UNIVERSITY OF LIVERPOOL

“Interactions between pathogenic and non-pathogenic Rickettsiales and the tick host”

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

ALAA AL-KHAFAJI

October 2018

Table of contents

Abstract

Background: Ticks are obligate haematophagous arthropods constituting one of the most important groups of vectors for parasitic, bacterial and viral diseases of humans and other animals. Ticks have complex microbiomes, but only a small proportion of the bacterial symbionts recorded from ticks are vertically transmitted. In this study, we focus on Rickettsiales bacteria found in ticks, some of which are vertebrate pathogens, and how they interact with their tick hosts by analysing different aspects of their ecology and molecular evolution, as well as the cellular responses they induce in host cells.

Methods: Three species of ticks, *Ixodes ricinus* from Western Europe, *Amblyomma variegatum* from Cameroon, and *Ixodes scapularis* from the Eastern USA, were examined in this study alongside several species of Rickettsiales found in each. Bacterial densities in each tick species were quantified by qPCR. For the symbiont of *I. ricinus*, *Candidatus* (*Ca*.) Midichloria mitochondrii, a multi-locus sequence typing (MLST) scheme for both the symbiont and host mitochondria was used to examine potential co-evolution across the UK and mainland Europe. For the vertebrate pathogen *Anaplasma phagocytophilum*, *in vitro* culture in tick cells coupled with next-generation genome sequencing and both shotgun and targeted proteomics was used to investigate host-pathogen interactions. Finally, geLC-MS was used to quantify protein expression from the symbiont of *I. scapularis, Rickettsia buchneri,* in tick ovaries and salivary glands.

Results: A quantitative survey of *Ca*. M. mitochondrii in *I. ricinus* was conducted in the UK for the first time, revealing a similar pattern of infection across different tick life stages as recorded for mainland Europe, except that the prevalence rate in nymphs and adult male ticks in Wales was unexpectedly high. The MLST analysis of tick and symbiont revealed that *Ca*. M. mitochondrii exhibits very low levels of sequence diversity, although a consistent signal of host-symbiont coevolution was apparent in Scotland. Moreover, the tick MLST scheme revealed that Scottish specimens form a clade that is partially separated from other British ticks, with almost no contribution of continental sequence types in this north-westerly border of the tick's natural range. The prevalence of two Rickettsiales bacteria in the tropical bont tick, *A. variegatum*, in the Adamawa Region of Cameroon was very low (3.1%) for *Ehrlichia ruminantium* (the causative agent of heartwater disease in ruminants), and very high (95.3%) for the vertically transmitted *Rickettsia africae* (the causative agent of African tick-bite fever in humans). A genome for a British isolate of *A. phagocytophilum* (strain Old Sourhope from sheep) was obtained for the first time. This was used to support a quantitative label-free mass spectrometry analysis to examine changes in the proteome of two tick cell lines (ISE6 and IRE/CTVM19) in response to infection at different time-points. Interestingly, we observed an apparent shared response involving cholesterol hijacking by the pathogen in both cell lines; however, the IRE/CTVM19 responses reflected the initial stages of infection, while most of the ISE6 responses were derived from the establishment phase of infection. Targeted proteomics was used to quantify expression of an *A. phagocytophilum* variable outer membrane protein family (Msp2/p44) *in vitro*. Although few significant differences in abundance between the two time-points in culture were apparent, the dominant Msp2/p44 we observed shared orthologues primarily with *Anaplasma* strains infecting ruminants, rather than other mammalian hosts. Finally, the presence of the non-pathogenic *R. buchneri* was demonstrated in salivary glands of *I. scapularis* for the first time, suggesting that it might generate an immune response of diagnostic significance in humans. Expression of biotin synthase from a *R. buchneri* plasmid was detected in tick ovaries, indicating a possible nutritional basis to the symbiosis.

Conclusions: Many tick species are potentially infected with multiple species of Rickettsiales. To date, the main research focus has been on those species that cause disease in vertebrates. However, non-pathogenic Rickettsiales are likely to interact with their pathogenic relatives and possibly other tick-borne pathogens in the tick host. Here, we have contributed to our understanding of the roles of three vertically-transmitted symbionts in three of the most important tick vectors worldwide: *Ca*. M. mitochondrii in *I. ricinus, R. africae* in *A. variegatum*, and *R. buchneri* in *I. scapularis*. The use of tick cells to examine Rickettsiales:tick interactions in vitro, as shown here with *A. phagocytophilum*, could be expanded in future to include pathogen-symbiont coinfections.

List of Figures

List of Tables

List of Abbreviations

ANOVA	Analysis of Variance
ATBF	African tick-bite fever
ATP	*ATPase 6 (ATP* synthase F0 subunit 6)
AT	Allele type
BLASTn	Nucleotide basic local alignment search tool
bp	Base pair
^{o}C	Celsius degree
Ca.	*Candidatus* (refers to provisional taxonomic status of unculturable bacteria)
CGR	Centre for Genomic Research, University of Liverpool
COI	Cytochrome C oxidase subunit 1 gene
COII	Cytochrome C oxidase subunit 2 gene
COIII	Cytochrome oxidase subunit 3 gene
CYTB	Cytochrome B
DDT	Dithiothreitol
dNTPs	Deoxynucleotide triphosphates
Dpi	Days post infection
EM	Electron microscope
E-value	Expected value (in BLAST)
F-actin	Filamentous actin
GC	Guanine-cytosine
gltA	Citrate synthase gene
groEL	Heat Shock protein
gyrb	DNA gyrase subunit B
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MLST	Multilocus sequence typing
ML	Maximum likelihood
µl	Microliter
µM	Micromolar
mM	Millimolar per litter
NCBI	National Center for Biotechnology Information
NH_4OH	Ammonium hydroxide
NJ	Neighbour joining
ng	Nanogram
p or p-value	Calculated value probability
qPCR	Quantitative polymerase chain reaction
16S rRNA	16 Svedberg ribosomal ribonucleic acid
SDC	Sodium deoxycholate
SFG	Spotted Fever Group
ST	Sequence type
SLV	Single locus variants
UK	The United Kingdom
USA	The United States of America
WGS	Whole genome sequencing
ffs	Fast-feeding stage
sfs	Slow-feeding stage
ifs	Intermediate-feeding stage

Acknowledgments

This doctoral project would not have been possible without the support and guidance of a multitude of people. First, I would like to express my deep and senior gratitude to Dr Ben Makepeace who guided me through my PhD project; thanks for his support and encouragement on a daily basis from the start of the project till date, and thanks to have such knowledgeable and enthusiastic support in difficult times.

Also, I would like to thank my two co-supervisors, Dr John McGarry and Dr Alistair Darby. Thanks to Dr Lesley Bell-Sakyi for all her knowledge and enthusiasm about ticks and for English copy-editing of this thesis.

I would like to acknowledge the financial support of my government and AL-Qadisyia University-Iraq for an invaluable opportunity to do my PhD at the University of Liverpool.

Heartfelt thanks should also be expressed to Simon Clegg (now at the School of Life Sciences, University of Lincoln) who supported me in MLST data analysis, Dr. Samriti Midha for her help in analysing whole genome data and Dr. Stuart Armstrong for his help and input in proteomics analysis. I would like to thank our collaborators and everyone who has provided me with samples for analysis.

I express my gratitude also to Dr Catherine Hartley, our senior technician in IC2 for her support in the laboratory work. Everyone in the lab in IC2 and office deserve a huge thank you for their help and encouraging words. Huge thanks to my colleagues Lisa, who has been so helpful, Tess, and John.

A massive thank you to my wonderful family; I would like to present my sincere thankfulness to my dear father, my brothers and sister who encouraged and supported me through my entire studies in every possible way.

I owe thanks to a very special person, my husband, Ali for his continued and unfailing love, support and understanding during my PhD degree that made the completion of thesis possible. You were always around at times I thought that it was impossible to continue, you helped me to keep things in perspective. I appreciate my twin's boys (Hasan and Hussein), my lovely Rawan and my little girl Rimas for abiding my

ignorance and for their patience, love and unconditional support they showed during all of my PhD project.

I would like to dedicate this work to my late mother Mrs. Maha Al-khafaji whose dreams for me have resulted in this achievement and without her loving upbringing and nurturing, I would not be where I am today and what I am today. My dreams of excelling in education would have remained mere dreams.

I thank my mother for all her untold number of sacrifices, everything that she taught me about perfection, passion, unflinching courage and conviction has always inspired me to do more, so that today I can stand proudly with my head held high.

I thank my mother with all my heart and I know she is up there, listening, watching over me and sending me her blessings constantly and is my guardian angel.

This one is for you Mom!

Author's Declarations

The following describes the contributions of other individuals to the work presented in this thesis alongside my own.

Chapter 2:

Ixodes ricinus samples for *Midichloria* and pathogen screening were supplied by Dr Jolyon Medlock and Kayleigh Hansford from Public Health England, Porton Down, UK. Additional specimens from Powis Castle Estate, Powys, were collected by my supervisors (Dr Ben Makepeace and Dr John McGarry), and with permission of the estate gamekeepers Tom Till and Peter Shaw. Samples of *Amblyomma variegatum* from Cameroon were collected by Germanus Bah and colleagues at the Institut de Recherche Agricole pour le Développement, Regional Centre of Wakwa, Ngaoundéré.

For this chapter, I performed all of the tick identifications, DNA extractions and qPCR assays with occasional assistance from senior technician.

Chapter 3:

In addition to the *I. ricinus* samples mentioned above, I received additional whole specimens from Dr Davide Sassera (University of Pavia, Italy), Prof. Lucy Gilbert (James Hutton Institute, Aberdeen) and Prof. Olivier Plantard (Oniris, Nantes, France). The French specimens were collected from roe deer captured in the Chizé forest with permission from Gilles Capron (Office nationale de la Chasse et de la Faune Sauvage) and the Office National des Forêts from the *Réserve biologique domaniale intégrale de la Sylve d'Argenson*; all other specimens were collected from the environment. DNA extracts from German ticks were provided by Prof. Edward Feil at the University of Bath.

The tick MLST was designed by Drs Frederik Seeling and Ruth Dinnis with Prof. Edward Feil at the University of Bath. The Midichloria MLST was designed by Dr Davide Sassera.

For this chapter, I performed tick identifications, DNA extractions, conventional PCRs (including improvements to the performance of the MLST schemes by optimising PCR conditions), and initial sequence alignments. Phylogenetic analyses were performed with assistance from Dr Simon Clegg in Liverpool. In addition to the new sequences generated in Liverpool, additional unpublished tick MLST sequences for Germany, Switzerland and Portugal were provided by Drs Gabriele Margos, Frederik Seelig and Ruth Dinnis (University of Bath).

Chapter 4:

Tick cell lines and training in tick cell culture were provided by Dr Lesley Bell-Sakyi when she was based at the Pirbright Institute. Initial *Anaplasma*-infected cultures were also provided by Dr Bell-Sakyi from *Anaplasma* strains supplied by Dr Zerai Woldehiwet (University of Liverpool).

Anaplasma strain OS genome sequencing was undertaken at the Centre for Genomic Research. My co-supervisor, Dr Alistair Darby, performed the initial assembly, quality checks and automatic annotation pipeline.

Proteomics of tick cell cultures was performed by Dr Stuart Armstrong in Liverpool from cell pellets provided by me. This included the parallel reaction monitoring design and the Progenesis, Pfam enrichment and GO analyses. The relevant Methods sections for proteomics were also provided by Dr Armstrong.

For this chapter, I performed the tick cell culture, including maintenance of *Anaplasma in vitro*. I also conducted the DNA extractions for genome sequencing, the Cytospin microscopy and qPCR assays. The comparative genomics analysis of *Anaplasma* and identification of Msp2/p44 family members were performed with assistance from Drs Samriti Midha and Xiaofeng Dong.

Chapter 5:

Ixodes scapularis specimens were provided by Drs Clotilde Carlow and Zhiru Li from New England Biolabs, Massachusetts, USA. Salivary glands and ovaries were dissected from the ticks by Stefano Gaiarsa and Prof. Sara Epis at the University of Milan.

Protein extractions and proteomic analysis (including comparison of ovary and salivary gland proteins in Progenesis) were performed by Dr Stuart Armstrong.

For this chapter, I conducted the conventional and qPCR assays, sequence alignments, and geLC-MS work including tryptic digestions.

I declare that the above is a true and accurate description of how the data presented in this thesis were obtained.

Signed: **Date: 25.10.2018**

Alaa Mohammed Abdul Razak Al-Khafaji

Chapter 1 – General Introduction

1. Introduction

Ticks are obligate, haematophagous ectoparasites that obtain nutrition via blood feeding. Although tick bites initiate multiple effects including local damage and risk of toxins that might be excreted within the tick saliva, the main danger is from the pathogens that ticks can transmit, of which there are many types. Indeed ticks are second only to mosquitoes regarding the range and importance of human diseases they can transmit, and like mosquitoes, they have a complex life cycle(Guglielmone et al. 2010; Jore et al. 2014; Mediannikov & Fenollar 2014). In the veterinary sphere, ticks are considerably more important than any other disease vector. All ticks (except soft ticks that have multiple nymph stages) go through four life stages: egg, six-legged larva, eight-legged stages as both nymph and adult; and they may feed on one two, or three hosts (Guglielmone et al. 2010). Ixodid ticks feed once during each developmental life stage but for extended periods of time; several days or weeks (Estrada-Peña 2015). Around 900 different tick species have so far been identified (Guglielmone et al. 2010). Of these, however, only 10% have been documented as being of medical or veterinary importance. Different tick species can transmit a multitude of different pathogens including viruses, bacteria and protozoa to wild and domestic animals, as well as humans (Jongejan & Uillenberg 2004; Hajdušek et al. 2013; Swanson et al. 2006; Noda et al. 1997; Filippova 2017). However, most tick-borne diseases employ similar mechanisms to induce infection and transmission (de la Fuente et al. 2017).

This study will focus on two important tick vectors. *Ixodes ricinus* infests livestock and a wide variety of other species including humans, and acts as a vector

for a large number of pathogens including bacteria such as *Anaplasma phagocytophilum* and *Borrelia burgdorferi* (Lyme disease), protozoa such as *Babesia divergens* (babesiosis), and viruses such as louping ill virus and tick-borne encephalitis virus (Lommano et al. 2012; Rizzoli et al. 2011; Medlock et al. 2013).

The second species that has been investigated in this study is the tropical bont tick *Amblyomma variegatum*, one of the most important livestock ticks that transmits *Ehrlichia ruminantium* (the causative agent of heartwater disease) in Africa and now in the Caribbean. In addition, *A. variegatum* transmits *Rickettsia africae* (the causative agent of African tick bite fever) which is the most widespread spotted fever group (SFG) rickettsia in sub-Saharan Africa (Ndip et al. 2004).

New pathogens and tick-pathogen relationships have emerged in the past 20 or more years, and several tick-borne diseases are considered to be new and emerging. For example, an increase in human outdoor activities and adventure travel which results in exposure to tick habitats has led to an increase in human tick-borne diseases such as Lyme disease (Rizzoli et al. 2011). Over this same timescale, there has been a recent surge in interest in ticks and tick-borne pathogens, to elucidate the factors that may influence the host-pathogen relationship and geographical/spatial distribution of ticks (Estrada- Peña, 2015). Furthermore, there has been an enormous interest in non-pathogenic bacteria including species of *Rickettsia* and obligate intracellular endosymbionts with which ticks share a mutualistic relationship (Clay et al. 2008; Harrus et al. 2011; Ahantarig et al. 2013). Although ticks may harbour both pathogenic and non-pathogenic microflora, little is known about how the diversity of the microflora within ticks may influence the transmission of pathogens. Research into the discovery of tick-pathogen and tick-endosymbiont interactions that influence

vector competence and pathogen transmission is fundamental to understanding the development of tick-borne diseases - new and old - and their control.

In general, ticks have complex microbiomes, but only a small proportion of the bacterial symbionts recorded from ticks have been shown to be vertically transmitted. One of the most widespread tick symbionts is *Candidatus* (*Ca*). Midichloria, which has been detected in all of the major ixodid tick genera of medical and veterinary importance (Sassera et al. 2008). In some species of *Ixodes*, such as the sheep tick *Ixodes ricinus* (infected with *Ca.* Midichloria mitochondrii) the symbiont is fixed in wild adult female ticks, suggesting an obligate mutualism. However, almost no information is available on genetic variation in *Ca.* M. mitochondrii or possible co-cladogenesis with its host across its geographic range. So, the evolutionary importance of *Ca*. Midichloria, and whether these bacteria might impact on transmission of pathogens by their tick hosts, as seen for *Wolbachia* in some insect hosts, remains largely a matter of speculation. Similarly, very little is known about whether interaction between co-infecting pathogens, such as *R. africae* and *E. ruminantium* in *A. variegatum*, might impact the transmission of either or both pathogens. Finally, substantial knowledge gaps exist concerning bacterial proteins involved in interaction with the tick vector, for example, the interaction at the protein level between *A. phagocytophilum* and its ixodid tick vectors.

1.1Tick Biology

1.1.1 Tick classification

Ticks belong to the phylum Arthropoda, Class Arachnida, Subclass Acari, order Parasitiformes, suborder Ixodida (Metastigmata). The suborder Ixodida comprises three families, two of which are of medical importance. These are the Argasidae (soft ticks) and the Ixodidae (hard ticks). Of the ~900 tick species described worldwide, 707 belong to the Ixodidae. Around 200 species of soft tick belong to the Argasidae including the genera *Argas*, *Ornithodoros*, and *Antricola* (Sonenshine 1992; Guglielmone et al. 2010). Nutalliellidae, which is monotypic, contains just one species *Nuttalliella namaqua*; however, little is known about it. The Ixodidae include the genera *Amblyomma, Dermacentor, Haemaphysalis, Hyalomma, Ixodes,* and *Rhipicephalus* (**Figure 1. 1**).

Phylum: Arthropoda
Subfamily: Chelicerata
Class: Arachnida
Subclass: Acari
Order: Parasitiformes
Suborder: Metastigmata (Ixodida)
Ixodidae……. *Amblyomma, Dermacentor, Haemaphysalis, Hyalomma, Ixodes, Rhipicephalus*
Argasidae……. *Argas*, *Ornithodoros*, *Otobius*
Nutalliellidae….. *Nuttalliella*

Figure 1. 1 Tick classification based on (Sonenshine 1992).

Modified according to recent identification (Lindquist et al. 2009). Genera of medical and veterinary importance are listed.

Early classifications of ticks were based on ecological, morphological and biological characteristics (Estrada-Peña, 2015). Further molecular analysis showed that the family Ixodidae could be divided into two groups: the Prostriata containing a single genus, *Ixodes*, comprising two groups (Australian-New Guinea *Ixodes* group and a group containing all other *Ixodes* spp.) and totalling around 243 species; while all other Ixodidae are Metastriata with 459 species (Walker et al., 2003; Guglielmone et al. 2010). (**Figure 1. 2**)

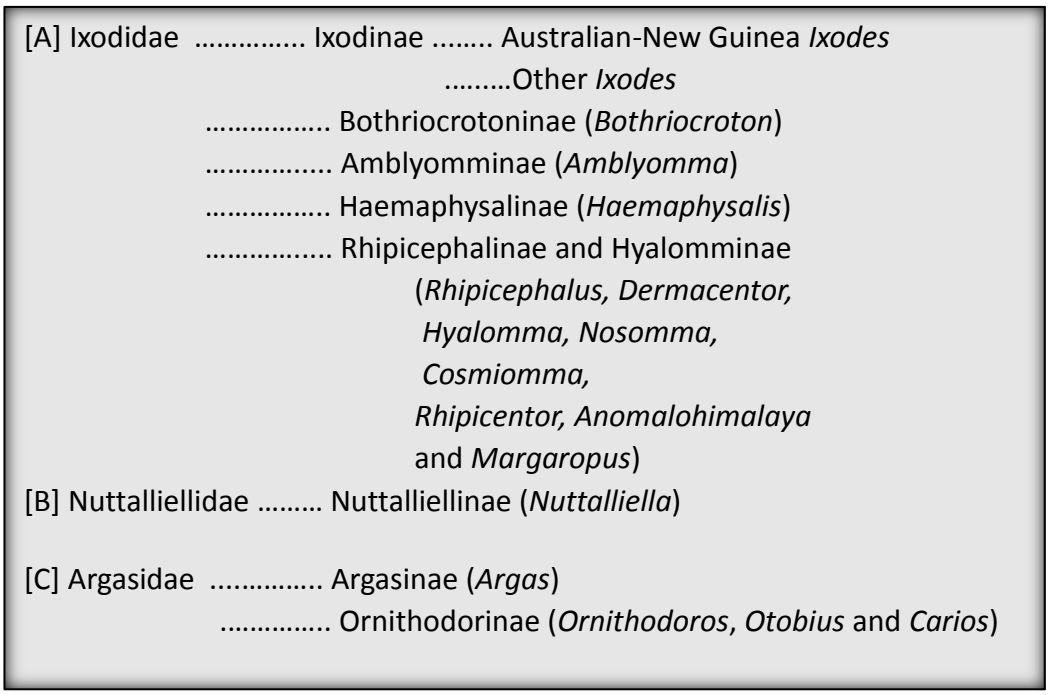


Figure 1. 2 Phylogenetic analysis of suborder Ixodidae adapted from (Barker & Murrell 2004).

1.1.2 Tick morphology and feeding patterns

The predominant morphological characteristics of ticks according to (Hillyard 1996) are the scutum and the visible mouthparts. The scutum is a plate-like shield on the dorsal surface, which covers the entire dorsal part of the body in males, while it covers only the anterior part of the iodosome in females, providing opportunity to expand during engorgement. Tick eyes, if present, are allocated on the sides of the

scutum. However, the larval stage has only three pairs of legs; all other stages have four pairs of legs.

On the other hand, lack of scutum and less distinguishable mouthparts are considered as are the main morphological features of soft ticks. Soft ticks (Parasitiformes order, Metastigmata suborder, Argasidae family) have a wrinkled body. Interestingly, the front portion of the tick's body hides the mouthparts, allowing nymphs and adults to engorge repeatedly. Soft ticks do not have a scutum and the males and females are quite similar in size (Hillyard 1996).

The other important difference between hard and soft ticks is tick feeding behaviour; while hard ticks require only one bloodmeal during each developmental stage, soft ticks feed frequently by taking small amounts of blood each time. Finally, female hard ticks are capable of laying thousands of eggs in a single mass, whereas soft ticks lay several hundred eggs in multiple batches (Gray 2002). Hard ticks may have one, two or three hosts during their development. One-host ticks belonging to the subgenus *Rhipicephalus* (*Boophilus*) have a one-host life-cycle, as all the life stages persist, feed and moult to the next stage while attached to the same host for several weeks (Estrada-Peña et al. 2006). In contrast, some ticks such as *Hyalomma anatolicum excavatum* are two-host ticks, as unfed larvae moult into the nymph developmental stage, which reattaches to the same individual host, then afterwards, detachment occurs and subsequent moulting to the adult stage occurs after engorgement (Ghosh et al. 2017). All *Ixodes* spp. ticks and *A. variegatum* are characterized by three host tick lifestyles (Medlock et al. 2013).

1.2 *Ixodes ricinus*

1.2.1 Classification of the *Ixodes ricinus* complex

Current classification of the *I. ricinus* species complex is based on molecular phylogenetic studies using 16S rDNA sequences of 11 different members of the *I. persulcatus/I. ricinus* species complex and 16 other species of *Ixodes* (Swanson et al. 2006). Bayesian methodology indicated that *I. ricinus* species are closely linked despite a wide geographical dispersal (Xu et al. 2003) (**Figure 1. 3**).

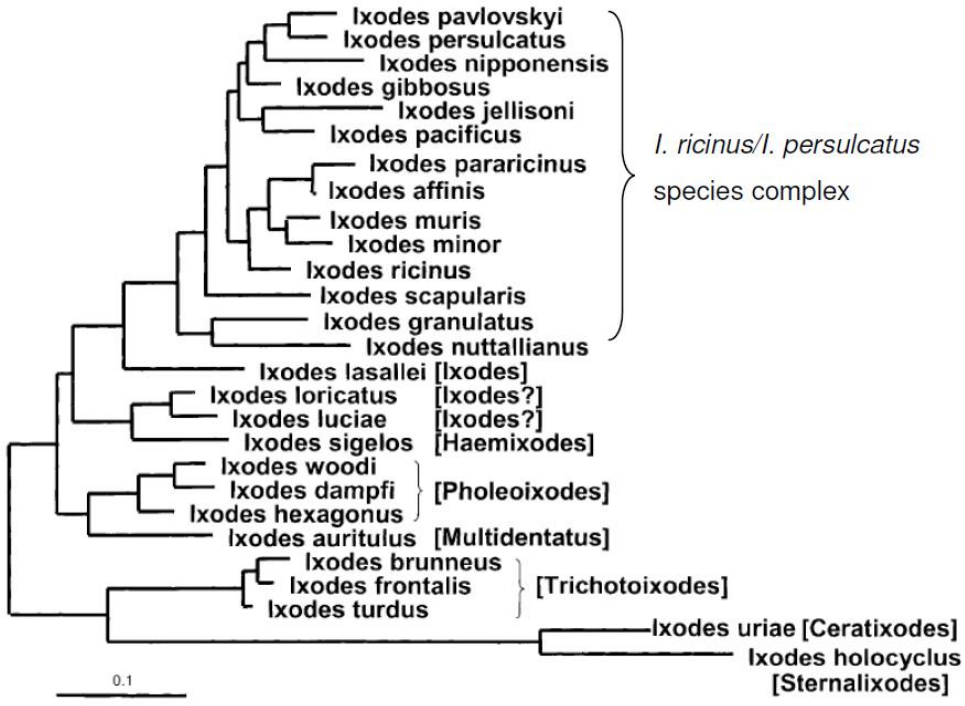


Figure 1. 3 Bayesian Tree of 27 *Ixodes* species; 16S rDNA analysis revealed eleven species associated with *I. persulcatus* and *I. ricinius.*
Square brackets indicate subgenera while ‘‘?’’ points to an uncertainty of the subgenus according to morphological characters (Xu et al. 2003). [Copyright permission to reproduce figure was approved from Allen Press Publishing Services].

1.2.2 The geographical distribution of *I. ricinus*

Distribution records demonstrate establishment of *I. ricinus* across Europe from Ireland to Russia, and they have been identified in deciduous forests of north and west areas of continental Europe, and from northern parts of Sweden towards North Africa (from north towards south) (Estrada-Peña 2001; Medlock et al. 2013). (**Figure 1. 4**). In general, numerous factors influence the distribution and dispersal of *I. ricinus*. Considerable research efforts have focused on the biological activity and geographical expansion of this species. In many of the areas where *Ixodes* ticks are established, the pathogens that they are known to transmit are endemic (Gassner et al. 2011). In the British Isles, different habitats where *I. ricinus* occur have been recorded, including grassland and coniferous woodland where *I. ricinus* is abundant, largely due to the optimal level of humidity (Pietzsch et al. 2005; Medlock et al. 2013).

The tick life cycle, distribution of animal hosts and ecological aspects should be taken into consideration as factors affecting *I. ricinus* expansion (Swanson et al. 2006). During tick questing, ticks climb up vegetation and wait for hosts to which to attach; loss of the required humidity leads to ticks descending down to the litter layer to avoid dehydration. This is known as the 'ambush' technique (Medlock et al. 2013). Tick survival requires around 80% humidity, so the expansion is always restricted to areas characterised by rainfall with vegetation that preserves the humidity (*i.e.,* leaf litter/mat layer to rehydrate). Climatic changes including temperature increases and different behavioural factors like diapause also have a direct impact on tick density and distribution (Fourie et al. 2001). In general, availability of required hosts, feeding behaviour and wooded or brushy habitats represent important limiting factors for tick dispersal within different habitats (Swanson et al. 2006; Gassner et al. 2011).

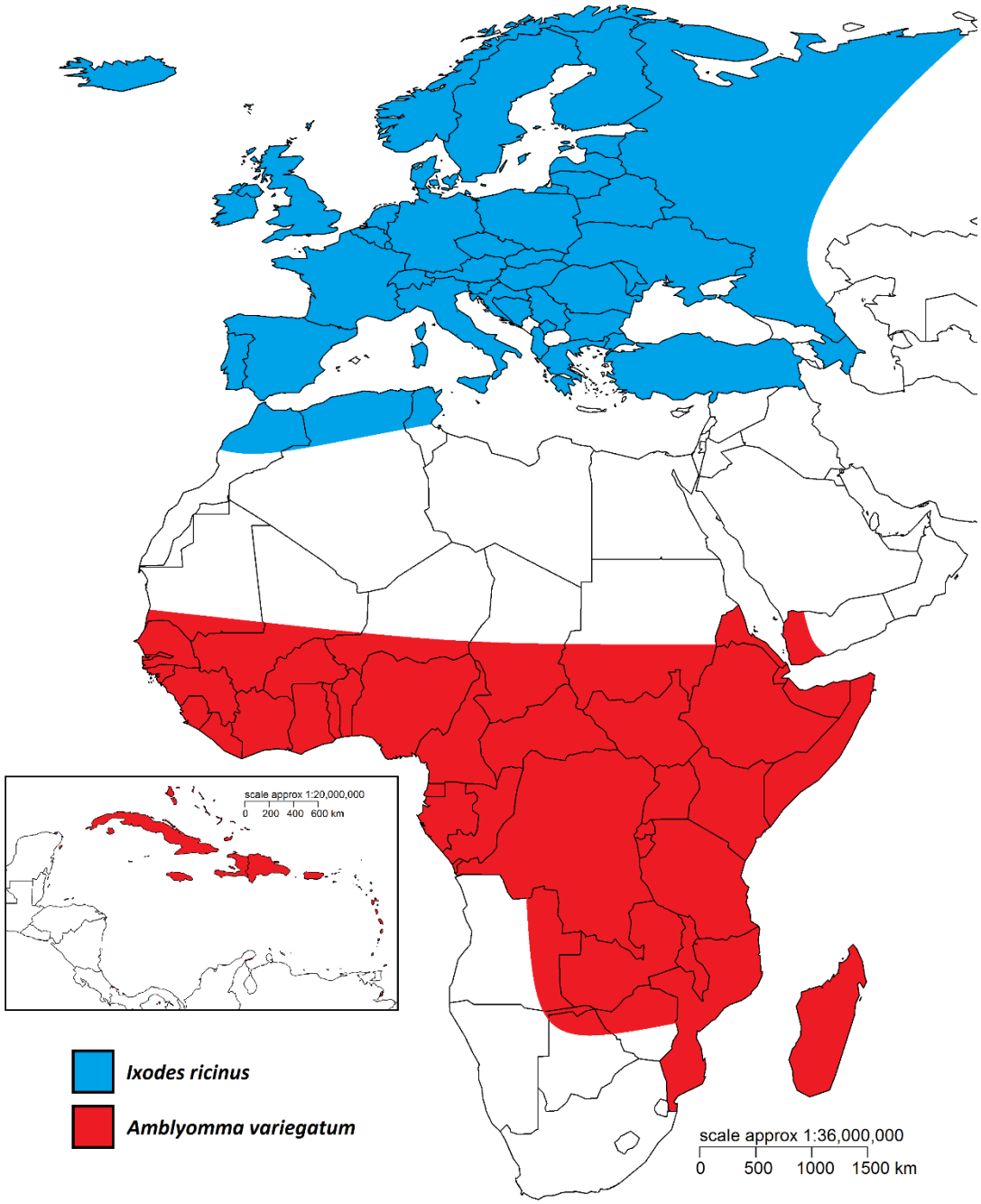


Figure 1. 4 Geographical distribution of *Ixodes species* and *Amblyomma variegatum* involved in transmitting a range of different tick-borne diseases.

1.2.3 *Ixodes ricinus* morphology

Ixodes ricinus is characterised by a red/brown scutum. The mouthparts (gnathosoma) including the hypostome, segmented palps and chelicerae are clearly

visible when the tick is viewed from above (reviewed by Hillyard, 1996). *Ixodes ricinus*, also known as the sheep tick, the castor bean tick, or the deer tick, can be identified by the length of its palpi, which are longer than they are wide. Additionally, one of the main features is the distinct anal groove that surrounds the anus. The size of unfed female *I. ricinus* is around 3.6 mm, while the fast-feeding female can measure up to 11 mm in length according to the engorgement level. In comparison, male *I. ricinus* ticks measure between 2.4 and 2.8 mm in length. The internal spur of coxa 1 (first leg segment) is elongated and distinct compared with coxa 2 and coxa 4, where the spur is external and unclear (Hillyard, 1996). (**Figure 1. 5**)

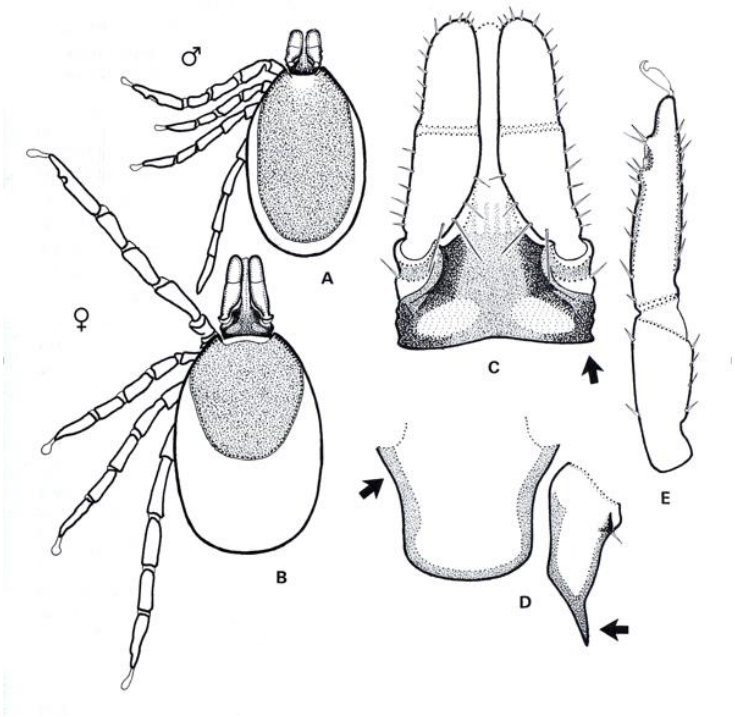


Figure 1. 5 Morphology of *Ixodes ricinus.*
A. Adult male; B. Adult female; C. Female capitulum; D. Ventral view of female basis and coxa I; E. Lateral view of female tarsus (reproduced from Hillyard, 1996). [Copyright permission was approved from the Linnean Society of London].

1.2.4 *I. ricinus* life cycle

Ixodes ricinus requires three hosts in order to complete its ife cycle (Medlock et al. 2013); the life cycle of *I. ricinus* (**Figure 1. 6**) usually takes two to four years to complete. Hatching of the eggs into the larval stage is the first step of the developmental cycle. Subsequently, the larvae seek a small host such as a rodent or a bird and then having found a favourable site to attach, they take a blood meal before dropping off. Ticks find hosts by moving up from the ground onto vegetation. After feeding and dropping off, the larvae moult into the nymphal stage. Temperature and humidity have an impact on this process because if they are not optimal, the ticks will not survive and disperse. After the cuticle has hardened, the nymph quests for a host, again often a small animal. The fed nymph detaches and moults to the adult stage. Adults again quest, attach and feed, this time on a larger host, *e.g*., a deer, then drop off the host. The female tick lays thousands of eggs and then dies (Medlock et al. 2013; Estrada-Peña, 2015) (See **Figure 1. 6**). In temperate climates, each stage of *I. ricinus* is usually active only once each year and the life cycle therefore takes up to 4 years.

The rate of development of the tick may be influenced by ecological factors. Across Europe*, I. ricinus* are facing a wide range of climatic changes and environmental conditions which appear to be assisting their development (Medlock et al. 2013). A significant variation in *I. ricinus* developmental rates was documented by (Esrada-Peña, 2015) who stated that winter temperature and environmental conditions have an impact on population dynamics of different *I. ricinus* life stages across Europe. According to German records, in 2006 during an unusually warm winter, questing adult *I. ricinus* were collected throughout the whole winter while

they are not commonly active in that period (Dautel et al. 2008). This suggests that tick developmental rate and host questing activity is closely associated with temperature.

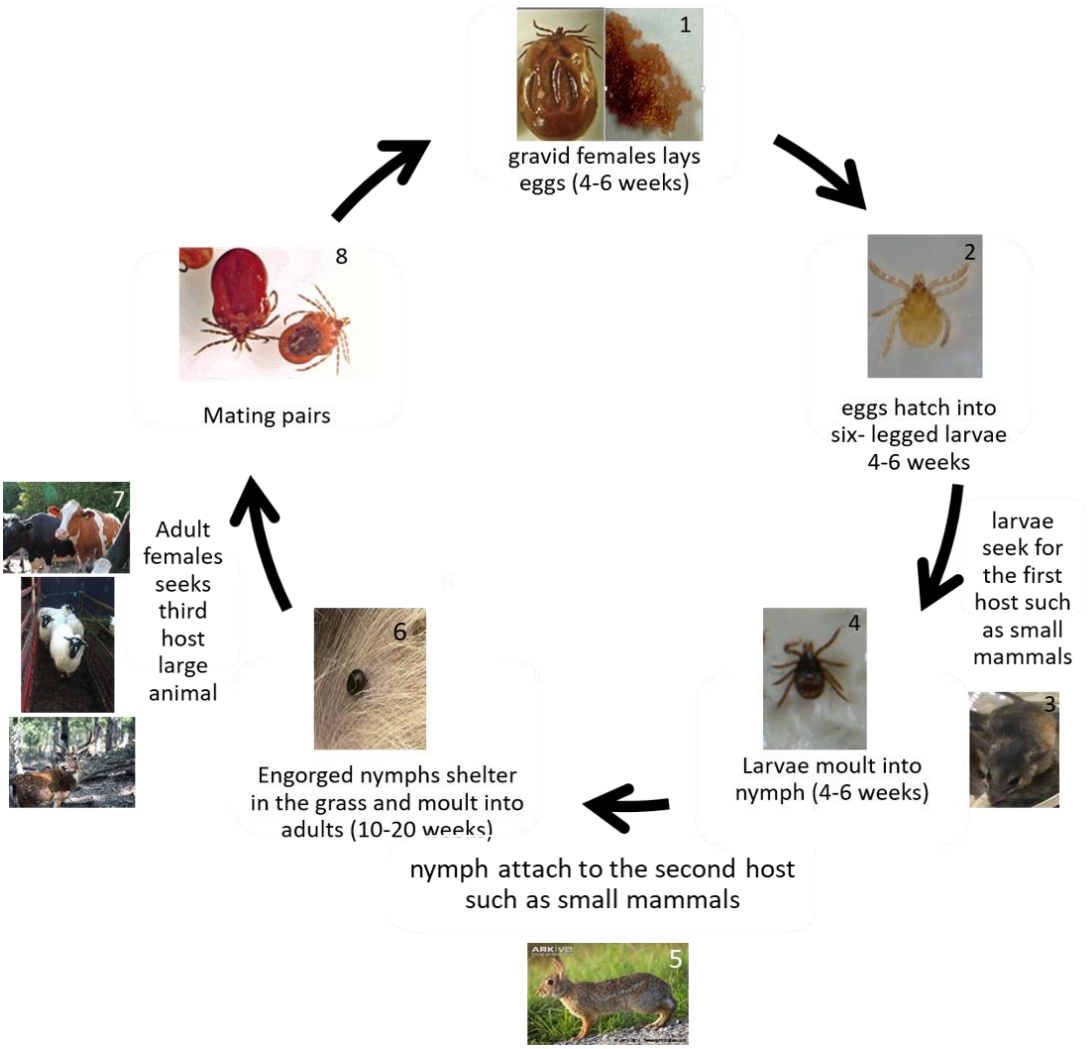


Figure 1. 6 Life cyle of *Ixodes ricinus* in the UK.
1, 2, 3, 4, 6, 8: courtesy of Dr.J.W. McGarry; 7: courtesy of Dr J. Graham-Brown; 5: courtesy of Drs J. Garry and D.B. Kaye CDC Life Science.

In terms of host effects and impact of latitude, previous studies highlighted deer as a main host for *I. ricinus*, as an individual deer is capable of supporting a large number of adult ticks. (Medlock et al. 2013). In another study, it was suggested that deer are playing a crucial role in tick dispersal and they may act as vehicles which disseminate ticks in the landscape. Moreover, the abundance of both deer and ticks are generally

correlated (Ruiz-Fons & Gilbert 2010). Recently, it has been suggested that the spatial distribution of ticks is controlled by both the local distribution of large hosts and the impact of climatic conditions (Medlock et al. 2013; Qviller et al. 2016).

1.3 *Amblyomma variegatum*

Amblyomma variegatum, also known as the tropical bont tick, is a vector of multiple pathogens of humans, livestock and wild ruminants including heartwater disease (cowdriosis) caused by *E. ruminantium* in animals across Africa (Robinson et al. 2009). Heartwater disease was also introduced into the Caribbean in imported cattle. *Amblyomma variegatum* is also responsible for transmission of the SFG rickettsiosis caused by *R. africae* (Ndip et al. 2004). *Amblyomma variegatum* also plays an important role in the epidemiology of bovine dermatophilosis and some protozoan parasites including *Theileria mutans* and *Theileria velifera* (Yonow, 1995; Akuffo et al. 2016). *Amblyomma variegatum* is the predominant vector of *R. africae* in the West, Central and Eastern areas of Africa, as well as in the eastern Caribbean; while *Amblyomma hebraeum* is the main vector of *R. africae* in Southern Africa (Lorusso et al. 2013; Kelly et al. 2010).

1.3.1 *Amblyomma variegatum* classification

Amblyomma variegatum, in common with *I. ricinus,* belongs to the subclass Acari, order Parasitiformes and suborder Ixodida, family Ixodidae. There are 129 species of *Amblyomma* ticks worldwide (Jongejan & Uillenberg 2004).

1.3.2 The geographical distribution of *Amblyomma variegatum*

Amblyomma variegatum is widely distributed in sub-Saharan Africa, and has become successfully established in areas which border the Indian Ocean, Yemen and in the Caribbean islands (Barré et al. 1995; Estrada et al. 2007; Kelly et al. 2010).

Human trade and transportation of cattle or small ruminants may have played an important role in expanding *Amblyomma* spp. distribution. Furthermore, migratory birds may act as long-distance transporters of ticks (Bram et al. 2002; Yssouf et al. 2011). The capability of *A. variegatum* to adapt to a wide range of hosts, habitats and climatic zones might explain the successful expansion of its distribution outside Africa (Barré et al. 1995; Estrada-Peña et al. 2008). *Amblyomma variegatum* has been detected in the islands of the Union of the Comoros, which is located off the southeastern coast of Africa, and is widespread on Madagascar (Yssouf et al. 2011). Temperature, relative humidity and rainfall are considered to be the most important factors which drive *Amblyomma* dispersion; *A. variegatum* prefers areas with relatively high annual rainfall and without prolonged seasonal droughts (Estrada-Peña et al. 2008). The distribution of *A. variegatum* in Africa is shown in (**Figure 1. 4**).

1.3.3 Morphology and life cycle of *Amblyomma variegatum*

Amblyomma variegatum is characterized by a brightly coloured, patterned scutum that makes it easy to differentiate; males are generally characterized by a yellow-gold colour, compared to females which are generally a brown color, (Walker et al. 2003). Furthermore, it possesses distinctly convex eyes which are separate from the margin of the scutum. The posterior lips of the *A. variegatum* female genital aperture forms a wide U shape; enameling of the festoons is absent; while lateral areas of enamel ornamentation on the scutum are small in males and the enamel of both sexes is pink-to-orange in colour (Walker et al. 2003). (**Figure 1. 7**)

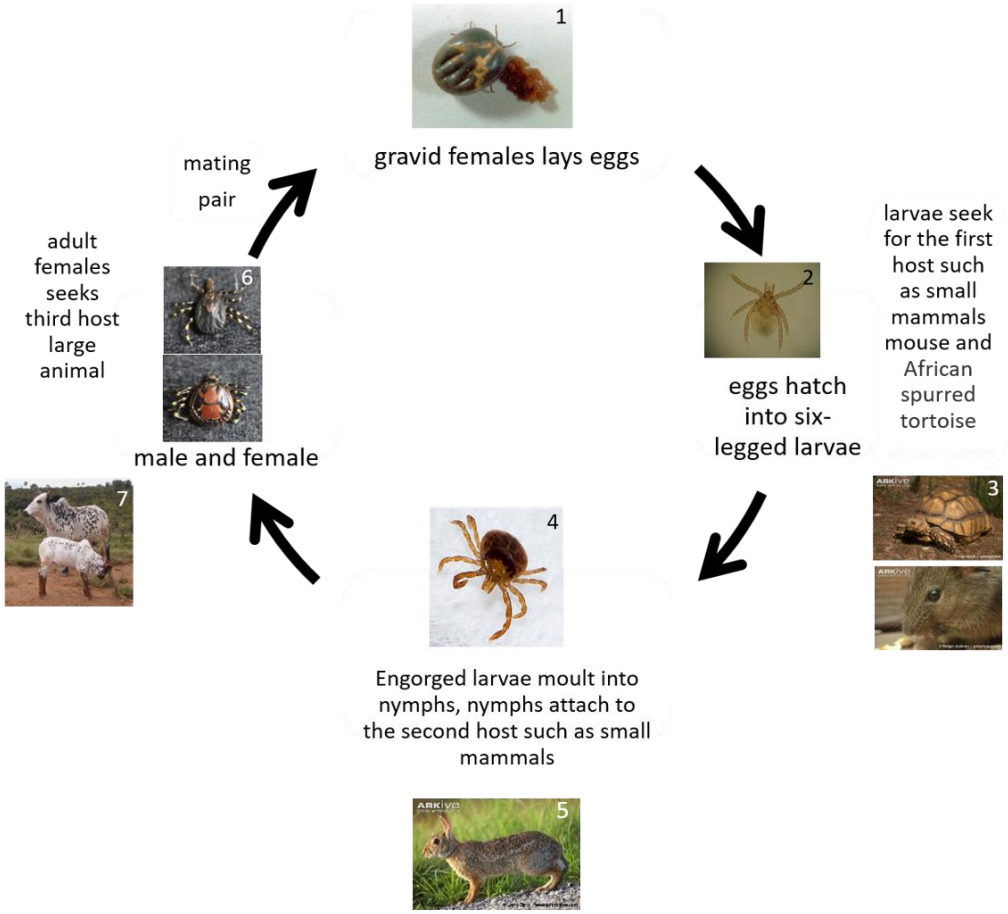


Figure 1. 7 Life cycle of *Amblyomma variegatum.*
1, 2: courtesy of Dr.J.W. McGarry; 7: courtesy of Dr B.L. Makepeace; 4, 6: courtesy of Dr J. Graham-Brown; 3, 5: courtesy www. arkive.orgarkive.org

Amblyomma variegatum is a three-host tick like *I. ricinus*. Its life cycle is characterized by two different phases: a parasitic phase (including larva, nymph and adult stages), when the tick needs to feed on blood from the host before dropping off, moulting and developing to the next life stage; and a non-parasitic phase, which takes place off the host when the larvae and nymphs moult on the ground (Barré et al. 1995).

