J., MONTENIERI, J. A., GRAHAM, C. B., WOODS, M. E., EISEN, R. J. & GAGE, K. L. 2011. Effects of temperature on early-phase transmission of *Yersinia pestis* by the flea, *Xenopsylla cheopis*. *Journal of medical entomology*, 48, 411-417.
- SILLERO, N. & BARBOSA, A. M. 2021. Common mistakes in ecological niche models. *International Journal of Geographical Information Science*, 35, 213-226.
- SMILEY, S. T., SZABA, F. M., KUMMER, L. W., DUSO, D. K. & LIN, J.-S. 2019. *Yersinia pestis* Pla protein thwarts T cell defense against plague. *Infection and immunity*, 87, e00126-19.
- SMITH, J. E., LONG, D. J., RUSSELL, I. D., NEWCOMB, K. L. & MUÑOZ, V. D. 2016. *Otospermophilus beecheyi* (Rodentia: Sciuridae). *Mammalian Species*, 48, 91-108.
- SOBERÓN, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology letters*, 10, 1115-1123.
- SPYROU, M. A., TUKHBATOVA, R. I., WANG, C.-C., VALTUEÑA, A. A., LANKAPALLI, A. K., KONDRASHIN, V. V., TSYBIN, V. A., KHOKHLOV, A., KÜHNERT, D. & HERBIG, A. 2018. Analysis of 3800-year-old *Yersinia pestis* genomes suggests Bronze Age origin for bubonic plague. *Nature communications*, 9, 1-10.
- STANFIELD, J. T., JACKSON, G. J. & AULISIO, C. 1985. *Yersinia enterocolitica*: survival of a pathogenic strain on milk containers. *Journal of food protection*, 48, 947-948.
- STAPP, P. 2007. Trophic cascades and disease ecology. *EcoHealth*, 4, 121-124.
- STENSETH, N. C., ATSHABAR, B. B., BEGON, M., BELMAIN, S. R., BERTHERAT, E., CARNIEL, E., GAGE, K. L., LEIRS, H. & RAHALISON, L. 2008. Plague: past, present, and future. *PLoS Med*, 5, e3.

- STENSETH, N. C., SAMIA, N. I., VILJUGREIN, H., KAUSRUD, K. L., BEGON, M., DAVIS, S., LEIRS, H., DUBYANSKIY, V., ESPER, J. & AGEYEV, V. S. 2006. Plague dynamics are driven by climate variation. *Proceedings of the National Academy of Sciences*, 103, 13110-13115.
- STRUBBE, D., BROENNIMANN, O., CHIRON, F. & MATTHYSEN, E. 2013. Niche conservatism in non-native birds in Europe: niche unfilling rather than niche expansion. *Global Ecology and Biogeography*, 22, 962-970.
- SUN, Y.-C., JARRETT, C. O., BOSIO, C. F. & HINNEBUSCH, B. J. 2014. Retracing the evolutionary path that led to flea-borne transmission of *Yersinia pestis*. *Cell host & microbe*, 15, 578-586.
- SUN, Y.-C., KOUMOUTSI, A. & DARBY, C. 2009. The response regulator PhoP negatively regulates *Yersinia pseudotuberculosis* and *Yersinia pestis* biofilms. *FEMS microbiology letters*, 290, 85-90.
- SUN, Z., XU, L., SCHMID, B. V., DEAN, K. R., ZHANG, Z., XIE, Y., FANG, X., WANG, S., LIU, Q. & LYU, B. 2019. Human plague system associated with rodent diversity and other environmental factors. *Royal Society open science*, 6, 190216.
- SUTTON, D. A. 1992. *Tamias amoenus*. *Mammalian species*, 1-8.
- SVENNING, J. C., NORMAND, S. & KAGEYAMA, M. 2008. Glacial refugia of temperate trees in Europe: insights from species distribution modelling. *Journal of Ecology*, 96, 1117-1127.
- SWETS, J. A. 1988. Measuring the accuracy of diagnostic systems. *Science*, 240, 1285-1293.
- TREVORS, J. 1996. Sterilization and inhibition of microbial activity in soil. *Journal of Microbiological Methods*, 26, 53-59.
- VALLÈS, X., STENSETH, N. C., DEMEURE, C., HORBY, P., MEAD, P. S., CABANILLAS, O., RATSITORAHINA, M., RAJERISON, M., ANDRIANAIVOARIMANANA, V. & RAMASINDRAZANA, B. 2020. Human plague: An old scourge that needs new answers. *PLoS neglected tropical diseases*, 14, e0008251.
- VAN VALEN, L. 1965. Morphological variation and width of ecological niche. *The American Naturalist*, 99, 377-390.
- VAN VEEN, J. A., VAN OVERBEEK, L. S. & VAN ELSAS, J. D. 1997. Fate and activity of microorganisms introduced into soil. *Microbiology and Molecular Biology Reviews*, 61, 121-135.
- VERTS, B. & CARRAWAY, L. N. 2001. *Tamias minimus*. *Mammalian Species*, 2001, 1-10.
- VON EIFF, C. 2008. *Staphylococcus aureus* small colony variants: a challenge to microbiologists and clinicians. *International journal of antimicrobial agents*, 31, 507-510.
- VOS, M., WOLF, A. B., JENNINGS, S. J. & KOWALCHUK, G. A. 2013. Micro-scale determinants of bacterial diversity in soil. *FEMS microbiology reviews*, 37, 936-954.
- WALSH, M. & HASEEB, M. 2015. Modeling the ecologic niche of plague in sylvan and domestic animal hosts to delineate sources of human exposure in the western United States. *PeerJ*, 3, e1493.
- WALTARI, E. & GURALNICK, R. P. 2009. Ecological niche modelling of montane mammals in the Great Basin, North America: examining past and present connectivity of species across basins and ranges. *Journal of Biogeography*, 36, 148-161.
- WANG, X., SINGH, A. K. & SUN, W. 2020. Protection and Safety Evaluation of Live Constructions Derived from the Pgm- and pPCP1- *Yersinia pestis* Strain. *Vaccines*, 8, 95.
- WARREN, D. L., GLOR, R. E. & TURELLI, M. 2008. Environmental niche equivalency versus conservatism: quantitative approaches to niche