Males attach to a host, and release pheromones which attract females for mating, and these remain on the same host for around 14 days, before the female engorges, detaches and lays her eggs on the ground (Barre & Garris 1989). Around

30,000 eggs can be produced by one female, and the air and ground temperature can influence the time taken to hatch, which can be between 50 and 100 days [warmer temperatures may accelerate egg hatching (Pegram & Banda 1990)]. Larvae are less active in cooler conditions, but they can congregate on vegetation and search for a suitable host such as cattle on which to attach for blood feeding, then engorge, drop off and finally moult into adult ticks followed by attachment to a third host. Adult ticks prefer cattle, but can also be found on camels, dogs and some wildlife. Adult *A. variegatum* feed mainly in the rainy season, while the immature ticks feed primarily during the dry season (Pegram & Banda 1990). Temperature is considered as the climate factor with the largest impact on the whole life cycle of the tick, as most of the life cycle takes place on the ground or on vegetation during the non-parasitic phase (Yonow, 1995).

1.4 Tick-borne diseases transmitted by *I. ricinus* and *A. variegatum*

Recent research has greatly improved our knowledge of individual tick-borne pathogen relationships and new and emerging zoonoses. However, our understanding of co-infections and the relationship of the tick endosymbionts to the tick-borne pathogenic bacteria remains quite unclear. Ticks are capable of transmitting a broad range of pathogenic agents, including viruses, bacteria and parasites. Different transmission routes have been described for the various pathogens transmitted by hard ticks. (Rajput et al. 2006). Selected tick-borne diseases of public health importance, including those transmitted by *I. ricinus* and *A. variegatum*, are overviewed in **Table 1. 1**.

Table 1. 1 Selected tick-borne pathogens of humans and domestic animals that are transmitted by *Ixodes ricinus* and *Amblyomma variegatum*

Vector	Pathogen	Disease caused	Geographical distribution*
Ixodes ricinus	*Borrelia burgdorferi* complex	Lyme Disease	Asia, Europe, North America
	Anaplasma phagocytophilum	Human granulocytic anaplasmosis Tick-borne fever in sheep Pasture fever in cattle Equine anaplasmosis in horse Canine anaplasmosis in dogs	Europe, North America
	Rickettsia monacensis	Spotted fever	Europe
	Candidatus Neoehrlichia mikurensis	Febrile illness in humans	Europe, Asia
	Babesia divergens,	Redwater fever in cattle	Europe, North America
	Babesia microti	Babesiosis in humans	North America, Europe
	Tick-borne encephalitis virus	Tick-borne encephalitis in humans and animals	Asia, Europe
	Louping ill virus	Neurological disease in sheep and humans	UK, Norway, Spain
Amblyomma variegatum	*Ehrlichia* (*Cowdria*) *ruminantium*	Heartwater disease (cowdriosis of ruminants)	Africa and Caribbean
	Rickettsia africae	African tick bite fever in human	Africa and Caribbean
	Theileria mutans	Benign African theileriosis of cattle	Sub-Saharan Africa

*Outside Western Europe, these diseases are transmitted by vectors other than *I. ricinus*.

Most attention has been paid to *Ixodes* ticks and their abilities to transmit multiple pathogens. *Ixodes ricinus* was recognised as a vector for erythema chronicum migrans (reported in Swanson et al. 2006). Further epidemiological studies linked this symptom to Lyme disease, caused by a number of different strains and species of *Borrelia* spirochaetes, but most commonly *B. burgdorferi*, which is transmitted primarily by *I. scapularis* in the United States and *I. ricinus* in continental Europe (Burgdorfer et al. 1982). Moreover, the first detection of human granulocytic anaplasmasis was in 1994 in Minnesota and Wisconsin, USA, caused by *Anaplasma phagocytophilum* (previously known as *Ehrlichia equi* and *Ehrlichia phagocytophila*) and shown to be transmitted by *I. scapularis*, *I. pacficus, I. persulcatus* and *I. ricinus.* Notably, in Europe, *A. phagocytophilum* is a major cause of disease in ruminants – and this is much more widespread and economically important than the human disease (Belongia et al. 2001). *Candidatus* Neoehrlichia mikurensis has been detected in questing *I. ricinus* ticks from North-West European countries (Jahfari et al. 2012), Austria (Glatz et al. 2014), and Sweden (Welinder-Olsson et al. 2010), where it causes human disease. However, although *Ca.* Neoehrlichia mikurensis causes a febrile illness in humans, it is currently uncertain whether it poses a more significant risk to public health (Glatz et al. 2014; Welinder-Olsson et al. 2010). A positive relationship has been suggested between the incidence in ticks of *Ca.* Neoehrlichia mikurensis and *A. phagocytophilum*, as both of these bacteria share an similar ecology in Central Europe (Derdáková et al. 2014; Glatz et al. 2014). Furthermore, amongst other *Rickettsia* spp., *Rickettsia monacensis* is transmitted by *I. ricinus* and causes a spotted fever -like illness in humans that has been reported in Europe (Simser et al. 2002). In addition, *Ixodes* ticks are known as a vector of parasites such as *Babesia microti* and *Babesia divergens,* causative agents of babesiosis in humans and cattle, respectively.

Ixodes ticks are capable of transmitting viruses to specific host species through their bite and cause diseases that are distributed on a regional or national geographical scale. An example of this is the genus *Flavivirus*, members of which can cause tick-borne encephalitis (TBE) disease in Central and Eastern Europe, causing signs of neurological disease in humans (Oksi et al. 1993; Estrada-Peña et al. 2012), and louping ill, which is a neurological disease of sheep and rare zoonosis in the UK, Norway and Spain (Jongejan & Uillenberg 2004).

In sub-Saharan Africa, A*. variegatum* is the predominant vector of *E. ruminantium* (which is transstadially transmitted) and *R. africae* (vertically transmitted) (Ndip et al. 2004). Co-infection with *R. africae* and *E. ruminantium* in *A. variegatum* has been identified (Allsopp 2015).

Ticks such as *Ixodes* and *Amblyomma* are capable of transmitting multiple pathogens, for instance common co-infections within a tick are borreliosis with either human granulocytic anaplasmosis or babesiosis (Swanson et al. 2006). Frequency of co-infections should be taken into consideration to allow for an evaluation of potential multiple infections following a tick bite on a human or animal (Lockwood et al. 2016). Moreover, high throughput techniques such as comparative genomic approaches and proteomic studies revealed a molecular interaction between ticks and pathogens suggesting that both ticks and pathogens co-evolved genetically in order to promote tick survival and pathogen transmission (Rikihisa 2011; de la Fuente et al. 2016).

1.5 Focusing on *Anaplasma phagocytophilum*, an exemplar pathogen of the Rickettsiales

A. phagocytophilum belongs to the family *Anaplasmataceae* in the order Rickettsiales (Dumler et al. 2001). This family includes five main genera, *Ehrlichia, Anaplasma*, *Neorickettsia, Aegyptianella* and *Wolbachia;* but *Ca.* Neoerhlichia has been recognized more recently (see below). All of these genera infect specific invertebrate hosts (acarines, insects, trematodes or nematodes). While *Wolbachia* and *Neorickettsia* can be transmitted vertically, *Anaplasma* and *Ehrlichia* cannot effectively pass from adults to offspring. The bacterium was originally described in the 1930s as *Rickettsia phagocytophilum* infecting sheep only, then was reported as *Cytoecetes phagocytophila* and *Ehrlichia phagocytophila* in ruminants and *Ehrlichia equi* in equine hosts (Bergey et al. 1974; Rikihisa 2011), and finally reclassified as *Anaplasma phagocytophilum* according to molecular studies (Dumler et al. 2001). This pathogen also causes the disease human granulocytic anaplasmosis (HGA). Although granulocytic anaplasmosis was previously characterized as having significant veterinary importance only (Rikihisa 1991), in the early 1990s this pathogen was identified as the causative agent of human disease and first recorded as human granulocytic ehrlichiosis (Chen et al. 1994). *Anaplasma phagocytophilum* is related to *E. ruminantium* that infects endothelial cells and *Ehrlichia chaffeensis* that causes human monocytic ehrlichiosis (infecting monocytes). However, *Anaplasma phagocytophilum* invades and replicates inside human polymorphic granulocytes (Chen et al. 1994).

Fatal HGA disease has been confirmed in the USA, but the disease is rare in Europe and usually clinically mild. However, at least two human cases has been detected in the UK (Sumption et al. 1995; Hagedorn et al. 2014). The disease is

characterized by flu-like symptoms in humans with fever, headache, and leukopenia. In sheep, anorexia, chills, and immunosuppression are observed, which can lead to tick pyaemia due to a secondary infection with *Staphylococcus* species (Woldehiwet 2006). *A. phagocytophilum* causes pasture fever or tick-borne fever in cattle and sheep, respectively; it also causes equine and canine granulocytic anaplasmosis. This disease can have serious economic impacts on the livestock industry (Stuen 2007).

The human disease is a significant public health concern, particularly in people who work outdoors regularly. The more extensive distribution of tick vector populations due to climate change and reservoir host movement has increased the emergence of tick-borne diseases in general, including anaplasmosis (Beugnet & Marié 2009). Moreover, the severity of HGA disease varies from asymptomatic to a potentially fatal disease (Lin et al. 2011; Woldehiwet et al. 2002; Rikihisa 2010). Doxycycline is the drug of choice for anaplasmosis, but there is not yet any available vaccine (Rikihisa 2010).

1.5.1 *Anaplasma phagocytophilum* transmission and life cycle

Anaplasma phagocytophilum depends on haematophagous ticks as vectors, including *I. scapularis* and *I. pacificus* in the USA; *I. ricinus* is the predominant vector in Europe, while anaplasmosis is transmitted in the Baltic States, Russia and Northern Asia by *I. persulcatus* (Lin & Rikihisa 2003; Burkot et al. 2001; Rikihisa 2010). The *Anaplasma* life cycle involves wild mammals and tick species, whereas humans acquire anaplasmosis accidentally via the bite of infected ticks (Rikihisa, 2011). White-tailed deer and white-footed mice in the USA act as a reservoirs for *Anaplasma* infection. In Europe, wild rodents (Ogden et al. 1998), as well as roe deer and red deer, act as reservoirs for *Anapalsma* infection (Alberdi et al 2016; Stuen et al. 2013).

Vertical or transovarial transmission has not been detected in *A. phagocytophilum*. Once larvae or nymphs are infected via blood feeding on infected animals, *Anaplasma* migrates through the gut to infect the salivary glands, then it is maintained in the secretory acini of the salivary glands to the next developmental life stage (Hodzic et al. 1998). The pathogen than replicates and is transported to the vertebrate host during blood-feeding; successeful infection occurs when the animal host is susceptible to the particular strain (Ogden et al. 1998; Rikihisa 2010). Once mammals are infected, *A. phagocytophilum* invades and proliferates within cells of the primary host immune system, targeting polymorphonuclear granulocytes in which it forms morulae (Lin et al. 2011). (**Figure 1. 8**)

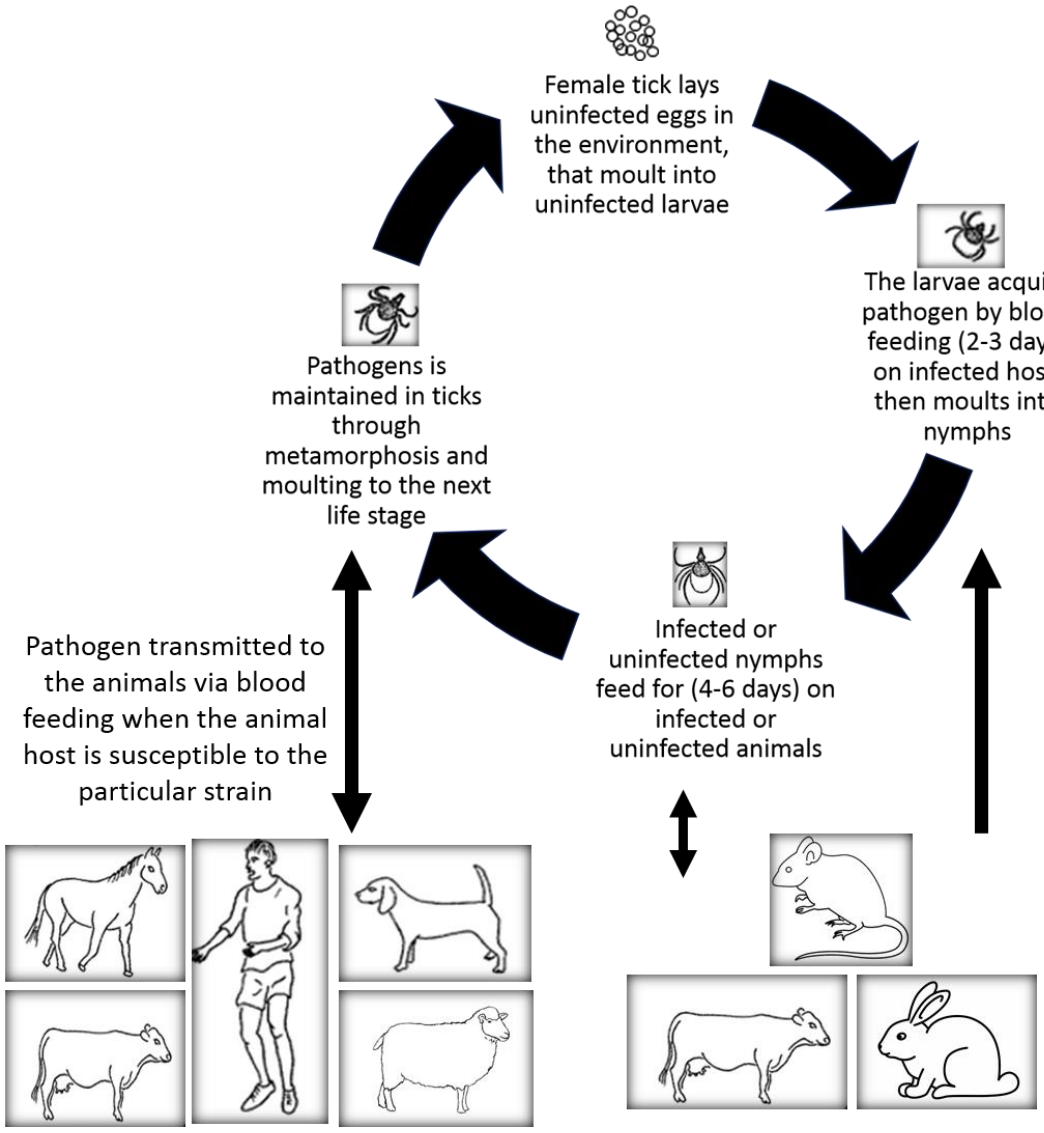


Figure 1. 8 Proposed transmission of cycle of *Anaplasma phagocytophilum* from (Rikihisa 2011).

1.5.2 *Anaplasma phagocytophilum* in the mammalian host

Although *A. phagocytophilum* infects different mammalian hosts, in this section we focus on humans, as almost all published work has been done with the human host. In humans, *A. phagocytophilum* exhibits a tropism for granulocytes; *A. phagocytophilum* is characterized by a unique ability to invade and multiply in these innate immune system cells (*i.e.*, neutrophils) by subversion of their powerful innate antimicrobial defences (Rikihisa 2006). In most of the research on *A.*

phagocytophilum, the human promyelocytic cell line HL- 60 has been used to identify the molecular mechanisms and specific domains that mediate binding with host receptors and invasion (Rikihisa, 2011; Seidman et al. 2015).

Neutrophils express a variety of cell receptors comprising nucleotide-binding oligomerization and transmembrane Toll-like receptors (TLRs) that encompass cytoplasmic protein receptors, which have an ability to interact with pathogen-associated molecular patterns (PAMPs). This results in multiple immune responses such as autophagy, the fusion of lysosomes, activation of cytokines/chemokines and mitogen-activated protein kinases against foreign microorganisms. However, two key bacterial PAMPs, peptidoglycan and lipopolysaccharides, are absent in *A. phagocytophilum,* so the bacteria can be engulfed, then replicate forming morulae that occupy most of the cytoplasm of infected host cells (Lin et al. 2011). Furthermore*, "*regulatory hijacking" is considered as another subversion mechanism of this pathogen involving histone modifying enzymes (Rikihisa, 2011). The *A. phagocytophilum* inclusion avoids combination with secretory vesicles harbouring nicotinamide adenine dinucleotide phosphate-oxidase (NADPH oxidase) in order to avoid NADPH oxidase-mediated killing during neutrophil infection (Carlyon et al. 2004), and it interferes with vesicular trafficking to avoid lysosomes (Mott et al. 1999). Subversion of neutrophils by *A. phagocytophilum* is implied by two different strategies. Firstly, involving autophagy induction*,* rapamycin stimulation enables *A. phagocytophilum* bateria to set up their infection in autophagosome-like compartments, away from lysosomes, to guarantee space for their growth and proliferation (Niu et al. 2008). Secondly, suppression of apoptosis results in upregulation of the expression of the anti-apoptotic bcl-2 family members (Pedra et al. 2005; Lee & Goodman 2006).

It has been reported that the tetrasaccharide sialyl Lewisx (sLex) on P-selectin glycoprotein ligand 1 (PSGL-1) is critical for *A. phagocytophilum* infection (Seidman et al. 2015). Induction of (PSG-1) resulted in promotion of tyrosine phosphorylation of ROCK1, an effector kinase of the GTPase RhoA, by spleen tyrosine kinase (Syk), which was confirmed by small interfering RNA (siRNA) knockdown targeting either Syk or ROCK1, which resulted in suppression of *A. phagocytophilum* infection (Rikihisa, 2011). A recent study identified three surface proteins, OmpA (outer major protein), Asp14 (14-kDa *A. phagocytophilum* surface protein) and AipA (*A. phagocytophilum* invasion protein A), that are essential for *A. phagocytophium* invasion and binding of host cells (Seidman et al. 2015). An antibody cocktail targeting these proteins impaired adhesion/ invasion activity between those proteins and α-2,3- sialic acid and α1,3-fucose to bind PSG-1-modelled glycopeptides, which impaired binding to myeloid cells (HL-60), thus revoking efficient *A. phagocytophilum* infection (Seidman et al. 2015).

Another study reported a successful infection with *A. phagocytophilum* of human microvascular endothelial cells (HMEC) (Severo et al. 2012). Incubation of *A. phagocytophilum*-infected HMEC-1 with neutrophils resulted in granulocyte infection, and that may explain the upregulation of HMEC-1 protein related to leucocyte adhesion, suggesting the initial role of infected endothelial cells after *A. phagocytophilum* infection and before neutrophil infection (Severo et al. 2012).

1.5.3 *Anaplasma phagocytophilum* compared with *Anaplasma marginale*

Anaplasma phagocytophilum and *A. marginale* are both transmitted by ticks and infect vertebrates (Zivkovic et al. 2009). However, *A. marginale* infects erythrocytes (Zivkovic et al. 2009), whereas *A. phagocytophilum* infects leukocytes (Rikihisa 1991). Both *A. marginale* and *A. phagocytophilum* can be propagated in the *I. scapularis* embryo-derived cell line ISE6. *In vitro* cultivation systems using continuous tick cell lines have been applied in different investigations, and represent a suitable tool for purification and isolation of pathogens and their subsequent propagation (Passos 2012; Bell-Sakyi et al. 2007).

Different tick gene expression profiles in response to infection with *A. phagocytophilum* and *A. marginale* in ISE6 cells have been identified (Zivkovic et al. 2009). Both GST and ferritin genes were found to be expressed in ISE6 due to infection of *A. marginale* and *A. phagocytophilum*; however, upregulation of GST and ferritin followed by downregulation was seen in *A. marginale* infection of ISE6 cells, while expression of the same genes showed opposite pattern in the case of *A. phagocytophilum*. *In vitro* gene expression patterns following infection with *A. phagocytophilum* were confirmed by analysis of whole *I. scapularis* ticks (Zivkovic et al. 2009). Variation in gene expression may reflect the differences in the life cycle of both pathogens; another explanation for this variation is that *I. scapularis* is not a natural vector of *A. marginale* (Zivkovic et al. 2009). Moreover, other genes were differentially expressed after *A. phagocytophilum* infection of ISE6 cells, including signal sequence receptor delta, ixodegrin-2A RGD containing protein, NADH-ubiquinone oxidoreductase, ubiquitin C variant 5-like, gamma actin-like protein, vonWillebrand factor, troponin I, and ribosomal protein L32 (Zivkovic et al. 2009),

which may play a similar role in pathogen replication or modulation of tick cell immune response as the role reported previously in *A. marginale* (de la Fuente et al. 2017).

1.5.4 Biological process involved in tick – *Anaplasma phagocytophilum* interaction

Like other members of the family *Anaplasmataceae*, *A. phagocytophilum* replicates in membrane-bound vacuoles within the cytoplasm of eukaryotic host cells, which make it different biologically compared with members of the family Rickettsiaceae, which replicate inside host cell cytoplasm (Rikihisa 2011).

As an intracellular bacterium, *Anaplasma* exhibits adaptations for this lifestyle. A lack of biosynthesis genes including peptidoglycan and lipopolysaccharides allows the bacterium to live inside a membrane-bound inclusion in the host cell cytoplasm. This feature also allows *Anaplasma* to exploit host cell cholesterol to support bacterial membrane integrity (Lin & Rikihisa 2003; Rikihisa 2011). Invasion of *A. phagocytophilum* depends on transglutaminase activity instead of microfilament activity, which involves signaling molecules such as transglutaminase, tyrosine phosphorylation, phospholipase C and consequently increases of cytosolic free Ca^{2+} required for cell invasion(Rikihisa 2003). *Anaplasma phagocytophilum* inclusion bodies are unique not only because of the exceptional invasion method that they employ, but also because of the unique ability of *A. phagocytophilum* inclusion bodies to avoid the lysosomes (Rikihisa, 2003). Different strategies are applied by *A. phagcocytophilum* by isolating its inclusion bodies away from lysosomes; as *A. phagocytophilum* inclusions are negative for cytoplasmic vacuolar-type H^+ ATPase and early endosomal antigen 1 (EEA1) (Mott et al. 1999), *A.*

phagocytophilum inclusions successfully segregate themselves away from host endocytic and exocytic vesicular traffic and lysosomal fusion (Rikihisa 2003).

Anaplasma infection can be characterised by a biphasic developmental cycle in mammalian and tick host cells. The adherent and infectious form of *A. phagocytophilum* is spheroid in shape and known as the dense-cored (DC) form. The DC form induces its receptor-mediated uptake into a host cell-derived vacuole; in other words, it adheres to the host cell surface and facilitates its own uptake, but it does not replicate. Within the host cell vacuole and taking between 4 and 8 hours, the DC form evolves into the non-infectious pleomorphic reticulate cell (RC) form, and subsequently a bacterial morula (bacterial cluster) is generated through replication by binary fission. The RC form is then converted back to the DC form within 24-36 hours; then DC organisms either extrude from or lyse the host cell and are released to initiate the next infection cycle (Seidman et al. 2015; Rikihisa 2011). Previous proteomics analysis, confirmed by transcriptional and electron microscopy observations, revealed that the hypothetical protein APH_1235 was the most abundant protein upregulated on the DC form while detection of APH_1235 on RC bacteria was minimal (Troese et al. 2011). This study found that one-quarter of the *A. phagocytophilum* proteins that were detected are involved in protein synthesis and are metabolically active within the RC form (Troese et al. 2011).

Currently, considerable attention has been focused on the host-pathogen interactions that occur during *Anaplasma* infection and replication. However, a few features have been identified. Interestingly, *A. phagocytophilum* utilizes sophisticated strategies and unique mechanisms critical for its invasion, survival, persistence and subsequent transmission from the tick vector to mammalian host

cells; a key component of this is the manipulation of host cellular events (Sultana et al. 2010).

1.5.4.1 Host sensing/initial invasion

The most common target of *A. phagocytophilum* during adherence and invasion is associated with host cytoskeletal remodelling (actin protein), resulting in changes in host cell morphology *in vivo* and *in vitro* (de la Fuente et al. 2017). *Anaplasma phagocytophilum* reorganizes the vector actin cytoskeleton to facilitate infection through selective regulation of gene transcription in association with the RNA polymerase II and the TATA-binding protein (Sukumaran et al. 2006). Specifically, tick cell signalling is modulated when *A. phagocytophilum* increases the expression of the G protein–coupled receptor Gβγ subunits, which activate the phosphoinositide 3-kinase (PI3K–p21-activated kinase (PAK1) signalling pathways (de la Fuente et al. 2017). This mechanism is obscured by elevation of actin phosphorylation: increased nuclear phosphorylated G-actin is followed by inhibition of F-actin (actin polymerization), which is correlated with stable induction of RNA polymerase II (RNAPII) transcription factor and TATA box–binding protein (TBP), which selectively controls and promotes *I. scapularis salp16* gene transcription (a gene crucial for *A. phagocytophilum* survival and salivary gland infection) (Sukumaran et al. 2006). This suggests a notable role for *A. phagocytophlilum* in altering gene transcription in the tick vector (Sultana et al. 2010). Moreover, in the vector midgut, the tick response to infection through spectrin-chain or fodrin activation is utilized by *A. phagocytophilum* to facilitate infection due to cytoskeleton remodelling (de la Fuente et al. 2016). Interestingly, a cytoskeletal remodelling mechanism through actin rearrangement is

also employed by *A. phagocytophilum* to facilitate the infection of vertebrate host cells (de la Fuente et al. 2016). (**Figure 1. 9**)

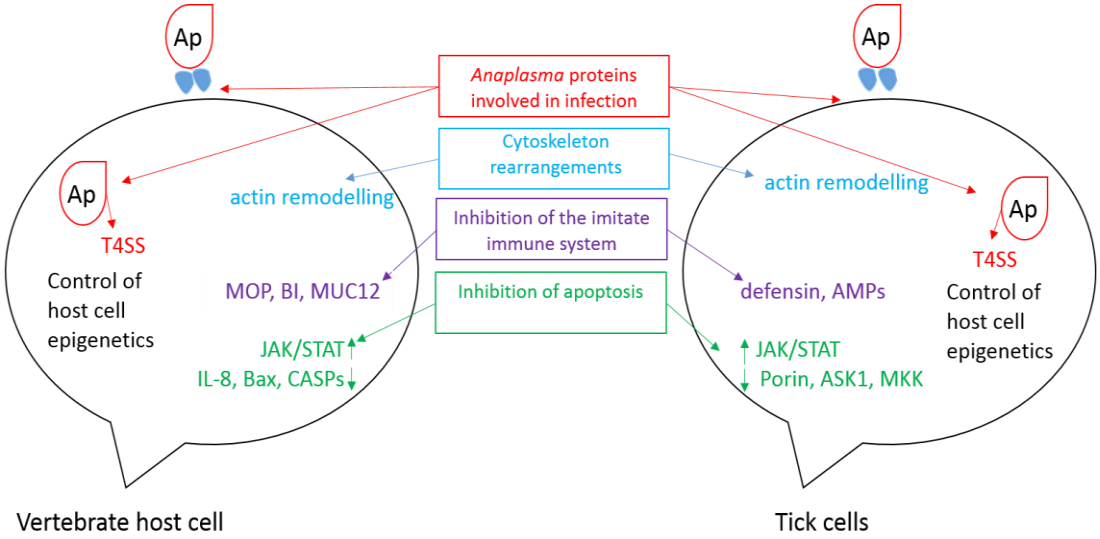


Figure 1. 9 Comparison of the common strategies used by *Anaplasma phagocytophilum* between tick and vertebrate.
Diagram is modified from de la Fuente et al. 2016. Cell invasion by *A. phagocytophilum* involves actin remodelling and manipulation of host apoptosis and host immune system. However, both actin rearrangement and apoptosis inhibition (involving up-regulation and down-regulation of variable pathways) facilitate *A. phagocytophilum* infection in both host and tick cells, but different pathways are involved in manipulation of host immune response between vertebrate and tick cells. (PAMPs): pathogen-associated molecular patterns; (JAK/STAT): Janus kinase/signal transducers and activators of transcription; (T4SS): type IV secretion system.

1.5.4.2 Inhibition of cell apoptosis

Following successful infection of host cells, inhibition of apoptosis is a key adaptation that *A. phagocytophilum* exploits to maintain infection (Ashida et al. 2011) (**Figure 1. 9** and **Figure 1. 10**) Infection with *A. phagocytophilum* produces complex tissue-specific alterations at the transcript and protein levels (Ayllón et al. 2015a).

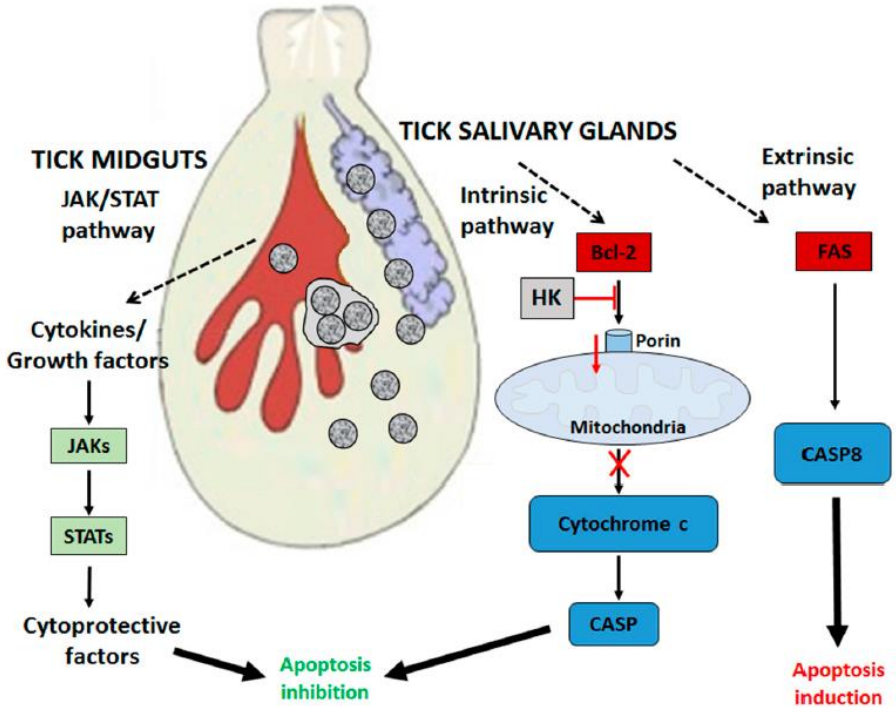


Figure 1. 10 Tissue specific modulation of apoptosis pathways in tick cells following *Anaplasma* infection.

In the tick midgut, *A. phagocytophilum* inhibits apoptosis via stimulation of the JAK/STAT pathway. In tick salivary glands, inhibition of porin results in the decrease of the cytochrome c release, consequently inhibiting the mitochondrially-induced intrinsic apoptosis pathway. On the other hand, induction of the extrinsic apoptosis pathway is associated with inhibition of fatty acid synthase (FAS) proteins. Downregulation in red, upregulation in green, CASP: caspase; HK: hexokinase (Alberdi et al. 2016a).

Apoptotic cell death is an intrinsic immune defence mechanism in response to microbial infection that results in reduction of infected cells for the benefit of the remaining cells (Ashida et al. 2011). Inhibition by *A. phagocytophilum* of the intrinsic apoptosis pathway in *I. scapularis* tick salivary glands is implied by down-regulation of mitochondrial porin expression at a tissue-specific level (Ayllón et al. 2015b), suggesting subversion of tick immune mechanisms to increase infection (**Figure 1. 9** and **Figure 1. 10**). Moreover, at the mRNA levels, porin expression down-regulation

in tick salivary glands results in the inhibition of cytochrome C release characterised by its localisation within the mitochondrion of infected cells, contrasting with a diffuse pattern of localisation in non-infected tick salivary gland cells (Ayllón et al. 2015b). On the other hand, up-regulation of cytochrome C has been detected in adult female midguts which corroborated with upregulation of porin expression in response to infection, suggesting that this impact is a part of the anti-apoptotic mechanism to further facilitate bacterial infection. Nevertheless, litte is known about differences in porin and cytochrome C protein levels between uninfected and *A. phagocytophilum*-infected tick salivary glands (Ayllón et al. 2015b) (**Figure 1. 11**).

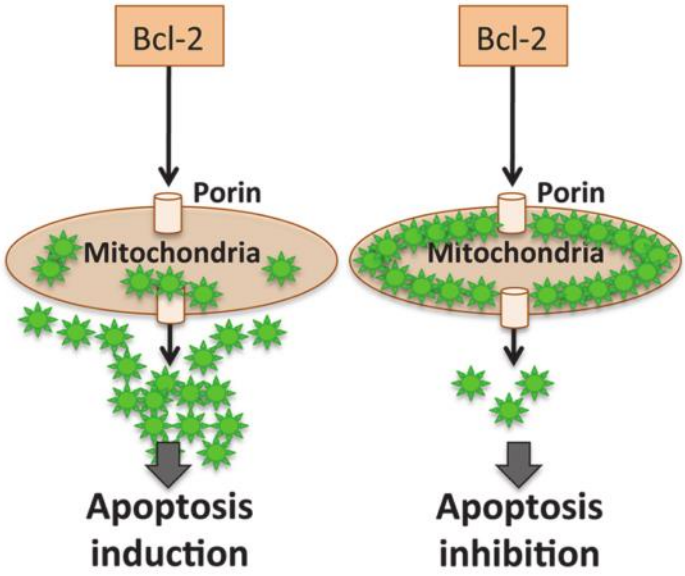


Figure 1. 11 Model of Porin-mediated inhibition of cytochrome c release highlights the anti-apoptosis mechanism.
This process facilitates *Anaplasma phagocytophilum* infection in tick salivary glands cells (Ayllón et al. 2015b).

On the other hand, ticks possess a defence mechanism comprising activation of the extrinsic pathway (death receptor mediated apoptosis) through activation of fatty acid synthase (FAS) ligand (FasL)/receptor that triggers apoptosis in order to limit, at least in part, bacterial infection in the salivary glands

(Bandyopadhyay et al. 2006). In other words, tick salivary glands may be responding to *A. phagocytophilum* infection by activation of the extrinsic apoptosis pathway, which may serve to counteract bacterial inhibition of the intrinsic apoptosis pathway (Ayllón et al. 2015b).

In addition, transcriptomic analysis revealed that *A. phagocytophilum* also inhibits cell apoptosis through induction of the tick Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway during infection in both midgut and salivary glands to facilitate bacterial infection (Galindo & de la Fuente 2012; Ayllón et al. 2015a). This suggests that a similar mechanism is used by *A. phagocytophilum* to induce infection by activating the JAK/STAT pathway to inhibit neutrophil apoptosis (Galindo & de la Fuente 2012; Ayllón et al. 2015a). So, transcriptome and proteome studies have demonstrated tissue-specific molecular pathways involved in pathogen infection, replication and transmission by ticks, which highlight the different strategies that *A. phagocytophilum* uses to modulate cell apoptosis and avoid destruction within host cells. While it has been reorted that inhibition of apoptosis caused by up-regulation of the JAK/STAT pathway is more likely to be restricted to the tick midgut, apoptosis inhibition by decreasing the intrinsic apoptosis pathway is more likely to be restricted to tick salivary glands in order to invade cells and manipulate host defence mechanisms (Ayllón et al. 2015a).

A recent study of *A. phagocytophilum* infection based on metabolomics (^{1}H nuclear magnetic) resonance (NMR), transcriptomics (mRNA), and proteomic studies using the ISE6 *in vitro* cultivation system documented a new metabolic pathway affected by *A. phagocytophilum* infection of tick cells (Villar et al. 2015b). It has been suggested that *A. phagocytophilum* infection influences protein processing in the endoplasmic reticulum (ER) by promoting protein misfolding (protein conformational

disorders) to establish and assist infection by limiting and impeding tick immune responses (Villar et al. 2015b). Misfolding applied by the bacteria induces protein targeting and degradation to avoid ER stress and cell apoptosis in the tick vector, consequently favouring *A. phagocytophiluim* infection (Villar et al. 2015a). In addition, decreased tick cell glucose metabolism as a result of phosphoenolpyruvate carboxykinase (PEPCK) inhibition, mitogen-activated protein kinase (MKK) and apoptosis signal-regulating kinase 1 (ASK1), decreases cell apoptosis, thus counteracting the tick cell response to infection. It is possible that ticks benefit from *A. phagocytophilum* infection by increased survival, suggesting a **n**ewly discovered tick- *Anaplasma* co-evolution strategy (Villar et al. 2015a).

Continuing with apoptosis, another mechanism used by *A. phagocytophilum* in vertebrates to subvert the host innate immune system and exploit the host cells is the use of a type IV secretion system (T4SS) (**Figure 1. 9**). The T4SS is a multi-protein complex arranged in a needle-shaped structure that translocates bacterial factors across membranes (Rikihisa 2003). The T4SS effector molecules or substrates include hypothetical *A. phagocytophilum* effector molecules, ankyrin-repeat-rich protein A (AnkA) and *Anaplasma* translocated substrate 1 (Ats-1), characterized by distinct subcellular localization and signalling in host cell (Rikihisa 2003). These are not toxic to the host cell and are involved in cell metabolism, facilitating bacterial infection by manipulation of the host cell and aiding the infection process (Niu et al. 2010; Lin et al. 2007). Ats-1 has a mitochondria-targeting presequence; it translocates into the host cell mitochondrion via its distinct ability to pass through five membranes (bacterial inner and outer membranes, inclusion membrane, and outer and inner membranes of the mitochondrion); subsequently Ats-1 inhibits apoptosis through inhibition of etoposide-induced cytochrome C release and poly ADP-ribose

polymerase cleavage (Niu et al. 2010; Rikihisa et al. 2010). Other bacterial anti-apoptotic mechanisms involving decrease of Bax translocation into the mitochondria, caspase (CASP)-3 processing, cleavage of pro-CASP-8, activation of CASP-8 and CASP-9, and degradation of the X-linked inhibitor of apoptosis protein (XIAP) have been identified (Ge et al. 2005; Lee et al. 2008) (**Figure 1. 9**).

The host apoptotic response *in vitro* differs to that of the *in vivo* mechanisms described above. This is further complicated by cell line-specific responses that have been reported *in vitro* (Alberdi et al. 2016a), which suggests that *A. phagocytophilum* utilises multiple mechanisms to establish and maintain infection (**Figure 1. 12**). Infection of *I. scapularis* ISE6 cells leads to a downregulation of cytochrome C, which inhibits the intrinsic apoptosis pathway (mitochondria mediated apoptosis pathway) (Alberdi et al. 2015; Ayllón et al. 2013); whereas *I. ricinus* embryonic IRE/CTVM19 cells respond to infection via an upregulation of both fatty acid synthase and the JAK/STAT pathway (Alberdi et al. 2016a). The transcriptional profile of *A. phagocytophilum* infection observed *in vitro* suggest that the ISE6 cell more closely represents the salivary gland apoptotic response, whereas IRE/CTVM19 cells more closely represent the midgut response to infection (Alberdi et al. 2016a) (**Figure 1. 12**). Finally, the regulation of anti-apoptotic genes coding for B-cell lymphoma 2 (Bcl-2), phosphoinositide kinase-3 (PI3K)/protein kinase B (Akt) signalling, the activated myeloid leukemia cell differentiation (Mcl-1) proteins, and interleukin 8 (IL-8) autocrine secretion are also involved in apoptosis inhibition by *A. phagocytophilum* in infected mammalian host cells (Ge et al. 2005; Lee et al. 2008).

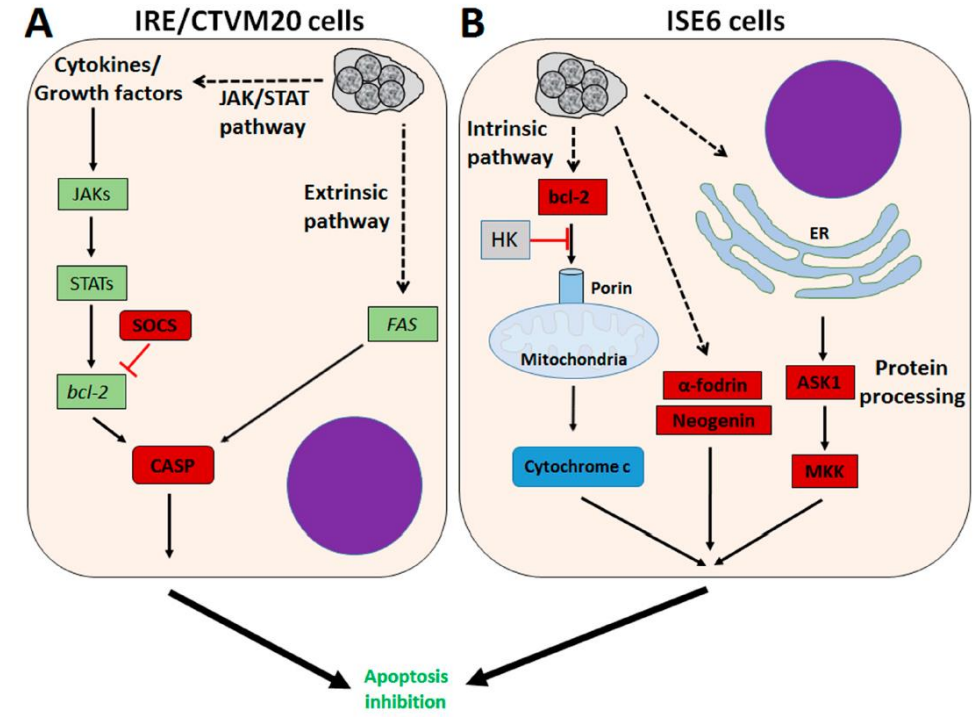


Figure 1. 12 The host response to *Anaplasma* infection differes between different tick cell lines, suggesting a species-specific interaction.
A- Apoptosis inhibition in IRE/CTVM20 via both fatty acid synthase and Janus kinase/signal transducers and activators of transcription (JAK/STAT) pathway. **B-** Apoptosis inhibition via decrease the intrinsic apoptosis pathway, inhibition of sectrin alpha chain or alpha-fodrin and Neognin, and modulation of glucose metabolism and mitochondrial porin and manipulates protein processing in Endoplasmic reticulum. Downregulation: red, up-regulation: green, ASK1: apoptosis signal-regulating kinase 1; MKK: mitogen-activated protein kinase; SOCS: suppressor of cytokine signalling (Alberdi et al. 2016a).

1.5.4.3 Manipulation of the immune response

In addition to cytoskeletal remodelling and modulation of host apoptotic processes, *A. phagocytophilum*, like other pathogens, subverts cellular innate immune responses in both ticks (Ayllón et al. 2015a) and vertebrate hosts (Ayllón et al. 2015b; de la Fuente et al. 2005; Lee et al. 2008). In addition, tick saliva contains different bioactive molecules that both facilitate tick feeding on vertebrates and

promote pathogen acquisition (Kotál et al. 2015). Upregulation of genes coding for Toll pathway proteins (Toll, Pelle, and MyD88) have been observed, while those encoding the immune deficiency (IMD) pathway (Caudal, Relish, TAK1, Caspar, and IAP2), JAK/STAT pathway (JAK, STAT, JAK receptor, and SOCS), defensin, and antimicrobial peptide proteins were down-regulated as a result of *A. phagocytophilum* infection in *I. scapularis* tick midgut. However, downregulation of defensin and antimicrobial peptide proteins have been reported in the salivary glands only (Ayllón et al. 2015a). Nevertheless, *A. phagocytophilum* infection induced upregulation of subolesin (a protein involved in the tick immune response to pathogen infection) in both tick midgut and salivary glands (Hajdušek et al. 2013). In contrast, neutrophils revealed downregulation of genes coding for proteins involved in bacterial killing (myeloperoxidase [MPO], transferrin (TF) and bactericidal/permeability-increasing (BPI) protein] and cell protection [mucin 12 (MUC12)], compared with upregulation of interferons, cytokines and chemokines (encoding immune-system-related genes), as a result of *A. phagocytophilum* infection (de la Fuente et al. 2005) ; Lee et al. 2008).

1.5.4.4 *Anaplasma phagocytophilum* infection benefits ticks

Although *A. phagocytophilum* infection is not essential for tick survival, *A. phagocytophilum* infection appears to benefit ticks. It has been reported that *A. phagocytophilum* infection contributes to tick survival and feeding fitness by increasing the ability of ticks to endure the cold through up-regulation of an *I. scapularis* antifreeze glycoprotein (ISAFGP) and heat shock proteins (HSPs) as elucidated through RNAi knockdown investigations (Neelakanta et al. 2010). In addition, HSP70 and HSP90 were overexpressed during stress such as high

temperature, toxicity and *A. phagocytophilum* infection; overexpression of HSPs counteracts stress and pathogen infection and assures tick survival (Villar et al. 2015a).

Other research suggested a compensatory mechanism between *A. phagocytophilum* and the tick vector through tick cell epigenetics controlled by T4SS or other secretion mechanisms, by which inhibition of host cell responses is employed by increasing the level of tick histones and histone modifying enzymes (HMEs) (Cabezas-Cruz et al. 2016). These HMEs inhibit cell apoptosis, thereby facilitating and guaranteeing *A. phagocytophilum* infection; this suggests a potential role in increasing tick fitness which assures survival for both pathogen and tick vector (Cabezas-Cruz et al. 2016). In general, different complex and distinct mechanisms are applied by *A. phagocytophilum* to survive and multiply in different hosts despite its small genome and obligatory intracellular lifestyle. Another example of a positive effect of pathogen infection on tick fitness is observed in *I. ricinus* infected by *B. burgdorferi*, in which the bacteria help the ticks to survive in highly desiccating conditions (de la Fuente et al. 2017).

However, although recent research has highlighted some examples that elucidate a beneficial effect of pathogen acquisition on tick fitness as explained above for *A. phagocytophilum* and *B. burgdorferi*; acquisition of pathogens by ticks is not necessarily beneficial. It has been shown that infection of ticks with *Babesia* parasites poses a negative effect on tick development and a harmful effect on the oviposition of *R. microplus* engorged females (Cen-Aguilar et al. 1998), thus ticks are proposed to have evolved defence mechanisms to control *Babesia* infection and to regulate their mutual interaction (Cen-Aguilar et al. 1998).

1.6 Tick-microbiome interaction

1.6.1 Ticks and endosymbiotic relationships

The powerful and flexible nature of next generation sequencing (NGS) technologies have revolutionized genome research in a broad spectrum of biological sciences. These NGS technologies have revealed the complex nature of the tick microbiome that consists of a combination of potential commensal microorganisms and notorious pathogenic agents, which may eventually affect tick-pathogen interactions and vector competence (Vayssier-Taussat et al. 2015). In common with many other arthropods, the advent of 16S rRNA profiling using NGS methods has revealed complex microbiomes in ticks (Bonnet et al. 2017). In general, bacterial symbionts that are transmitted vertically are classified into 10 genera [*Coxiella*-like endosymbionts (LE), *Rickettsiella*, *Arsenophonus*, *Francisella*-LE, *Cardinium*, *Spiroplasma*, *Lariskella*, *Midichloria*, *Rickettsia* and *Wolbachia*], while obligate tick symbionts (*i.e*., those present in most, if not all, individuals of a given species across both temporal and geographical scales) are identified in only five genera: *Coxiella* and *Rickettsiella* (order *Legionellales*), *Rickettsia* and *Ca.* Midichloria (order *Rickettsiales*), and *Francisella* (order *Thiotrichales*) (Duron et al. 2017). Co-cladogenesis has been documented between *Rhipicephalus* and their *Coxiella*-LE, which is remarkable (Duron et al. 2015). Co-evolution between ticks and their bacterial symbionts shows dynamic evidence of losses, gains and horizontal transfers between tick species, with the widespread *Coxiella*-like endosymbionts being substituted by members of the other four symbiont genera in multiple lineages (Duron et al. 2017).

Analysis of *I. ricinus* bacterial 16S rRNA sequences showed that besides the predominant *Ca.* Midichloria mitochondrii, it harbours many commensal bacteria

such as W*olbachia, Spiroplasma*, *Stenotrophomonas*, *Pseudomona*s, *Rhodococcus* and *Proprionibacterium* (Carpi et al. 2011). In addition, *I. ricinus* transmits the *B. burgdorferi* complex, *A. phagocytophilum* and *Rickettsia* spp. (spotted fever group) in continental Europe and *A. phagocytophilum* and *B. burgdorferi* complex in the UK (Carpi et al. 2011). This same study reported that the ecological analysis revealed that the bacterial community structure differed between the examined geographic regions and tick life stages. This finding suggests that the environmental context (abiotic and biotic factors) and host-selection behaviors affect the tick microbiome (Carpi et al. 2011).In cattle ticks *Rhipicephalus microplus* (collected from Zebu [*Bos indicus*] in Zimbabwe, Coxiella- LE was the predominant endosymbiont, while *Francisella* was detected at very low levels, and no *Midichloria* or *Rickettsia* were detected (Duron et al. 2015). Both *I. scapularis* and *I. pacificus* are predominant vectors of *B. burgdorferi* and *A. phagocytophilum* in the USA; however, *I. scapularis* harbours non-pathogenic *R. buchneri*, while *Ca.* Cryptoplasma californiense was detected in *I. pacificus* only (Eshoo et al. 2015).

There is evidence to suggest that symbiosis and interference between tick bacterial communities happens frequently within the internal tick environment. Cooperative or competitive interactions exist among these pathogen communities for space and resources, and this could impact the natural transmission of disease, which highlights the need to elucidate possible implications of these interactions and their potential impact on human and animal health (Vayssier-Taussat et al. 2015). Taking this into consideration, an integrated view of the pathobiome (*i.e.*, all bacteria within a tick, both pathogenic and commensal) rather than solely the identified pathogenic agents only, is crucial to elucidate vector competence and pathogen transmission.

From an applied perspective, tick symbionts are of interest for at least four reasons. First, the origin of some pathogens of humans (including *Coxiella burnetii,* the spotted-fever group rickettsiae and possibly *Francisella tularensis*) and other vertebrates can be traced to symbiotic bacteria that were originally restricted to arthropods (Duron et al. 2015; Scoles 2004). This is an ongoing process that can lead to the emergence of diseases, such as Q-fever and tularaemia, which no longer require ticks in order to be transmitted between vertebrate hosts. Second, symbionts might affect the transmission of related (*i.e*. other intracellular bacteria) or unrelated (*i.e*. viruses or helminths) pathogens by the vector, as has been demonstrated for *Wolbachia* infections in mosquitoes under certain conditions (Kambris et al. 2009; Walker et al. 2011). Third, symbionts have the potential to affect the reproductive fitness of their hosts, and this could be exploited for vector control as has been attempted with *Wolbachia* symbionts in mosquitoes (Kambris et al. 2009). Finally, symbionts might interfere with the diagnosis of infections caused by related bacteria, or more positively, could provide a means to identify biomarkers of tick exposure (Mariconti et al. 2012).

Endosymbionts can be classified into primary endosymbionts that are transovarially transmitted (Braendle et al. 2003) where the survival of both organisms are linked with the existence of them together in a obligate mutualistic relationship (Duron et al. 2017), and secondary endosymbionts that are not obligatory for tick survival (Moran et al. 2008). Symbionts might play an important role in adaptation of the tick to counteract environmental stress, or might confer a beneficial effect in the development of the bacteria or confer fitness advantage to their hosts in a facultative relationship (Moran et al. 2008).

Different tick species harbour a variety of different endosymbionts. For example, *Amblyomma americanum* has a *Coxiella*-LE in all of the different developmental stages and even in laboratory-reared specimens, suggesting a primary relationship where the endosymbiont is potentially providing essential nutrients for host fitness and survival, thereby proposing an obligatory relationship (Zhong et al. 2007). Recent studies have revealed that several pathways involved in vitamin B biosynthesis and their cofactors such as biotin (B7) and riboflavin (B2) are present within the *Coxiella* genome. As a blood-based diet has insufficient amounts of these vitamins, it may be the case that the tick and *Coxiella*-LE mutualistic relationship enables the tick to specialize on an unbalanced blood diet, with the endosymbiont providing some of the vitamins which are at low levels (Duron et al. 2017). Moreover, one study revealed that elimination of *Coxiella*–LE as a result of antibiotic treatment led to a decrease in production of tick progeny, further indicating the importance of this mutualistic relationship (Zhong et al. 2007).

Obligate mutualists that represent a strict co-cladogenesis can be also be found in aphids, leafhoppers and tsetse flies, generating congruent host–symbiont phylogenies suggesting co-evolution of the organisms (Chen et al. 1999). On the other hand, facultative symbionts can occasionally exhibit a horizontal transmission by co-feeding on the same animal and may confer a fitness benefit and/or protection against natural stress upon their tick carriers (Ahantarig et al. 2013). It has been documented that infection with a rickettsial endosymbiont considerably increased questing motility in the tick *Dermacentor variabilis* (Kagemann & Clay 2013). *Rickettsia peacockii* has been recorded as a symbiotic bacterium in some Rocky Mountain wood tick (*Dermacentor andersoni)* populations (Niebylski et al. 1997), while *I. ricinus* (Mediannikov et al. 2012), *D. andersoni* and *D. variabilis* were found

to harbour *Arsenophonus*-type bacteria (Grindle et al. 2003), which suggests a possible secondary symbiotic relationship. Furthermore, identification of *Ca.* M. mitochondrii*,* a vertically transmitted endosymbiont described from the sheep tick *I. ricinus,* in different ticks such as *Rhipicephalus sanguineus* and *Hyalomma* spp., suggests the possibility of horizontal transmission as well (Harrus et al. 2011). The impact of these bacteria on their tick host remains unknown but has been suggested to be a facultative symbiotic relationship (Lo et al. 2006).

Additionally, *Wolbachia* spp., which are widely distributed in arthropods, represent an example of a defensive endosymbiont in insects as they seem to protect their arthropod carrier against other microbial infections. Previous studies highlighted how the antiviral effect in *Wolbachia*-infected *Drosophila melanogaste*r functions to protect insects from different RNA viruses (Hedges et al. 2008). *Wolbachia* also plays an important role as a reproductive manipulator in *D. melanogaster* and many other arthropods by distorting the sex ratio of the offspring through cytoplasmic incapability (Hurst et al. 2002). Some studies proposed that infection of *A. americanum* and *D. variabilis* with endosymbiotic bacteria may influence host-seeking success by decreasing tick motility (Kagemann & Clay 2013).

1.6.2 Relationships between symbionts and pathogens

Disease processes may be affected by the interaction between pathogenic micro-organisms and their associated endosymbiont bacteria which may have an impact on vector competence. As a result, an increasing body of research has taken place investigating the pathobiome of ticks (Vayssier-Taussat et al. 2015).

Research highlighted the population dynamics of defensive symbionts in pea aphids. Aphids completely depend on vitamins and amino acids supplied by their

obligatory endosymbiont *Buchnera aphidicola* (Oliver et al. 2008). However, the facultative bacterial symbiont *Hamiltonella defensa* of the pea aphid *Acyrthosiphon pisum*, appears to provide a protection against the pathogenic wasp *Aphidius ervi*, suggesting that the level of resistance is determined by presence of the symbiont (Oliver et al. 2008; Oliver et al. 2005).

Another interesting example has been suggested in fruit flies: a potential rapid effect on defensive symbiosis could be highlighted by the adaptation of maternally-inherited *Spiroplasma* that occur in *Drosophila neotestacea*. Maternally-inherited *Spiroplasma* rescues females from the sterilizing effects of nematode parasitism. Recent spread of *Spiroplasma* the from east to west of North America highlights an alteration to a symbiont-based mode of protection against nematode parasites (Jaenike et al. 2013). This potential phenomenon would suggest new measures for nematode control. If the *Spiroplasma* was capable of preventing development of filarial nematodes in *Drosophila*, this would affect their vector competency (Jaenike et al. 2013). This suggests a potential defensive mechanism that would play a significant role in the ecology of species interaction.

In ticks, previous studies pointed out an interaction between *B. burgdorferi* and rickettsial endosymbionts in male *I. scapularis* ticks. The infection rate of *B. burgdorferi* in male *I. scapularis* ticks decreased considerably in ticks which were co-infected with a rickettsial endosymbiont when compared with uninfected males (Simser et al. 2001). In contrast, a *Francisella*-LE, a *D. andersoni* symbiont closely related to *Francisella tularensis* (the causative agent of the zoonotic disease tularemia), does not appear to have excluded infection of *D. andersoni* with either a second closely-related proteobacterial symbiont or with a second less closely-related proteobacterial symbiont (Scoles 2004). However, the presence of vertically-

transmitted *Amblyomma*-associated *Coxiella*-LE in tick salivary glands seems to interfere with maintenance or transmission of the causative agent of human monocytic ehrlichiosis (*E. chaffeensis*) (Klyachko et al. 2007).

Another suggested potential mode of interaction is intra-family_bacterial cross-immunity in tick-endosymbiont populations. The presence of some *Rickettsia* endosymbionts leads to an alteration of the capability of other *Rickettsia* spp. to be transmitted transovarially, which might be due to intra-family bacterial cross-immunity and suggests that the bacterial microbiome might aid tick defence against pathogens. A reverse relationship between the non-pathogenic *R. peacockii* (restricted to the ovarian and Malpighian tubule tissues of the tick) and pathogenic *Rickettsia rickettsii* has been reported in the Rocky Mountain wood tick *D. andersoni* (Simser et al. 2001). This suggests that a high prevalence of *R. peacockii* in the tick ovaries may impair transovarial transmission of the virulent *R. rickettsia* to tick progeny, ultimately suggesting an exclusion mechanism between Rickettsiales (Simser et al. 2001). The authors suggested that the molecular expression of the oocyte may be influenced by the presence of some *Rickettsia* spp. which leads to an exclusion of a secondary infection with further *Rickettsiae* bacteria (Ahantarig et al. 2013).

Interestingly, the possibility for biocontrol by decreasing pathogen susceptibility has been suggested (Gall et al. 2016). Microbiome alteration by antibiotic treatment in *D. andersoni* ticks followed by exposure to pathogens elucidated an endosymbiont-pathogen reverse relationship (Gall et al. 2016). A higher density and quality of *Rickettsia bellii* in the microbiome inhibited *Anaplasma marginale* load; this could have been due either to competition for resources, or as a

result of specific molecules, such as a nutrient, that *Rickettsia bellii* released that would inhibit the susceptibility of ticks to *A. marginale* (Gall et al. 2016).

On the other hand, a positive interaction was seen between increasing *Francisella* endosymbionts and the infection level with the pathogen *Francisella novicida*; possibly a higher load of *Francisella* endosymbionts inhibits the immune response, and promotes *F. novicida* infection (Gall et al. 2016). A negative side of symbiosis could also be highlighted in some species of mosquitoes between the endosymbiont *Wolbachia* and the pathogen *Plasmodium relictum* (Zele et al. 2014). *Wolbachia* increases the susceptibility of *Culex pipiens* mosquitoes to *P. relictum,* considerably increasing the sporozoite load in the salivary glands, suggesting that the capability of mosquitoes naturally infected with *Wolbachia* to transmit avian malaria is greater compared with *Wolbachia*-free individuals (Zele et al. 2014). Moreover, in spite of the protective role of some *Wolbachia* spp. in *Drosophila melanogaster* against different RNA viruses (Hedges et al. 2008), *Wolbachia* make the African armyworm host more susceptible to baculovirus infection, historically referred to as 'wilting disease' (Graham et al. 2012). Finally, activation of the Toll and Imd immune pathways in *D. melanogaster* has been shown to increase the titer of *Spiroplasma* endosymbionts, which results in an increased host susceptibility to certain pathogens, including *Erwinia carotovora carotovora* 15 and *Enterobacter cloacae* (Herren & Lemaitre 2011).

These mechanisms suggest that in spite of the protective effect of some endosymbiotic bacteria in insects that could defend against natural enemies, refractoriness to other pathological agents would suggest a negative impact and might increase susceptibility of *Drosophila* to certain Gram-negative pathogens (Hedges et al. 2008).

Finally, it is important to highlight the recent growing interest in emerging or re-emerging tick-borne pathogens that originated as symbiotic bacteria. It has been suggested that *Coxiella burnetii*, the causative agent of acute Q fever and chronic endocarditis in humans, was derived from a *Coxiella*-LE (Weller et al. 1998), while other studies suggested that *C. burnetii* in an infected vertebrate was the main origin of the *Coxiella*–LE in ticks (Noda et al. 1997). It is now known that tick symbionts were the evolutionary source of the vertebrate pathogen (Duron et al. 2015). In addition, a lack of many hypothetical proteins (proteins unrecognised function) associated with DNA repair (Jasinskas et al. 2007) in the *Coxiella*-LE bacterial genome when compared with the *C. burnetii* genome may elucidate the reason for the pathogenic impact of C. *burnetii,* in contrast with the closely related, but apathogenic *Coxiella*–LE (Ahantarig et al. 2013).

This suggests a potential mechanism that would allow endosymbiotic organisms to re-emerge as pathogens, that the interface between different endosymbionts and pathogens might be widely variable, and that in future, microbiome manipulation may provide a possible method for biocontrol by decreasing pathogen susceptibility of ticks. Therefore, elucidation of the existence and role of endosymbionts in the diverse tick communities is crucial to identify their possible roles in pathogen transmission. Indeed, this could eventually promote new approaches to the control of ticks and the serious tick-borne diseases that they carry by targeting their endosymbionts.

1.7 Overall aims and objectives of the research reported in this thesis

The overall aim of this thesis is to further our understanding of how Rickettsiales bacteria (whether they are pathogenic to vertebrates or not) interact with their tick hosts at both the epidemiological and molecular level. The thesis focuses on several tick species in which both pathogenic Rickettsiales and symbiotic Rickettisales (which are not known to cause disease in animals, including man) coexist. Each chapter has a specific objective:

Chapter 2. To quantify co-infecting pathogenic and symbiotic Rickettsiales in two different tick species in two countries for the first time as baseline data for epidemiological studies on interactions between them.

Chapter 3. To define the population structure and potential co-evolutionary relationship between *I. ricinus* and its symbiont *Ca.* Midichloria mitochondrii in different parts of the UK and continental Europe through the application of multi-locus sequence typing.

Chapter 4. To describe the molecular interactions of a British ovine isolate of *A. phagocytophilum* with *Ixodes* cells *in vitro* for the first time.

Chapter 5. To determine if a symbiotic species of Rickettsiales is present in an anatomical location in the tick that could expose it to the vertebrate host's immune system.

CHAPTER TWO

Quantification of co-infecting pathogenic and symbiotic Rickettsiales in two different tick species

2.1 Introduction

2.1.1 Prevalence of the symbiont *Candidatus* Midichloria mitochondrii and pathogenic Rickettsiales in *Ixodes ricinus* from different parts of the UK

Ixodes ricinus, the sheep tick, is the vector of many causative agents of human and animal disease, including *A. phagocytophilum* and the *B. burgdorferi* complex, and is widely distributed across Europe, causing serious health, welfare and economic impacts. Although ticks carry and can transmit these dangerous pathogens, they possess a defence mechanism to protect themselves against them, and in order for a tick to act as a vector, the pathogen must be able to overcome and evade the tick's immune system (Hajdušek et al. 2013). Furthermore, different tick species carry different symbiotic bacteria that might influence tick biology and their competence to transmit different pathogens (Moutailler et al. 2016).

In addition to its role as a vector of multiple pathogens, *I. ricinus* also harbours the bacterium *Ca.* M. mitochondrii as an endosymbiont. Whether infection with *Ca*. M. mitochondrii has any beneficial or detrimental impacts on *I. ricinus* is still unknown, but it has been proposed that this symbiotic bacterium is a facultative mutualist, as it has been demonstrated in 100% of wild female ticks with no obvious detrimental effect (Rizzoli et al. 2014).

High pathogen co-infection rates in adult *I. ricinus* have been demonstrated recently, which highlights the possibility of co-transmission of various pathogens to humans and livestock, and taking into consideration that tick endosymbionts may

interfere with pathogen transmission (Moutailler et al. 2016), it is possible that they may also have an impact on disease severity, diagnosis and treatment efficacy.

2.1.1.1 *Candidatus* Midichloria mitochondrii, a unique endosymbiont of ticks

Rickettsia-like bacteria in the ovaries of *I. ricinus* were described for the first time by Lewis (Lewis 1979). However, it took many years for the unique life cycle of this symbiont to be characterised, most distinctively its ability to invade and destroy the mitochondria of ovarian host cells (Lo et al. 2006). This symbiont of *I. ricinus* was identified by electron microscopy in the mitochondria of ovarian cells of adult female ticks, as well as in the primordial ovarian tissue of female larvae and nymphs (Zhu et al. 1992). Moreover, ultrastructural studies employing *in situ* hybridization suggested that they enter the mitochondria in a similar way to the bacterium *Bdellovibrio bacteriovorus;* that is, between the inner and outer membrane of mitochondria (Sacchi et al. 2004). The first description of *Ca.* M. mitochondrii was in 2006 (Sassera et al. 2006). Prior to that, the identity of this symbiont in *I. ricinus* was not known, even though some earlier microscopic and molecular studies had been published. The research revealed that this bacterium was detected in different ovarian cell types including luminal cells, funicular cells and oocytes (Beninati et al. 2004; Ninio et al. 2015).

For *I. ricinus*, the presence of *Ca.* M. Midichloria mitochondrii has been shown in 100% of blood-feeding females, and in 44% of males across 12 countries: Sweden, Russia, Ireland, England, Germany, Czech Republic, Austria, Switzerland, Italy, Turkey, Algeria and Tunisia (Sassera et al. 2006; Lo et al. 2006). The same study also highlighted that male ticks harboured relatively lower concentrations of the bacteria compared to the females of the species and eggs had 100% prevalence of the

bacteria. Ovarian tissue has been found to have the highest concentrations of *Ca.* M. mitochondrii*,* supporting the findings of previous studies suggesting that vertical transfer of *Ca.* M. mitochondrii is a major factor for survival of the bacterium (Sassera et al. 2006).

Other reports have demonstrated the diverse host range of *Ca.* Midichloria by detecting its presence in numerous different tick species. This has brought to light the suggestion of horizontal transmission and infection of ticks from feeding on the blood of an infected host, but possible mechanisms for this are currently unknown (Epis et al. 2008). Further study confirmed the distribution of the bacterium in *I. ricinus* in different stages of the tick's life cycle using fluorescence in situ hybridization (FISH), and it was proposed that *Ca*. M. mitochondrii is important in the biology and survival of the tick (Epis et al. 2013). In addition, a recent paper suggested that certain species of ticks with a particularly high prevalence of *Ca.* Midichloria could have a stronger physiological link with the bacterium (Cafiso et al. 2016). Furthermore, a review paper highlighted that many endosymbionts in haematophagous insects are important for the acquisition of vitamins that are scarce in the blood meal (Rio et al. 2016), and this might also apply to ticks. Alteration of this mechanism resulting in starvation of the vector would reduce the transmission of a variety of infectious vector-borne diseases (Rio et al. 2016).

It has been confirmed that, for *Ca.* M. mitochondrii*,* the lifecycle or feeding stage of *I. ricinus* does not affect the presence of the bacterium, as even unfed larval samples tested positive for *Ca.* M. mitochondrii (Granquist et al. 2014). Due to the different prevalence of *Ca.* Midichloria in different species of ticks, it was proposed that the bacterium may not have as significant a role in the physiology and interactions of certain tick species as it perhaps does in *I. ricinus,* which as stated, has

a very high infection level in females (Granquist et al. 2014). *Ca.* Midichloria has been detected in different *Ixodes* species including *Ixodes colasbelcouri, Ixodes frontalis, Ixodes ventalloi* and *Ixodes uriae;* it has also been detected in other Ixodidae: *Amblyomma* spp., *Rhipicephalus* spp., *Dermacentor* spp. (Epis et al. 2008), *Hyalomma anatolicum* and *Hyalomma truncatum* (Epis et al. 2008). *Ca.* Midichloria was also present in all tested Nigerian *Ixodes aulacodi* ticks (Cafiso et al. 2016). Furthermore, *Ca.* Midichloria was detected in a French population of the soft tick *Ornithodoros maritimus* (Cafiso et al. 2016).

With respect to effects on the host, studies confirmed that *Ca.* M. mitochondrii DNA and antibodies against the recombinant flagellar protein FliD (Mariconti et al. 2012) can be detected in the blood of animals, *e.g*. dogs and horses (Sassera et al. 2008; Bazzocchi et al. 2013), and also in humans following exposure to ticks. Hosts tested seropositive for these bacteria, which suggests *Ca.* M. mitochondrii could act as a tick-bite biomarker for tick immunology investigations, and it is also possible that it might cause pathological changes in a host species (Bazzocchi et al. 2013). Researchers reported that *Ca.* M. mitochondrii has been lost in colonies of *I. ricinus* maintained in the laboratory. However, tick colonies continued to survive and reproduce, suggesting that *Ca.* M. mitochondrii is lost due to the stable laboratory conditions (such as a constant temperature), or it may be affected by a number of antibiotics that have been used in laboratory animals (Sassera et al. 2008).

2.1.1.2 Pathogens transmitted by *Ixodes ricinus*

Ixodes ricinus co-infected with *B. burgdorferi* and emerging pathogens such as *A. phagocytophilum, Candidatus* Neoehrlichia mikurensis and *Rickettsia helvetica* have been found in urban tick populations in Europe, for example in Austria (Rizzoli

et al. 2014; Glatz et al. 2014). Other studies reported the co-occurrence in *I. ricinus* of *A. phagocytophilum* and *R. helvetica* in Sweden (Severinsson et al. 2010) and Germany (Tappe & Strube 2013), and of *A. phagocytophilum* and *Ca.* Neoehrlichia mikurensis in Central Europe, suggesting a high prevalence of co-infections (Derdáková et al. 2014). These latter three bacteria, the subject of our study, are all non-spore forming, obligatory intracellular organisms that belong to the Rickettsiales and are transmitted exclusively by ticks.

2.1.1.2.1 *Anaplasma phagocytophilum*

Anaplasma phagcocytophilum, the causative agent of pasture fever in cattle and tick-borne fever in sheep, was recorded in Scottish sheep for the first time in 1940 (cited in Woldehiwet et al. 2002; Woldehiwet & & Scott 1993). Later, *A. phagcocytophilum* was identified in northern and southwest England, Wales and Ireland, as well as in Scandinavia and other European countries (Hudson 1950; Stuen et al. 2005; Woldehiwet 2006; Zintl et al. 2017). *Anaplasma phagocytophilum* infects a wide range of mammals and different species of tick vectors (Bown et al. 2009). In addition to *I. ricinus* and *I. scapularis* (Ladbury et al. 2008) , *A. phagocytophilum* was detected in other tick species such as *D. marginatus* and *I. persulcatus* ticks from northern China (Cao et al 2003) and *Ixodes trianguliceps* in northwest England (Bown et al. 2003). It is worthy of note that a higher infection rate with *A. phagocytophilum* was detected in adult ticks compared to nymphs. This is because transovarial transmission of this bacterium has not yet been detected, so adults have a double chance of receiving *A. phagocytophilum* via the blood meal (reviewed in Stuen et al. 2013).

Identification of *A. phagocytophilum* in the fast-feeding phase of *I. trianguliceps* (a nidicolous tick that is specialized to feed on small mammals) and in the blood of their hosts including wood mice and bank voles suggests that European wild rodents are competent reservoirs (Ogden et al. 1998). Roe deer (*Capreolus capreolus*) are considered an important natural mammalian reservoir of *A. phagocytophilum* highlighting their epidemiological role in maintenance of the pathogen (Alberdi et al. 2000). Then following studies highlighted an interesting finding; while one *Anaplasma* genotype was detected in infected roe deer and *I. ricinus* vectors in northern England, a different *Anaplasma* genotype was detected in the vole species that serve as a reservoir host (Bown et al. 2009). The fact of *A. phagocytophilum* strain variability was suggested as well when 16S rDNA variants of *A. phagocytophilum* were found to infect Norwegian red deer (*Cervus elaphus*) (Stuen et al. 2013).

Another species of *Anaplasma*, *Anaplasma ovis* (Dumler et al. 2001), infects European sheep, goats and red deer; these mammals might play role in the epidemiology of this *Anaplasma* species as reservoir hosts (Hornok et al. 2011). Also, *Anaplasma mesaeterum* was detected in sheep blood in Northern Europe (Uilenberg et al. 1979). While thousands of HGA disease cases have been confirmed in the USA (some fatal), it is rare to find symptomatic human infections in Europe. Moreover, although some serological evidence suggests human exposure in the UK, only two human cases have been documented (Sumption et al. 1995; Hagedorn et al. 2014). This suggests that different biological characteristics of strains of *A. phagocytophilum* may have an impact on their pathogenicity and host-specificity, as well as their ability to cause disease, and that there may be notable genetic diversity between European strains of *Anaplasma* (Woldehiwet 2006).

2.1.1.2.2 *Candidatus* Neoehrlichia mikurensis

Candidatus Neoehrlichia mikurensis causes an emerging vector-borne zoonosis across the European landmass, particularly Western Europe (Lommano et al. 2012; Jahfari et al. 2012; Richter & Matuschka 2012). This bacterium was identified for the first time in 1999 in *I. ricinus* ticks from the Netherlands ('Schotti variant') (Schouls et al. 1999). Phylogenetic analysis of *Candidatus* Neoehrlichia mikurensis isolated from wild rats and found in *Ixodes ovatus* ticks in Japan, revealed a close relation to *Ehrlichia*-like microorganisms (Kawahara 2004). The prevalence and distribution of *Ca.* Neoehrlichia mikurensis in Europe, its natural life cycle and reservoir hosts remain unclear. However, this bacterium was found to be distributed across the Netherlands and Belgium (Jahfari et al. 2012), and more recent studies showed the presence of this bacterium in Hungary (Hornok et al. 2013). In addition, it has been detected in questing *I. ricinus* ticks from Austria (Glatz et al. 2014), and in *I . ricinus* from Sweden (Andersson et al. 2013), where they have had a number of human cases of infection (Wenneras 2015; Pekova et al. 2011).

Analysis of 16S rDNA sequences, citrate synthase (*gltA*) and the heat shock protein (*groEL*) genes has been used to identify *Ca.* Neoehrlichia mikurensis in other tick species (Kawahara 2004); for instance it was identified in *I. persulcatus* and *I. ovatus* from Japan (cited in Yabsley et al. 2008). It has previously been known as *Candidatus* Ehrlichia walkerii in Italy (Brouqui et al. 2003). Different rodent reservoirs have been reported across the different countries where infected ticks have been found (Jahfari et al. 2012). Although *Ca.* Neoehrlichia mikurensis was reported to cause a febrile illness for the first time in humans in 2010, but it is currently uncertain whether *Ca.* Neoehrlichia mikurensis poses a significant risk to public health (Welinder-Olsson et al. 2010; Richter & Matuschka 2012). Molecular analysis

confirmed a positive association between the appearance of *Ca.* Neoehrlichia mikurensis and *A. phagocytophilum*, as both of these bacteria share a similar ecology in Central Europe across a wide ecological spectrum of habitats in Slovakia, the Czech Republic and Austria (Derdáková et al. 2014; Glatz et al. 2014).

2.1.1.2.3 Tick-borne Rickettsioses

Phylogenetic analysis revealed that *Rickettsia* species are classified into nine groups (Pilgrim et al. 2017). There are multiple different species of *Rickettsia* which vary in their dispersal, ecology and pathogenicity (Eremeeva et al. 2006). *Rickettsia* can be symbionts in a variety of non-haematophagous hosts, and ticks are infected by rickettsial symbionts and pathogens from multiple clades; also, there are *Rickettsia* transmitted by other arthropods besides ticks, including fleas and mites. Thus, *Rickettsia akari* is transmitted by mites, while Typhus group rickettsiae including *Rickettsia typhi* and *Rickettsia prowazekii* are transmitted by fleas and lice, respectively (Uchiyama 2012). In Europe, many different *Rickettsia* spp. have been documented (Portillo et al. 2015). (**Table 2. 1**)

Rickettsia can be transmitted vertically to offspring or horizontally after ingestion of an infected blood meal by ticks. Moreover, *Rickettsia* can be transferred to vertebrates through salivary secretions of ticks, and are often pathogenic to humans; however *Rickettsia* transmitted by other arthropod groups tend to be excreted in arthropod faeces and scratched into bite wounds (Scarpulla et al. 2016). Major rickettsial pathogens include the agents of Mediterranean spotted fever (MSF) and Rocky Mountain spotted fever, caused respectively by *Rickettsia conorii* (transmitted by the brown dog tick, *Rhipicephalus sanguineus*) and *Rickettsia rickettsii* (transmitted mainly by *Dermacentor* spp.), which are human pathogens and

cause eruptive fevers due to their ability to replicate in endothelial cells. *Rickettsia prowazekii*, the causative agent of epidemic typhus, is transmitted by the human body louse (Fournier et al. 1998; Portillo et al. 2015). In France, Switzerland and Italy, *Rickettsia helvetica* was identified in *I. ricinus* ticks, causing MSF-like symptoms (Beninati et al. 2002). This pathogen was also detected in *Ixodes acuminatus* (Beninati et al. 2002). Furthermore, *R. helvetica* has been described in cases of sudden death with clinical signs of meningitis and perimyocarditis in Sweden only (Portillo et al. 2015).

Two cases of infection with *Rickettsia monacensis,* which is transmitted by *I. ricinus* and causes a MSF-like illness in humans, have been reported in Spain (Portillo et al. 2015). *Rickettsia massiliae* has also been identified in European ticks (Oteo & Portillo 2012). Some rickettsial species such as *Rickettsia slovaca*, *R. helvetica* and *Rickettsia aeschlimannii* were originally reported as non-pathogenic bacteria, but recent reports suggest that they could be pathogenic (Tijsse-Klasen et al. 2013). In addition to pathogenic *Rickettsia*, non-pathogenic *Rickettsia* in tick vectors including *R. buchneri* in *I. scapularis* (Kurtti et al. 2016) and *R. peacockii* in *D. variabilis* (Felsheim et al. 2009), suggested the possibility of *Rickettsia* spp. interactions which might affect vector-borne pathogen transmission (Bonnet et al. 2017).

Table 2. 1 Classification and vectors of *Rickettsia* species that are present in Europe. Modified from (Portillo et al. 2015).

Vector(s)	*Rickettsia* species	Disease name
Unspecified ixodid tick	***Ca.*** Rickettsia kotlanii	None (identified in hard ticks only)
Ixodes ricinus	***Ca.*** Rickettsia vini	
Ixodes arboricola		None (identified in hard ticks only)
Ixodes ricinus	*Rickettsia helvetica*	
Ixodes ricinus	*Rickettsia monacensis*	Not named yet
Ixodes lividus	*Rickettsia* sp. strain Davousti	None (identified in ticks only)
Ixodes persulcatus	***Ca.*** Rickettsia tarasevichiae	Not named yet
Rhipicephalus bursa	***Ca.*** Rickettsia barbariae (Rickettsia PoTiRb 169)	None (identified in ticks only)
	Rickettsia	None (identified in hard ticks only)
Rhipicephalus turanicus	***Ca.*** Rickettsia siciliensis	Mediterranean spotted fever
Rhipicephalus sanguineus	*Rickettsia conorii* subsp. *conorii*	Indian tick typhus
	Rickettsia conorii subsp. *indica*	
	Rickettsia conorii subsp. *caspia*	Astrakhan fever
	Rickettsia conorii subsp. *israelensis*	Israeli tick fever
Rhipicephalus sanguineus, *Rhipicephalus turanicus*	*Rickettsia massiliae*	Not named yet

Dermacentor marginatus	*Ca.* Rickettsia rioja *Rickettsia raoultii* *Rickettsia slovaca*	Tick-borne lymphadenopathy/*Dermacentor*-borne necrosis erythema lymphadenopathy
Ixodes ricinus, *Rhipicephalus turanicus*, *Rhipicephalus sanguineus*, *Rhipicephalus bursa*, *Hyalomma marginatum*, *Hyalomma excavatum*, *Hyalomma rufipes*, *Haemaphysalis punctata*, *Haemaphysalis inermis*	*Rickettsia aeschlimannii*	Not named yet
Amblyomma variegatum, *Amblyomma hebraeum*	*Rickettsia africae*	African tick bite fever
Hyalomma anatolicum, *Hyalomma marginatum*,	*Rickettsia sibirica* subsp. *Mongolitimonae*	Lymphangitis-associated rickettsiosis
Haemaphysalis punctata, *Haemaphysalis sulcate*	*Rickettsia hoogstraalii*	None (identified in ticks only)
Ornithodoros erraticus	*Rickettsia lusitaniae* sp. nov	None (identified in soft ticks only)
Liponyssoides sanguineus (mite)	*Rickettsia akari*	Rickettsialpox
Pediculus humanus corporis (louse)	*Rickettsia prowazekii*	Epidemic typhus
Ctenocephalides felis (flea)	*Rickettsia felis*	Flea-borne spotted fever
Xenopsylla cheopis, *C. felis* (fleas)	*Rickettsia typhi*	Murine or endemic typhus

2.1.2 *Rickettsia africae* and *Ehrlichia ruminantium* in *Amblyomma variegatum*

Amblyomma variegatum, the tropical bont tick, poses a major threat to human and livestock health across its geographical distribution in sub-Saharan Africa. Following cattle transport and bird migration, this tick is also prevalent in some islands of the Caribbean. *A. variegatum* is considered a vector and reservoir of the SFG pathogen *R. africae*, the agent of African tick-bite fever (ATBF). The disease is characterized by mild to severe clinical signs, particularly for elderly people, including headache and multiple eschars and more complicated symptoms of myocarditis and subacute neuropathy (Robinson et al. 2009; Kelly et al. 2010). Acording to previos results, around 16 – 75% of *A. variegatum* along the coastal region of Cameroon are infected with *R. africae* (Ndip et al. 2004). However, ATBF disease is quite common across Africa and some Caribbean islands, and has been reported in passengers or tourists returning to Europe from Africa (Ndip et al. 2004).

Furthermore, heartwater, a major disease of ruminants (domestic ruminants and wild animals) caused by *E. ruminantium*, is transmitted by *A. variegatum* and *A. hebraeum* (Robinson et al. 2009; Dumler et al. 2001). Diverse strains of *E. ruminantium* have been reported in sub-Saharan Africa (Esemu et al. 2011). While a commercial vaccine against heartwater had been used for many years in South Africa, the main control strategy elsewhere is characterised by hand-picking of ticks and use of acaricides (Ndi et al. 1998). In spite of the fact that heartwater disease poses a substantial economic threat because it causes significant losses of domestic ruminants in Africa south of the Sahara (Robinson et al. 2009; Steyn et al. 2008; Dumler et al. 2001), and has also expanded into several islands in the Carribean

(Robinson et al. 2009), little is known about *E. ruminantium* infection in the tick vector, *A. variegatum*.

2.1.3 The aims of the work described in this chapter

The aims of the work undertaken in this chapter were to assess, for the first time, levels of *Ca.* M. mitochondrii in different feeding stages and both sexes in British specimens of *I. ricinus* and to investigate further the prevalence of three pathogenic members of the Rickettsiales: *A. phagocytophilum*, *Ca.* Neoehrlichia mikurensis and *Rickettsia* spp. Levels of symbiotic bacteria were calculated using quantitative polymerase chain reaction (qPCR), and were then compared in different feeding stages and sexes of *I. ricinus* ticks. The rationale for this study was to determine if there is a negative or positive correlation between the pathogens listed above and the symbiotic bacterial load. The second aim was to examine interactions between *R. africae* and *E. ruminantium* by assessing their levels in *A. variegatum* using quantitative real-time PCR assays and to determine if there is a correlation between them due to displacement or facilitation.

2.2 Materials and Methods

2.2.1 *Ixodes ricinus* sampling

Adult and nymphal ticks in various states of engorgement were removed from three freshly-culled fallow deer (*Dama dama*) belonging to the Powis Castle estate, Powys, Wales (**Table 2. 2**). Questing nymphs were collected by Public Health England, using the flagging method, from five field sites in southern England comprising Porton Down, Salisbury (urban habitat), Bentley Wood (Wiltshire), Dartmoor, Exmoor, and the New Forest (**Figure 2. 1**). Specimens were stored in 70% ethanol at 4°C. Tick species were identified using a standard key (Hillyard 1996). Any damaged ticks were

discarded. In order to confirm species identification, PCR targeting mitochondrial 16S rRNA was applied in order differentiate *I. ricinus* from *Ixodes inopinatus* (a recently described, closely related species) according to the method of Black & Piesman 1994.

Nymphs are similar in appearance to an unfed female but much smaller and their sex cannot be determined at this stage. Larvae are smaller still and only have six legs. For different levels of engorgement, a slow-feeding stage female is flat, a fast-feeding stage female is much larger and round, and an intermediate-feeding stage female is between these sizes. Ticks were placed into six categories and quantified (**Table 2. 2**).

Figure 2. 1 Location of tick collection sites in the UK.
Each triangle represents the location sites of this study. Powis Castle (po); Exmoor (ex); Salisbury (sa); Bentley Wood (be); New Forest (ne); Dartmoor (da).

Table 2. 2 British *Ixodes ricinus* specimens used for the estimation of *Ca.* Midichloria density.

Stage	NO. of ticks assayed	Field sites	County	Country
Fast-feeding stage	20	Powis Castle	Powys	Wales
Intermediate-feeding stage	25			
Slow-feeding stage	14			
Male	48			
Nymph	44			
	100	Salisbury	Wiltshire	England
	100	Bentley Wood	Wiltshire	
	100	New Forest	Hampshire/ Wiltshire	
	100	Exmoor	Somerset/Devon	
	100	Dartmoor	Devon	

2.2.2 DNA extraction and quantification

Ticks were rinsed with distilled water three times, then dried on clean filter paper to remove any ethanol, which could inhibit the PCR. The ticks were individually placed on a Petri dish under a stereo microscope and bisected from the palps to the anal groove using a disposable scalpel blade. Quadrisection was used for larger ticks [fast-feeding stage (ffs), and intermediate-feeding stage (ifs)]; **Table 2.2**) and bisection for the smaller ticks [slow-feeding stage (sfs); male and nymph; **Table 2.2**]. All parts of each individual tick were placed together into a separate sample tube and labelled with a unique number and the appropriate category. A new scalpel blade and new Petri dish were used for each tick to reduce the risk of cross-contamination of the DNA.

A variety of methods were tested for optimal physical disruption of the hard, chitinous tick exoskeleton, alongside a series of commercial and non-commercial DNA

isolation protocols (**Table 2. 3**) in order to obtain the highest yield of DNA (Ammazzalorso et al. 2015).

Table 2. 3 Methods of DNA extraction and physical disruption of tick samples.

DNA extraction method	**Alterations to manufacturer's protocols**	**Physical disruption**	**Speed and duration**
Thermo GeneJET Genomic DNA Purification Kit (cat. no. K0722)	Overnight incubation at 56 ◦C in lysis buffer/proteinase K. Elution in 100 µl dsH_20.	Bisection (nymphs).	5,000 rpm for 4 min
MoBio PowerSoil DNA Isolation Kit (cat. no. 12888)	Elution in 100 µl of dsH_20 with 5 min room temperature incubation.	Bisection (nymphs), followed by vortexing with MoBio garnet beads.	3,200 rpm for 10 min
NH_4OH (Guy & Stanek, 1991; Pichon et al., 2003).	Initial volume of 150 µl NH_4OH. Final volume of 70 – 100 µl.	Bisection (nymphs). Quadrisection (adults).	N/A

DNA extraction and processing were evaluated with the Thermo GeneJET Genomic DNA Purification Kit (Thermo Scientific, Waltham, Massachusetts, USA); MoBio PowerSoil DNA Isolation Kit (MoBio, Carlsbad, California, USA) and a simple ammonium hydroxide (NH_4OH) boiling method. For the ammonium hydroxide method, 150 µl of ammonium hydroxide (14.5 M) was added into a 1.5-mL snap cap tube containing an individual bisected tick, then the samples were boiled at 100°C for 20 min using a heating block. The lids of the tubes were then opened at 100°C for 18 min inside a fume hood to evaporate the ammonium hydroxide to around 70 - 100 µl, then centrifuged for 10 min at 10,000 × *g.* The supernatant was collected and stored at 4°C. In order to increase DNA concentration in nymphal samples only, samples were centrifuged at 5,000 x *g* in a Nanosep centrifugal concentrator device (Nanosep®

Centrifugal Device, Pall life Sciences, USA) (30 kDa), retaining 30 µL of supernatant. This was collected and returned to the original tube and stored at 4℃. All DNA samples were quantified using a fluorescence labelling method, the Quant-iT Picogreen dsDNA kit (Invitrogen), and read in an Infinite F200 microplate fluorimeter with Magellan-Data Analysis Software (TECAN). The DNA concentration was calculated by comparison to a 0 – 1,000 ng/mL Lambda DNA standard curve.

2.2.3 *Ca.* M. mitochondrii qPCR

A previously-published SYBR Green-based qPCR targeting the *gyrB* gene was used to detect *Ca.* M. mitochondrii in tick lysates. The method was modified by using a long oligonucleotide (146 bases; synthesised by Eurogentec) to represent the standard amplicon sequence instead of quantification using cloned standards) (Sassera et al. 2008).

PCR cycling conditions were as follows: Initial denaturation at 95°C for 2 min; then 40 cycles of denaturation at 95°C for 15 sec, and annealing and extension at 60°C for 30 sec each. Following amplification, a melt-curve from 55°C to 95°C with increasing increments of 0.5°C per cycle was used to confirm the specificity of the products. All reactions were performed in 20 µl reaction volumes, comprising final concentrations of 1x SensiMix SYBR (Bioline), 400 nM of each primer and 1 µl of DNA template. The qPCR reactions were performed using a CFB-3220 DNA Engine Opticon 2 System. Quantitative data analysis was conducted in MJ Opticon Monitor Analysis software version 3.1. Although the number of genome copies per cell for *Ca.* M. mitochondrii has never been determined, we assumed that the copy number of *gyrB* (a single-copy gene) would be roughly equivalent to bacterial numbers or to at least correlate with them as previously described (Sassera et al. 2008). The synthetic

standards were diluted from $5 \times 10^6 - 5 \times 10^{-1}$ copies/µl in 100 ng/µl yeast tRNA (Invitrogen) to prevent aggregation and were included in duplicate (alongside nuclease-free water as no-template controls) in every run (**Figure 2. 2**) (**Table 2. 4**).

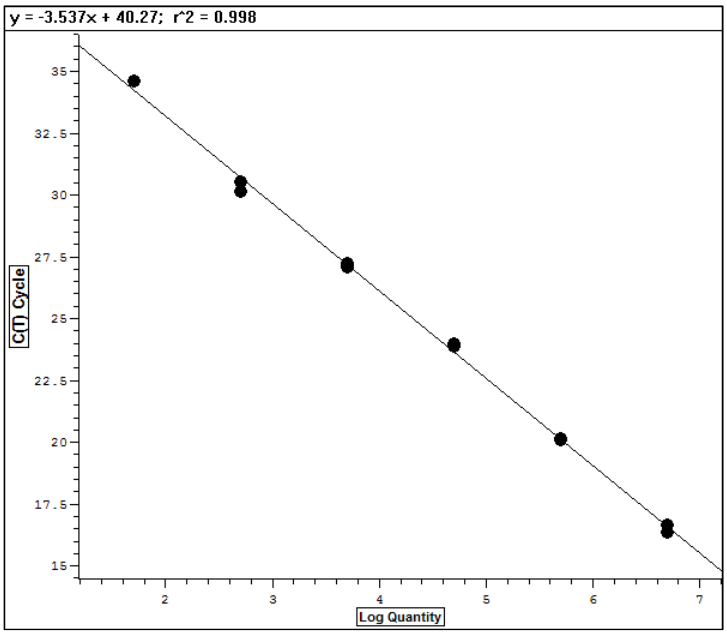


Figure 2. 2 Standard graph of qPCR targeting the *gyrB* gene to calculate *Ca.* M. mitochondrii density in ticks.