- evolution. *Evolution: International Journal of Organic Evolution*, 62, 2868-2883.
- WARREN, D. L., GLOR, R. E. & TURELLI, M. 2010. ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, 33, 607-611.
- WARREN, D. L., MATZKE, N. J., CARDILLO, M., BAUMGARTNER, J. B., BEAUMONT, L. J., TURELLI, M., GLOR, R. E., HURON, N. A., SIMÕES, M. & IGLESIAS, T. L. 2021. ENMTools 1.0: An R package for comparative ecological biogeography. *Ecography*, 44, 504-511.
- WARREN, D. L. & SEIFERT, S. N. 2011. Ecological niche modeling in Maxent: the importance of model complexity and the performance of model selection criteria. *Ecological applications*, 21, 335-342.
- WARSAWSKI, L., FRIELER, K., HUBER, V., PIONTEK, F., SERDECZNY, O. & SCHEWE, J. 2014. The inter-sectoral impact model intercomparison project (ISI-MIP): project framework. *Proceedings of the National Academy of Sciences*, 111, 3228-3232.
- WEBB, C. T., BROOKS, C. P., GAGE, K. L. & ANTOLIN, M. F. 2006. Classic flea-borne transmission does not drive plague epizootics in prairie dogs. *Proceedings of the National Academy of Sciences*, 103, 6236-6241.
- WEBER, M. M., STEVENS, R. D., DINIZ-FILHO, J. A. F. & GRELLER, C. E. V. 2017. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. *Ecography*, 40, 817-828.
- WENGER, S. J. & OLDEN, J. D. 2012. Assessing transferability of ecological models: an underappreciated aspect of statistical validation. *Methods in Ecology and Evolution*, 3, 260-267.
- WIECHMANN, A. 2016. *Quorum sensing and the regulation of multicellular behaviour in Yersinia pseudotuberculosis*. University of Nottingham.
- WILSCHUT, L. I., ADDINK, E. A., HEESTERBEEK, H., HEIER, L., LAUDISOIT, A., BEGON, M., DAVIS, S., DUBYANSKIY, V. M., BURDELOV, L. A. & DE JONG, S. M. 2013. Potential corridors and barriers for plague spread in central Asia. *International journal of health geographics*, 12, 1-15.
- WILSCHUT, L. I., LAUDISOIT, A., HUGHES, N. K., ADDINK, E. A., DE JONG, S. M., HEESTERBEEK, H. A., REIJNIERS, J., EAGLE, S., DUBYANSKIY, V. M. & BEGON, M. 2015. Spatial distribution patterns of plague hosts: point pattern analysis of the burrows of great gerbils in Kazakhstan. *Journal of Biogeography*, 42, 1281-1292.
- WORLD HEALTH ORGANIZATION. Blueprint for R&D preparedness and response to public health emergencies due to highly infectious pathogens. Workshop on Prioritization of Pathogens, 2015.
- WORLD HEALTH ORGANIZATION. 2017. *Plague Fact Sheet* [Online]. <https://www.who.int/en/news-room/fact-sheets/detail/plague>. [Accessed 2022].
- WREN, B. W. 2003. The yersiniae—a model genus to study the rapid evolution of bacterial pathogens. *Nature Reviews Microbiology*, 1, 55-64.
- XU, L., LIU, Q., STIGE, L. C., ARI, T. B., FANG, X., CHAN, K.-S., WANG, S., STENSETH, N. C. & ZHANG, Z. 2011. Nonlinear effect of climate on plague during the third pandemic in China. *Proceedings of the National Academy of Sciences*, 108, 10214-10219.
- XU, L., SCHMID, B. V., LIU, J., SI, X., STENSETH, N. C. & ZHANG, Z. 2015. The trophic responses of two different rodent–vector–plague systems to climate change. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20141846.
- XU, L., STIGE, L. C., KAUSRUD, K. L., BEN ARI, T., WANG, S., FANG, X., SCHMID, B. V., LIU, Q., STENSETH, N. C. & ZHANG, Z. 2014. Wet climate

- and transportation routes accelerate spread of human plague. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20133159.
- YERSIN, A. 1894. La peste bubonique à Hong-Kong. *Ann. Inst. Pastur.*, 2, 428-430.
- YUAN, Z., YANG, Z., YAN, D. & YIN, J. 2017. Historical changes and future projection of extreme precipitation in China. *Theoretical and Applied Climatology*, 127, 393-407.
- ZHANG, Z., XU, L., GUO, C., WANG, Y. & GUO, Y. 2010. Effect of ENSO-driven precipitation on population irruptions of the Yangtze vole *Microtus fortis calamorum* in the Dongting Lake region of China. *Integrative Zoology*, 5, 176-184.