To normalise numbers of symbiont copies between different ticks stages, a new assay was designed which targeted a tick single-copy nuclear gene. This amplified a 77-bp fragment from exon 2 of the ribosomal protein L6 gene (*rpl6*), based on the sequence from *I. scapularis* (NCBI Reference Sequence: XP_002400555.1). All reactions were performed in 20 µl reaction volumes, comprising final concentrations of 1x SensiMix SYBR (Bioline), 200 nM of each primer (**Table 2. 4**), and 1 µl of DNA template. PCR cycling conditions were as follows: Initial denaturation at 95°C for 10 min, then 35 cycles of denaturation at 95°C for 15 sec, annealing at 55°C for 30 sec, and extension at 72°C for 15 sec; followed by a melt-curve from 55°C to 95°C with increasing increments of 0.5°C per cycle. Quantification was performed using a long

oligonucleotide (synthesised by Sigma-Aldrich) representing the full-length amplicon as described above for *gyrB* (**Figure 2. 3**).

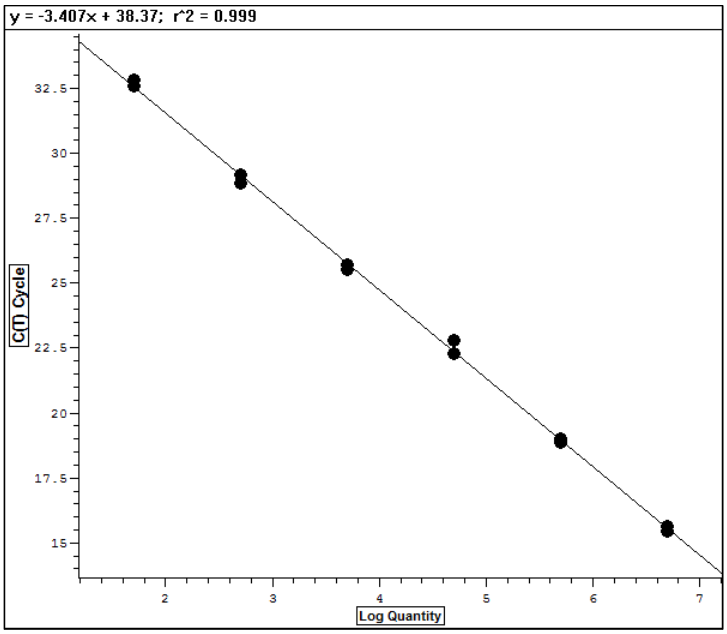


Figure 2. 3 qPCR targeting the tick ribosomal protein L6 nuclear gene for normalising bacterial density in ticks.

Symbiont density was assessed by serial dilution of samples from each life stage subset. The females (F) of each engorgement level (ffs, ifs and sfs) and adult male (M) tick samples to be tested for *Ca.* M. mitochondrii were all diluted with nuclease-free water (Ambion), with F(ffs) and F(ifs) ticks diluted to 1/100 and 1/1,000, and F(sfs), nymphs and M diluted to 1/10 and 1/100, to reduce the concentration of PCR inhibitors from blood in ffs specimens and/or to reduce a generally high concentration of tick DNA.

Table 2. 4 Gene targets used for amplification and sequencing.			
Gene target	**Detection**	**Primer or probe sequence**	**Reference**
gyrB (forward)	*Ca*. M. mitochondrii	5’ CTTGAGAGCAGAACCACCTA 3’	(Sassera et al. 2008)
gyrB (reverse)		5’ CAAGCTCTGCCGAAATATCTT 3’	
Isrpl6 (forward)	*Ixodes* ribosomal protein L6 nuclear gene	5’ CCGGTCCAAGATGTTCCACA 3’	This study
Isrpl6 (reverse)		5’ TGCGCTTCCTCTTCTCCTTG 3’	
Amspf2 (forward)	*A. phagocytophilum*	5’ATGGAAGGTAGTGTTGGTTATGGTATT 3’	(Jahfari et al. 2012)
Amspf2 (reverse)		5’ TTGGTCTTGAAGCGCTCGTA 3’	
Amspf2 (probe)		5’-FAM-TGGTGCCAGGGTTGAGCTTGAGATTG-BHQ1-3’	
CS-F (forward)	*Rickettsia* spp.	5’ TCGCAAATGTTCACGGTACTTT 3’	(Stenos et al. 2005)
CS-R (reverse)		5’CACAATGGAAAGAAATGCACGA 3’	
CS-P (probe)		5’-6-FAM-TGC AAT AGC AAG AAC CGT AGG CTG GAT G-BHQ-1-3’	
GroEL (forward)	*Ca. N. mikurensis*	5’CCTTGAAAATATAGCAAGATCAGGTA‘3	(Jahfari et al. 2012)
GroEL (reverse)		5’CCACCACGTAACTTATTTAGTACTAAAG ‘3	
GroEL (probe)		5’-TET-CCTCTACTAATTATTGCTGAAGATGTAGA AGGTGAAGC-BHQ2-‘3	

2.2.4 *Rickettsia* qPCR

For *Rickettsia*, specific primers that amplify the citrate synthase gene *(gltA)*, a highly conserved gene for the majority of SFG rickettsial bacteria, were used (Stenos et al. 2005) in a modification of cycles number from 60 cycles into 40 cycles. The primers amplified a 74 base-pair fragment (**Table 2. 4**). Quantitative PCR was performed using a CFB-3220 DNA Engine Opticon 2 System with a 6-carboxyfluorescein (FAM) and Black-Hole Quencher (BHQ1)-labelled TaqMan probe [synthesized by Eurofins MWG Operon (Ebersberg, Germany)] and SensiMix II Probe No-ROX mastermix (Bioline) (Stenos et al. 2005). Each reaction contained final concentrations of 1 × qPCR SensiMix, *gltA* primers and probe at 200 nM each, and 1 µl DNA to give a total reaction volume of 20 µl.

Serial dilutions of DNA standard control (Sigma-Aldrich) from 5 x 10^6 to 10^{-1} copies in duplicate were prepared for use on each plate as described above (**Figure 2. 4**). The reactions were conducted with an initial holding temperature of 50°C for 3 min, followed by 95°C for 5 min and then 40 cycles of 95°C for 20 sec and 60°C for 40 sec. Fluorescence was monitored at the end of every 60°C annealing step on the FAM channel.

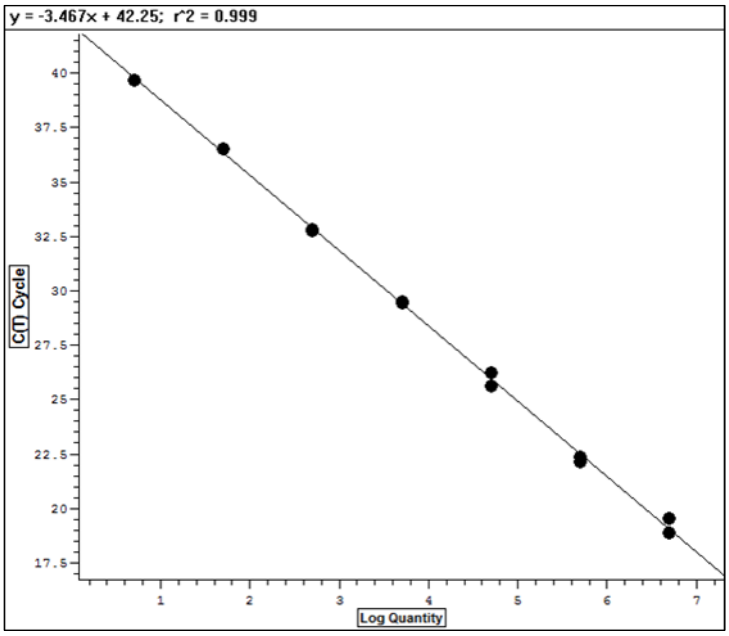


Figure 2. 4 Standard graph of qPCR to detect *Rickettsia* spotted fever group using citrate synthase gene.

In order to differentiate *Rickettsia* species by sequence analysis, positive qPCR samples were examined by conventional PCR for the *Rickettsia* spp. *gltA* gene using primers *Rp*CS.409d (5′-CCTATGGCTATTATGCTTGC-3′ [forward]) and *Rp*CS.1258n (5′-ATTGCAAAAAGTACAGTGAACA-3′ [reverse]) (Roux et al. 1997). Each reaction contained final concentrations of 1× BioMix Red and 10 pmol of each primer, and 1.5 µl of DNA template in a total reaction volume of 50 µl.

Amplification was carried out under the following conditions: a 2-min denaturation step at 95°C was followed by 40 cycles of denaturation at 95°C for 30 sec, annealing at 45°C for 30 sec, and extension at 65°C for 55 sec. Amplification was completed by incubation for 3 min at 72°C to allow complete extension of the PCR products. To verify the results of amplification, 10 µl of the PCR products were resolved using a 1.5% molecular agarose gel (Bioline) incorporating SYBR® Safe DNA gel stain (Invitrogen) and run at 150 V for 30 min in a Bio-Rad gel electrophoresis set. The gels were visualised using a G:Box Gel Documentation System (Syngene).

2.2.5 *A. phagocytophilum* and *Ca.* N. mikurensis qPCR

q PCR was performed in a multiplex format according to a published protocol (Jahfari et al. 2012), using a CFB-3220 DNA Engine Opticon 2 System in a total reaction volume of 20 µl. Amplification was performed using fluorescently-labelled TaqMan probes with final concentrations of 1× SensiMix II Prob No-ROX (Bioline), primers at 250 nM each; a FAM-BHQ1 probe to detect *A. phagocytophilum* at 125 nM, a tetrachlorofluorescein- and BHQ1-labelled probe to detect *Ca.* N. mikurensis at 250 nM, and 1 µl of template DNA **(Table 2. 4**). Serial dilutions of DNA standard control from 5 x 10^{6} to 10^{-1} copies in duplicate were prepared for use on each plate as described above. **(**See **Figure 2. 5**). Cycling conditions were 95°C for 5 min, followed by 40 cycles of denaturation at 95°C for 5 sec and a 35 sec annealing-extension step at 60°C (Jahfari et al. 2012).

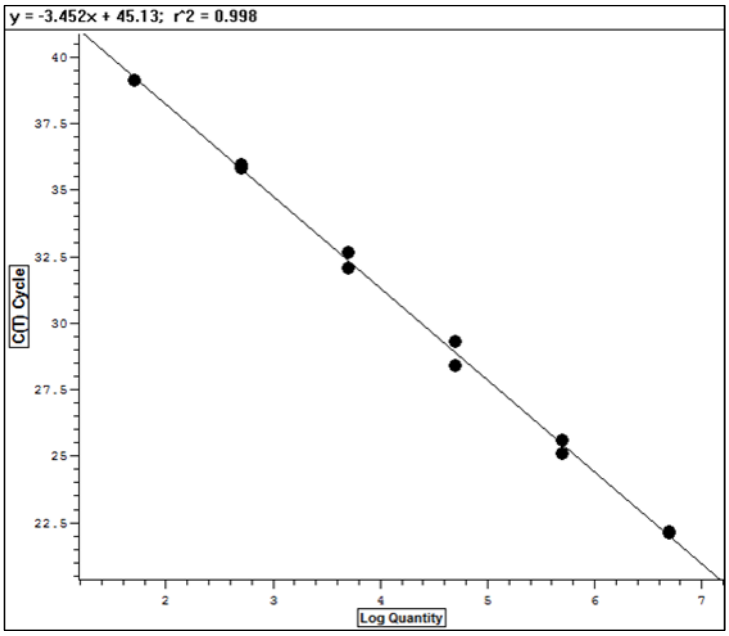


Figure 2. 5 Multiplex qPCR standard graph using an *Amsp2* gene and *groEL* gene for *A. phagocytophilum* and *Ca.* N. mikurensis respectively.

To confirm *A. phagocytophilum*-positive samples, they were checked by nested PCR using primers targeting the *Anaplasma* 16S ribosomal RNA (Massung et al. 1998). Final concentrations comprised 1× BioMix Red and primers ge3a 5'-CACATGCAAGTCGAACGGATTATTC-3' [forward] and ge10r 5'-TTCCGTTAAGAAGGATCTAATCTCC-3' [reverse] at 0.5 µM each. Reactions used 2.5 µl of purified DNA as the template in a total volume of 25 µl in the first round. Cycling conditions involved an initial 2-min denaturation at 95°C, followed by 40 cycles, each comprising a 30 sec denaturation at 94°C, a 30 sec annealing at 55°C, and a 1 min extension at 72°C. These 40 cycles were followed by a 5-min extension at 72°C. Reaction products were maintained at 4°C until they were analysed by agarose gel electrophoresis or used as templates for the second round of nested PCR reactions. In the second-round amplifications, 1 µl of the primary PCR product was used as a template in a total volume of 50 µl comprising 1× BioMix Red and 0.2 µM of each primer ge9F 5'-AACGGATTATTCTTTATAGCTTGCT-3'[forward] and ge2R 5'-

GGCAGTATTAAAAGCAGCTCCAGG-3' [reverse]. Nested cycling conditions were as described for the primary amplification, except that 30 cycles were used instead of 40 cycles (Massung et al. 1998).

Reaction products were subsequently maintained at 4°C until they were analysed by agarose gel electrophoresis using a 1.5% agarose gel as described above for confirmatory *Rickettsia* PCRs. PCR products were purified using a QIAquick PCR purification Kit (Qiagen) for DNA sequencing following the manufacturer's protocol. Amplicons were sequenced by Eurofins MWG (Germany), and then assembled using BioEdit software (Ibis Bioscience, Carlsbad). Online BLASTn (National Centre for Biotechnology Information) was used to compare results with published sequences in the GenBank sequence database.

2.2.6 Statistical analysis of bacterial densities

Statistical calculations were performed in Microsoft Excel 2016 to produce a ratio of bacteria to tick, allowing comparison of bacterial concentrations in each life cycle stage. *Ca*. M. mitochondrii copy numbers normalised against *IsRPL6* were compared between tick stages (for the Welsh samples) and collection location (for the English nymphs) in IBM SPSS Statistics v. 24 (IBM Corp.). Data were log-transformed and Levene's homogeneity of variance test was run to verify that variances were not significantly different between groups. Where variances were equal, a one-way ANOVA was performed with Tukey's post-hoc test. If log-transformation failed to equalise the variances, a non-parametric ANOVA (Kruskal-Wallis test) was conducted. The critical probability was set as $P < 0.05$.

2.2.7 *Amblyomma variegatum:* Tick collection, DNA extraction and quantification

A total of 192 *A. variegatum* adult samples [(142 females, 70 ifs, 72 ffs) and 50 males] were kindly provided by Germanus Soh Bah as three batches: one batch in October 2016 (males and females) and two batches in October 2017 (first batch males and females, while the second batch had ffs females only) collected from cattle around Ngaoundéré in the Adamawa Region Cameroon. Ticks were stored for species identification and DNA extraction in 1.5 ml vials containing molecular-grade 70% ethanol. Tick samples were transported to the University of Liverpool on gel-ice packs (4 - 8°C) and sorted individually by engorgement stage and sex on arrival. Tick species were identified using a standard key (Walker et al. 2003). Any damaged ticks were discarded.

DNA extraction and quantification were done using non-commercial ammonium hydroxide and a fluorescent labelling method respectively as described above in (2.2.2).

2.2.8 *Rickettsia africae* qPCR

All *A. variegatum* samples were tested individually for *R. africae* by qPCR using specific primers targeting the citrate synthase gene *gltA* (which is a highly conserved gene in the majority of SFG *Rickettsia*) as described above in (2.2.4); and normalised against the tick single-copy nuclear gene (*IsRPL6*) as described in (2.2.3) (see Error! eference source not found.).

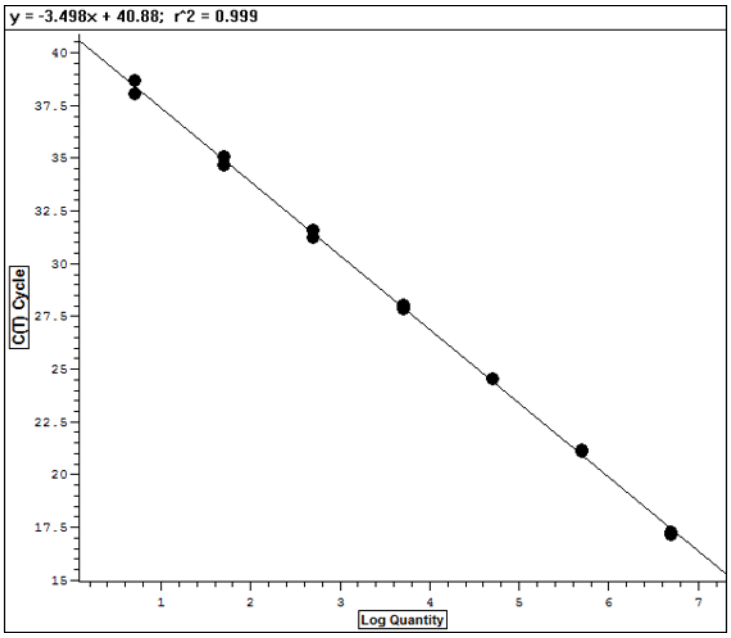


Figure 2. 6 Standard graph of a qPCR to detect *R. africae* using the *gltA* gene as a target.

The presence of *R. africae* was confirmed by conventional PCR targeting three different genes (*ompB*, *12S* and *sca4*) according to (Alberdi et al. 2012).

2.2.9 *Ehrlichia ruminantium* pCS20 PCR

A conventional PCR assay for *E. ruminantium* targeting pCS20 [a conserved, single-copy region of the genome (Steyn et al. 2008)] was applied using 1× BioMix Red (Bioline), primers HH1F and HH1R (**Table 2. 5**) at 400 nM (final concentration) each, and 2 µl of DNA template in a final volume of 25 µl to amplify a 903-bp product. The positive control DNA was provided by Dr Lesley Bell-Sakyi (the Pirbright Institute, UK); DNA was extracted by Dr Lesley Bell-Sakyi from a culture of *I. scapularis* embryonic cells (ISE6) infected with the Ball 3 strain of *E. ruminantium* (Haig 1952); successful growth of the Ball3 strain in ISE6 cells was previously confirmed (Moniuszko et al. 2014). Nuclease-free water was used as a negative control. Thermal cycling comprised initial heating at 94°C for 2 min, then 35 cycles of denaturation at 94°C for 1 min, annealing at 62°C for 2 min, and final extension at 72°C for 10 min.

Using a Safe Imager transilluminator (Invitrogen), 10 µl PCR products were visualised in a 1.5% agarose gel stained with SYBR Safe DNA Gel stain (**Figure 2. 7**). The PCR products were then purified using a QIAquick PCR Purification Kit (QIAGEN) following the manufacturer's protocol and sequenced by Eurofins MWG Operon (Ebersberg, Germany). The sequences were compared using BLASTn (https://blast.ncbi.nlm.nih.gov/Blast.cgi) with data available in GenBank for *E. ruminantium*.

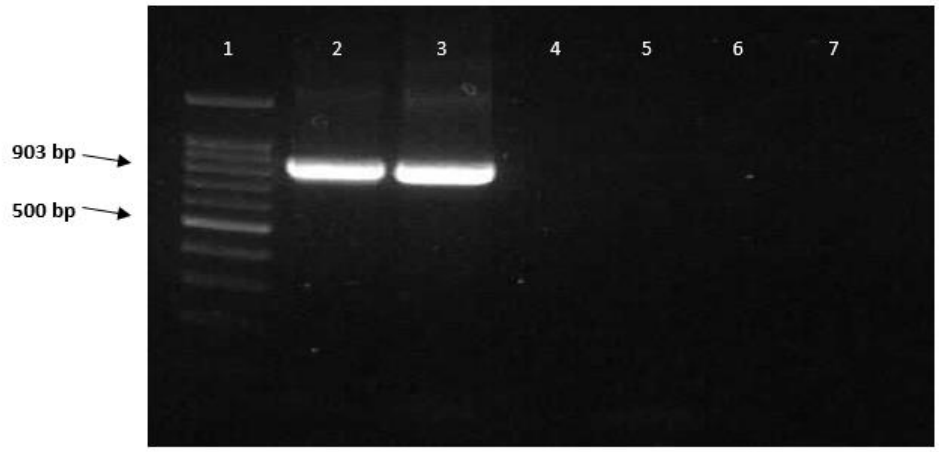


Figure 2. 7 Gel electrophoresis visualisation of pCS20 region products.
Amplified from DNA extracted from *Ehrlichia ruminantium* (Ball3)-infected ISE6 cells (positive control sample) for cloning to generate a pCS20 plasmid. 1.5% agarose gel; lane 1, 100bp DNA ladder; lanes 2 and 3, 903 bp product amplified from the pCS20 region.

A positive control PCR amplicon band was excised from the gel and purified using a PureLink® Quick Gel Extraction Kit (Invitrogen, California, USA) according to the manufacturer's instructions, then subjected to cloning in the pGEM®-T Easy Vector System (Promega). PCR fragments were ligated overnight at 4°C into the pGEM-T Easyplasmid with a 1:1 insert-plasmid ratio. The ligants were transformed into JM109 *E. coli* competent cells (Promega). Plasmid DNA was extracted from the

transformed cell pellets using the Wizard Plus SV Minipreps DNA Purification Kit (Promega), following the manufacturer's protocol. To verify that the plasmid was cloned successfully, plasmid was cut by restriction enzyme (echoR1) (Thermofisher code is FD0274 FastDigest EcoRI), at 37 ^{o}C and for 5 minutes incubation.

Finally, the plasmid DNA samples were excised from the Gel and purified using a PureLink® Quick Gel Extraction Kit (Invitrogen, California, USA) according to the manufacturer's instruction (**Figure 2. 8**) and sent for Sanger sequencing with pUC/M13 forward and reverse primers to Source Bioscience Ltd, UK. Nucleotide Basic Local Alignment Search Tool with BLASTn online platform (https://blast.ncbi.nlm.nih.gov) was used to align and compare DNA sequences to the nucleotide collection database.

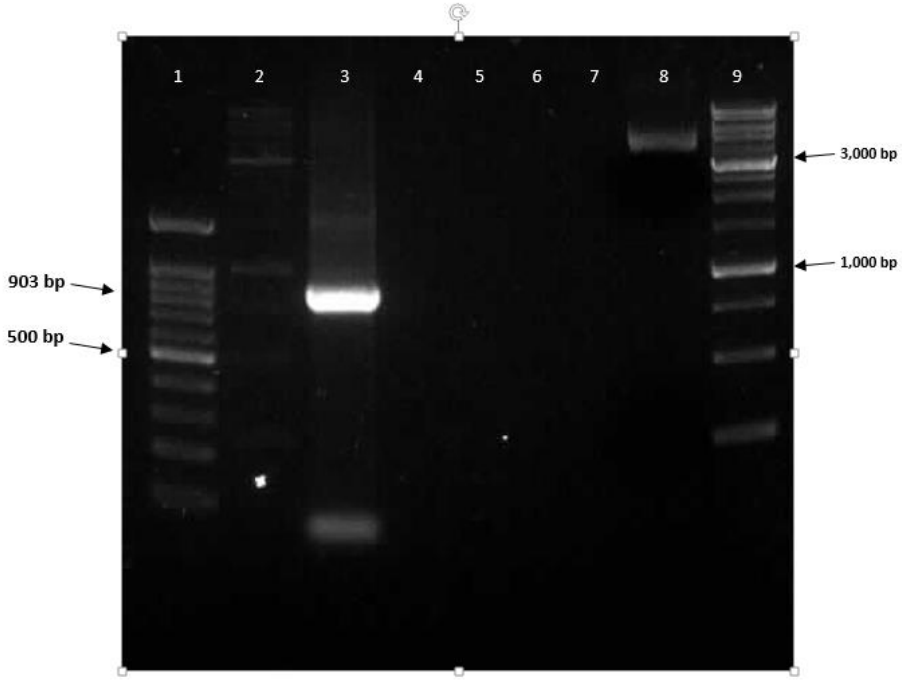


Figure 2. 8 Gel electrophoresis confirmation of generation of pCS20 plasmid. Before and after extraction. 1.5% agarose gel. Lane 1,100 bp DNA ladder; lane 3, a 903-bp pCS20 product released from digested plasmid; lane 8, plasmid prior to digestion; lane 9, 1kb DNA ladder .

2.2.10 Generation of the pCS20 plasmid standard curve to calculate *Ehrlichia ruminantium* copy numbers

In order to measure *E. ruminantium* DNA concentration and convert to genome copy numbers, generation of a pCS20 plasmid standard curve was performed using pGEM®-T-pCS20 comprising the 903-bp pCS20 region. A 10-fold dilution serial dilution of plasmid DNA (5×10^{-1} to 5×10^{8}) was prepared (starting concentration 2.22 ng/µl) in 100 ng/µl yeast tRNA solution (Invitrogen; to prevent aggregation) for standard curve generation. The DNA concentration was converted into number of pCS20 copies/µl using Avogadro's constant.

2.2.11 *Ehrlichia ruminantium* qPCR

A published TaqMan probe assay was applied using CowF forward and CowR reverse primers to amplify a 226-bp fragment of the conserved pCS20 region (Steyn et al. 2008). A TaqMan probe (**Table 2. 5**) was designed 120 bp downstream from the forward primer and synthesized by Eurofins MWG Operon (Ebersberg, Germany). In a total reaction volume of 20 µl, each reaction contained 1 × Sensimix Probe II mastermix (Bioline), 0.5 µM of each primer and 0.4 µM probe. Nuclease-free water was used as a negative control. Thermal cycling conditions comprised 10 min at 40°C as the initial incubation, then 40 cycles of 10 min denaturation at 95 °C, annealing at 48 °C for 20 sec, and a 30 sec elongation at 58 °C. Standards were included as described above in (2.2.3). As we could not confirm qPCR positive products by conventional PCR, qPCR products were separated and visualised on a 1.5% agarose gel stained with SYBR Safe gel stain, bands were excised and purified using a PureLink® Quick Gel Extraction Kit (Invitrogen, California, USA) according to the manufacturer's instructions.

Table 2. 5 Primers and Probe sequences used in conventional and qPCR of pCS20 region of *Ehrlichia ruminantium* (Steyn et al. 2008).	
Primer name	**Sequence**
HH1	5'-CCC TAT GAT ACA GAA GGT AAC CTC GC-3'
HH2	5'-GAT AAG GAG ATA ACG TTT GTT TGG-3'
CowF	5'-CAA AAC TAG TAG AAA TTG CAC A-3'
CowR	5'-TGC ATC TTG TGG TGG TAC-3'
Cow probe	5'-FAM TCC TCC ATC AAG ATA TAT AGC ACC TAT TA NT-BHQ1-3'

2.3 Results

2.3.1 Confirmation of species identification and DNA yield

Maximum likelihood trees (bootstrapped 10,000 times) for the mitochondrial 16S rRNA gene confirmed identification of our samples as *I. ricinus*, comparing with *I. inopinatus and I. persulcatus* sequences (**Appendix A**).

As extraction of DNA from ticks can be challenging due to their small size and high level of inhibitors such as haem from blood, different methods for DNA extraction were trialled. This study showed that alkaline lysis is a rapid, effective and economical method to purify genomic DNA from ticks and also provided the highest DNA yield. The highest average of DNA yield was 158.6 μg, while the average was 33 using the Thermos Gene Jet kit; however, no DNA recovery was achieved using the MoBio Power Soil kit.

2.3.2 *Ca.* Midichloria mitochondrii density in different *Ixodes ricinus* life stages from Wales.

The ticks collected from deer in Wales exhibited a very high prevalence of infection with *Ca.* M. mitochondrii, with only four specimens (three males and one ifs

female) testing negative for the symbiont (**Table 2. 6**). In all cases, the tick nuclear gene *IsRPL6* was successfully amplified. The density of *Ca*. M. mitochondrii was significantly higher in female ticks of every engorgement stage compared with both males and nymphs (one-way ANOVA, $P < 0.01$; **Figure 2. 9**); whereas levels between males and nymphs were not statistically distinguishable. Neither was the apparent positive trend between symbiont density and female tick engorgement stage statistically significant.

Table 2. 6 Number of *Ca.* Midichloria mitochndrii positive and negative samples for each categoryof tick.

Category	**Positive for *Ca.* M. mitochondrii**	**Negative for *Ca.* M. mitochondrii**	**Prevalence of *Ca.* M. mitochondrii**
Fast-feeding stage	20	0	100%
Intermediate-feeding stage	24	1	96%
Slow-feeding stage	14	0	100%
Male	45	3	93.8%
Nymph	44	0	100%

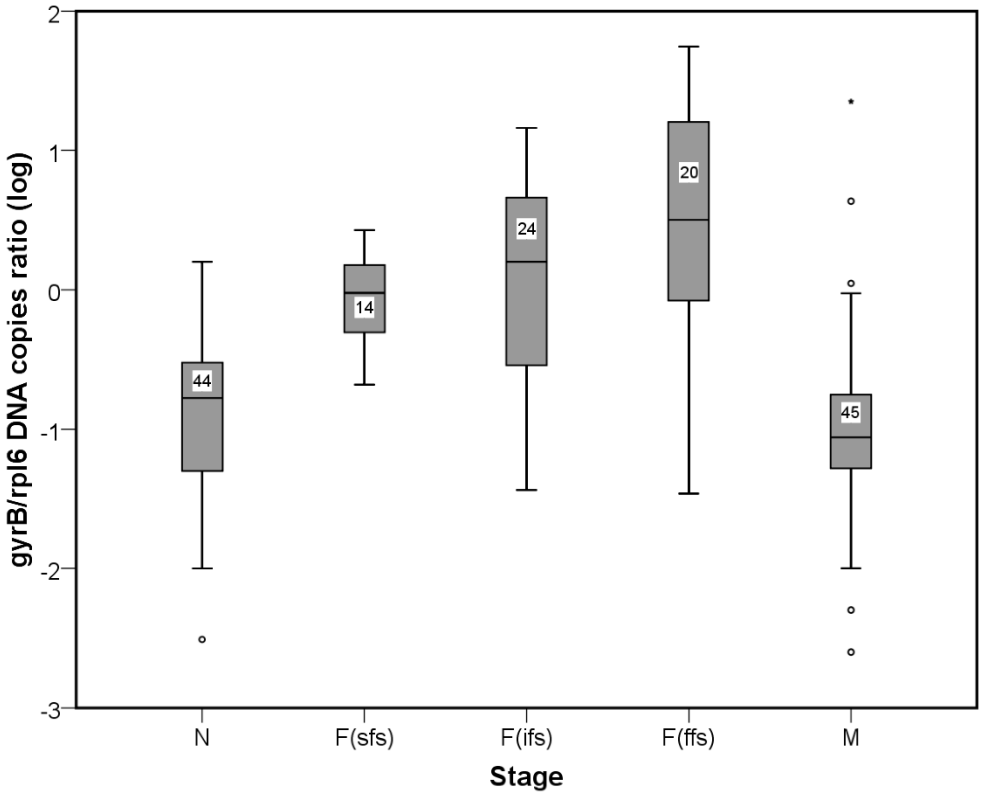


Figure 2. 9 Density of *Ca.* Midichloria mitochondrii in *I. ricinus* from Wales.
Density of the symbiont in different stages of *I. ricinus* collected from deer at Powis Castle estate, Wales (box-plot). Note that the copy number ratio has been subjected to log10 transformation. Sample sizes (n) are displayed within each box-plot. N = nymph; F(sfs) = slow feeding stage; F(ifs) = intermediate-feeding stage ; F(ffs) = fast-feeding stage; M = male.

2.3.3 *Ca.* Midichloria mitochondrii density in *Ixodes ricinus* nymphs from England

In contrast with the partially-fed Welsh nymph samples, only approximately three-quarters of questing nymphs in England were positive for the symbiont, and this prevalence was remarkably consistent between locations (**Figure 2. 10**). While the range of symbiont densities in the English samples was very wide, the median was not significantly different between locations (Kruskal-Wallis test, $P = 0.092$).

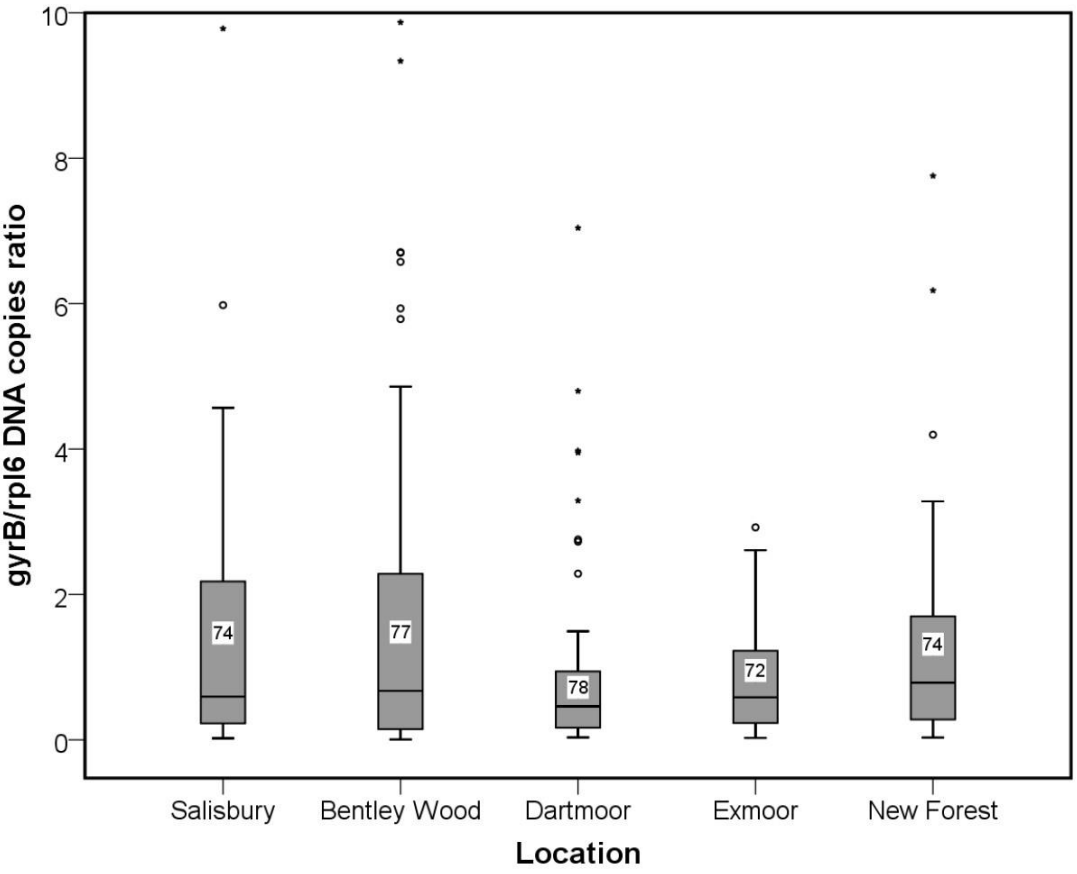


Figure 2. 10 Density of *Ca.* Midichloria mitochondrii in questing nymphs from field sites in southern England.
For the location of these sites, the copy number ratio has been subjected to log10 transformation. Sample sizes (n) of Midichloria-positive specimens are displayed within each box-plot; 100 specimens were assayed per location.

2.3.4 Pathogenic bacterial quantification

Quantification of bacterial pathogens in British *I. ricinus* revealed that overall, 1.4% (7/500) of collected English nymphs tested positive for *A. phagocytophilum* using qPCR with confirmation by nested PCR and sequencing (**Figure 2. 11**).

In total, 155 Welsh *I. ricinus* ticks of different life stages, blood-feeding stage and sex were analysed for the presence of *A. phagocytophilum* (**Table 2. 7**). The overall prevalence was low, with only five ticks being infected (**Table 2. 7**; **Figure 2. 11**). The

nested PCR alignment of *A. phagocytophilum* sequences showed one nucleotide difference between the Welsh samples and English nymphs. (**Appendix B**)

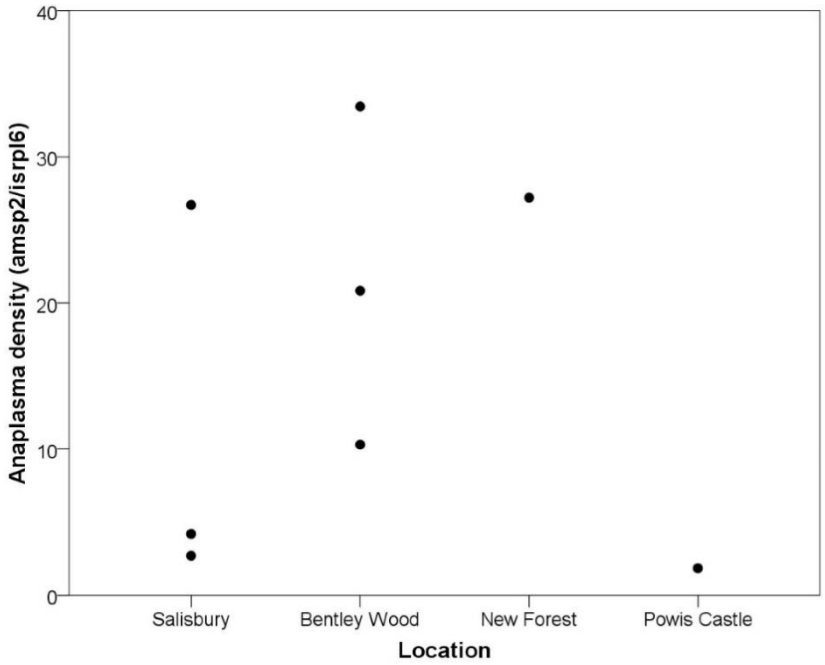


Figure 2. 11 *Anaplasma phagocytophilum* copy numbers determined by qPCR in nymph samples only across different locations in the UK.

Table 2. 7 *Anaplasma phagocytophilum* prevalence from all life stages of ticks collected in Wales.

Category	Number tested	Positive for *Anaplasma phagocytophilum* (%)
Female (fast-feeding stage)	22	0
Female (intermediate-feeding stage)	25	0
Female (slow-feeding stage)	14	2 (14.2 %)
Male	48	2 (4.2 %)
Nymph	46	1 (2.2 %)

For *Rickettsia* spp, only 0.8% (4/500) of English nymphs were positive with very low copy numbers by qPCR, but no *Rickettsia* infections could be confirmed by sequencing. Additionally, all of the samples were negative for *Ca.* N. mikurensis.

2.3.5 *Rickettsia africae* and *E. ruminantium* quantification in *A. variegatum*

qPCR analysis revealed a high prevalence of *R. africae* in *A. variegatum*. However, the density of infection seemed to be low, and variable between different collection batches; the highest was observed in the ffs samples of the third collection. Overall of 183 ticks positive for *R. africae*, 136 (95%) of 142 female ticks were found to contain *Rickettsia* byq PCR, while 47 (94%) of 50 males were positive for *R. africae*. Overall, the density was higher in female ticks compared with male ticks.

A comparison of male ticks between the first two collections revealed higher density in the second collection **(Figure 2. 12**). In female ticks, while there was no difference between the first and the second collections, the density of *R. africae* in the third collection was highly variable (**Figure 2. 13**). The presence of *R. africae* was confirmed in 5 tick specimens per collection batch by sequencing of three genes following conventional PCR. The top BLAST hits for each of these genes were *Rickettsia africae* isolate 30966 outer membrane protein B (*ompB*) gene, accession number KU721071.1; *Rickettsia africae* cell surface antigen (*sca4*) gene, accession number AF151724.2; and *Rickettsia africae* isolate 30966 16S ribosomal RNA gene, accession number KU721068.1.

A quantitative real-time PCR TaqMan probe assay based on a sequence in the pCS20 region of *E. ruminantium* detected six positive samples; these were positive for *R. africae* as well (**Figure 2. 14**). As we could not confirm the real-time PCR positive results by conventional PCR, qPCR products were separated and visualised on a 1.5% agarose gel stained with SYBR Safe gel stain. Bands were excised and purified using a PureLink® Quick Gel Extraction Kit (Invitrogen, California, USA) according to the manufacturer's instructions (**Figure 2. 15**), and sequenced by Eurofins MWG Operon

(Ebersberg, Germany). They were identified using BLAST (https://blast.ncbi.nlm.nih.gov) and compared with sequences available in GenBank for *E. ruminantium*. However, the full-length conventional PCR for pCS20 did not produce an amplicon with these qPCR-positive samples. The top three BLAST hits for sequence confirmation were *Ehrlichia ruminantium* Sankat (Pcs20) region ribonuclease III gene, accession number AY236065.1; *Ehrlichia ruminantium* Gardel (pCS20) region ribonuclease III gene, accession number AY236061.1 and *Ehrlichia ruminantium* isolate DAMAN-HS ribonuclease III (pCS20) gene, accession number MG564188.1.

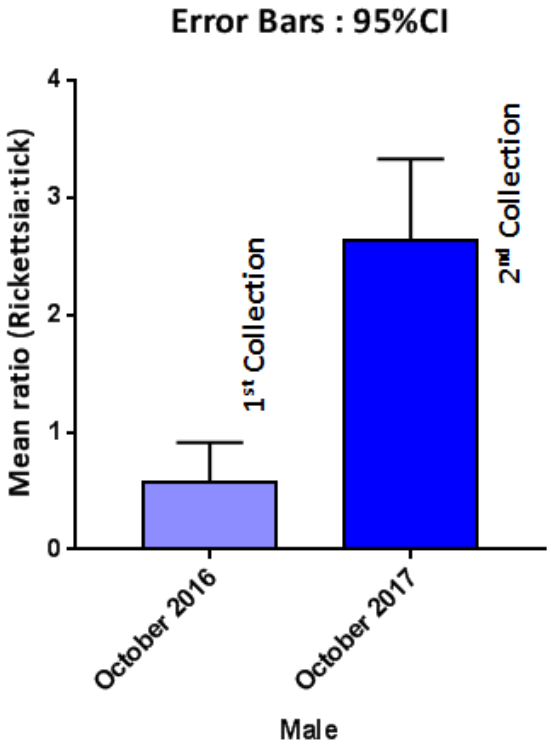


Figure 2. 12 A comparison of *R. africae* density in *A. variegatum* tick (males) from two collections (2016 and 2017) in Cameroon.

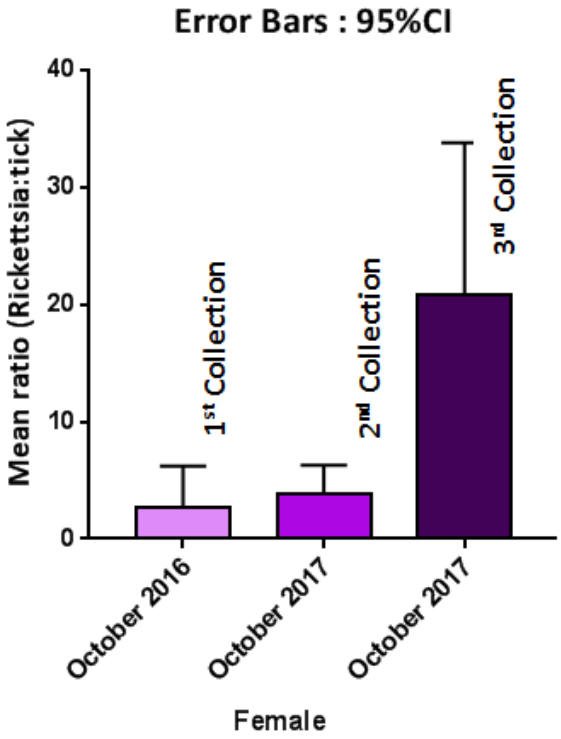


Figure 2. 13 A comparison of *R. africae* density in *A. variegatum* tick (females) from three collections in 2016 and 2017 in Cameroon.

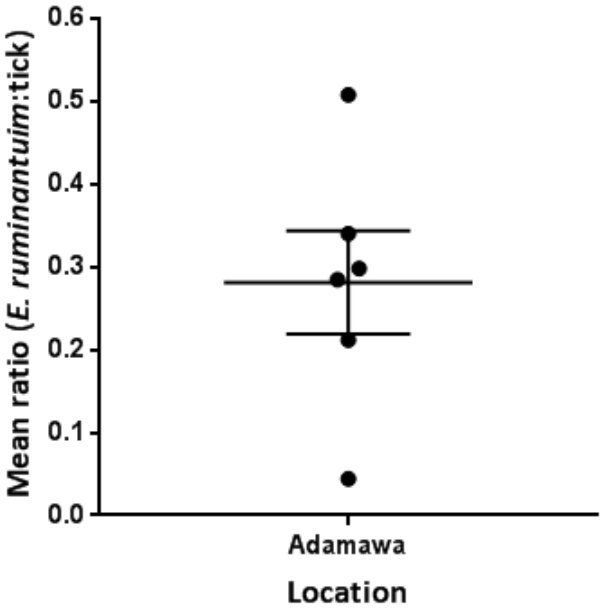


Figure 2. 14 The density of *E. ruminantium* in the five positive *A. variegatum* ticks based on a sequence in the pCS20 region of *E. ruminantium.* Mean with standard error.

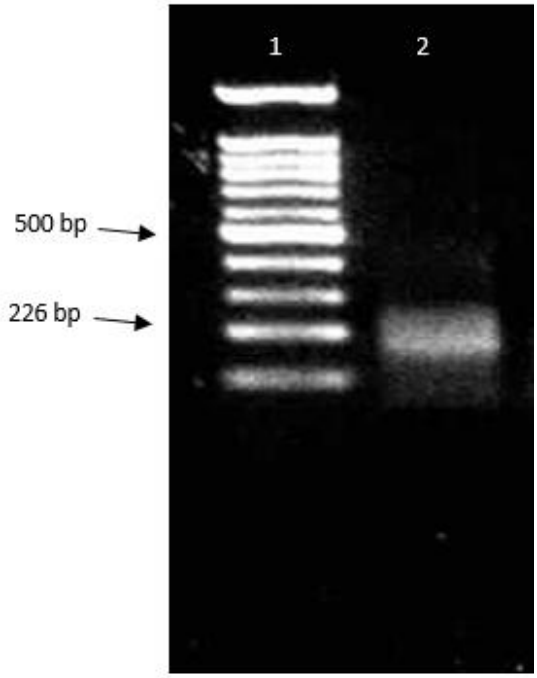


Figure 2. 15 Gel electrophoresis illustrating qPCR products amplified from the pCS20 conserved region. Lane 1, (100 bp) ladder; lanes 2 = a 226-bp product from the pCS20 region of one representative positive sample.

2.4 Discussion

Ixodes ricinus is known to transmit some serious bacterial, parasitic and viral infections to humans and animals, including *B. burgdorferi*, the causative agent of Lyme disease, and *A. phagocytophilum*, the causative agent of tick-borne fever in ruminants. In addition to pathogenic bacteria, *I. ricinus* ticks are known to harbour endosymbiotic bacteria, including the recently described *Ca*. M. mitochondrii.

Although *Ca.* M. mitochondrii has been reported from British ticks in one study only (Smith et al. 2013), this was not a quantitative study and did not attempt to link symbiont and pathogen levels. Also, it did not report *Ca*. Midichloria infection by geographic location or by life stage. This work presents the first results of quantitative screening for DNA of the *Ca.* M. mitochondrii bacterium and its association with *A. phagocytophilum*, *Rickettsia* spp. and *Ca.* N. mikurensis in *I. ricinus*

ticks from different parts of the UK. The present study indicates that all stages of the lifecycle, although not every individual specimen, showed evidence of *Ca.* M. mitochondrii infection as previously reported for *I. ricinus* populations from continental Europe (Sassera et al. 2008). The comparison of the *gyrB/IsRPL6* gene ratio indicated that relative levels of *Ca.* M. mitochondrii in *I. ricinus* ticks were generally increased with engorgement, suggesting that the bacteria may obtain nutrients from the blood meal within the tick. The results revealed that 98% of Welsh female *I. ricinus* were positive for *Ca.* M. mitochondrii, which is not quite compatible with the previous studies that reported 100% prevalence of *Ca*. M. mitochondrii in European female *I. ricinus* (Sassera et al. 2006), although only one negative sample reduced the bacterium prevalence rate in the present study. Interestingly, 93.8% of Welsh male ticks were infected with the bacterium, which highlights a higher prevalence rate in the Welsh males compared with only 44% previously recorded in continental Europe (Sassera et al. 2006). This suggests that male *I. ricinus* have a variable rate of *Ca.* M. mitochondrii infection in different countries. The high prevalence in Welsh nymphs supports the data for the adult males (if both sexes are almost 100% infected in Wales, then we would expect nymphs to be nearly 100% infected too).

The overall prevalence of *Ca.* M. mitochondrii was approximately 70 - 80% in English unfed nymphs. Although nymphs cannot be sexed morphologically, this prevalence estimate closely matches published data for adult ticks from continental Europe if a 1:1 sex ratio among nymphs is assumed. However, the density of *Ca.* M. mitochondrii in English unfed nymphs was apparently more variable than that previously reported from the Czech Republic (Sassera et al. 2006), although our sample size was much larger (500 in the current study *v*. 18 in Sassera's study). Taking

the above results into consideration emphasises the need to investigate a larger number of UK *I. ricinus* to assess if the higher prevalence in both males and nymphs in Wales could be significant.

The data from the present study indicate that as the engorgement level of the tick increases, so does the relative concentration of *Ca.* M. mitochondrii. This finding supports other studies which show that due to high levels of bacteria residing in the ovarian tissue, females will contain a larger concentration of bacteria than male ticks (Sassera et al. 2006). The range of values for bacterial concentration was very wide for ffs females, with some having very low concentrations and some with exceptionally high numbers.

In this study, *A. phagocytophilum* was detected at a low level in *I. ricinus* nymphs. Overall, 1.47% (8/546) of nymphs collected from different locations within the UK were infected with *A. phagocytophilum*, which is compatible with other *Anaplasma* prevalence rates reported from continental Europe (Henningsson et al. 2015; Silaghi et al. 2008). Low infection rates are linked to the absence of transovarial transmission of the bacteria in *I. ricinus* ticks, and therefore the reservoir host is required to maintain the lifecycle of the pathogen (Rizzoli et al. 2014). Another interesting finding presented here is associated with *A. phagocytophilum* 16S rRNA gene sequences from Welsh samples that showed one nucleotide difference compared with English nymphs. This could suggest diversity between different strains of *A. phagocytophilum* based on geographical location, so analysing a higher number of *A. phagocytophilum* positive samples and examining more genes is required and highlights an area for future research.

In the UK, the infection rate of pooled *I. ricinus* nymphs and adults ranged from >0.25 to 2% at four sites in Scotland (Alberdi et al. 1998). Furthermore, a

previous study revealed that the infection rate with *A. phagocytophilum* was higher in adult *I. ricinus* in upland heath, grassland and bog habitats (North Wales) compared with woodland habitat (southern England); which is consistent with the present findings, suggesting that *A. phagocytophilum* infections acquired by larvae may be maintained through two moults (Ogden et al. 1998). Also differences were demonstrated between sequences from *I. ricinus* collected from North Wales compared with previous sequences of *A. phagocytophilum* isolated from sheep in Scotland (Anderson et al. 1991), suggesting that Scottish sheep might act as competent reservoirs for several genotypes of *A. phagocytophilum* (Ogden et al. 1998). Previous reports from Great Britain demonstrated *A. phagocytophilum* infection of both roe deer and the tick vector, suggesting an important natural mammalian reservoir in the UK (Alberdi et al. 2000). In northern England, *A. phagocytophilum* was detected in around 47.3% of roe deer; however only 0.7% of tested nymphs were infected compared with 3.4% and 2.5% of examined adult females and males respectively (Bown et al. 2009).

In mainland Europe, the average rate of *A. phagocytophilum* infection is around 1.6 - 5% in adult *I. ricinus* (Hartelt et al. 2004); for instance, the prevalence of *A. phagocytophilum* in free-living ticks is around 1 – 5% in Germany (Hartelt et al. 2004). Moreover, characterization of tick borne fever disease in Norwegian lambs inoculated experimentally with a red reer (*Cervus elaphus*) isolate of *A. phagocytophilum* suggested that deer might serve as reservoir hosts for ovine infections (Stuen et al. 2010). A subsequent study concerned with fallow deer (*Dama dama*) and associated ticks supported this suggestion; the authors detected *A. phagocytophilum* infection in 15 - 28% of fallow deer by serology and PCR, as well as in around 7.3% of associated ticks by PCR- in Italy (Veronesi et al. 2011). Thus, fallow

deer might be considered as competent hosts for *A. phagocytophilum* and consequently may play a role in the epidemiological perspective of the infection across its geographical distribution (Veronesi et al. 2011), In addition, large animals such as sheep, goats, cattle and horses (Hulínská et al. 2004) and small rodents such as *Apodemus flavicollis*, *Apodemus sylvaticus* and *Myodes glareolus* (Hildebrandt et al. 2002), have been described as reservoir animals for anaplasmosis in Germany.

Although *Ca*. M. mitochondrii infection in domestic and wild ruminants was not investigated in the present study, another study detected circulating *Ca*. M. mitochondrii DNA in up to 11% of sampled horses (n = 46), 29% of dogs (n= 62) and 9% of sheep (n = 11) in Italy (Bazzocchi et al. 2013). Previous studies have reported a 55% serum prevalence of *A. phagocytophilum* tested lambs in the north-west coast of Norway (Grøva et al. 2011) and a 18% serum prevalence of *B. burgdorferi* in examined sheep in the west coast of Norway (Grøva et al. 2011). Furthermore, another study indicated *Ca*. M. mitochondrii in the ticks within the same geographical region, suggesting co-infection with *A. phagocytophilum*, *B. burgdorferi* and *Ca*. M. mitochondrii in the same tick vector (Granquist et al. 2014).

In Europe, *Ca. N. mikurensis*, from the family Anaplasmataceae, is referred to as an emerging human pathogen, since its first detection in patients with febrile illness was reported in 2010. However, this microorganism has been identified in ticks from Europe, Asia and Africa, suggesting that more human cases are likely to be detected in the future (Glatz et al. 2014). Previous studies have shown that *Ca. N. mikurensis* and *A. phagocytophilum* share a similar ecology, with large foci found within areas of Central Europe (Glatz et al. 2014). All samples were negative for Ca. *N. mikurensis* in the present study, which agreed with a previous study of adult *I. ricinus* which did not find this pathogen in the UK (Jahfari et al. 2012). However, as

temperatures increase due to global warming, it may allow the geographical range of this pathogen to increase, so regular surveillance is required, particularly as *Ca.* N. mikurensis has been detected in the Netherlands and Belgium (Jahfari et al. 2012).

This study revealed that only 0.8% (4/500) of *I. ricinus* questing nymphs were positive for *Rickettsia* spp. according to real-time PCR. However, samples positive for *Rickettsia* were not confirmed by sequencing, perhaps due to very low copy numbers. The low rate of *Rickettsia* infection of ticks from different geographical locations may be due to different detection limits of the techniques used (Tappe & Strube 2013).

Increasing evidence of the pathogenic nature of rickettsiae in Europe led to increasing concerns about tick-borne rickettsial diseases in the UK such as those caused by *R. helvetica* and *R. raoultii* (Tijsse-Klasen et al. 2011). Both *R. helvetica* and *R. raoultii* have been reported as human pathogenic agents; *R. helvetica* transmitted by *I. ricinus*, was detected for the first time in Switzerland (Beati et al. 1993), and then in France (Parola et al. 1998); *R. helvetica* was recorded in a patient with meningitis in Sweden (Nilsson et al. 2006); also *R. conorii*, the agent of Mediterranean spotted fever (MSF), is endemic across the Mediterranean area (reported in Parola et al. 1998). Subsequently, wildlife-associated *Rickettsia* spp. were demonstrated in *I. lividus* ticks collected from sand martin birds in northwest England (Graham et al. 2010). Potentially pathogenic SFG rickettsiae were detected for the first time in the UK by analysing ticks of two different species from eight counties across the UK; 6.5% of *I. ricinus* were infected with *Rickettsia* spp, while the infection rate was 9.7% in *Dermacentor reticulatus* (Tijsse-Klasen et al. 2011). Whereas *R. raoultii* was detected in England and Wales only, *R. helvetica* was widespread, as both sfs and ffs specimens of *I. ricinus* were infected across the sampling area (Tijsse-Klasen et al. 2011). However, although serological reactivity against both *A. phagocytophilum* and

Rickettsia spp. in humans is rare, the coexistence of both microorganisms in European ticks has been recorded frequently (Hildebrandt et al. 2010).

Unfortunately, due to the low prevalence of pathogens in this study, the statistical power was not sufficient to determine if symbiont and pathogen densities were positively or negatively correlated. However, the results obtained from this study highlight many areas for future investigation, including potential research on other pathogenic bacteria, such as *B. burgdorferi*, the causative agent of Lyme disease. If a positive or negative association between endosymbiotic and pathogenic bacterial density in ticks could be determined, then there is a possibility for manipulation of the microbiome of ticks transmitting pathogenic bacteria, leading to control of some major mammalian pathogenic agents. This has been hypothesized previously (Gall et al. 2016; Macaluso et al. 2002).

The high prevalence of infection with *R. africae* in *A. variegatum* ticks obtained from the Adamawa Region in Cameroon was anticipated and corroborated previous studies. This could be explained if there is efficient transovarial transmission of *R. africae* in *A. variegatum* as shown for *A. hebraeum* (Robinson et al. 2009). The higher prevalence of *R. africae* in females compared with males might be because ovaries are the predominant location for these bacteria although, *Rickettsia* have also been detected in Malpighian tubules of *Amblyomma* ticks (Cowdry 1925). Moreover, host blood meal acquisition increases bacterial load and multiplication (Robinson et al. 2009). An alternative explanation is that *A. variegatum* may acquire this pathogen during blood feeding on infected animals as suggested previously (Kelly et al. 1991), although both sexes of the tick feed as adults.

An early histological investigation by Cowdry (1925) reported presence of pleomorphic *Rickettsia*-like bacteria (distinct from *E. ruminantium*, at that time

referred to as *Rickettsia ruminatium*) in 100% of larval, nymphal and adult *A. hebraeum* and confirmed its transovarial transmission. A later study using a haemolymph test revealed a high prevalence of rickettsia-like organisms in ticks in Zimbabwe (Raoult 1995), then a molecular study demonstrated *R. africae* within *A. variegatum* in Mali and Niger (Parola et al. 2001). These results correlated with subsequent serological studies that highlighted rickettsial antibodies in cattle and humans in the northern region of Cameroon and in other animals in the south of the country; as well as an immunofluorescence assay detecting antibodies to *R. africae* in Cameroonian patients (cited in Ndip et al. 2004). Following that, a real time PCR study observed *R. africae* in patient blood and in *A. variegatum* ticks in coastal Cameroon (in the south of the country). The organism was identified in 68% of the 22 female ticks and 47 77% of the 61 male ticks (Ndip et al. 2004). Finally, a more recent study, using the same quantitative real-time PCR that was used in the present study, demonstrated *R. africae* in 92.6% of adult *A. variegatum* ticks collected from domestic ruminants in western Kenya, whereas no evidence of the pathogen was found in blood samples from domestic animals (Maina et al. 2014). These results are consistent with the current findings of high prevalence of *R. africae* in the Adamawa Region, which confirmed the existence of *R. africae* in different Cameroonian regions, as well as extending the known range of *R. africae* into the northern part of Cameroon.

The use of a TaqMan probe for real-time PCR based on the pCS20 sequence region of *E. ruminantium* confirmed the sensitivity of this assay to detect the pathogen in *A. variegatum* tick DNA compared with conventional pCS20 PCR. The low prevalence of *E. ruminantium* in *A. variegatum* that we observed could be because transovarial transmission of *E. ruminantium* has not been detected in *Amblyomma*

species, or could be because of the very low infection rate that cannot be picked up by ticks (Bezuidenhout 1987). Blood feeding on infected animals is the only known route for *A. variegatum* to be infected by *E. ruminantium*. Tick co-feeding [hypothesized for *I. scapularis* infected with *B. burgdorferi* (Patrican 1997; Ogden et al. 2008)] might be another possible route of infection, but this has never been investigated for *E. ruminantium*. It is worth noting that failure to amplify full-length pCS20 in the positive samples suggests either the sequence is not as conserved as previously thought, or it is difficult to amplify the full-length sequence if the bacterial density is low.

In spite of the small number of *E. ruminantium*-infected ticks in the present study, a higher infection rate with *E. ruminantium* was observed in males, corroborating previous findings using the nested pCS20 PCR assay (Ndip 2013). It is possible that higher infection rates in male ticks could be associated with *A. variegatum* feeding pattern differences between male and female; whereas the female attaches to the host for a few days prior to the start of engorgement, the male attaches for a longer period that increases the chance of infection (Jordaan & Baker 1981). Thus, based on the current and previous findings, endemic stability of heartwater in Cameroon could be suggested, agreeing with a previous serological survey of heartwater (cowdriosis) in sheep and goats in north Cameroon that highlighted a high prevalence of antibodies where the density of *A. variegatum* was higher (Awa 1997).

2.5 Conclusions

Overall, all stages of the tick life cycle examined, from nymph to ffs female, contained *Ca*. M. mitochondrii bacteria, but at varying prevalence. In general, the greater the degree of engorgement of the female tick, the higher the level of *Ca*. M. mitochondrii*,* with males and nymphs showing a generally low infection density. However, although the prevalence in English nymphs was as expected, a 93.8% prevalence was detected in adult male specimens from Wales (which is more than twice the prevalence found in continental Europe). In addition, the prevalence of *A. phagocytophilum* was similar to previous reports in the UK. In agreement with previous data from the UK, all tick samples were negative for *Ca.* N. mikurensis.

The present study spotlights multiple areas for future research and suggests further studies to determine *Ca.* M. mitochondrii prevalence and density in different regions of the UK, in case this might affect vector competence. Also, it highlights the need for larger sample sizes to address the pathogen-symbiont interaction (especially partially-fed ticks, which are likely to have a higher pathogen prevalence compared to nymphs).

In addition, the tropical bont tick has a high prevalence of infection with *R. africae* in the Adamawa Region of Cameroon, posing a threat to the health of the local people and tourists. However, the infection rate was significantly higher in females (particularly in ffs females) compared with male ticks; also, as the predominant location of the bacteria is the ovaries, this suggests a similar role for *R. africae* in its tick vector to that of *Ca*. M. mitochondrii in *I. ricinus,* and that could explain increasing of bacterial load during engorgement.

Several ticks feeding on cattle were also infected with *E. ruminantium*, providing evidence for local transmission of the heartwater agent around the

Adamawa Region in Cameroon. Unfortunately, due to the low prevalence of *E. ruminantium* in this study, the statistical power was not sufficient to determine if *R. africae* and *E. ruminantium* densities are positively or negatively correlated.

CHAPTER 3

Co-evolutionary analyses of the symbiosis between *Ixodes ricinus* and the symbiont "*Candidatus* Midichloria mitochondrii" by multi-locus sequence typing

3.1 Introduction

3.1.1 Phylogenetic analysis of populations of *Ixodes ricinus* in the UK and Europe

Ixodes ricinus (the predominant vector of tick-borne diseases in the UK and Europe) harbours *Ca*. M. mitochondrii as described in chapter 2. The role of *Ca.* M. mitochondrii in the biology of its host has not been clarified; more importantly, defining the phylogeography of *I. ricinus* across its range has proved controversial, probably because the substantial size and repeat content of its genome has precluded genome-wide population studies. On the other hand, vector mobility, behaviour and transportation of ticks by large host animals such as deer and birds are closely related to phylogenetic structure and diversity of ticks (Roed et al. 2016). In general, absence of co-cladogenesis between *Midichloria* and ixodid ticks has been reported (Duron et al. 2017).

Originally, it was thought that limited genetic variability exists among *I. ricinus* ticks, on both a national and perhaps even an international scale, although this variability might increase due to changing environmental conditions resulting from global warming (Ishida & Taylor 2007). In previous studies, limited genetic diversity was found on a localised level, such as within a country. For instance, European *I. ricinus* ticks did not seem to show strong phylogeographical structures when comparing mitochondrial DNA sequences (Casati et al. 2008), and analysis of five polymorphic microsatellite loci of *I. ricinus* revealed no significant differentiation between Swiss

populations (Meeus et al. 2002). However, increased diversity across a wider range separated by geographical barriers has been reported in some, but not all, international studies (Noureddine et al. 2011; Dinnis et al. 2014; Roed et al. 2016; Norris et al. 1996). Two mitochondrial and four nuclear genes were used to discriminate Eurasian populations of *I. ricinus* and showed a general lack of genetic variation across Europe. However, divergent clades were identified when comparing European ticks with those taken from North African populations (Noureddine et al. 2011), although North African *I. ricinus* may in fact be a different species, *I. inopinatus* (Estrada-Peña et al. 2014).

Recent studies suggest that geographical barriers such as bodies of water or mountain ranges play an important part in limiting the global spread of ticks, probably through limiting the range and travel of their large mammal hosts. (Estrada-Peña et al. 2013). Few comparative studies have taken place within the UK, but considerable genetic variation was reported between 506 questing nymphs in Great Britain and Latvia using six protein-encoding and ribosomal genes from the mitochondrial genome (Dinnis et al. 2014). In addition, a recent spatial analysis, using a mitochondrial control region gene and the cytochrome b gene of ticks from Northern Europe, Great Britain and the Scandinavian Peninsula, revealed a distinct separation between continental European and British clades. The reasons for this separation could include geographical barriers against mammal migration such as the sea separating Great Britain from the European continent *(*Roed et al. 2016).

Studies of host-associated specialization in *I. ricinus* using 11 microsatellite markers revealed a significant level of genetic variability in a large sample of ticks collected from trapped animals across Western and Central Europe. This study found genetic variability in local populations among both types of host animal and within

ticks infesting an individual host animal, suggesting that *I. ricinus* host usage is not random (Kempf et al. 2011). However, population structure of its close relative *I. scapularis* based on mitochondrial 16S rDNA analysis revealed two distinct lineages between the southern and northern regions of the USA (Rich et al. 1995). Moreover, Norris et al. (1996) produced phylogenetic analyses based on both 16S and 12S mitochondrial genes of ticks from two different locations in the United States. Two distinct populations were reported, one clade included ticks from North-Eastern region, while the other clade was restricted to South-Eastern and Middle-Eastern regions of the USA.

Different methods have been used to investigation the genetic variation within bacterial and fungal pathogens, but multi locus sequence typing (MLST) has great promise for monitoring the spread of pathological agents (Enright et al. 2000). MLST was first developed to identify different strains of *Neisseria meningitidis*, the causative agent of meningococcal meningitis and septicemia in humans (Maiden et al. 1998). Since then, the technique has been applied to many different bacterial species involving those that are difficult to culture.

Our study is the first combined population genetic analysis of *I. ricinus* and *Ca*. M. mitochondrii across a wide swathe of their geographic range using multi-locus sequence typing (MLST) of tick mitochondrial markers and bacterial housekeeping genes.

3.1.2 Phylogenetic analysis of *Ca.* Midichloria, the endosymbiont in *Ixodes ricinus* ticks

As we reported in the previous chapter, the hard tick *I. ricinus* harbours an endosymbiont (*Ca*. M. mitochondrii) with the peculiar capacity to invade and multiply in the mitochondria. This bacterial endosymbiont was previously identified based on 16S rRNA gene sequence data (Beninati et al. 2004; Lo et al. 2006) and an investigation of the genetic diversity of the endosymbiont was carried out for two or three representatives of populations across its geographic distribution by sequencing of the 16S rRNA gene and *gyrB* gene. All 16S rRNA gene sequences from the different populations were found to be almost identical to that originally sequenced from *I. ricinus* from Italy; however, Algerian and Tunisian bacterial sequences shared one base-pair difference when compared to the other sequences. Furthermore, *gyrB* sequences revealed two synonymous substitutions within the Algerian and Irish bacterial populations (Lo et al. 2006).

As the *Ca.* M. mitochondrii endosymbiont was only recently identified (Lewis 1979; Sassera et al. 2006) there have been limited studies on the molecular variation of the bacterium, and indeed there are no studies which begin to examine the potential of co-evolution of the *I. ricinus* host with the *Ca.* M. mitochondrii endosymbiotic bacteria. The present study was designed to investigate the co-evolution of the tick hosts with the endosymbiont using MLST techniques and represents the most in-depth study with the largest number of *Ca*. M. mitochondrii sequences obtained to date.

3.1.3 The aims of the work described in this chapter

The aims of the work carried out here were to develop a MLST scheme for *Ca.* M. mitochondrii using five housekeeping genes (*NUF2, PPDK, ADK, LIPA* and *SECY* genes) and to analyse *I. ricinus* population structure in different parts of the UK and continental Europe for phylogenetic differences using a previously published MLST scheme (Dinnis et al. 2014). The PCR assays used in the study of (Dinnis et al. 2014) were further optimised with the aim of enhancing sensitivity. The final aim was to reconstruct phylogenetic relationships between *I. ricinus* and its endosymbiont *Ca*. M. mitochondrii as co-evolutionary analyses of *Ca*. M. mitochondrii and its tick host could begin to elucidate details of this symbiosis – whether parasitic, mutualist or commensal - as well as its origin and spread.

3.2 Materials and Methods

3.2.1 Sample collection for MLST

Questing or partially-fed adult female *Ixodes ricinus* were collected from Powis Castle Estate, Wales, as described in section 2.1 of chapter 2 (*n* = 12). Ticks collected from Aberdeenshire, Scotland (*n* = 12) were provided by Prof. Lucy Gilbert (the James Hutton Institute Aberdeen); those from Chizé, France (*n* = 16) were supplied by Prof. Olivier Plantard (Oniris, Nantes); and ticks collected from Appennino Tosco-Emiliano, Italy (*n* = 12) and the Zermatt valley, Switzerland (*n* = 12) were kindly provided by Dr. Davide Sassera (University of Pavia, Italy). In addition, 12 samples of *I. ricinus* genomic DNA originating from Bonn, Germany, were kindly provided by Prof. Edward Feil, University of Bath, from an archived collection (stored at -20˚C) used for borreliosis studies (Vollmer et al. 2013). Specimens were identified as *I. ricinus* using morphological criteria (Hillyard 1996). All locations sampled for the MLST analyses

are shown in **Figure 3. 1**). DNA extraction was done using alkaline hydrolysis as explained previously in section 2.2.2. It should be note that the subsequent MLST analysis also includes sequences from the literature; the source of those MLST sequences is highlighted in **Table 3. 1** and the locations are shown in **Figure 3.1.**

Table 3. 1 Locations for existing sequences used in the MLST analysis from the literature.

Location	Country	Reference
Riga	Latvia	Dinnis et al. 2014
Inverness	Scotland	
Exmoor	England	
New Forest		
Richmond Park		
Thurlbear Woods		
Lennestadt-Meggen	Germany	Vollmer et al. 2013
Susten	Switzerland	
Grândola and Mafra	Portugal	

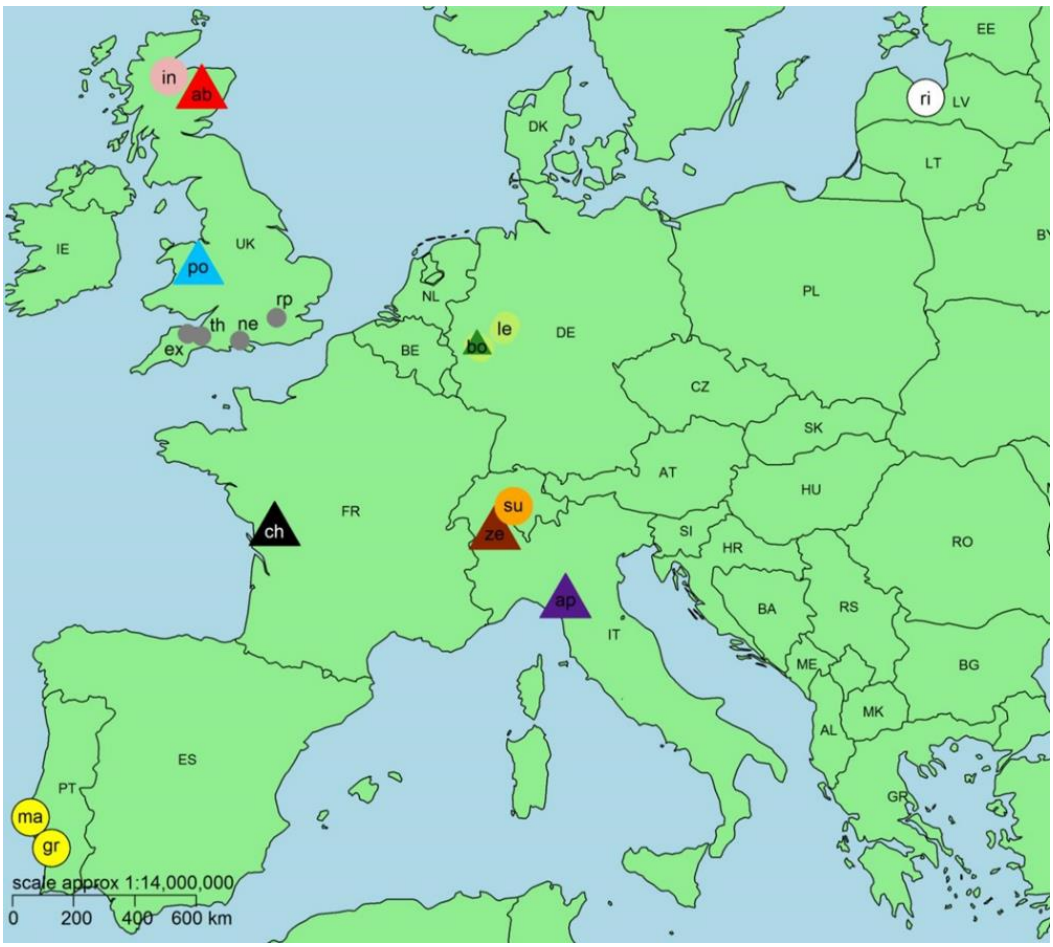


Figure 3. 1 Tick samples locations used in the MLST analysis.

Samples from sites marked by triangles were newly sequenced for this study for both *Ca.* M. mitochondrii and *I. ricinus*. Sites marked by circles represent locations where sequences were obtained previously from an MLST study of the tick only (Dinnis et al. 2014) or during an epidemiological study on Lyme borreliosis (Vollmer et al. 2013).

Key: ab – Aberdeenshire and in - Inverness (Scotland); po - Powis Castle (Wales); ch - Chizé (France); bo – Bonn and le - Lennestadt-Meggen (Germany); ap - Appennino Tosco-Emiliano (Italy); ze - Zermatt valley and su – Susten (Switzerland); ri - Riga (Latvia); ex – Exmoor, ne - New Forest, rp - Richmond Park and th - Thurlbear Woods (England); gr – Grândola and ma – Mafra (Portugal).

3.2.2 MLST design

In order to analyse the evolution of *Ca*. M. mitochondrii using an MLST scheme, oligonucleotides were developed to amplify five housekeeping genes: *NUF2* (kinetochore protein), *ADK* (adenylate kinase), *PPDK* (pyruvate orthophosphate dikinase), *LIPA* (lysosomal acid lipase) and *SECY* (protein translocase subunit) (**Table 3. 2**). These primers were designed by Dr Davide Sassera.

Table 3. 2 PCR oligonucleotides designed to amplify housekeeping genes from *Ca.* Midichloria mitochondrii

Locus	Putative encoded protein	Predicted size (bp)	Gene position [a]	Forward primer (5' to 3')	Reverse primer (5' to 3')
NUF2	Kinetochore protein	648	968094-968609	CTTTATGGACAAGATAGTGCTG	CAGTACGCCTCATAATGGC
PPDK	Pyruvate orthophosphate dikinase	532	219742-220086	GTAAATCCATTCTAGGAGGCAA	ACCAGCATGTTGTAAGACGA
ADK	Adenylate kinase	510	27193-27540	GCGAAATACTTAGGAATGAGGT	AAATCAATCGTCTTATCTCCATCA
LIPA	Lysosomal acid lipase	550	393462-393707	GATATTAGGAAGTGTCTGCAC	GCTGTAGATATTGTCCAATCG
SECY	Protein translocase subunit	550	26532-26870	AAAGTTTATGCAGGAGATTCAAC	GTGAGGAAATAGGTTTGGATTC

a. Gene positions are relative to those from *Ca*. M. mitochondrii strain *IricVA* (Genbank accession number NC_015722)

In order to analyse the evolution of the *I. ricinus* tick host, a previously described MLST scheme (Dinnis et al. 2014) was used to allow for comparison with published datasets. This MLST scheme amplified six different housekeeping genes: *COI* (cytochrome oxidase I), *COII* (cytochrome oxidase II), *COIII* (cytochrome oxidase III), *ATP6* (ATPase 6), *12S* (small RNA subunit) and C*ytB* (cytochrome B) (**Table 3. 3**).

Table 3. 3 PCR oligonucleotides from Dinnis et al. (2014) were used to amplify to amplify housekeeping genes from *I. ricinus* ticks.

Locus	Putative encoded protein	Predicted size (bp)	Gene position [a]	Forward primer (5' to 3')	Reverse primer (5' to 3')
ATP6	ATPase 6	659	3826-4233	AYAAAYYTWTTTTCWATTTTTGATCC	TTAAATTTCRTTWGTRTAWARDGA
CO I	Cytochrome oxidase I	780	1293-1739	ATTTTACCGCGATGAYTWTWCTC	TCCTGTGRAAACARATRATATGGGA
CO II	Cytochrome oxidase II	661	2952-3350	TTTTTCCATGACCATTCAATAATAA	ATAAAGTGGTTTAAGAGACCAATGC
CO III	Cytochrome oxidase III	779	4548-5051	ATGATATTYCAYCCWTTTCAYATAG	AWAYTCATCATTATATRAAWGTAAATA
CYTB	Cytochrome B	780	10247-10616	CCATTCAAATGGAGCATCAA	ACAGGGCAAGCTCCTAAGAA
12S	Small RNA subunit	600	13306-13784	CCCTAATGCAAAAGGTACCCTAA	GCCGCGGTTATACAAGTGAA

a. Gene positions are relative to those from *Ixodes persulcatus* mitochondrion (Genbank accession number KU935457)

3.2.3 PCR assays

3.2.3.1 *Ca*. Midichloria mitochondrii PCR assay

All PCR primers (Invitrogen) were supplied in dehydrated and salted form. Conventional touchdown PCR was performed using 10 µl of 2X BioMix Red (BIOTAQ DNA polymerase, 2 mM dNTPs, 32 mM $(NH4)_2SO_4$, 125 mM Tris-HCL, 0.02% Tween 20, 4 mM $MgCl_2$, stabilizer, inert dye (Bioline, UK) and each primer (2 µl, final concentration = 1 µM), 5 µl nuclease free water (H_2O) and extracted DNA (1 µl) to a total reaction volume of 20 µl. Cycling conditions for the touchdown phase involved initial heating at 94°C for 2 min, DNA melting at 94°C for 30 sec, followed by primer annealing at 65°C for 30 sec, decreasing by 1°C per cycle for 10 cycles (from 65°C to 56°C), followed by extension at 72°C for 30 sec. The amplification phase comprised initial heating at 94°C for 2 min, followed by DNA melting at 94°C for 30 sec, primer

annealing at 55°C for 30 sec, and extension at 72°C for 30 sec for an additional number 29 cycles. A final extension was conducted at 72°C for 10 min.

3.2.3.2 Tick MLST assay

Conventional PCR was performed using 12.5 µl of 2X BioMix Red (BIOTAQ DNA polymerase, 2 mM dNTPs, 32 mM $(NH4)_2SO_4$, 125 mM Tris-HCL, 0.02% Tween 20, 4 mM $MgCl_2$, stabilizer, inert dye (Bioline, UK), 3 µl each primer (final concentration, 1.2 µM), 4.5 µl nuclease free water (H_2O) and extracted DNA (2 µl) to a total reaction volume of 25 µl. Cycling conditions for amplification of loci from *I. ricinus* are shown in **Table 3. 4**.

Table 3. 4 PCR cycling conditions for amplification of loci from *I. ricinus*.
[Note that these have been optimised from a published protocol (Dinnis et al. 2014), for example the increased annealing temperature for *12S* and *COIII* loci].

Loci	Number of Cycles	PCR step				
		Initial heating	Denaturation	Primer annealing	Extension	Final extension
		Temperature (˚C), time (min)				
COI	35	94°, 2:00	94°, 1:00	45°, 1:00	72°, 2:00	72°, 10:00
COII	40		94°, 1:00	57°, 1:30	72°, 2:00	
COIII	40		94°, 1:00	45°, 1:30	72°, 2:00	
CYTB	40		94°, 0:30	52°, 1:30	72°, 1:00	
ATP6	40		94°, 1:00	47°, 1:30	72°, 2:00	
12S	40		94°, 0:30	45°, 1:00	72°, 1:30	

Despite optimisation attempts, in some cases it was not possible to produce a single PCR band, and as a result, bands had to be excised using a PureLink® Quick Gel Extraction Kit (Invitrogen, California, USA) according to the manufacturer's instructions. PCR amplicons were visualised using a UV transilluminator, and purified

with a QIAquick PCR purification kit (Qiagen, Manchester, UK) following the manufacturer's instructions.

For both assays, PCR products were resolved using a 1.5% agarose gel (Bioline) incorporating SYBR® Safe DNA gel stain (Invitrogen, Thermo- Fisher Scientific) and run at 150 V for 30 min in a Bio-Rad gel electrophoresis set. The gels were visualised using a G:Box Gel Documentation System (Syngene). DNA purification was carried out using a Qiaquick PCR purification kit (Qiagen, Germany) according to the manufacturer's instructions.

3.2.4 Sequencing and sequence analysis

Amplified and purified PCR amplicons were sequenced commercially by Eurofins MWG (Germany) and Macrogen (Amsterdam, The Netherlands), using the same forward and reverse primers as were used in the PCR amplification (**Table 3. 2** and **Table 3. 3**). Forward and reverse chromatograms were aligned, and assembled using BioEdit [Ibis Bioscience, Carlsbad; (Hall 1999)]. Traces were checked for errors manually. Sequences of high quality were aligned using Molecular Evolutionary Genetic Analysis (MEGA6) version 6.0 software and aligned using CLUSTALW (Tamura et al. 2013). Alleles and sequence types (STs) were assigned manually and analysed using eBurst (Feil et al. 2004). Individual loci from *Ca*. M. mitochondrii and *I. ricinus* were concatenated separately, manually.

To infer phylogeny for all individual loci, the sequence alignment was subjected to Modeltest as implemented in Topali (Milne et al. 2009), which indicated that the best fitting model for the phylogeny was TN93 (Tamura et al. 2013). In order to analyse phylogenetic relationships of each gene for the bacterial endosymbiont

and the tick host, nucleotide and amino acid maximum likelihood phylogenetic trees (with bootstrap values based on 10,000 iterations) were produced.

Individual gene alignments were concatenated and realigned. Phylogenetic inferences were made as described above using the same TN93 model as predicted by Topali, and maximum likelihood trees were again drawn using 10,000 bootstrap iterations.

Minimum spanning distance trees were drawn using Phylovix (Francisco et al. 2012), and pie charts were produced in Microsoft Excel and transposed onto the minimum spanning tree manually to indicate the origin of each of the samples in each node. Full DNA alignments were screened for recombination using SplitsTrees4 (Huson & Bryant 2006) and were also screened for positive and negative selection using GARD and SLAC as available through the Datamonkey web server (Pond & Frost 2005).

3.3 Results

3.3.1 *Ixodes ricinus*

3.3.1.1 Phylogenetic analysis of *Ixodes ricinus*

The PCR amplification and sequencing of all six gene loci was successfully achieved for 64 ticks from 76 sampled from six different European countries. Of the six genes sequenced, the 12S gene was found to be the most variable, with 171 different sequence types observed across this study and the previous study by Dinnis et al. (2014), compared to 106 different alleles seen within the *COII* gene. However, within our study only, the most variable gene was *CYTB*, which had 43 different alleles; but *COII* was still the least variable gene. Interestingly, the *COI* gene showed the highest level of amino acid mutations, suggesting that most substitutions were

non-synonymous, whereas *COII* was the least mutated at the amino acid level. Despite numerous mutations being seen within the *CYTB* gene in this study, the majority of these were synonymous, resulting in only 27 new amino acid STs (AA STs) (**Table 3. 5**). All gene sequences were submitted to Genbank, under the Accession numbers ([*ATP*- MH334375 - MH334945]; [*COI*- MH336088 - MH336658]; *[COII*- MH334946 - MH335516]; [*COIII*- MH336659 - MH337229]; [*12S*- MH333804 - MH334374] and [*CYTB*- MH335517 - MH336087]).

The sequences obtained for the six individual tick loci (*COI* [cytochrome oxidase I], *COII* [cytochrome oxidase II], *COIII* [cytochrome oxidase III], ATP6 [ATPase 6], *12S* [small RNA subunit] and C*ytB* [cytochrome B]) were used to make a concatenated sequence alignment of 2,410 bp. Selection pressure analysis showed several negatively selected residues within the ATP6 gene (residues 60, 71, 93). Selection pressures appear to play a part in the evolution of the *COI* gene, as one residue was positively selected and residue 27 was negatively selected, despite mainly non-synonymous mutations being observed in this gene. Residues 12 and 72 were negatively selected in *COII*, residues 36 and 39 were under negative selection pressure in in *COIII* and residue 72 was negatively selected in *CYTB*. In contrast, there was no evidence of selection pressures in the 12S gene. Overall, a generally high level of recombination was observed using SplitsTrees analysis (data not shown).

3.3.1.2 Comparison of MLST *I. ricinus* tick loci

Comparison of the sequence data revealed differing levels of variability at all loci. The average dissimilarity between loci of *I. ricinus* was 11%, with a range from 9.8% in *COI* to 11.9% for *COIII* and *12S* (**Table 3. 5**). At the amino acid level, the average dissimilarity between loci was 22%, with a range from 11.3% in *COII* to 26.8% in *COIII*.

The number of new alleles detected within this study ranged from 16 (COII) to 43 (CYTB) (Table 3. 5). Sequence types (STs), assessed based on the MLST allelic profiles revealed 64 STs, of which none were previously reported in Dinnis et al 2014. Of these STs, only one was seen more than once, with the remaining STs being singletons.

3.3.1.3 Variation amongst *I. ricinus* ticks

A maximum-likelihood (ML) phylogenetic tree with 10,000 bootstrap replicates, prepared from the concatenated sequences obtained from the 64 ticks, revealed moderate variability of tick mitochondrial genes **(Figure 3. 2**). However, some geographic clustering of ticks from Scotland and Wales was observed. Indeed, most of the Scottish sequences grouped separately from the European sequences, indicating that they formed a distinct clade. Nucleotide variability was seen at each of the six different loci tested, and no areas of notable hyper-variability were seen within any gene (**Table 3. 5).**

Table 3. 5 Comparison of variation at dfferent loci for 64 *I. ricinus* ticks sequenced as part of this study.

Analysis at both the nucleotide and amino acid levels are shown.

Gene Name	Nucleotide								Amino Acid				
	Number of Alleles*	Number of new alleles*	Length (bp)	Gene product	GC content	Polymorphic sites %	Parsimony informative sites	dN/dS	No. of AA ST*	Number of new AA ST*	Length (aa)	Polymorphic sites %	Parsimony informative sites %
ATP6	127	29	411	ATPase 6	21.1% (87/411)	11.4% (47/411)	4.1% (17/411)	0.29602	62	22	137	22.6% (31/137)	10.2% (14/137)
COI	132	29	447	Cytochrome oxidase I	29.7% (133/447)	9.8% (44/447)	4.5% (20/447)	0.92570	114	33	149	24.2% (36/149)	12.0% (18/149)
COII	106	16	399	Cytochrome oxidase II	29.3% (117/399)	13.0% (52/399)	5.8% (23/399)	0.70581	33	7	133	11.3% (15/133)	4.5% (6/133)
COIII	131	23	504	Cytochrome oxidase III	26.9% (136/504)	11.9% (60/504)	5.4% (27/504)	0.65842	87	14	168	26.8% (45/168)	13.0% (22/168)
12S	171	19	285	Small RNA subunit	20.0% (57/285)	9.8% (28/285)	3.9% (11/285)	1.30853	103	39	95	20.0% (19/95)	10.5% (10/95)
CYTB	154	43	371	Cytochrome B	24.5% (91/371)	10.0% (37/371)	4.0% (15/371)	0.149389	21	27	123	25.2% (31/123)	11.3% (14/123)

*Indicates total sequence types including those previously reported (Dinnis et al. 2014) and new STs sequenced as part of this study.

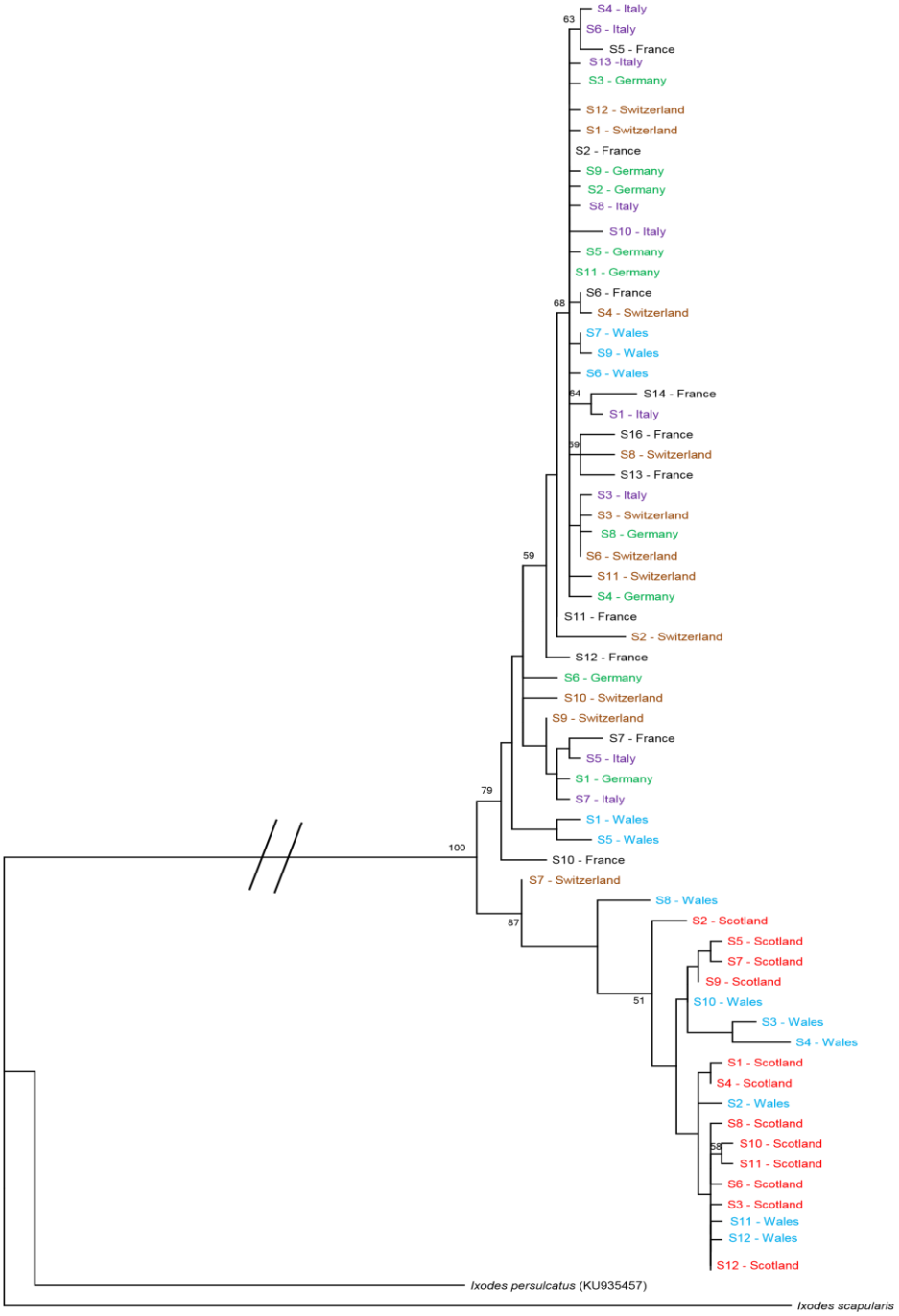


Figure 3. 2 Maximum-likelihood phylogenetic tree based on concatenated mitochondrial genes from 64 *I. ricinus* ticks.

The corresponding mitochondrial DNA sequences of *I. persulcatus* obtained from NCBI (accession no. KU935457) were used to root the tree. There was a total of 2,410 nucleotide positions in the final dataset.

3.3.1.4 Comparison of *I. ricinus* tick sequence data generated here with previous studies

Phylogenetic comparison of the concatenated loci produced as part of this study with those from a previous study (Dinnis et al. 2014) revealed a generally low level of diversity across the 64 sequences (Figure 3. 3). In total, 63 unique STs were seen (**Appendix D**).

Clustering of Scottish and Welsh sequences from this study with those previously sequenced was observed (Dinnis et al. 2014). There was a separation into two separate clades, one of which contained mainly Welsh and Scottish sequences, while the other contained mainly European and English sequences. However, not all of the English isolates clustered together with Welsh and Scottish isolates (**Figure 3. 3**).

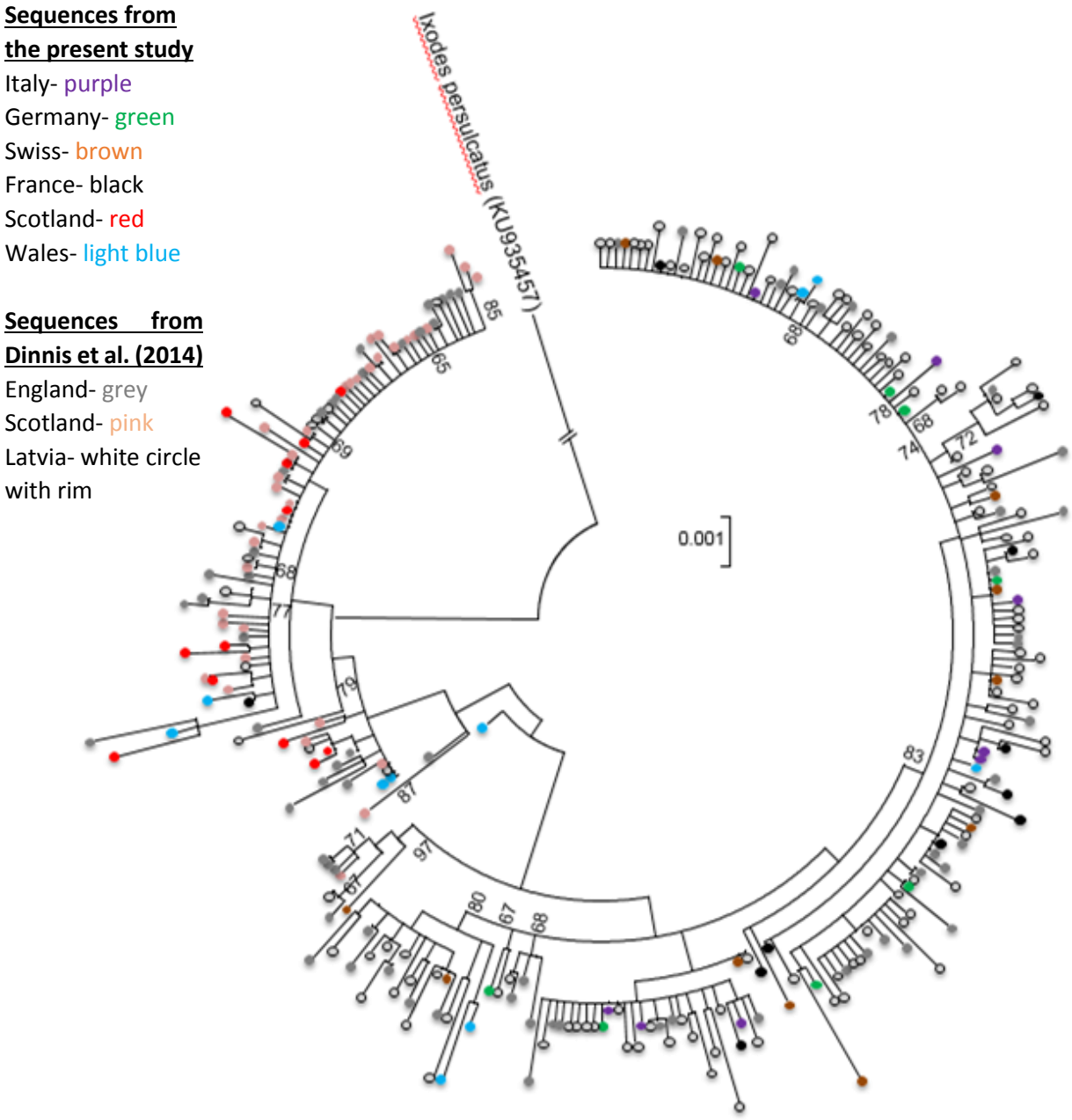


Figure 3. 3 A circularised maximum likelihood tree.

Drawn using the general time reversible model with 10,000 bootstrap iterations for the tick sequences generated as part of this study compared to each unique tick sequence from a previous study (Dinnis et al. 2014). Sequence names were replaced with colour coded circles corresponding to the country/countries where the tick sequences were obtained: Scotland – red; France – black; Switzerland - brown; Italy – purple; Germany -, green; Wales – light blue; Dinnis et al (2014) comprising: England- grey; Latvia – white circle with rim; and Scotland-pink.

3.3.1.5 Minimum spanning trees and Scottish and Welsh tick sequence segregation

To allow for comparison of similarities of different STs and their relatedness, a minimum spanning tree was drawn to compare the sequences generated as part of this study. Sequences which are close to each other on the tree are generally only different at a single locus, whereas the more distant sequences have fewer loci in common.

Minimum spanning distance data suggest that ST 459 was the linker ST between the mainland European cluster and the Welsh/Scottish cluster. Another cluster, composed mainly of European samples, centred around the founder ST (ST 478), with another 27 single STs radiating out from this. The Welsh and Scottish sequences formed eleven single locus variants (SLV) STs, which centred around the founder ST (ST 461). In general, Scottish and most of the Welsh sequences form a distinct cluster away from mainland European sequences, which in turn were usually separated from the English samples (**Figure 3. 4)** and (**Appendix D**)

Interestingly, the only ST which was seen more than once (ST 478) (as described above) was the founder sequence of a large European clade (**Figure 3. 4**). Again, there was a notable distinction of European sequences away from the Welsh and Scottish ones (**Figure 3. 4**), further suggesting a degree of genetic separation between the two.

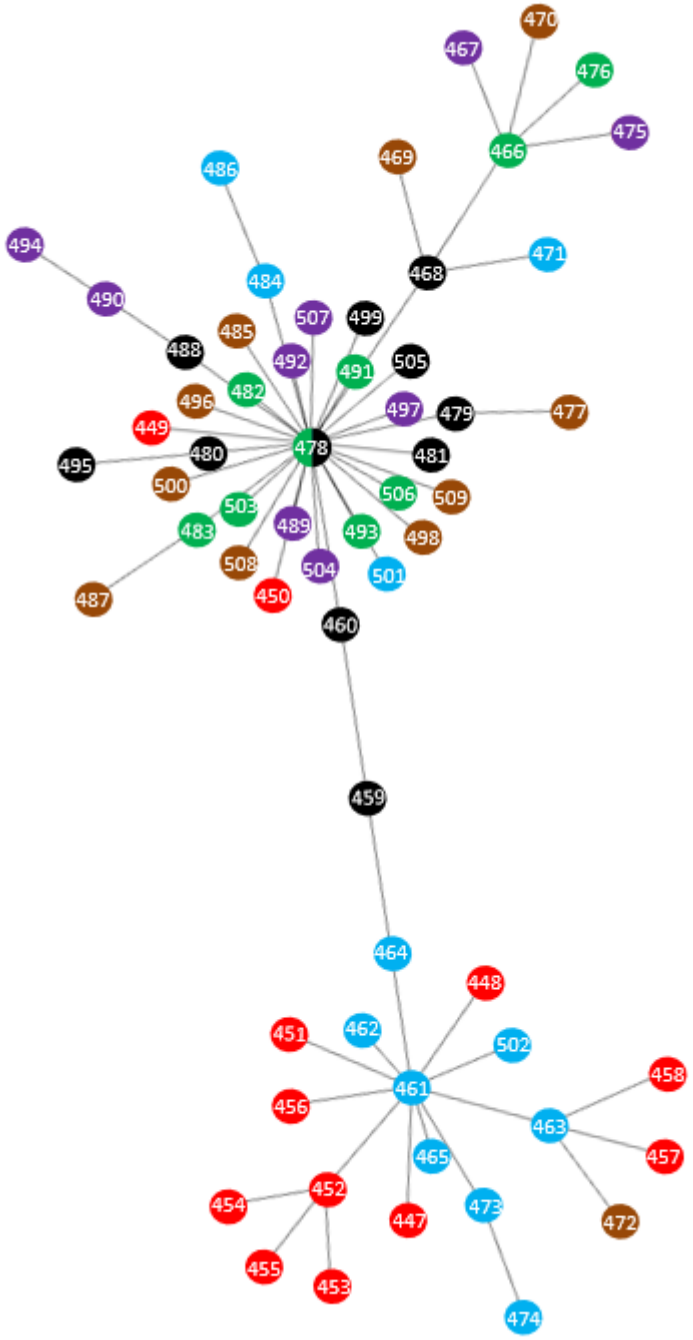


Figure 3. 4 Minimum-spaning distance tree of 64 *I. ricinus* specimens.
The ST numbers are shown inside the nodes as pie charts, representing the different country of origin of those sequences. France – black, Italy – purple, Germany – green, Switzerland – brown, Wales - blue, Scotland - red. A full list of STs is provided in (**Appendix D)** for the tick.

3.3.1.6 Comparison of ST data with previous studies

As there was evidence of clustering of Welsh and Scottish sequences within this study, it was of interest to compare them to previous studies and see if this was confirmed. Minimum spanning trees were drawn for the 570 sequences from this study and a previous study (**Figure 3. 5).** Once again, there was a distinct separation between mainland European (right side of figure 3.5) and, in this case, British (English, Scottish and Welsh) sequences on the left of Figure 3.5. However, a few sequences from Europe were found in the British clade, and vice versa, although this was relatively rare.

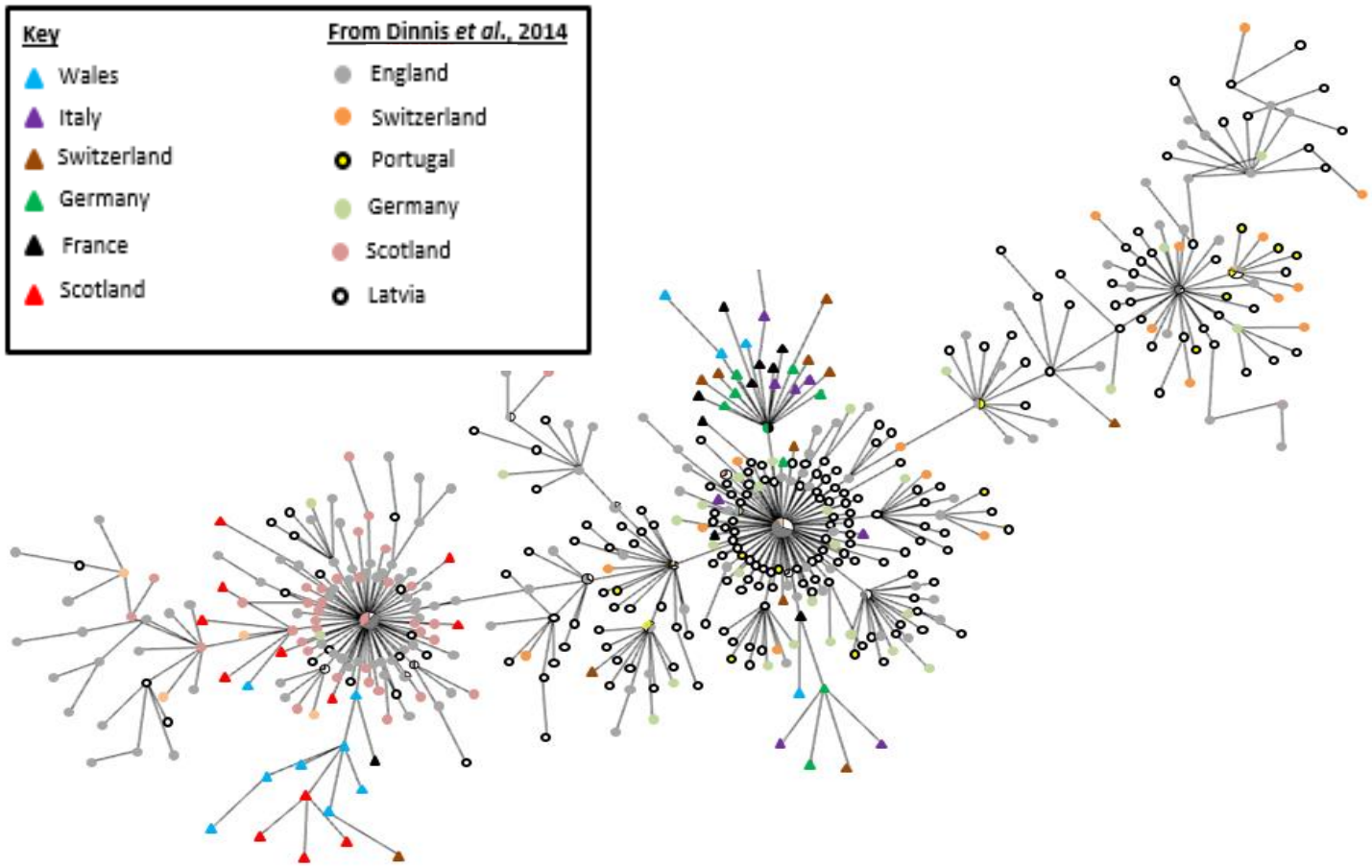


Figure 3. 5 Minimum-spaning distance tree of *I. ricinus* mitochndrial sequences. Previously-acquired data (circles), some of which has been published (Dinnis et al. 2014), are incorporated into the tree alongside newly-obtained sequences (triangles, this study).

3.3.2 *Ca.* Midichloria mitochondrii

3.3.2.1 Phylogenetic analysis

In order to estimate the genetic divergence of endosymbiont *Ca.* M. mitochondrii across its *I. ricinus* tick hosts, PCR amplification was achieved for all five genes (*NUF2* [Kinetochore protein], *ADK* [Adenylate kinase], *PPDK* [Pyruvate orthophosphate dikinase], *LIPA* [Lysosomal acid lipase] and *SECY* [Protein translocase subunit]) for 64 samples out of 76 ticks, and these were included in the eventual analysis. Of these five genes, *PPDK* was found to be the most variable, with five different sequence types (STs) observed, and *LIPA* was found to be the least variable, with all sequenced samples identical. Additionally, there were four different AA STs for the *PPDK* gene, suggesting that most of these changes were non-synonymous mutations (**Table 3. 6).**

The sequences obtained for the five *Ca*. M. mitochondrii loci were concatenated into a sequence of 1,853 bp. There was no evidence of positive or negative selection (below p value of 0.2) for any of the *Ca.* M. mitochondrii sequences. *Ca*. Midichloria mitochondrii recombination analysis was attempted, but was found to be quite limited (data not shown). All gene sequences were submitted to Genbank, under the Accession numbers of ([*ADK*- MH295309 - MH295373]; [*LIPA*-MH295374 - MH295438]; [*NUF2*-MH295439 - MH295503]; [*PPDK*- MH295504 - MH295568]; [*SECY*- MH295569 - MH295633]).

A maximum-likelihood (ML) phylogenetic tree with 10,000 bootstrap replicates is shown in Figure 3.6. Analysis of the MLST concatenated gene sequences for *Ca*. M. mitochondrii indicated a generally low level of diversity, with notable geographic clustering of bacteria from ticks collected in Scotland and Wales. Some of the Scottish sequences form a distinct cluster away from other sequences suggesting

possible localised evolution. Sequence variation at most loci (except *LIPA*) was identified, but no hypervariable regions were observed within any gene (**Table 3. 6).**

3.3.2.2 Comparison of *Ca*. M. mitochondrii MLST loci

Comparison of the sequence data (as described above) revealed differing levels of variability at all loci, except for *LipA* which was wholly conserved throughout all sequences.

The average dissimilarity between loci of *Ca.* M. mitochondrii was very low at 0.54%, with a range from no variation within the *LipA* locus to 1.2% seen within *PPDK* (**Table 3. 6)**. At the amino acid level, only two loci (*ADK* and *PPDK*) showed any diversity, suggesting that these genes are highly conserved, and that the majority of DNA mutations are synonymous (**Table 3. 6**).

The number of DNA alleles within this study ranged from one (*LipA*) to five (*PPDK*) (**Table 3. 6**). Sequence types assessed based on the MLST allelic profiles revealed 14 STs, of which five were seen more than once, with the remaining STs being singletons. The most common ST was ST 8 which contained 28 (44%) of the 64 endosymbiont sequences, with STs 3 and 7 containing 14 and 8 sequences, respectively. Less variation was observed with the *Ca.* M. mitochondrii MLST sequences compared to those seen in the *I. ricinus* tick hosts.

Table 3. 6 Comparison of variation at different loci for the *Ca.* M. mitochondrii endosymbiont populations sequenced as part of this study. Analysis at both the nucleotide and amino acid levels are shown.

Gene name	Nucleotide							Amino Acid			
	Number of DNA alleles	**Length (bp)**	**Gene product**	**GC content (%)**	**Polymorphic sites (%)**	**Parsimony informative sites (%)**	**dN/dS**	**No. of AA alleles**	**Length (aa)**	**Polymorphic sites %**	**Parsimony informative sites %**
ADK	3	349	Adenylate kinase	36.7% (128/349)	0.6% (2/349)	0	n/a	3	116	1.7% (2/116)	0
LipA	1	247	Lysosomal acid lipase	38.7% (106/247)	0	0	n/a	1	82	0	0
NUF2	4	516	Kinetochore protein	42.2% (218/516)	0.6% (3/516)	0.4% (2/516)	5.00E-09	1	172	0	0
PPDK	5	337	Pyruvate orthophosphate dikinase	38.3% (129/337)	1.2% (4/337)	0.6% (2/337)	0.780116	4	112	2.7% (3/112)	1.8% (2/112)
SecY	2	339	Protein translocase subunit	34.5% (117/339)	0.3% (1/339)	0	n/a	1	113	0	0

3.3.2.3 Variation of *Ca*. M. mitochondrii

Phylogenetic analysis of these concatenated *Ca*. M. mitochondrii sequences revealed localised clustering of tick sequences taken from Scotland and Wales, compared to a generalised mixing of sequences from continental Europe (Italy, France, Germany and Switzerland) (**Figure 3. 6**).

Sequence types (STs) were assigned based on the MLST allelic profiles. Comparison of allelic profiles revealed a total of 14 unique STs, of which ST8 was dominant, containing 44% (28/64) of the total endosymbiont sequences (**Appendix C**), while ST3 was the next most dominant, containing 21% (14/64) of the sequences **(Appendix C**), implying low sequence variation.

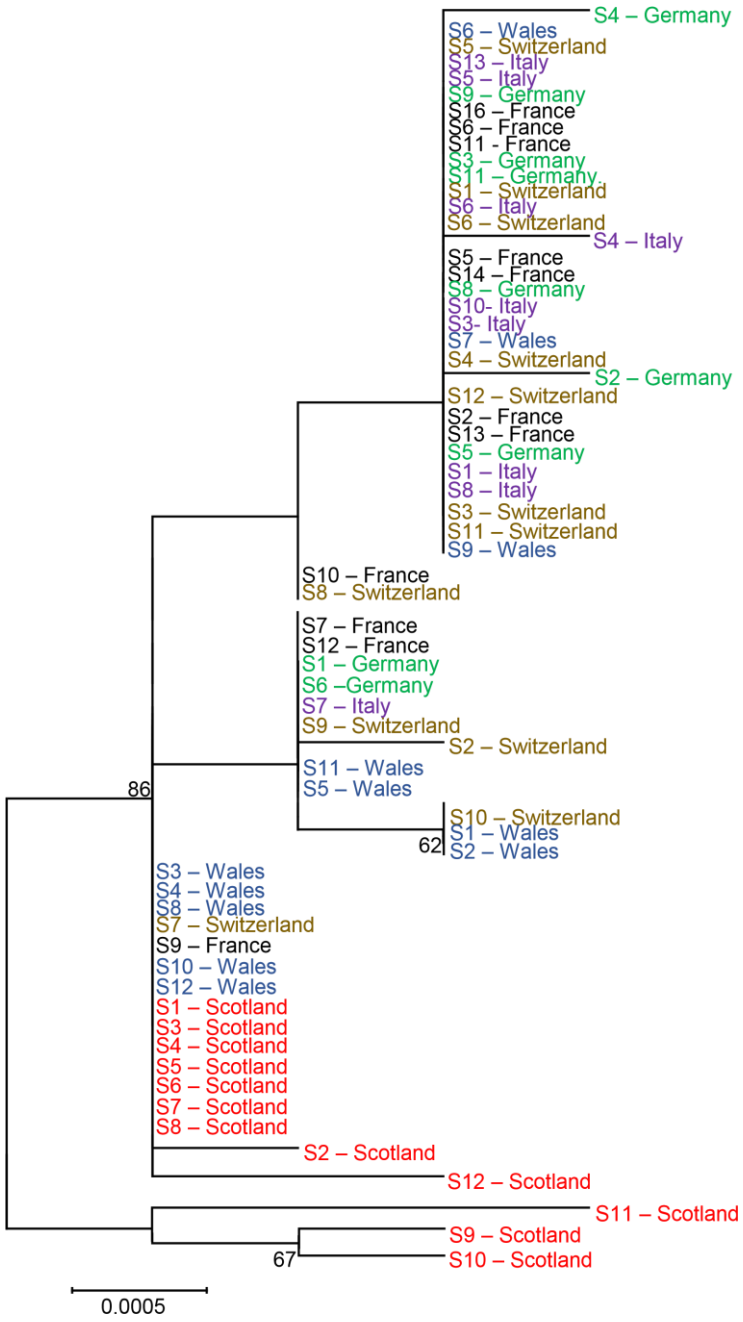


Figure 3. 6 Phylogenetic reconstruction inferred from the concatenated endosymbiont houskeeping gene fragments.

Five loci (LIPA, NUF2, ADK, PPDK and SECY) of *Ca*. M. mitochondrii endosymbionts from 64 tick samples in four mainland European countries and two locations in the UK using the Maximum-Likelihood method (10,000 replicates). There were a total of 1,853 nucleotide residues in the final dataset. The scale bar represents 0.05% nucleotide difference. Bootstrap values are shown on the branch nodes. Bar: 0.0005 inferred substitutions per site. Note that samples are colour coded based on their origin- blue- Wales; red- Scotland; black- France; brown- Switzerland; purple- Italy.

3.3.2.4 Minimum spanning trees confirm Scottish and Welsh tick sequence segregation

Minimum spanning trees compare similarities among different STs and how closely related they are. Thus, as explained above, samples situated close to each other on a tree are generally diverse at a single MLST locus, whereas more distant isolates have fewer loci in common. The *Ca*. M. mitochondrii minimum spanning distance tree (**Figure 3. 7**) suggested that three STs (ST8, 3 and 7) were founder sequences, with ST 7 and ST 3 being the two which gave rise to the cluster of Scottish sequences. ST 8 clustered with three other STs as SLV of this founder; the neighbouring ST 7 had eight SLV and ST 3 had two SLV evolved from them. As seen with *I. ricinus* ticks, there was a notable grouping of Scottish *Ca*. M. mitochondrii sequences (**Figure 3. 7**) (**Appendix C**).

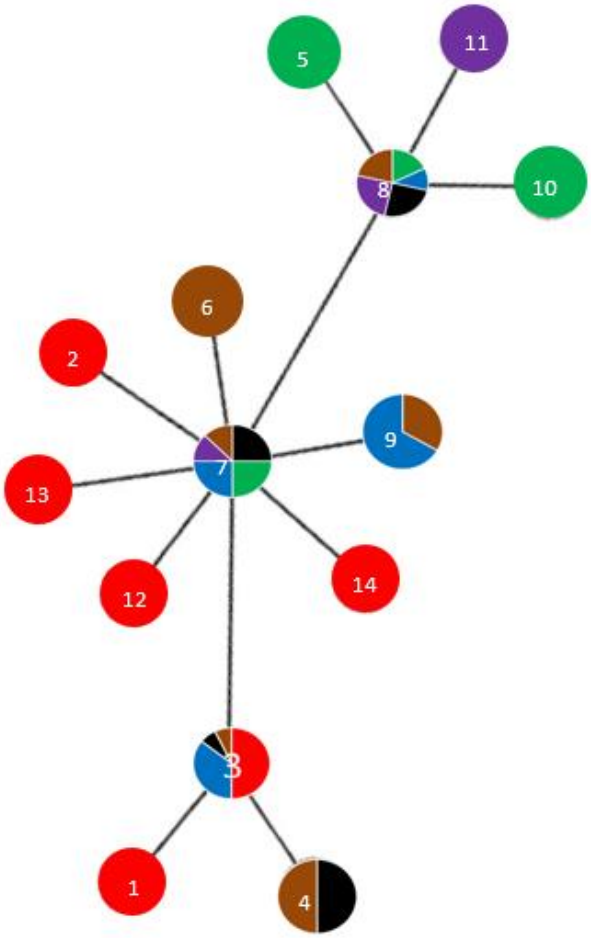


Figure 3. 7 Minimum spanning distance tree of sequence types obtained from 64 *Ca*. M. mitochondrii samples sequenced as part of this study.
Sequence type numbers are shown inside the nodes which are pie charts, representing the different country of origin of those sequences. Further details of the STs are shown in (Appendix D). France - black; Italy - purple; Germany -, green; Switzerland - brown; Wales - light blue; Scotland - red.

Applying a minimum-spanning phylogenetic analysis to the *Ca*. M. mitochondrii data revealed that three STs (8, 3 and 7) were founder sequences, with STs 7 and 3 representing those which gave rise to the cluster of Scottish sequences (**Figure 3. 7**).

3.3.3 Co-evolution of *I. ricinus* and *Ca*. M. mitochondrii

As *Ca*. M. mitochondrii is an endosymbiont of *I. ricinus,* it was of interest to compare the phylogeny of the two organisms. Individual maximum likelihood trees were placed next to each other, and compared, allowing tree topologies for the tick host and endosymbiotic bacteria to be examined for possible co-cladogenesis. (**Figure 3. 8**)

In total, 54 out of 64 sequences showed similar positions on the two trees, suggesting a significant degree of co-evolution. Scottish sequences clustered together, revealing probable coevolution between the tick and symbiont, but geographic clustering was not strong for the samples from the other geographical locations. The seven incongruent pairs in the tree included one specimen from Scotland, two from Wales and three from Switzerland, which provides evidence for occasional horizontal transmission. (**Figure 3. 8**)

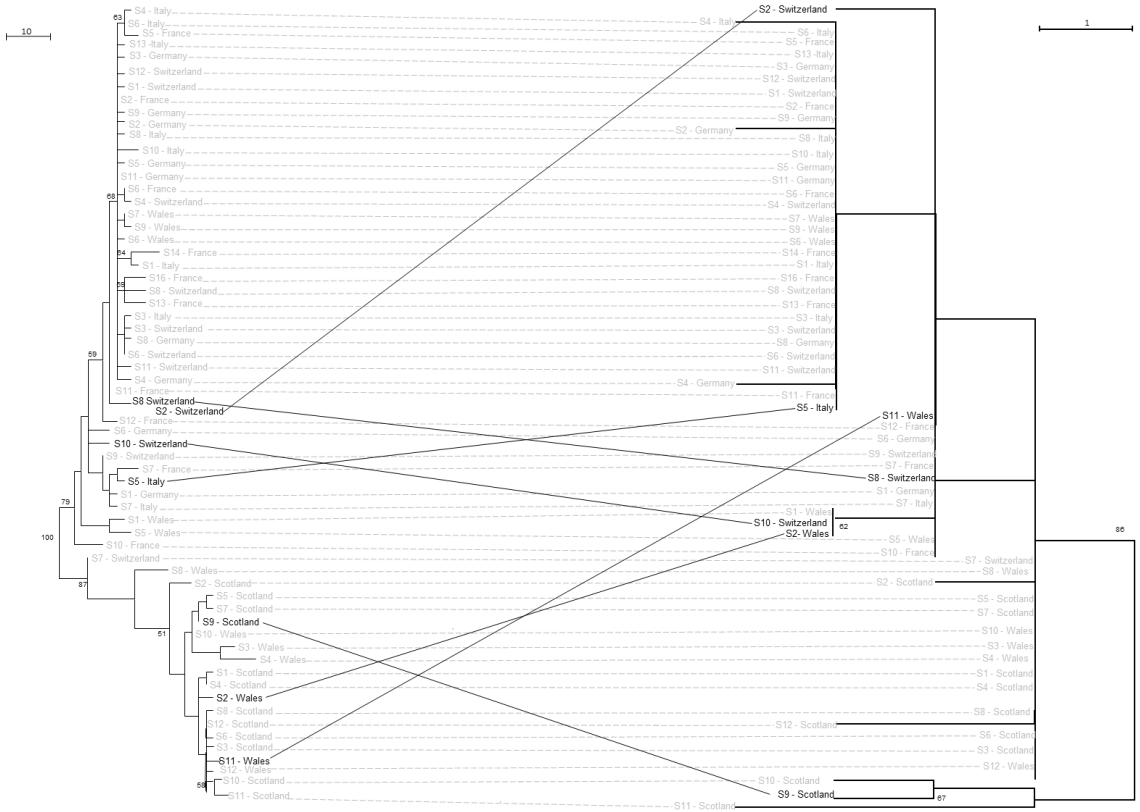


Figure 3. 8 Comparison of the phylogenetic reconstruction of 64 *I. ricinus* tick samples and their *Ca*. M. mitochondrii endosymbionts.

Tick (lefthand side), and *Ca.* M. mitochondrii endosymbiont (righthand side) trees are shown. Dashed lines join endosymbionts and tick hosts on the respective trees.

3.4 Discussion

Overall a generally low level of genetic diversity was seen among the *I. ricinus* ticks, which is in agreement with previous studies (Dinnis et al. 2014; Roed et al. 2016). There was notable separation of sequences obtained from ticks in Scotland and Wales in this study when compared with those obtained from mainland continental Europe, including Germany, Switzerland, France and Italy, where there was no evidence of geographical clustering. This suggests genetic admixture of tick populations within Europe, but inter breedingwith populations in the British Isles is limited due to the English Channel. As these ticks are usually carried on large mammals including sheep and deer (Roed et al. 2016; Pietzsch et al. 2005; Walker 2001), this lack of mixing across a large body of water seems unsurprising. This lack of genetic clustering agrees with other studies (Dinnis et al. 2014; Lo et al. 2006; Roed et al. 2016).

Although there is no current data regarding the evolution or variation of Ca. M. mitochondrii*,* within this study, the symbiont exhibiting very low levels of sequence diversity, although a consistent signal of host-symbiont coevolution was apparent in Scotland. The low level of variation among *Ca*. M. mitochondrii sequences suggests that vertical transmission of the endosymbiont may be the main route of spread (Beninati et al. 2004). However, it has been reported that *Ca*. M. mitochondrii can be released with the tick saliva into hosts during blood feeding of the tick (Mariconti et al. 2012). Additionally, it has been shown that these endosymbiotic bacteria can stimulate an immune response in the host (Sassera et al. 2008). Therefore, should horizontal transmission via blood feeding occur, it would be

likely that less geographical clustering would occur as the bacteria would routinely spread by tick bite (Epis et al. 2008).

For *I. ricinus*, there was evidence of genetic clustering of sequences from Scotland and Wales, but this was not seen elsewhere across Europe, adding further weight to the view that spread by large mammals between the British Isles and Europe is limited because of the marine barrier. Immature stages including larvae and nymphs infest small mammals and birds, while adult stages feed on larger animals such as deer, and to allow for optimal tick survival, both types of animals are required (Pietzsch et al. 2005).

For both *I. ricinus* and *Ca*. M. mitochondrii*,* the clustering of Scottish and Welsh samples was confirmed using both maximum likelihood and minimum spanning phylogenetic tree analyses, and was further confirmed with comparison of over 500 different sequences from a previous study (Dinnis et al. 2014). However, previous English isolate sequences (Dinnis et al. 2014) appear to be dispersed along with the European samples and do not cluster together with the present Welsh and Scottish isolates.

The data presented here and by Dinnis et al. (2014) may suggest a degree of genetic separation of ticks that previously invaded from Europe to develop a new clade on the British Isles. This coupled with local environmental conditions and host selection may be leading to genetic drift to form a new subspecies of tick within Britain. Additionally, the perceived lack of spread of ticks from mainland Europe to the British Isles suggests that there is limited risk of the spread of diseases across the English Channel and formation of infection foci or epidemics with Britain, and vice versa. However, the risks of tick importation by companion animals may increase now

that there is no requirement for pets taken on holiday from the UK to Europe to be acaricide-treated before they return.

Examination of both mitochondrial and nuclear markers at various scales across Europe (both on the continent and in the British Isles), North Africa, and Western Asia failed to find an association between haplotype and geographic origin, with the exception of clearly differentiated populations in North Africa (Noureddine et al., 2011). However, these North African specimens are now suspected to be from a different tick species, *I. inopinatus* (Estrada-Pena et al. 2014). In contrast, analysis of microsatellite markers using several hundred ticks has provided evidence for distinct "races" of *I. ricinus* collected from different host species in France, Belgium and Slovakia; for instance, on roe deer and wild boar (Kempf et al. 2011). A whole mitochondrial genome (hereafter ''mitogenome") analysis of two *I. ricinus* populations in northern Italy, within a narrow spatial scale over which these samples were obtained, revealed four highly divergent lineages but no geographical structuring (Carpi et al. 2016). Overall, our findings correlated with the conclusion of Dinnis et al. (2014) and Road et al. (2016) in demonstrating the marine barrier to gene flow between the British Isles and continental Europe. In addition, our data agreed strongly with those of Road et al. (2016), who suggested STs of British origin are much more common among the continental European clades than vice-versa, reflecting greater easterly than westerly gene flow. This finding could be related to spring migrations of birds between the British Isles and continental Europe (Randolph et al. 2002), during which considerably more migrating birds carrying attached immature tick stages travel towards continental Europe compared with the opposite direction during autumn, after which successful tick establishment in the UK would be rarer due to winter attrition. While the success rate of moulting to the next stage in the

lifecycle has been estimated to be only 10% for *I. ricinus* (Randolph et al. 2002), it is possible that the spring migration of birds into the UK is accompanied by >1 million immature ticks (Pietzsch et al. 2008), therefore potentially contributing ~10,000 adult ticks from overseas that were transported by birds as larvae or nymphs.. Migrating birds may spread ticks over relatively long distances between feeding and migration destinations (Pietzsch et al. 2008; Comstedt et al. 2006). It is likely that the survival rate for this will be low, and this could be a potential reason for the low level of genetic mixing between Europe and the British Isles. However, it would be possible for tick nymphs to be carried short distances within a country or between neighbouring countries, perpetuating localised spread and dispersal.

It has been suggested that larger mammals such as sheep and deer are more important in spreading of ticks locally than smaller hosts (Randolph et al. 2002). Alternatively, Kempf et al. (2011) recently suggested that it was mainly male ticks which were carried long distances, and males were not included in the present MLST study, but this should be part of future studies.

Multi locus sequence typing of *Ca*. M. mitochondrii in the present study indicated that the ticks of Scotland formed a separate clade that contains STs from other parts of Britain, but very little contribution from continental Europe. This suggests that the evolutionary history of the tick-symbiont relationship in these parts of the British Isles has been subject to either genetic drift caused by a population bottleneck, or selection on ticks and their symbionts for traits that are important for reproductive success in a certain location.

Scotland has a number of biogeographical features that separate it from the rest of the British Isles. The border region between England and Scotland, the Southern Uplands, is a hilly landscape and further north, the Grampian Mountains are a

substantial barrier between the Central Lowlands and the Northwest Highlands. Nevertheless, these ranges of hills and mountains do not reach the lateral extent or altitude of the Alps and other mountain ranges in continental Europe where gene flow between *I. ricinus* populations appears to occur unhindered by barriers. Taking these into consideration would highlight the significant importance of the prehistory of the region compared with its current topology, because the locations where the Scottish ticks were sampled (Aberdeenshire and Inverness-shire) were covered by ice-sheets as recently as 15,000 years ago - some 10,000 years after the British-Irish Ice Sheet began to retreat at its southerly margins - and may have experienced re-advances even beyond this date (Clark et al. 2012).

The establishment in post-glacial Scotland of *I. ricinus* is likely to have been slow and irregular, perhaps leading to strong founder effects at this north-western extent of the species' range. Phylogeographic studies of red deer (*Cervus elaphus*) suggest that populations in Western Europe, including the British Isles, originate from a refugium located on the Iberian Peninsula, with little evidence that human actions (particularly translocations) have impacted considerably on natural migrations (Skog et al. 2009). Even currently, when populations of both native and imported deer species are increasing considerably across Britain (Putman et al. 2011), densities of red deer in the Scottish Highlands significantly exceed those in most other parts of the country (Edwards & Kenyon 2013).

According to the above, migrations of red deer brought *I. ricinus* to Scotland after the glacial retreats, and differences in the most abundant larger hosts between Scotland and other parts of Europe may have acted to select a "race" of *I. ricinus* that is better adapted for feeding on *C. elaphus* (Kempf, et al. 2011). In addition, other factors including the longer winter at higher latitudes resulting in higher attrition rates when

ticks are dormant, and furthermore the reduced duration of the questing season, may have also played a role in selecting for certain tick and symbiont genotype combinations.

Alternatively, this strong founder effect in Scottish tick populations might be the result of bird migrations. Most records of *I. ricinus* on birds in the UK come from passerines such as the blackbird (*Turdus merula*), the willow warbler (*Phylloscopus trochilus*), the whitethroat (*Sylvia communis*), the song thrush *Turdus philomelos* and the dunnock *Prunella modularis* (Pietzsch et al. 2005), some of which migrate, suggesting random introductions from bird migrations. A link between tick infestation and a tendency to forage on the ground was reported from one bird study in Scotland (James et al. 2011). However, while these common songbirds differ significantly in their migratory habits, none are restricted to Scotland.

Nevertheless, during the early post-glacial period, the transport of ticks by migratory birds into Scotland may have had a greater impact compared to other parts of the British Isles, perhaps supplanting existing clades through drift. In addition, some British avian *I. ricinus* records are from rare migrants that have almost never been observed outside Scotland, such as the Pechora pipit (*Anthus gustavi*) (Pietzsch et al. 2005); although as most sightings are from Shetland, it is very unlikely that such rarities have impacted on the mainland tick population.

Obligate mutualist symbionts that are transmitted vertically down the maternal line are expected to show strict co-cladogenesis with their host as previously suggested for *Buchnera aphidicola* in aphids (Funk et al. 2000); however, while our findings correlated with this pattern of co-evolution, the poor phylogenetic signal hindered efforts to clearly identify co-cladogenesis in the populations from continental Europe. Our findings agreed consistently with the previous report which

examined just two molecular markers (Lo et al. 2006), and revealed low diversity in the symbiont, which suggests a recent selective sweep through *I. ricinus* populations, at least on mainland Europe. On the other hand, this scenario would be incompatible with complete mitogenome data obtained in northern Italy, in which the four identified lineages were estimated to coalesce around 427,000 years ago, suggesting that *I. ricinus* has maintained a large population size since the Pleistocene (Carpi et al. 2016).

As mitochondria and vertically-transmitted symbionts are generally considered to be almost exclusively transmitted by the same route, a symbiont-mediated selective sweep would be expected to markedly reduce mitochondrial polymorphism (Cariou et al. 2017). The strong congruent relationship highlighted in the present study reflected significant co-cladogenesis between *Ca*. Midichloria symbionts and their tick hosts. Nevertheless, a critical role for horizontal transmission in this *Ca*. M. mitochondrii sweep might suggest a level of contradiction. Horizontal transmission of the symbiont, possibly by co-feeding (Cafiso et al. 2016; Epis et al. 2008), is compatible with data indicating the presence of the bacterium in the salivary glands and transmission to the host via the blood meal (Bazzocchi et al, 2013; Mariconti et al. 2012). Although only single genotypes of *Ca*. Midichloria symbionts were detected in the present study, co-infection with >1 genotype of a bacterial species should be taken into consideration, as genetic heterogeneity has been suggested before between *Borrelia* genospecies during co-infections (James et al. 2014).

CHAPTER FOUR

Identifying *Anaplasma* proteins involved in host-bacteria interactions *in vitro*

4.1 Introduction

Anaplasma phagocytophilum is transmitted during blood feeding of (primarily) *I. scapularis* ticks in the Eastern USA and *I. ricinus* in the UK and Western Europe. *Anaplasma phagocytophilum* is the causative agent of human granulocytic anaplasmosis (HGA) and is a significant public health concern. The human disease is predominantly prevalent within the USA, and is found alongside canine granulocytic anaplasmosis in dogs and equine granulocytic anaplasmosis in horses (Zivkovic et al. 2009; Rikihisa, 2011). In Europe, anaplasmosis is mostly seen as tick-borne fever in domestic ruminants (for classification, clinical signs and treatment, see chapter one, subsection 5).

The natural life cycle of *A. phagocytophilum* is dependent upon the presence of infected vertebrate reservoir hosts and ixodid tick vectors (Severo et al. 2012). *Anaplasma phagocytophilum* reproduce in a vacuole or morula in the cytoplasm of both tick and vertebrate host cells. Detection of *A. phagocytophilum* can be performed through examining the morulae in vertebrate neutrophils; however, detection in the tick host is more challenging because the pathogen is pleomorphic. This suggests that distinct strategies are utilized by this bacterium when colonizing distinct hosts (Villar et al. 2015a). The objective of this study is to identify *Anaplasma* key proteins involved in host-bacteria interactions *in vitro*, and identify *A. phagocytophilum* proteins that are modulated during infection in cultured tick cells. Use of high-throughput “omics” technologies is crucial for the identification of the

Anaplasma proteome, which in turn contributes to our understanding of vector-pathogen interactions, allowing us to develop novel control strategies for *A. phagocytophilum* and its associated diseases (Rikihisa 2011).

4.1.1 *Anaplasma phagocytophilum* genomics

The *A. phagocytophilum* genome is relatively small, with the whole genome size of the human strains (HZ, HZ2) being around 1.4 Mb. However, this bacterium is capable of synthesizing most vitamins and cofactors, although it lacks lipopolysaccharides and peptidoglycans, which makes *A. phagocytophilum* very sensitive to mechanical stresses such as sonication and freezing. These features add challenges for *Anaplasma* purification and genome sequencing for different strains (Severo et al. 2012). *Anaplasma phagocytophilum* is characterized by distinct ecotypes suggesting varying host tropisms across its geographical distribution (Foley et al. 2008). Human anaplasmosis is a potentially life-threatening public health issue in the USA; while in Europe, it is uncommon in humans, but causes significant economic losses to the livestock industry (Woldehiwet 2006), suggesting variable epidemiological factors that correlate to strain variation (Dugat et al. 2014). The mechanisms of *A. phagocytophilum* host tropism mostly remain unknown; however, as previously reported, the horse (MRK) pathogenic strain in the USA does not infect ruminants (Stannard et al. 1969), while European strains are not infectious in horses (Pusterla et al. 1999). Host tropism was demonstrated by experimental infection, showing that the white-footed mouse is the reservoir of the American human Ap-ha strain, while white tailed deer is the reservoir for Ap-variant 1 strain (Massung et al. 2006; Massung et al. 2005), and Ap-variant 1 is not infectious for mice (Massung et al. 2003). Moreover, the acute phase of the disease has been reported in humans and

horses, whereas chronic disease was identified in sheep and rodents (Foley et al. 2008).

Further genomic analysis comprising North American and European *Anaplasma* strains revealed that the the *drhm* gene was deleted consistently in ApDog (dog), and ApHZ or ApJM (human) strains in the USA, but maintained in the more distantly related USA Ap-variant 1 (ruminant), ApMRK (horse), and ApNorV2 (sheep) strains, which suggets a significant epidemiological feature influencing human virulence (Al-Khedery & Barbet 2014).

The infection of *A. phagocytophilum* (an obligate intracellular bacterium), similar to *B. burgdoferi*, *Neisseria gonorrhoeae,* and even the extracellular African trypanosome, *Trypanosoma brucei,* which expresses variable surface glycoproteins (Brayton 2012), requires a persistence phase in host reservoirs, which assures pathogen transmission. This suggests evasion of the immune system involving genetic recombination or conversion of outer membrane protein (Omp) genes. In *A. phagocytophilum*, the surface protein gene family undergoing variation is Msp2/p44, which permits the creation of new variants, allowing escape from the immune system (Caspersen et al. 2002; Barbet et al. 2003; Brayton et al. 2001).

The mechanism of antigenic variation has been more intensely studied in the related bovine pathogen *Anaplasma marginale* than in *A. phagocytophilum* itself. In *A. marginale*, Msp4 is a conserved single-copy Omp, while Msp2 and Msp3, together with some other Omps classified under the PF01617 signature, are identified as immunodominant, antigenically variable proteins that are involved in evading the host immune response. Each of the variable proteins contains conserved amino- and carboxy- termini flanking a central hypervariable region (HVR). New variants arise throughout acute and persistent infection; consequently, antibody produced at a

certain time point will not recognize variants found at a later time point, contributing to pathogen persistence within the intracellular situation (Caspersen et al. 2002; Chávez et al. 2012).

In *A. phagocytophilum,* Msp2 (also known as p44, 44-kDa antigen), is an immunodominant family of paralogous genes encoding up to 113 Msp2/p44 variants in the HZ strain. Each gene contains a central HVR region of around 280 bp, flanked by conserved sequences of 100 to 500 bp (Zhi et al. 1999; Lin et al. 2003; Zhi et al. 2002). These were initially identified as "functional pseudogenes", representing truncated copies of the full-length gene that cannot in their existing location be transcribed, but can create new variants when recombined into the expression site of the gene (Rikihisa 2011; Lin et al. 2003). Gene conversion (**Figure 4.1**) is the main mechanism that leads to p44 expression site (p44ES) recombination, by which the DNA sequence information of the donor p44 gene (which remains unchanged) is copied and transferred to the p44 ES, the sequence of which is replaced (Lin et al. 2006; Lin et al. 2003; Severo et al. 2012). Previous reports documented that the Msp2/p44 of *A. phagocytophilum* lacks the homologous RecBCD recombination pathway that repairs bacterial DNA breaks but uses the RecFOR pathway at a single expression locus for homologous recombination and gene conversion. This allows for the rapid exchange of the p44 fragments during transcriptional regulation because the same promoter is maintained (Rikihisa 2011). Proteomics analysis targeting a human promyelocytic leukemia cell line (HL-60) infected with *A. phagocytophilum* revealed that 110 P44 proteins are expressed, and 88 of them are encoded by pseudogenes (Lin et al. 2011). Previously, contradictory explanations have been proposed; some authors suggested a similar segmental recombination mechanism

that *A. marginale* uses to generate genetic variation (Barbet et al. 2003; Barbet et al. 2000; Brayton et al. 2001), while *in vitro* experiments of *A. phagocytophilum* infected HL-60 cells and RT-PCR analysis suggested RNA splicing as a mechanism by which these *A. phagocytophilum* donor alleles can be expressed as complete p44 genes (Zhi et al. 2002).

The context of the controversy is that *A. marginale* is characterised by the expression of complex mosaic Msp2 alleles during persistent infection (Brayton et al. 2001; Palmer et al. 2006). In contrast with *A. phagocytophilum*, *A. marginale* genomes are characterized by only seven to twelve Msp2 donor alleles (functional pseudogene regions). While simple variants are generated by the gene conversion mechanism when a single unique sequence (oligonucleotide) within the pseudogene HVR recombines with the expression site during initial infection, subsequent recombination of multiple HVRs from pseudogenes regions leads to the generation of complex mosaics that cannot be recognised by antibodies at the time of emergence. This evasion mechanism increases the opportunity for lifelong persistence through sequential cycles of bacteraemia dominated by different variants, allowing continuous disease transmission (Palmer et al. 2006; Chávez et al. 2012) (**Figure 4. 2).**

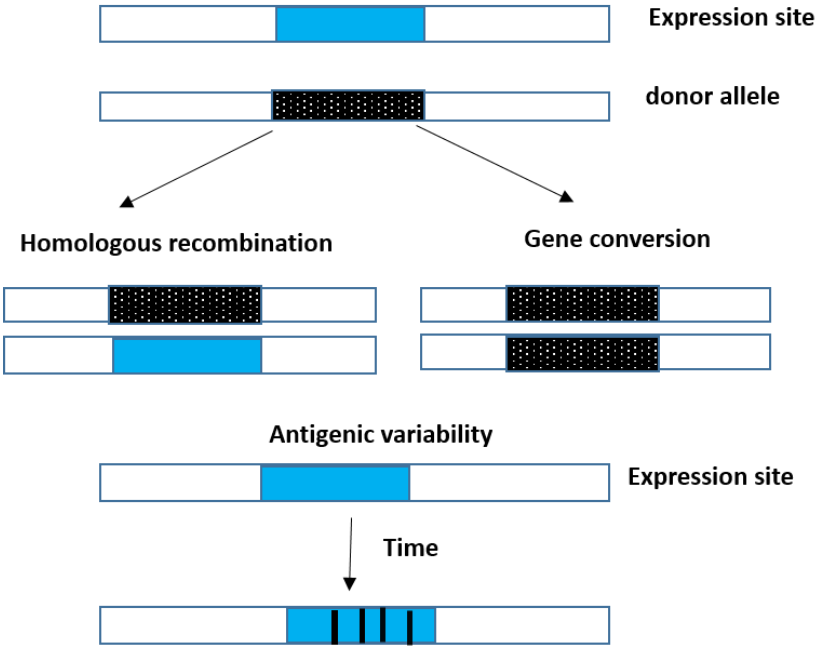


Figure 4. 1 Schematic diagram describing three mechanisms of antigenic variation

In *A. phagocytophilum* and *A*. *marginale*, rapid antigenic variation proceeds by gene conversion. This is distinct from classical homologous recombination, because the donor allele repertoire is maintained by "copy and paste" rather than exchanged with the gene in the expression site (Brayton 2012). In addition, antigenic variability (bottom) occurs more slowly by a secondary process of accumulation of non-synonymous mutations.

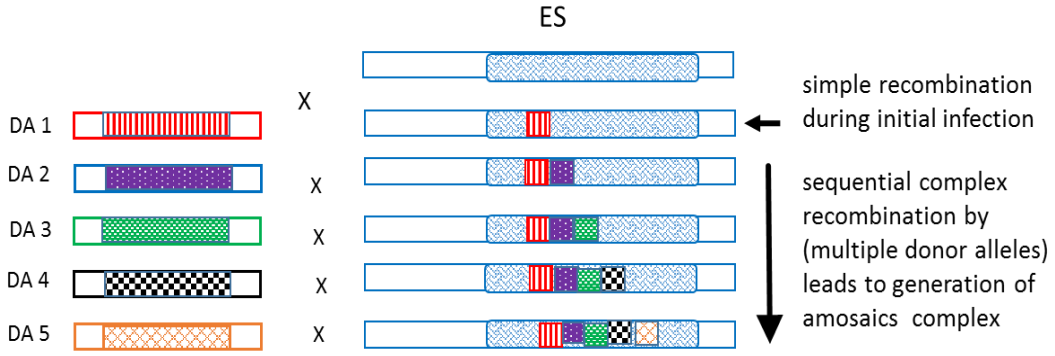


Figure 4. 2 Modified schematic diagram illustrates the generation of both simple and mosaic complex Msp2 variants of of *A. marginale.*

Right: bacterial Msp2 expression site (ES); left: multiple pseudogenes (donor alleles, DA). Initial recombination explains the generation of a simple variant by gene conversion mechanism, while subsequent recombination between the expression site and multiple donor alleles (different pseudogene HVR) highlights the generation of a variant mosaic complex (Palmer et al. 2006).

Subsequently, it was stated that defining full donor allele repertoires by whole genome sequencing is crucial for resolving the similarities and differences in antigenic variation between *A. marginale* and *A. phagocytophilum* (Brayton 2012). Even though the mosaic complex phenomenon appears to be lacking in *A. phagocytophilum* Msp2/p44, the larger Msp2/p44 repertoire (donor alleles) in its genome provide a valuable source for genetic variation (Zhi et al. 1999; Brayton et al. 2001; Barbet et al. 2003). In an epidemiological context, the extraordinary mechanisms that enable *A. phagocytophilum* to invade and replicate inside mammalian leukocytes and tick midgut and salivary glands successfully has been a major research focus in the field, highlighting the critical factors for *Anaplasma* infection and transmission from tick vector to different mammalian hosts. Further exploration of *A. phagocytophilum* genomes is crucial to understand the genetic diversity of this pathogen and to gain a deeper insight into *Anaplasma* phylogeny and ecotypes.

4.1.2 *Anaplasma phagocytophilum* - host protein interaction

Use of high-throughput techniques involving proteomics analysis has paved the way for *A. phagocytophilum* protein identification and quantitation to characterise the *A. phagocytophilum*-host protein interface, which may help in identifying protein targets to block *A. phagocytophilum* transmission. However, contradictory proteomic findings have been reported, which may be due to the difficulties or even impossibilities of *A. phagocytophilum* isolation from the host. It has been suggested that the challenge of intermixed bacterial-host proteins inhibits the sensitivity of most proteomics studies (Lin et al. 2011).

To partially overcome this, nano-liquid chromatography combined with tandem MS/MS (nano-LC–MS/MS)-based proteomic approaches have increased the

sensitivity of proteomics analysis and could be used to study the host-pathogen interactions of *Anaplasma*. Label-free protein quantitation based on LC–MS peptide peak intensity information allows for the quantitative analysis of multiple conditions without stable isotope labelling (Zimmer et al. 2006). So far, most of the *A. phagocytophilum* proteomics studies have been based on the human strain (HZ). Thus, despite the considerable economic losses that *A. phagocytophilum* causes across its European distribution in livestock, proteomics analysis using ruminant *A. phagocytophilum* strains has not received as much attention as human strains. Moreover, there is only one European sheep strain of *A. phagocytophilum* for which a whole genome is available.

In general, the *A. phagocytophilum* surface proteins Msp1, Msp2 (p44), Msp3, and Msp4 may be considered as key for bacterial infection in both the vector and mammalian reservoir hosts (de la Fuente et al. 2001; Villar et al. 2015b). In addition to the role of p44 in achieving successful persistence as described above (Barbet et al. 2003), the interaction of *A. phagocytophilum* MSP4 at the tick cell interface activates secretion of vesicles at the phagocytic cup and facilitates bacterial invasion and adhesion (Barbet et al. 2003). Bacterial proteins exported into the host cell via the type IV secretion system (T4SS) control host epigenetics and global DNA methylation, allowing nuclear effectors of *A. phagocytophilum* to manipulate host chromatin and gene expression to establish infection in both vertebrate and tick hosts (Sinclair et al. 2015).

Proteomic analysis targeting a human promyelocytic leukemia cell line (HL-60) infected with *A. phagocytophilum* HZ (human strain) revealed that chaperones, enzymes involved in energy metabolism and protein biosynthesis, and outer membrane proteins including the p44 family proteins were over-represented (Lin et

al. 2011). However, human (HL-60) cell responses involved cytoskeletal rearrangements and actin polymerization, vesicular transport and cell signalling induction correlated with bacterial invasion and multiplication; whereas TLR1 and mannose receptor 2 (one of the recognition receptors), which are associated with host innate immunity, were downregulated (Lin et al. 2011). Moreover, numerous isoforms of sarcoplasmic/endoplasmic reticulum calcium ATPase (SERCA) were up-regulated due to *A. phagocytophilum* infection, suggesting proteins that play a role in intracellular [Ca^{2+}] regulation like phospholipase C and transglutaminase are essential in bacterial infection (Lin et al. 2002). Further microarray research identified genes involved in similar gene expression profiles in terms of upregulation and downregulation in different host cell types (Pedra et al. 2005).

A recent study suggested that some proteins produced by *A. phagocytophilum* may contribute to changes in the metabolome, as both *A. phagocytophilum* glutamate dehydrogenase (GDG) protein and pyruvate kinase (PK) were over-represented in infected tick cells. This was correlated with a decrease in host glucose metabolism; furthermore, reducing levels of hexokinase and phosphoenolpyruvate carboxykinase (PEPCK) is another potential mechanism for glucose metabolism inhibition, which consequently decreases host cell apoptosis and facilitates tick infection (Villar et al. 2015a).

The aim of the current study was to investigate, for the first time, interactions between a British ruminant strain of *A. phagocytophilum* and tick cells *in vitro* using proteomic technologies. This required genome sequencing of this strain to define the p44 antigenic protein repertoire for targeted proteomic studies (parallel reaction monitoring).

4.2 Materials and Methods

4.2.1 Tick cell line culture

The tick cell lines ISE6 and IRE/CTVM19 derived from embryonic *I. scapularis* and *I. ricinus*, respectively (Kurtti & al 1996; Bell-Sakyi et al. 2007), were provided by Lesley Bell-Sakyi of the Tick Cell Biobank. These cell lines were chosen for this study as they are appropriate models for growth of *A. phagocytophilum* (Munderloh et al. 1996; Woldehiwet et al. 2002; Massung et al. 2007; Pedra et al. 2010; Sultana et al. 2010). The ISE6 cell line was maintained in L-15B medium (Munderloh & Kurtti 1989) (**Appendix E**). This was supplemented with 10% foetal calf serum (FCS) (Sigma Aldrich, Poole, Dorset, England), 10% tryptose phosphate broth (Sigma), 0.1% bovine lipoprotein concentrate (MP Biomedicals) and 2 mM L- glutamine. The pH of the medium was then adjusted to 7.5 by adding 1 M NaOH dropwise. The IRE/CTVM19 cells were maintained in L-15 (Leibovitz) supplemented with 10% FCS, 10% tryptose phosphate broth and 2mM L-glutamine (Růžek et al. 2008).

Uninfected tick cells of both cell lines (as control cells) were cultured with penicillin (100 units/ml) and streptomycin (100 µg/ml) until the cells became confluent. The propagation of cell monolayers was conducted in sealed flat-sided tissue culture tubes (Nunc, product no. 156758), incubated in an atmosphere of ordinary air in a dry incubator at 32˚C. Confluent monolayers were maintained for up to 6 weeks, with medium replaced weekly (Bell-sakyi et al. 2018).

4.2.2 *Anaplasma* culture

An ovine strain of *Anaplasma phagocytophilum*, Old Sourhope (OS), isolated for the first time by Drs Foster and Cameron from a Cheviot ewe grazing in the

Scottish Borders (Foster & Cameron 1970), and the Feral Goat strain (isolated in Galloway, Scotland) (Scott & Horsburgh 1983), were propagated through infection of ISE6 cultures in a modified L-15B medium (Munderloh et al. 1999) referred to as *Anaplasma* culture medium.

Prior to inoculation of ISE6 and IRE/CTVM19, medium was replaced with antibiotic-free medium buffered by adding 15 mM HEPES buffer and 0.1% $NaHCO_3$. Inoculation of ISE6 and IRE/CTVM19 was then performed by the addition of *Anaplasma*-infected ISE6 cells into each cell line, which were maintained at 34˚C in a different incubator separate to uninfected cells. Passages were continued every 2 - 3 weeks by adding 0.1 to 0.2 mL cell suspension from infected tubes to uninfected tubes (Woldehiwet et al. 2002). Young cultures are subcultured mostly 1:2 (Bell-Sakyi et al. 2012).

4.2.3 Cytocentrifuge smears and staining for microscopy

To monitor the progression of infection in tick cell line cultures, cell morphology, viability and infection density were examined regularly. This was performed by removing a small volume (0.2 - 0.5 mL) every 2 weeks and depositing an aliquot of the cell suspension on microscope slides with a cytocentrifuge (Cytospin, Shandon, Runcorn, Cheshire, UK). The smears were air dried, fixed in methanol for 3 min and transferred directly to 10% Giemsa stain for 20 min. Fixed cells were rinsed quickly three times with Giemsa buffer (pH 7.2) and examined at ×500 or ×1,000 magnification using a light microscope with oil immersion. Photomicrographs of the smears were taken using brightfield illumination on a Zeiss Axio Imager M2 microscope with ZEN 2011 imaging software (Carl Zeiss, Germany)

4.2.4 *Anaplasma phagocytophilum* infection for proteomics and genome sequencing

To identify suitable time points for further proteomic studies, preliminary experiments were performed. As the *A. phagocytophilum* Feral Goat strain has not been cultured extensively in tick cell lines, a time course experiment was performed to monitor the progression of infection and to identify a suitable time point to maximise the rate of infection. In all preliminary experiments, tick cell lines were used at 10^6 cells per tube. Duplicates of *A. phagocytophilum* (Feral Goat strain)-infected and uninfected ISE6 cells were sampled at four time-points; 24, 48, 120 and 192 hours post infection. Duplicates of *A. phagocytophilum* (OS strain)-infected and uninfected IRE/CTVM19 were sampled at 3 and 6 days post-infection (dpi).

The density of *Anaplasma* infection was examined using the cytospin protocol described above. The percentage of cells infected with *A. phagocytophilum* was calculated by examining at least 200 cells using a Plan Neofluar 100× oil immersion objective and Zeiss Axio Imager M2 microscope. The cells were harvested by scraping, then concentrated by centrifugation at 200 × *g* for 10 min. The supernatant was removed, and the cell pellets were washed in 1 mL PBS (Sigma). A final centrifugation at 200 × *g* for 5 min was used to concentrate the cell pellet, which was stored at -80^{o}C for protein and DNA extraction. Successful establishment and identification of *A. phagocytophilum* at different time points enabled selection of the following parameters for use in the main experiment.

Four independent cultures (with approximately 10^6 cells each) were sampled at two time-points after infection with *A. phagocytophilum* OS. In ISE6, 4 dpi as low infection (10 – 15% infection rate) and 8 dpi as high infection (60 - 70% infection rate) were used; whereas in IRE/CTVM19, days 10 and 20 were used. For proteomics and

qPCR quantification of infection, cell pellets were stored at -80^{o}C for protein and DNA extraction as in the preliminary experiments.

For genome sequencing, *A. phagocytophilum* OS was cultivated in ISE6 (Munderloh et al. 1999) as described above. Infections of tick cell lines were performed by adding 0.5 mL of *A. phagocytophilum*-infected ISE6 cells into three different tubes (one tube of ISE6 and two tubes of IRE/CTM19 cells).

4.2.5 Sample preparation for mass spectrometry

Infected ISE6 and IRE/CTVM19 cells were harvested alongside uninfected controls in quadruplicate (as reported above in 4.2.4). Cell pellets were lysed by sonication in 1% (w/v) sodium deoxycholate (SDC) in 50 mM ammonium bicarbonate. Samples were heated at 80°C for 15 min before centrifugation at 12,000 $\times$ *g* to pellet debris. Protein concentrations were determined using the Pierce Coomassie Plus (Bradford) Protein Assay (Thermo Scientific). Proteins were reduced with 3 mM dithiothreitol (DTT) (Sigma) at 60°C for 10 min, cooled, then alkylated with 9 mM iodoacetamide (Sigma) at room temperature for 30 min in the dark; all steps were performed with intermittent vortex-mixing. Proteomic-grade trypsin (Sigma) was added at a protein:trypsin ratio of 50:1 and incubated at 37°C overnight. The SDC was removed by adding trifluoroacetic acid (TFA) to a final concentration of 0.5% (v/v). Peptide samples were centrifuged at 12,000 x g for 30 min to remove precipitated SDC. An aliquot of peptide sample was stored for mass spectrometry, and the remaining material was used for DNA extractions.

4.2.6 DNA extraction and quantification

DNA was extracted using a DNeasy Blood and Tissue (QIAGEN) Kit according to the manufacturer's protocol into 100 µl of nuclease-free water. All DNA samples were quantified using a fluorescent labelling method, the Quant-iT Picogreen dsDNA kit (Invitrogen), and measured in an Infinite F200 microplate fluorimeter with Magellan Data Analysis Software (TECAN, Männedorf, Switzerland). DNA concentrations were calculated by comparison to a Lambda DNA standard component (0 – 1,000 ng/ml) curve.

4.2.7 qPCR (*Anaplasma phagocytophilum*) infection rate

Anaplasma phagocytophilum DNA levels were characterized using a 16S ribosomal RNA gene assay originally designed for quantification of another member of the Anaplasmataceae, *Wolbachia* (Makepeace et al. 2006). For normalisation of bacterial gene copy numbers against the tick (*Ixodes*) host, a new qPCR assay targeting a tick single copy-nuclear gene was designed (*IsRPL6*), as reported in chapter 2 (2.2.4). Normalized Ct values were compared between infected and uninfected cells, and a decrease in Ct values over time confirmed *A. phagocytophilum* multiplication. The total reaction volume used was 20 µl comprising 1× SensiMix (SensiMix SYBR, Bioline), amplified by PCR with 200 nM each 16S rRNA primer Wol16S (S) 5′ TTGCTATTAGATGAGAGCCTATATTAG 3′ (forward), and Wol16S (AS) 5′ GTGTGGCTGATCATCCTCT 3′ (reverse), and 1 µL of DNA. The PCR cycling conditions were as follows: initial denaturation 95°C for 10 min (enzyme activation), 35 cycles of denaturation at 95°C for 15 sec, annealing at 55°C for 30 sec, and extension at 72°C for 15 sec; followed by a melt curve from 55°C to 95°C with increasing increments of 0.5°C per cycle. All PCR assays were carried out in two dilutions, with the best dilution

selected on a CFB-3220 DNA Engine Opticon 2 System and an average result was calculated. Plate set-up, preparation of standards, and data collection were as reported in chapter 2 (2.2.4). Infection densities determined by qPCR were used to select and validate the time-points chosen for proteomic analysis and whole genome sequencing.

4.2.8 Enrichment of *A. phagcoytophilum* strain OS DNA for genome sequencing

Three methods were used in attempts to enrich *A. phagocytophilum* DNA relative to tick DNA. Two cell lines, IRE/CTVM19 (*I. ricinus*) and IDE8 (*I. scapularis*), both infected with *A. phagocytophilum* OS strain, were used as source material. The genome copy ratio between *Anaplasma* and tick cells was calculated according to qPCR copy numbers for the 16S rRNA gene of *Anaplasma* and the tick *IsRPL6* nuclear gene.

In the first method, no specific enrichment was attempted apart from removal of low molecular weight DNA fragments. The *A. phagocytophilum* ratio was calculated via qPCR after using a Nanosep centrifugal device (PALL, Life sciences, 30K), to increase the DNA concentration and remove degraded sequences. The ratio of *A. phagocytophilum* to haploid tick genomes was 18.22 with a yield of 2.63 µg DNA.

In the second method, we attempted to select *A. phagocytophilum* DNA on the basis of its circular structure, as host DNA is linear (except for mitochondrial DNA). This "Mseek" method uses exonuclease V to specifically digest linear DNA in a protocol originally designed to enrich for mitochondrial DNA (Jayaprakash et al. 2015).

According to the protocol from New England Biolabs (NEB), and in order to inactivate proteinase K that might remain from the DNA extraction protocol, eluted DNA was

heated at 70°C for 30 min. A first digestion was performed by adding NEB4 10× Buffer (6 µl), 10 mM ATP (12 µL), exonuclease V (NEB cat. no.M0345S; 4 µL), and H_2O (3 µl) to the total DNA sample (2.3 µg in 35 µl). The digest was incubated at 37°C for 48 hours, then heat-inactivated again at 70°C for 30 min and purified using SeraPure beads on a magnetic stand (Beckman Coulter). A second digest was carried out as before on the treated DNA sample, but incubated at 37°C for 16 hours this time. Heat-inactivation and purification were also performed for a second time. Finally, the qPCR assay was conducted using 16S rRNA and *IsRPL6* primers.

In the third method, the NEBNext Microbiome DNA Enrichment KIT was used to facilitate the enrichment of microbial DNA from samples containing methylated host DNA by selective binding and removal of CpG-methylated host DNA.

According to the qPCR results, the ratio of *A. phagocytophilum* to tick cell haploid genomes in this sample was 14.17 in 2 µg. In one tube, 32 µl of MBD2-Fc protein and 320 µl of protein A magnetic beads were added to 2 µg DNA in 57.1 µL. Then 14.27 µl of Bind /Wash buffer (5X) were added. The reaction was incubated at room temperature for 15 min with rotation. After brief spinning, the tube was put on a magnetic rack for 5 min until the beads were collected to the wall of the tube and the solution was clear. The supernatant was then carefully removed without disturbing the beads. This supernatant contained the target microbial DNA. Purification and clean-up was carried out using SeraPure beads.

To check the integrity of DNA before sequencing, DNA was loaded onto a 0.5% agarose gel in 1X TAE buffer at 30 V for 16 hours in a cold room (4˚C). A 1:10 dilution of 1 kb extension ladder (Sigma, Aldrich) (5 µl) was used to assess DNA integrity.

4.2.9 Genome sequencing, assembly and annotation

These two samples were submitted for sequencing to the Centre for Genomic Research, University of Liverpool, and used to prepare Illumina TruSeq PCR-free libraries with 350 bp inserts. Paired-end sequencing was conducted with 2 × 250 bp reads on an Illumina MiSeq system.

Raw data was analysed by Dr. Alistair Darby. The raw reads (approximately 4.1 million for the NEBNext Microbiome Kit sample and 3.2 million for the unenriched sample) were trimmed and normalised against host genome and mapped to the *A. phagocytophilum* str. Norway variant2 chromosome (GenBank assembly accession GCA_000689635.2). After discarding the reads mapping to host, assembly was done using the Discovar *de novo* assembler (Love et al. 2016). Thereafter, Blobplot (Laetsch & Blaxter 2017) was generated to visualise the contents of the genome assembly data against diverse taxonomic phyla. The assembled contigs mapped mostly against *Ixodes* with a proportion against *Anaplasma*, so only *Anaplasma* contigs were extracted and analysed further.

Genome annotation was performed by Dr. Alistair Darby using PROKKA (rapid prokaryotic genome annotation) (Seemann 2014); a command-line software tool to fully annotate a draft bacterial genome and search for gene functions.

4.2.10 Phylogeny and genome comparisons

The following analyses were performed under the guidance of Dr Samriti Midha. Using 29 other *A. phagocytophilum* genomes available in the NCBI database (**Table 4. 1**), the strain OS was compared with the others to investigate the evolutionary relationship of these *A. phagocytophilum* strains. Blast-based Average Nucleotide Identity amongst 30 isolates was calculated using JSpecies v. 1.2.1

(Richter & Rossello-Mora 2009) and visualized by generating a heat map using GENE-E software (https://software.broadinstitute.org/GENE-E/). All 30 genomes were compared to identify their core genome using Bacterial Pan Genome Analysis pipeline in which incorporated the KEGG and COG analyses as well (Chaudhari et al. 2016). Orthologous gene clusters shared by all 30 strains were grouped and identified as core genes. Each strain showed the presence of some genes that were not shared with any other strain. These genes were listed as unique genes. Other orthologous genes that were shared among some, but not all 30 *A. phagocytophilum* genomes, were identified as accessory genes.

The protein sequences encoded by the core genome were aligned and a maximum likelihood phylogenetic tree was constructed using MEGA7 with 500 bootstrap replicates (Kumar et al. 2016). To compare the genome structure of strain OS with other sheep isolates and a human isolate, CONTIGuator software was used (Galardini et al. 2011). CONTIGuator generates a pseudocontig from a draft genome based on mapping and then produces an Artemis (Carver et al. 2012) comparison with the reference strain.

Protein sequences belonging to the Msp2 family were identified by Dr Samriti Midha using tblastn analysis against the annotated sequences from the HZ strain. Furthermore, all *Anaplasma* proteins from strain OS were screened for Pfam family PF01617 members ("Surface_Antigen_2") using an HMM Database search (Eddy 2011). Finally, orthologous clusters of Msp2/p44 genes between 30 *A. phagocytophilum* strains (including OS) and 15 *A. marginale* strains were generated using orthoMCL software (Li et al. 2003) by Dr Xiaofeng Dong.

Table 4. 1 *Anaplasma* genome sequence accession numbers, genome statistics, host and location

	Strain	Host	Location	%GC	Genome size (Mb)	Number of genes	Reference
1	HZ	Human (*Homo sapiens*)	USA	41.60	1.47128	1,411	PRJNA336
2	HZ2	Human (*Homo sapiens*)	USA	41.60	1.47758	1,295	PRJNA163167
3	HGE1	Human (*Homo sapiens*)	USA	41.60	1.4696	1,188	PRJNA171710
4	JM	Meadow jumping mouse (*Zapus hudsonius*)	USA	41.60	1.4816	1,302	PRJNA158483
5	Dog2	Dog (*Canis lupus familiaris*)	USA	41.60	1.4733	1,304	PRJNA163169
6	CRT38	Tick (*Ixodes scapularis*)	USA	41.70	1.50655	1,202	PRJNA183838
7	CRT35	Tick (*Ixodes scapularis*)	USA	41.60	1.44702	1,148	PRJNA217037
8	MRK	Horse (*Equus caballus*)	USA	41.60	1.47923	1,155	PRJNA216999
9	Bovine 10_179	Cow (*Bos taurus*)	France	41.50	1.37082	1,041	PRJEB6582
10	Norway variant2	Sheep (*Ovis aries*)	Norway	41.70	1.5452	1,174	PRJNA217033
11	Cow_1	Cow (*Bos taurus*)	France	41.93	1.68232	1,519	PRJEB13033
12	Cow_2	Cow (*Bos taurus*)	France	42.23	1.64135	1,463	PRJEB13033
13	Cow_3	Cow (*Bos taurus*)	France	42	1.56165	1,417	PRJEB13033
14	Cow_4	Cow (*Bos taurus*)	France	42.12	1.60442	1,448	PRJEB13033
15	Cow_5	Cow (*Bos taurus*)	Germany	43.7	1.71862	1,979	PRJEB13033

16	Horse_1	Horse (*Equus ferus caballus*)	France	41.9	1.16875	1,657	PRJEB13033
17	Horse_2	Horse (*Equus ferus caballus*)	France	41.58	1.77221	1,997	PRJEB13033
18	Roe_deer_1	Roe deer (*Capreolus capreolus*)	Germany	42.72	1.58518	1,912	PRJEB13033
19	Webster	Human (*Homo sapiens*)	USA	41.6	1.47941	1,181	PRJNA212447
20	Annie	Horse (*Equus ferus caballus*)	USA	41.8	1.51652	1,225	PRJNA212493
21	Ap W11	Human (*Homo sapiens*)	USA	41.6	1.49707	1,201	PRJNA212467
22	ApNYW	Human (*Homo sapiens*)	USA	41.6	1.50309	1,217	PRJNA212468
23	HGE2	Human (*Homo sapiens*)	USA	41.7	1.48205	1,175	PRJNA212466
24	NCH-1	Human (*Homo sapiens*)	USA	41.6	1.50275	1,226	PRJNA212448
25	HGE1-mutant	Human (*Homo sapiens*)	USA	41.6	1.49367	1,204	PRJNA212450
26	CR1007	Tick (*Ixodes scapularis*)	USA	41.7	1.50202	1,213	PRJNA212464
27	CRT53	Tick (*Ixodes scapularis*)	USA	41.8	1.5709	1,275	PRJNA212465
28	ApNP	Dog (*Canis lupus familiaris*)	Austria	41.7	1.52158	1,247	PRJNA212452
29	ApMUC09	Dog (*Canis lupus familiaris*)	Netherlands	41.7	1.5204	1,247	PRJNA212451
30	**OS**	**Sheep (*Ovis aries*)**	**UK**	**41.7**	**1.43102**	**1.268**	**This study**

4.2.11 Protein analysis

4.2.11.1 NanoLC MS ESI MS/MS analysis

Peptides were analysed by on-line nanoflow LC using the Ultimate 3000 nano system (Dionex/Thermo Fisher Scientific). Samples were loaded onto a trap column (Acclaim PepMap 100, 2 cm × 75 µm inner diameter, C18, 3 µm, 100 Å) at 5 µl min−1 with an aqueous solution containing 0.1% (v/v) TFA and 2% (v/v) acetonitrile. After 3 min, the trap column was set in-line an analytical column (Easy-Spray PepMap® RSLC 50 cm × 75 µm inner diameter, C18, 2 µm, 100 Å) fused to a silica nano-electrospray emitter (Dionex). The column was operated at a constant temperature of 30°C and the LC system coupled to a Q-Exactive mass spectrometer (Thermo Fisher Scientific). Chromatography was performed with a buffer system consisting of 0.1 % formic acid (buffer A) and 80 % acetonitrile in 0.1 % formic acid (buffer B). The peptides were separated by a linear gradient of 3.8 – 50 % buffer B over 90 minutes at a flow rate of 300 nl/min. The Q-Exactive was operated in data-dependent mode with survey scans acquired at a resolution of 70,000. Up to the top 10 most abundant isotope patterns with charge states +2 to +4 from the survey scan were selected with an isolation window of 2.0Th and fragmented by higher energy collisional dissociation with normalized collision energies of 30. The maximum ion injection times for the survey scan and the MS/MS scans were 250 and 50ms respectively, and the ion target value was set to 1E6 for survey scans and 1E5 for the MS/MS scans. MS/MS events were acquired at a resolution of 17,500. Repetitive sequencing of peptides was minimized through dynamic exclusion of the sequenced peptides for 20s.

4.2.11.2 Parallel reaction monitoring (PRM) design

A library of Msp2 proteins (Fasta format file) was obtained from the *A. phagocytophilum* predicted proteome (section 4.2.11.1 above). Skyline (MacCoss lab software, v 4.1) was used to create a spectral library with previous MS/MS discovery runs (preliminary proteomic experiments; section 4.2.11.1 above). A background proteome of *I. scapularis* and *I. ricinus* was used to avoid possible shared peptide sequences. Msp2 proteins with peptide evidence were retained. For each protein, at least two unique peptides were chosen (including the detected peptide). Some peptides were shared by many Msp2 proteins and these were included in the PRM design as Msp2 'universal' peptides (**Figure 4. 3**). Synthetic peptides were made by a commercial supplier (JPT Peptide Technologies GmbH) in 'light' (non-labelled) and 'heavy' (^{13}C, ^{15}N; labelled lysine or arginine) forms. The peptides were reconstituted in 50 mM ammonium bicarbonate, 20% acetonitrile, then mixed equally and diluted to 100 fmol/µl. The synthetic peptides were run on the Q-Exactive in discovery mode (DDA) using the LC conditions for the PRM (2h) to obtain the retention times for each peptide (**Figure 4. 3**).

Figure 4. 3 Alignment of the MSP2/p44 proteins from *A. phagocytophilum* strain OS included in the PRM experiment.

The peptides selected for PRM analysis are highlighted. Purple colouration refers to the universal peptides (shared by more than one protein); the other colours refer to the unique (discriminatory) peptides for each protein.

4.2.11.3 PRM protocol

The PRM analysis was performed on the same LC-MS/MS platform as the shotgun analysis (section 4.2.11.2 above; the same columns and buffers were used). The PRM method combined two scan events starting with a full scan event followed by up to 46 targeted MS/MS scans for charge 2+ and 3+ precursor ions from a scheduled precursor inclusion list, using a ±5 min retention time window (**Table 4. 2**). The LC gradient settings were the same as above (section 4.2.11.2). Full scans (m/z 300–2000) and PRM scans were acquired with a resolution of 17,500 at 200 m/z. The maximum ion injection times for the survey scan and the MS/MS scans were 200 and 120 ms respectively, and the ion target value was set to 1×10^6 for survey scans and 3×10^6 for the MS/MS scans. Targeted peptides were isolated using a 2.0 m/z window. Fragmentation was performed with a normalised collision energy of 27.

Table 4. 2 Inclusion list for *Anaplasma* PRM on the Q Exactive platform.

Mass [m/z]	Formula [M]*	Formula type*	Species*	CS [z]	Polarity	Start [min]	End [min]	(N)CE	(N)CE type*	MSX ID*	Comment
564.272				2	Positive	38.03	48.03	27%			FEVEVGYER
569.2762				2	Positive	38.03	48.03	27%			FEVEVGYER[+10]
450.7793				2	Positive	31.62	41.62	27%			TAALAAALAK
454.7864				2	Positive	31.62	41.62	27%			TAALAAALAK[+8]
568.2907				2	Positive	33.4	43.4	27%			LSYDLSPQGR
573.2949				2	Positive	33.4	43.4	27%			LSYDLSPQGR[+10]
408.7267				2	Positive	16.26	26.26	27%			TTPAIGEK
412.7338				2	Positive	16.26	26.26	27%			TTPAIGEK[+8]
446.2686				2	Positive	44.35	54.35	27%			GIVAGLFSK
450.2757				2	Positive	44.35	54.35	27%			GIVAGLFSK[+8]
499.7451				2	Positive	32.25	42.25	27%			LFYDESPK
503.7522				2	Positive	32.25	42.25	27%			LFYDESPK[+8]
671.3121				2	Positive	31.68	41.68	27%			DFTIQESSGETK
675.3192				2	Positive	31.68	41.68	27%			DFTIQESSGETK[+8]
606.811				2	Positive	37.4	47.4	27%			AIYPYISEEK
610.8181				2	Positive	37.4	47.4	27%			AIYPYISEEK[+8]
545.8032				2	Positive	49.67	59.67	27%			TEDLISALTK
549.8103				2	Positive	49.67	59.67	27%			TEDLISALTK[+8]
600.8146				2	Positive	44.23	54.23	27%			DAQDGDVLVLR
605.8187				2	Positive	44.23	54.23	27%			DAQDGDVLVLR[+10]
812.4694				2	Positive	53.52	63.52	27%			LLTSESVVPVDPVLR
817.4736				2	Positive	53.52	63.52	27%			LLTSESVVPVDPVLR[+10]
524.3135				2	Positive	48.64	58.64	27%			LYPVSLLSR
529.3176				2	Positive	48.64	58.64	27%			LYPVSLLSR[+10]

461.7477				2	Positive	27.89	37.89	27%			VEVEVGYK
465.7548				2	Positive	27.89	37.89	27%			VEVEVGYK[+8]
436.2534				2	Positive	18.18	28.18	27%			VNIAGGAAAK
440.2605				2	Positive	18.18	28.18	27%			VNIAGGAAAK[+8]
677.3491				2	Positive	45.2	55.2	27%			ANIASYGFNIGAR
682.3533				2	Positive	45.2	55.2	27%			ANIASYGFNIGAR[+10]
765.881				2	Positive	32.53	42.53	27%			QALSASAGYSTGYVR
770.8851				2	Positive	32.53	42.53	27%			QALSASAGYSTGYVR[+10]
593.7906				2	Positive	39.87	49.87	27%			EYYISLENR
598.7947				2	Positive	39.87	49.87	27%			EYYISLENR[+10]
460.2534				2	Positive	33.89	43.89	27%			AFGVLEQR
465.2576				2	Positive	33.89	43.89	27%			AFGVLEQR[+10]
601.3248				2	Positive	35.36	45.36	27%			TIEGGEVVEIR
606.3289				2	Positive	35.36	45.36	27%			TIEGGEVVEIR[+10]
757.7113				3	Positive	60.2	70.2	27%			AGLSYQLSPEISAFAGGFYHR
761.0474				3	Positive	60.2	70.2	27%			AGLSYQLSPEISAFAGGFYHR[+10]
752.3755				2	Positive	39	49	27%			VVGDGVYDDLPAQR
757.3797				2	Positive	39	49	27%			VVGDGVYDDLPAQR[+10]
994.4168				2	Positive	25.93	35.93	27%			YAVYAEETDNGNNDGNNK
998.4239				2	Positive	25.93	35.93	27%			YAVYAEETDNGNNDGNNK[+8]
503.2536				2	Positive	22.27	32.27	27%			NWPTSTTAK
507.2607				2	Positive	22.27	32.27	27%			NWPTSTTAK[+8]

*These columns are intentionally empty due to instrument software requirements.

4.2.11.4 Protein identification and quantification

Thermo RAW files were imported into Progenesis QI for proteomics (version 4.1, Nonlinear Dynamics). Runs were time aligned using default settings and an auto-selected run as a reference. Peaks were picked by the software using default settings and filtered to include only peaks with a charge state between +2 and +7. Spectral data were converted into .mgf files with Progenesis QI for proteomics and exported for peptide identification using the Mascot (version 2.3.02, Matrix Science) search engine. Tandem MS data were searched against translated ORFs from *I. scapularis* (Uniprot, June 2017, 23,786 sequences; 6,510,768 residues), *I. ricinus* (Uniprot, June 2017, 42,237 sequences; 13,503,354 residues), an *A. phagocytophilum* pan proteome (Uniprot reference proteome UP000001943, June 2017, 4,610 sequences; 760,258 residues) or an in-house *A. phagocytophilum* database updated with protein predictions from strain OS October 2017, 1,230 sequences; 288,102 residues), and a contaminant database (cRAP, GPMDB, February 2015) (119 sequences; 40,423 residues). In the main proteomic experiments, results were compared when searching against *I. scapularis* only (for which a partial genome is available at (Gulia-Nuss et al. 2016) or against *I. scapularis* and *I. ricinus* (for the latter, protein predictions are derived from transcriptomic datasets only). The search parameters were as follows: the precursor mass tolerance was set to 10 ppm and the fragment mass tolerance was set as 0.05Da. Two missed tryptic cleavages were permitted. Carbamidomethylation (cysteine) was set as a fixed modification and oxidation (methionine) set as variable modification. Mascot search results were further validated using the machine learning algorithm Percolator embedded within Mascot. The Mascot decoy database function was utilised and the false discovery rate was <1%, while individual percolator ion scores >13 indicated identity or extensive

homology ($p < 0.05$). Mascot search results were imported into Progenesis QI for proteomics as XML files. Peptide intensities were normalised against the reference run by Progenesis QI for proteomics and these intensities were used to highlight relative differences in protein expression between sample groups. Only proteins with ≥2 identified peptides were included in the dataset. Statistical analysis (ANOVA) of the data was performed using Progenesis QI for proteomics to identify significantly [$p < 0.05$, $q \leq 0.05$, relative fold change (FC) ≥2] differentially expressed proteins.

4.2.11.5 PRM data analysis

Raw data was imported into Skyline and the spectra manually curated. For correct peak identification, as many co-eluting sequence information ions (fragments or transitions) as possible that correlate in intensity with a reference peptide were required (at least 4). For quantification, only transitions with a good signal-to-noise ratio and free from interferences were used. The ratio to the heavy peptide standard was used for normalisation. Group comparisons were performed using the MS stats function (t-test) in Skyline (Choi et al. 2014).

4.2.11.6 GO analysis

Enrichment of GO terms was performed using the agriGO online tool (http://bioinfo.cau.edu.cn/agriGO/) (Tian et al. 2017; Du et al. 2010). GO terms for the *Ixodes* proteins were obtained by entering the protein sequence into the GOAnna tool or the Uniprot accession into the GOretriever tool (both AgBase, http://www.agbase.msstate.edu/) (McCarthy et al. 2006). Proteins with significant changes in abundance compared to controls (>2 FC, $p < 0.05$) were compared to a whole proteome background (*I. scapularis* and *I. ricinus* individually or combined) using singular enrichment analysis in the agriGO tool, with a hypergeometric statistical test method incorporating a Yekutieli (false discovery rate) multi-test adjustment. The significance level was set at a p-value <0.05.

4.2.11.7 PFAM enrichment

Enrichment of protein domains was assessed using Pfam [EBI, v. 27.0 (Finn et al. 2014)] as previously described (Armstrong et al. 2014) using the gathering threshold as a cut-off. Briefly, a hypergeometric test for enrichment of Pfam domains in the observed proteome (for identifications supported by ≥2 unique peptides only) relative to the complete search database was performed using R (phyper). The Benjamini & Hochberg step-up false discovery rate-controlling procedure was applied to the calculated p values (Benjamini & Hochberg 1995) and enrichment was considered statistically significant where $p < 0.01$.

4.3 Results

4.3.1 *Anaplasma phagocytophilum* infection confirmation

Cytospin analysis by light microscopy confirmed *A. phagocytophilum* morulae inside ISE6 cells (**Figure 4. 4**). Subsequently, this was used as a method to monitor infection alongside qPCR, but *Anaplasma* were only visible when infection density was relatively high.

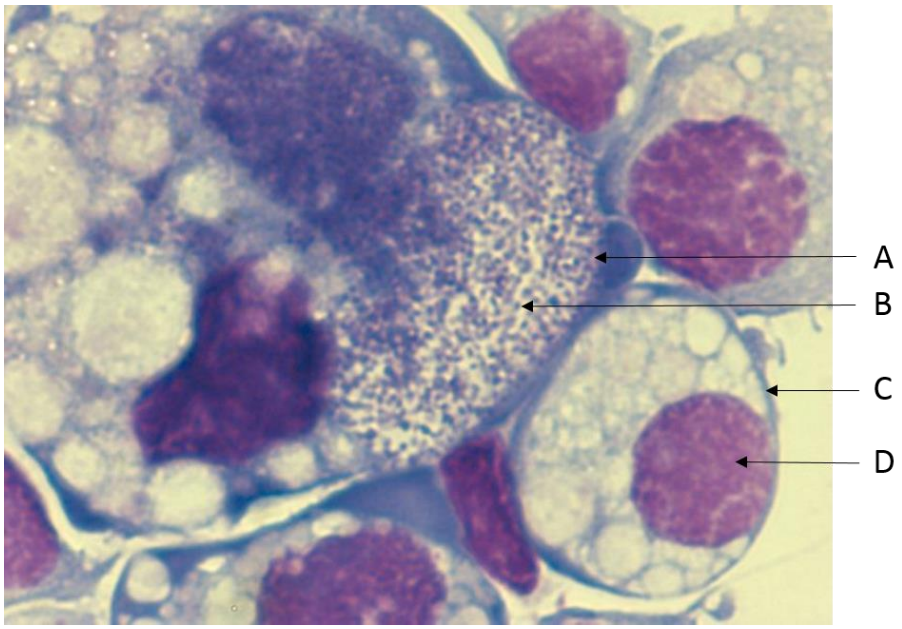


Figure 4. 4 Infection of ISE6 cells shows *Anaplasma phagocytophilum* infection

A: ISE6 infected cell at 8 dpi; B: *A. phagocytophilum* morula inside membrane bound vacuole; C: uninfected cell; D: nucleus. Cytocentrifuge smears for cell lines used in this experiment were stained with 10% Giemsa solution and infection density was examined using brightfield microscopy (Carl Zeiss, Germany).

4.3.2 Preliminary *in vitro* experiments

4.3.2.1 qPCR of infected *ISE6* cell lines identified suitable time points for proteomic studies

A qPCR assay was used to confirm the successful establishment of infection and quantify the infection rate of *A. phagocytophilum* (strain Feral Goat) in ISE6 tick cells at different time points. The findings here consequently contributed to the experimental design to determine the parameters for the main proteomic experiments.

Using duplicates samples for each time point, evidence of infection was detected in ISE6 cells 24 h after inoculating the bacteria. The copy number of *A. phagocytophilum* fluctuated over time, but was higher at 8 dpi compared with 24 h (**Figure 4. 5 A**). In contrast, tick genome copy numbers decreased by 48 h as a result of infection (**Figure 4. 5 B**). The overall ratio (*Anaplasma*:tick) increased by >3 logs between 5 and 8 dpi (**Figure 4. 5 C**). These findings suggest that there is a reduction in bacterial numbers after a proportion of the inoculum successfully invades the host cells. Following intracellular replication, host cells lyse, reducing viability of the cultures. Thus, we decided to test whether proteomics was sufficiently sensitive to detect changes in *Anaplasma* protein expression during the host cell invasion process.

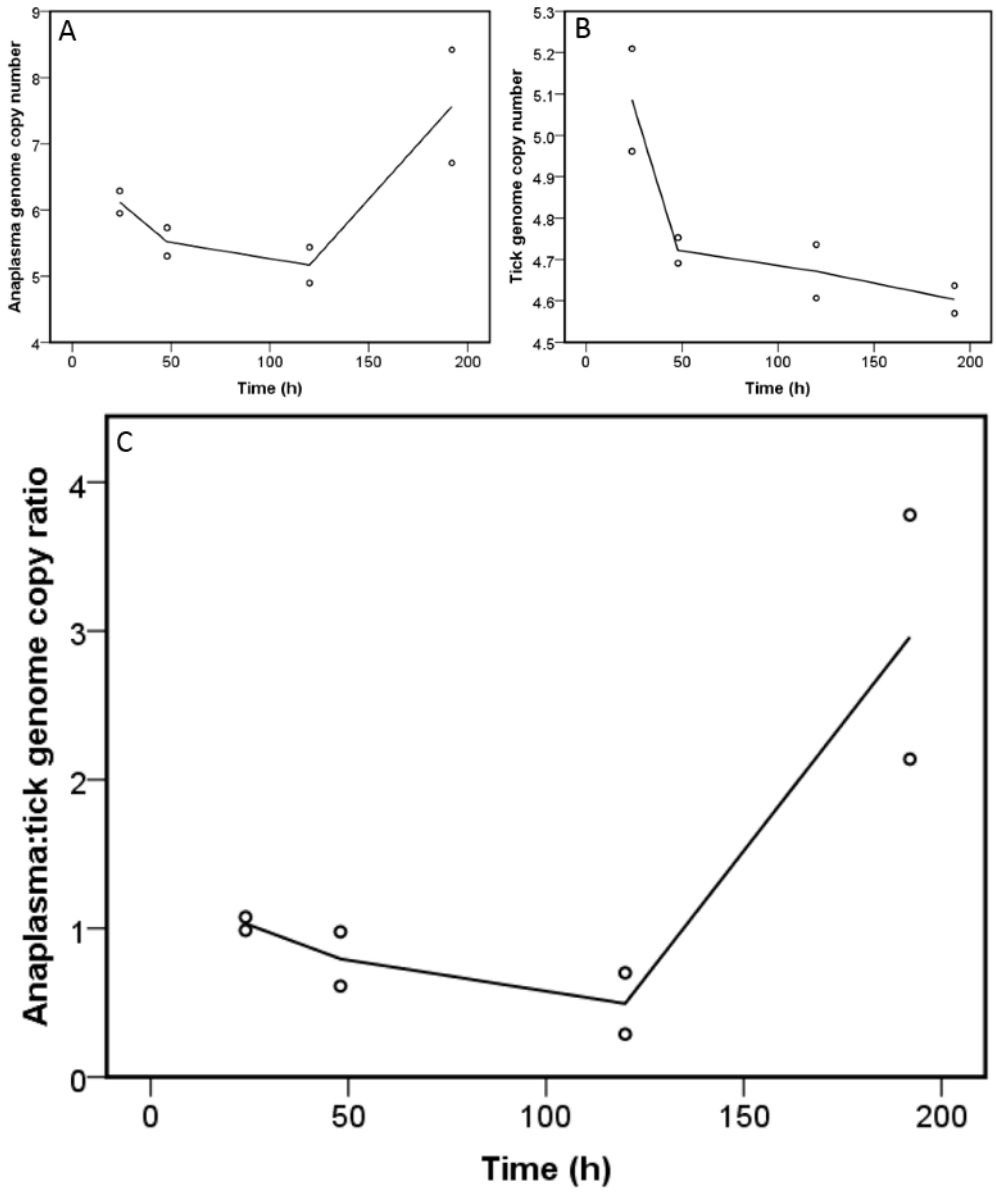


Figure 4. 5 (A) 16S rRNA copy numbers of *A. phagocytophilum* in tick cells in diffrent time points.

(B) *IsRPL6* tick gene copy numbers in different time points. **(C)** The ratio of *A. phagocytophilum* and tick cells in different time points using q-PCR. Data were obtained from duplicate cultures per time-point.

4.3.2.2 Label-free mass spectrometry detected a small number of *Anaplasma phagocytophilum* proteins after 3 dpi

An initial experiment was performed to determine if *Anaplasma* proteins could be detected under the "best-case scenario"; that is, at 8 dpi when the *Anaplasma*:tick ratio reached its highest point. For this, we used a single culture tube of ISE6 (*I. scapularis*) cells infected with *A. phagocytophilum* Feral Goat strain. Considering only high-confidence protein identifications obtained by shotgun MS/MS (≥2 unique peptides), 138 *Anaplasma* proteins and 1,323 tick proteins could be quantified. To investigate whether bacterial proteins were detectable at low levels of infection (*i.e.*, during potential host-cell invasion), the experiment was repeated with quadruplicate cultures of infected and uninfected cells, which were harvested at day 3. The numbers of high-confidence identifcations at this time-point were 22 proteins for *Anaplasma* and 1,660 proteins for the tick cells. Although the number of *Anaplasma* protein identifications was small, 423 downregulated and 291 upregulated tick proteins were detected by Progenesis (inclusion criteria: $\geq$ 2 unique peptides, $\log_2$ FC $\geq$1 or -1, and q value <0.05). This enabled a preliminary Pfam enrichment analysis to be conducted on the regulated proteins, revealing 8 downregulated and 5 upregulated protein families (**Table 4. 3**).

Table 4. 3 Preliminary proteomic response of tick cells to *Anaplasma* infection at 3 dpi using quadruplicate cultures of infected and uninfected cells. Pfam enrichment analysis.

Pfam_ID	Short name	Sorted enrichment (infected *v*. uninfected)	P-value	Putative function
PF08542	Rep_fac_C	-35.9	8E-04	Sliding clamp (DNA replication)
PF01798	Nop	-35.9	1E-02	Ribonucleoproteins involved in RNA methylation
PF08075	NOPS	-35.9	1E-02	Ribonucleoproteins involved in RNA methylation
PF00689	Cation_ATPase_C	-26.9	4E-02	Transmembrane ion transporters
PF01105	EMP24_GP25L	-23.9	7E-03	Intracellular transport
PF01145	Band_7	-13.5	3E-03	Cation conductance
PF00071	Ras	-6.5	2E-04	Signal transduction
PF00271	Helicase_C	-4.8	4E-02	Processing of mRNA 5'-UTRs
PF01920	Prefoldin_2	+41.8	1.00E-03	Actin assembly
PF01096	TFIIS_C	+39.1	2.00E-02	Zinc finger motif in transcription factor IIs
PF03357	Snf7	+23.2	1.00E-02	Multivesicular body formation
PF00412	LIM	+11.1	3.00E-03	Zinc finger domain; cytoskeletal organisation
PF00076	RRM_1	+5.6	1.00E-02	RNA recognition motif; RNA processing

Although some bacterial proteins could be identified without the need for additional enrichment or purification, the quantities were insufficient to examine regulation of protein expression in the pathogen. Therefore, different combinations of *Anaplasma* strains and tick cell lines were evaluated to see if the density of bacteria in cultured cells could be increased. The British sheep strain, OS, is of particular interest because it has been used in a number of published *in vitro* and *in vivo* (Woldehiwet et al. 2002) studies. This strain was supplied by Dr Lesley Bell-Sakyi (the Pirbright Institute) in a cell line called IRE/CTVM19, which is derived from the European vector of *A. phagocytophilum*, *I. ricinus*. Additional cultures of strain OS were also provided in IDE8 cells, an alternative *I. scapularis* cell line. Preliminary proteomic analyses of *Anaplasma* infections in these cell lines at 3 and 8 dpi showed that the number of bacterial proteins identified was even lower than for culture in ISE6 cells (zero high-confidence identifications). This suggested that a longer culture duration might be required to maximise the number of infected cells, depending on the kinetics of growth for a specific pathogen-cell line combination. This was explored further by qPCR.

4.3.3 Quantification of *A. phagocytophilum* infection via q PCR

As a result of the findings reported from the preliminary proteomics experiment, both Cytospin and qPCR was utilised again to examine infection rates of strain OS after a longer duration of infection in the IRE/CTVM19 cell line, and compared with infection of ISE6 cells. The density of *A. phagocytophilum* was higher in ISE6, with genome copy ratios of 5.38 and 11.08 after 4 and 8 dpi respectively, compared with 0.26 and 2.28 in IRE/CTVM19 after 10 and 20 dpi, respectively (**Figure 4. 6** and **Figure 4. 7**). It was deemed likely that these copy ratios would be too low to

examine regulation of protein expression in *Anaplasma*. However, these two cell lines from different tick species, which seem to allow growth of the pathogen at very different rates, provided an opportunity to compare the host cell response over time using shotgun proteomics.

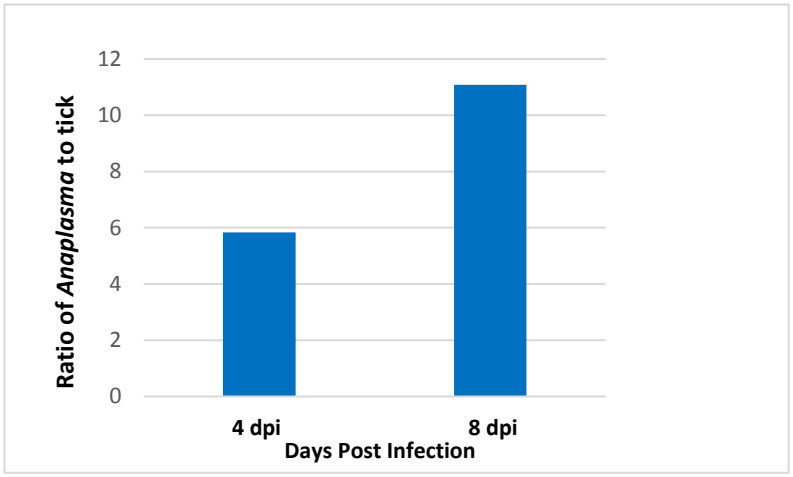


Figure 4. 6 The ratio of *A. phagocytophilum 16S rRNA* and *IsRPL6* genes in ISE6 cells at two time-points post-infection using qPCR (mean of duplicates for each time point).

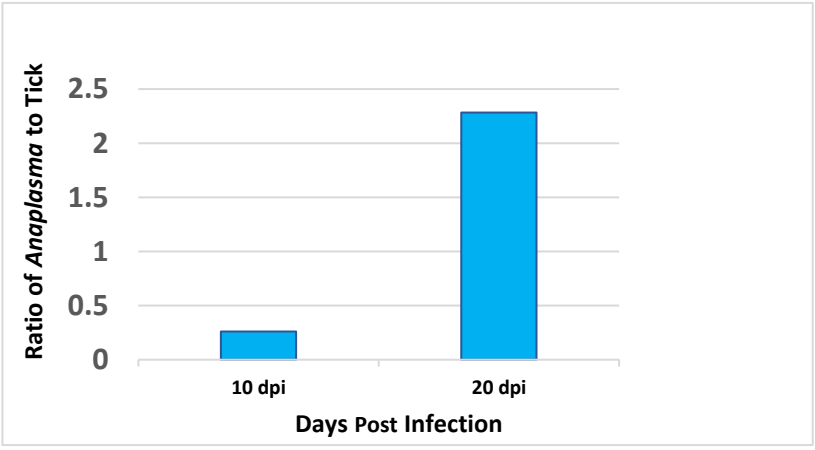


Figure 4. 7 The ratio of *A. phagocytophilum 16S rRNA* and *IsRPL6* genes in IRE/CTVM19 cells at two time-points post-infection using qPCR (mean of duplicates for each time point).

4.3.4 Proteomics results

4.3.4.1 Tick cell response to *A. phagocytophilum* infection is characterized by proteomic analysis

The main proteomic experiments used strain OS in ISE6 and IRE/CTVM19 cells, comparing infected cultures with control (uninfected) cultures in quadruplicate at 4 and 8 dpi (ISE6) or 10 and 20 dpi (IRE/CTVM19). Identification of regulated proteins in Progenesis and Pfam enrichment analysis was performed with the same inclusion criteria as for the preliminary experiment above (section 4.3.2.2). The number of identified and differentially expressed *A. phagocytophilum* and tick proteins in both cell lines at both time points is shown in (**Table 4. 4**).

Table 4. 4 Number of *A. phagocytophilum* and tick proteins identified or differentially expressed in two tick cell lines.

	Cell line	Protein identifications (≥ 2 peptides)[a]	Differential proteins[b]	
			T1[c]	T2[c]
Anaplasma	ISE6	71	67	57
	IRE/CTVM19	3	1	1
Tick	ISE6	1,460	55	135
	IRE/CTVM19	1,301	166	291

[a]Combined from four replicates each for infected and uninfected samples at each time point; note that MS spectra were searched against a combined *I. scapularis* and *I. ricinus* database.
[b]With ≥2 unique peptides and $\log_2$ FC ≥1 or -1 at $p < 0.05$. The average and variance was calculated at the individual protein level during ANOVA analysis in Progenesis.
[C]At each time-point compared with uninfected control cells.

Table 4. 5 Proteomic response of tick cells to *Anaplasma* infection: Pfam enrichment analysis[a].				
I. scapularis ISE6 cell line				
Pfam_ ID	**Pfam name**	**Enrichment**	***P* values**	**dpi**
PF16577	UBA domain	+949.7	7.00E-03	4
PF00878	Mannose 6-phosphate receptor	+569.8	7.00E-03	4
PF00564	PB1 domain	+219.2	4.00E-02	4
PF00262	Calreticulin protein family	+189.9	5.00E-02	4
PF00011	Hsp20	+92.9	7.00E-03	4
PF00076	RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain)	-10	2.00E-03	8
I. ricinus IRE/CTVM19				
PF01105	emp24/gp25L/p24 family/GOLD	+60	4.00E-02	10
PF01145	SPFH domain / Band 7 family	+57.6	1.00E-04	10
PF00735	Septin	-39.1	4.00E-02	20
PF00435	Spectrin repeat	-25.3	2.00E-02	20
PF00307	Calponin homology domain	-15.3	2.00E-02	20

[a]Analysis was based on four replicates each of uninfected and infected cultures.

Identification of *Anaplasma* proteins within infected IRE/CTVM19 cell cultures was low, even at 20 dpi. Consequently, proteomic analyses were focused on host cell responses to infection and not differential expression of bacterial proteins during infection. The Pfam enrichment analysis showed that five host protein families in ISE6 and two protein families in IRE/CTVM19 underwent significant upregulation during *A. phagocytophilum* infection. One protein family in ISE6 and three protein families in IRE/CTVM19 were downregulated (**Table 4. 5**).

4.3.4.2 Proteomic response of ISE6 and IRE/CTVM19 host cells to *Anaplasma* infection based on Gene Ontology

To further characterize differentially expressed proteins in response to *A. phagocytophilum* infection, gene ontology (GO) analysis was used to identify enriched host functions or processes during infection. UniP IDs were converted into GO IDs and were compared to the entire proteome (combined for *I. ricinus* and *I. scapularis*, or *I. scapularis* alone) using the agriGO tool. Infected samples were compared with their respective uninfected controls at both time points.

GO terms cover three domains; cellular component, molecular function and biological processes. Analysis of the latter two terms were utilised in this chapter; however, due to low numbers of proteins regulated at earlier sampling times (4 dpi/10 dpi), GO analysis was performed on proteins regulated at the later time points (8 dpi/20 dpi) only. Analysis of molecular function terms showed differential modulation between ISE6 and IRE/CTVM19, however modulations in biological processes were similar in both cell lines.

Significant upregulation of proteins associated with protein and calcium ion binding and a downregulation of proteins associated with RNA binding and nucleic acid binding and structural activity was detected in ISE6 cells at 8 dpi (**Figure 4. 8**). Interestingly, proteins associated with RNA binding and nucleic acid binding were also downregulated in IRE/CTVM19 as in ISE6 cells. Significant downregulation was also detected in cellular, protein catabolic process, actin cytoskeleton organisation, actin filament organisation and cellular protein assembly in IRE/CTVM19 (**Figure 4. 11**); however, no significant upregulation of proteins associated with other molecular functions was identified in IRE/CTVM/19 (**Figure 4. 10**).

The analysis of biological processes or biological function showed that the most modulated term in both cell lines corresponded to cellular processes. Additionally, significant downregulation was detected in gene expression, metabolic, biosynthesis processes and translation functions in ISE6 cells at 8 dpi (**Figure 4. 9**).

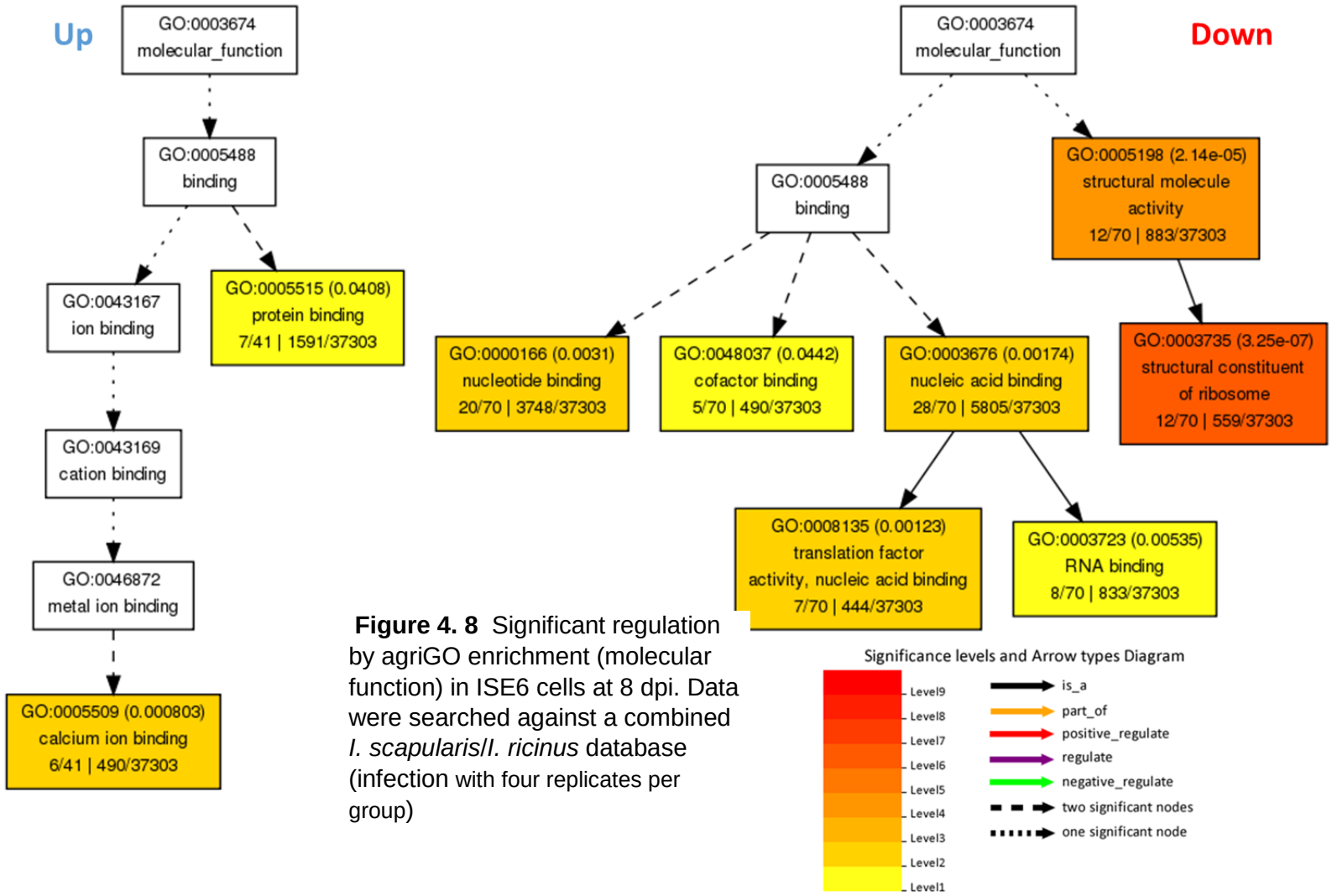


Figure 4. 8 Significant regulation by agriGO enrichment (molecular function) in ISE6 cells at 8 dpi. Data were searched against a combined *I. scapularis*/*I. ricinus* database (infection with four replicates per group)

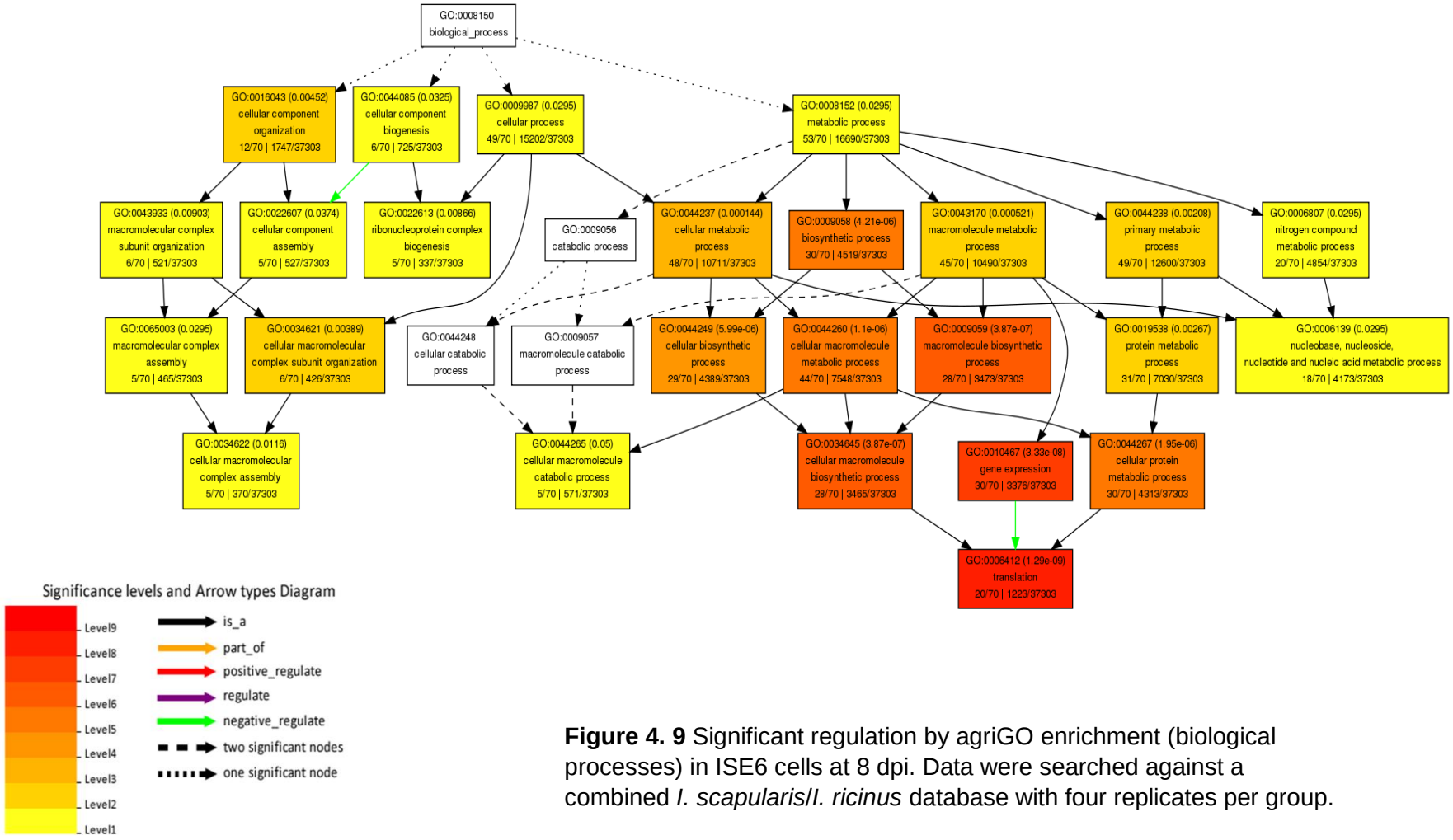


Figure 4. 9 Significant regulation by agriGO enrichment (biological processes) in ISE6 cells at 8 dpi. Data were searched against a combined *I. scapularis*/*I. ricinus* database with four replicates per group.

(no sig. upregulation)

Down

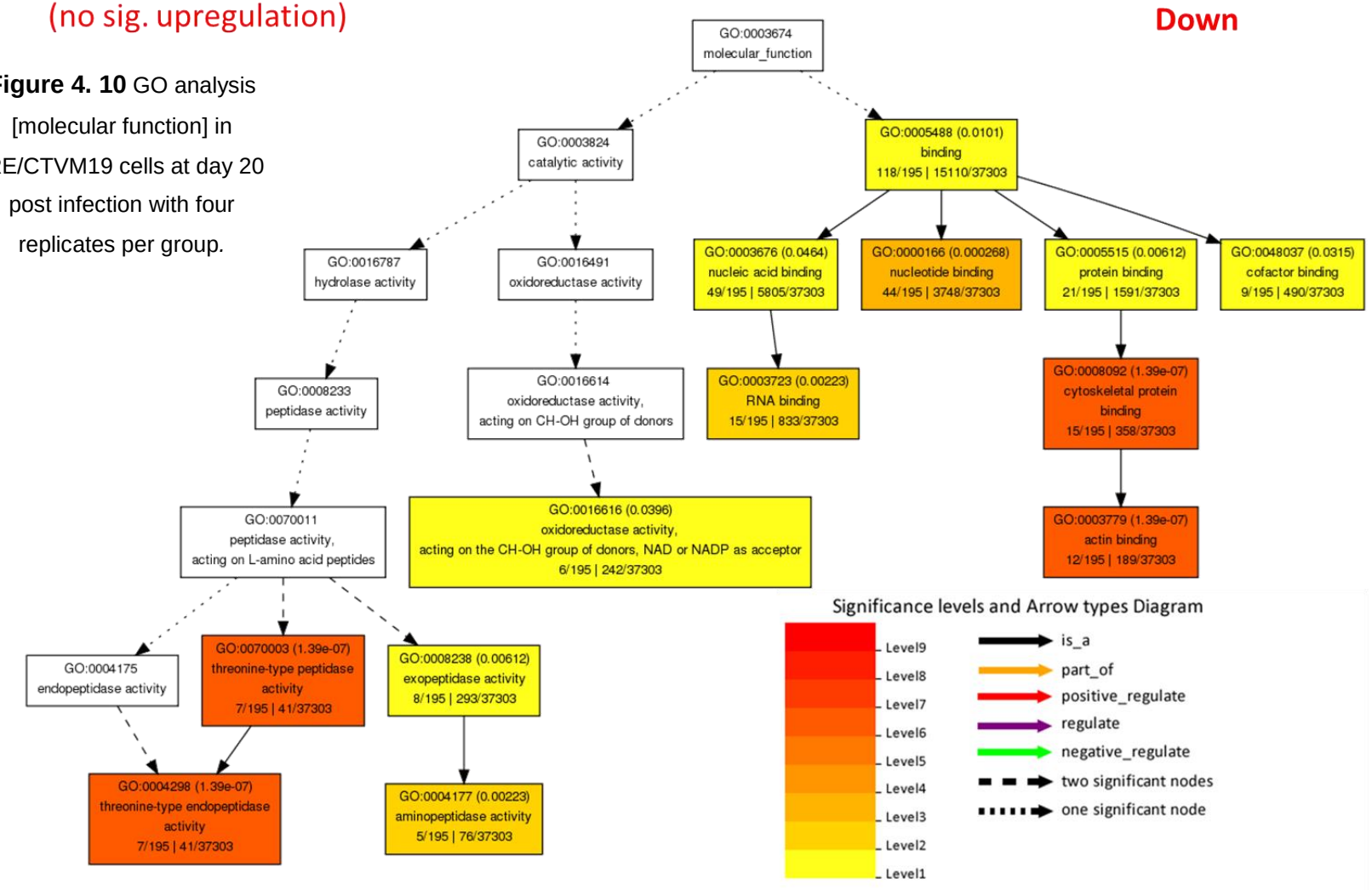


Figure 4. 10 GO analysis [molecular function] in IRE/CTVM19 cells at day 20 post infection with four replicates per group.

Figure 4. 11 Gene Ontology analysis [Biological process] in IRECTVM/19 cells at day 20 post infection with four replicates per group.

4.3.5 British *A. phagocytophilum* sample enrichment and draft genome results

The most successful method was the NEBNext Microbiome Kit (**Table 4. 6**) using material grown in IRE/CTVM19 cells. For the Mseek method, very high losses of DNA were encountered during the protocol such that the final sample was below the limits of detection. However, even the unenriched sample from IDE8 cells had a relatively high *Anaplasma*:tick ratio and was also used for sequencing alongside the material purified using the NEBNext Microbiome Kit. The DNA integrity is shown in (**Figure 4. 12**)

Table 4. 6 DNA concentration and *A. phagocytophilum* to tick cell genome ratio before and after bacterial purification.

Sample number and purification method types	DNA concentration	Ratio of *A. phagocytophilum* to tick cells before purification	Ratio of *A. phagocytophilum* to tick cells after purification
S1	2.63 µg	18.22	_
S2 Digestion using exonuclease V enzyme ("Mseek")	2.3 µg	20.05	N/A
S3 Enrichment using NEBNext Microbiome Kit	2 µg	14.17	39.05

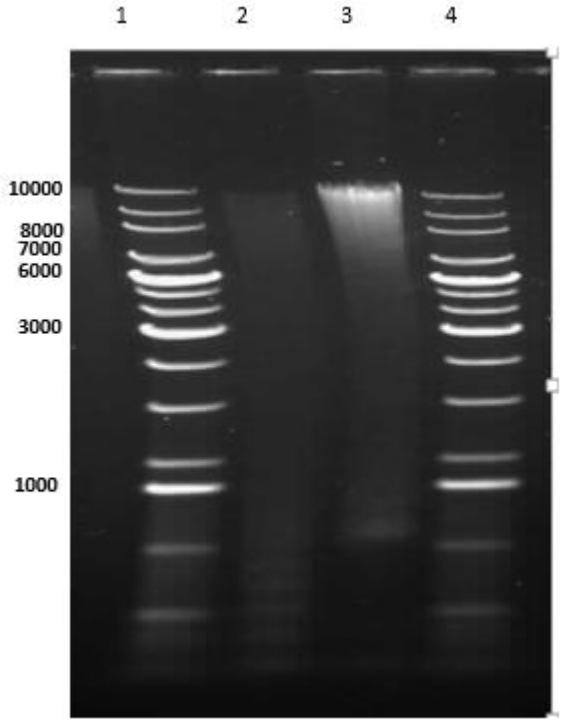


Figure 4. 12 Gel electrophoresis illustrates DNA integrity in 0.5% agarose.

Gel was run in 1X TAE buffer at 30V for 16 hours at 4°C. Lane 1, 4 = (1 Kb DNA ladder); lane 3 = *Anaplasma* sample.

By combining *Anaplasma* sequence reads obtained from both submitted samples, the strain OS genome was recovered in 268 contigs (**Figure 4. 13**). Blobplot analysis showed that coverage of the *Anaplasma* genome was quite low, even though the molar ratio of bacteria-to-tick was 20 - 40-fold (**Table 4. 6**). This reflects the fact that the tick genome is >2,000 times larger than the *Anaplasma* genome (Gulia-Nuss et al. 2016). The GC content of the genome of*A. phagocytophilum* OS strain was 41.7%.

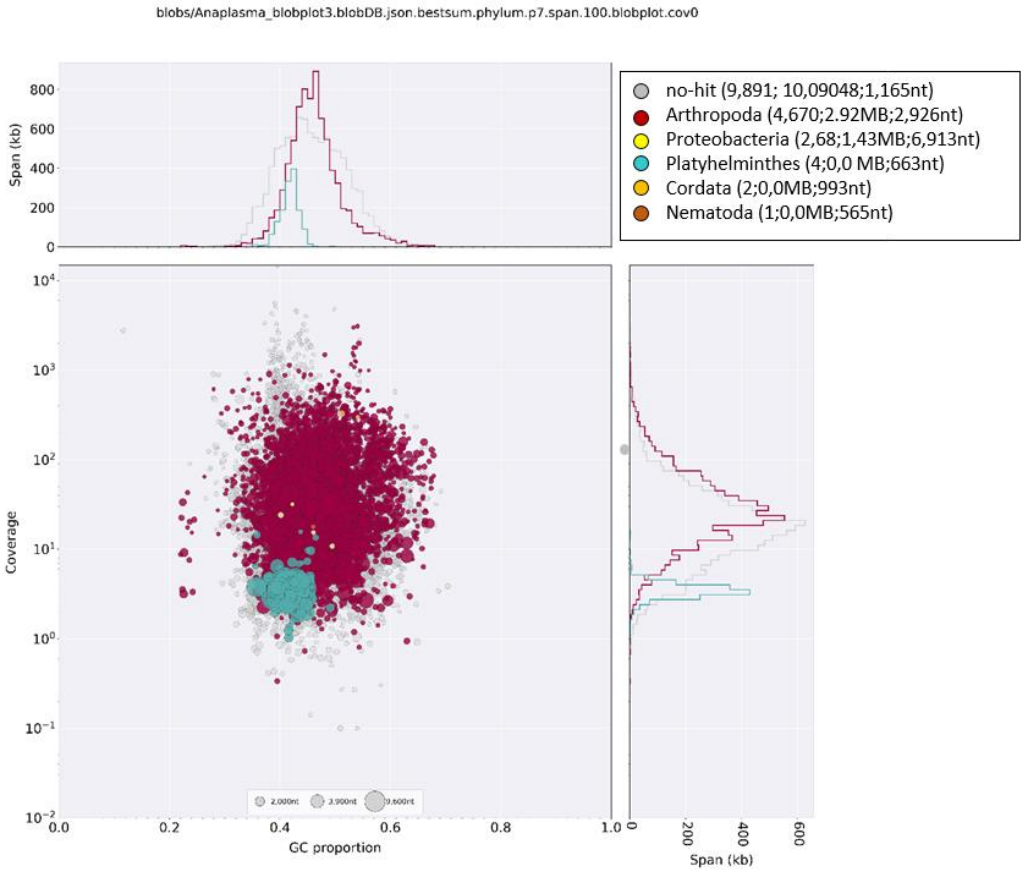


Figure 4. 13 Blobplot of genomic contigs with taxonomic assignment. "Proteobacteria" (in blue) represents *A. phagocytophilum* contigs.

4.3.6 Comparison of *A. phagocytophilum* strain OS genome to other *Anaplasma* strains.

A maximum likelihood phylogenetic tree based on the comparison of *A. phagocytophilum* strain OS with 29 other strains indicated that this British strain was clustered with the European bovine strains; whereas human, equine, canine and other strains of North American origin cluster away from the British and other European strains (**Figure 4. 14**); for host and location see **Table 4. 1**)

Comparison of the gene content (orthologous clusters) of the other 29 A. phagocytophilum genomes investigated in this study revealed that a total of 341 gene clusters comprise the core genome of *A. phagocytophilum*. The pan and core genome sizes were plotted against the number of genomes under study (**Figure 4.15**).

The remaining 578 gene clusters comprise the accessory genome, which also includes the species-specific unique genes. Furthermore, 123 genes were identified as unique genes for *A. phagocytophilum* OS strain, which is high compared to most of the other strains, although this could be due to a low minimum protein length in our annotated genome. These unique proteins have variable lengths (29-205 AA), and more than half of them are less than 60 AA. In addition, 79 (64.2%) were hypothetical proteins (**Appendix F** and **H**). However, only 31 genes which are absent from strain OS are present in all others, and 5 of these were hypothetical proteins (**Appendix G**).

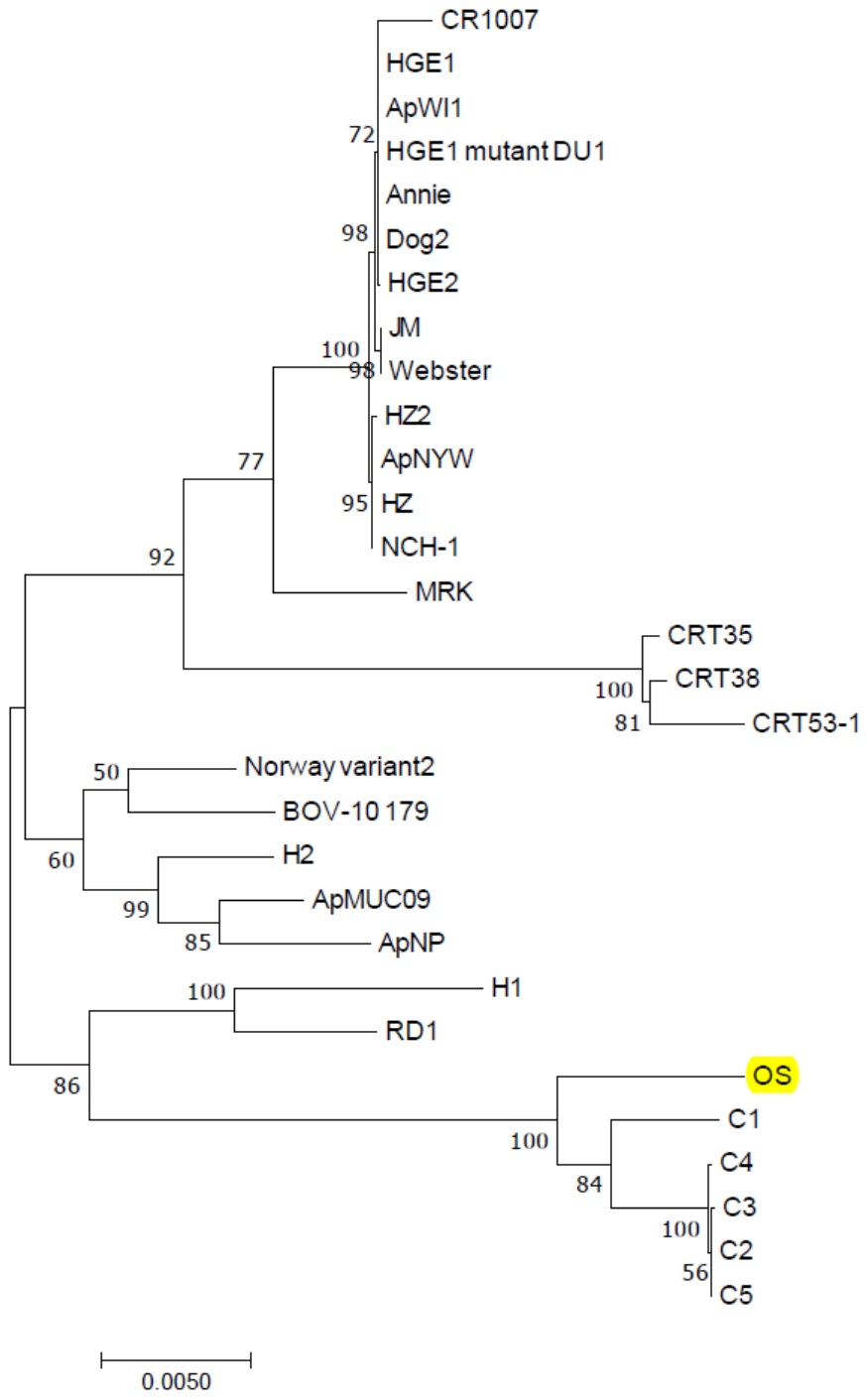


Figure 4. 14 Maximum likelihood phylogenetic tree of 30 *A. phagocytophilum* genomes including the British strain OS (highlighted in yellow). Geographic and host origin of the other genomes is presented in **Table 4. 1**.

The Cluster of Orthologous Groups (COG) and Kyoto Encyclopaedia of Genes and Genomes (KEGG) distributions were determined (**Figure 4. 16** and **Figure 4. 17**) for the core, accessory and unique proteins. This showed that the unique genes were dominated by the COG "Information storage and processing" and KEGG pathways associated with human diseases. In contrast, the core and accessory genes tended to be dominated by "Metabolism" according to both COG and KEGG. The relationship between *A. phagocytophilum* strains was further explored in a heatmap, which in agreement with the phylogenetic tree, clustered strain OS with bovine strains from France and Germany (**Figure 4. 18).** On the basis of searches against the HMM Database, 138 Msp2/p44-like protein sequences were identified in strain OS using the Pfam signature PF01617. These proteins were assessed for feasibility of detection by targeted proteomics (PRM).

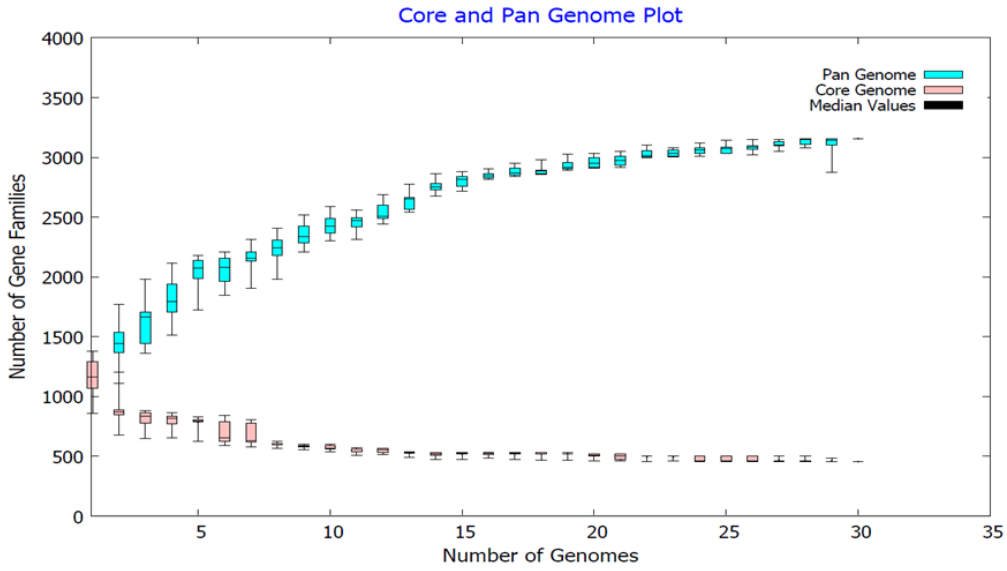


Figure 4. 15 Pan and core genome plot of *A. phagocytophilum* OS isolate compared with 29 other strains. Pan genome-blue; core genome- pink and median value –black.

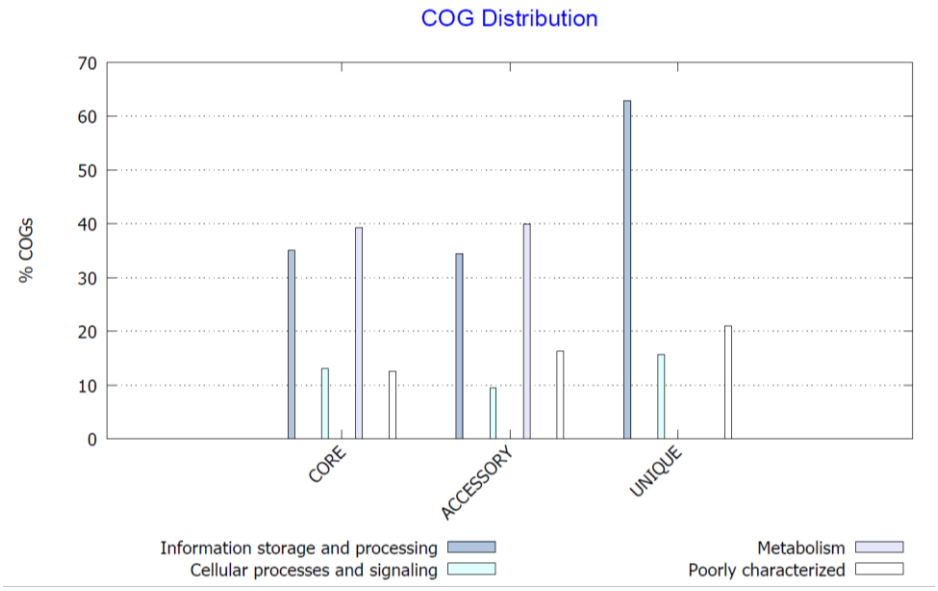


Figure 4. 16 COG distribution plot of core, accessory and unique genes.

Protein functions are associated with “Metabolism”, “Cellular process and signalling”, “Information storage and processing”, and “Poorly characterized”.

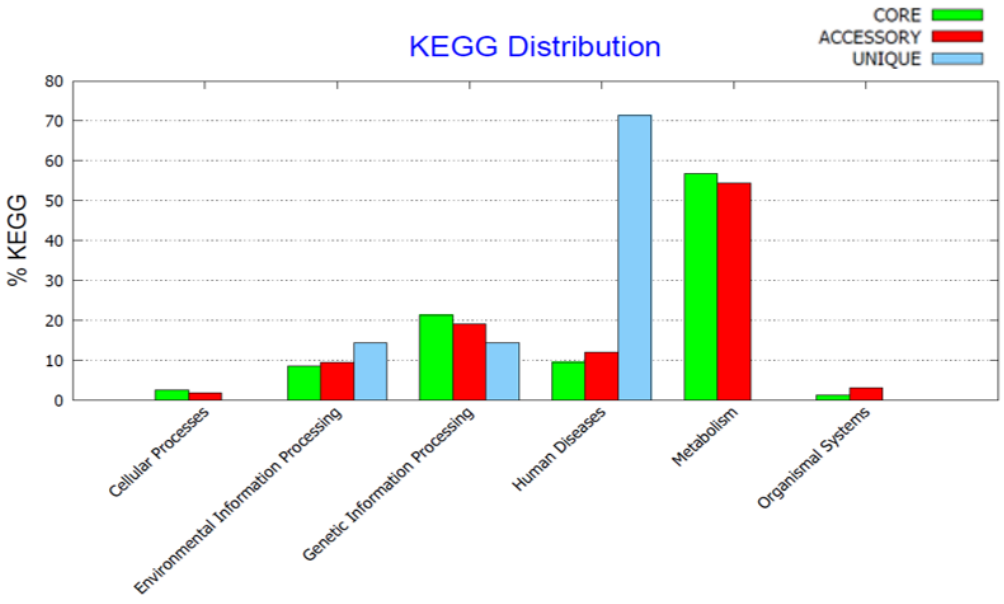


Figure 4. 17 KEGG distribution indicates the potential functions of genes in the *A. phagocytophilum* strain OS genome.

Pathways are associated with “Cellular processes”, “Environmental information processing”, “Genetic information processes”, “Human diseases”, “Metabolism”, and “Organismal systems”. Core (green bar), accessory (red bar) and unique (blue bar) genes.

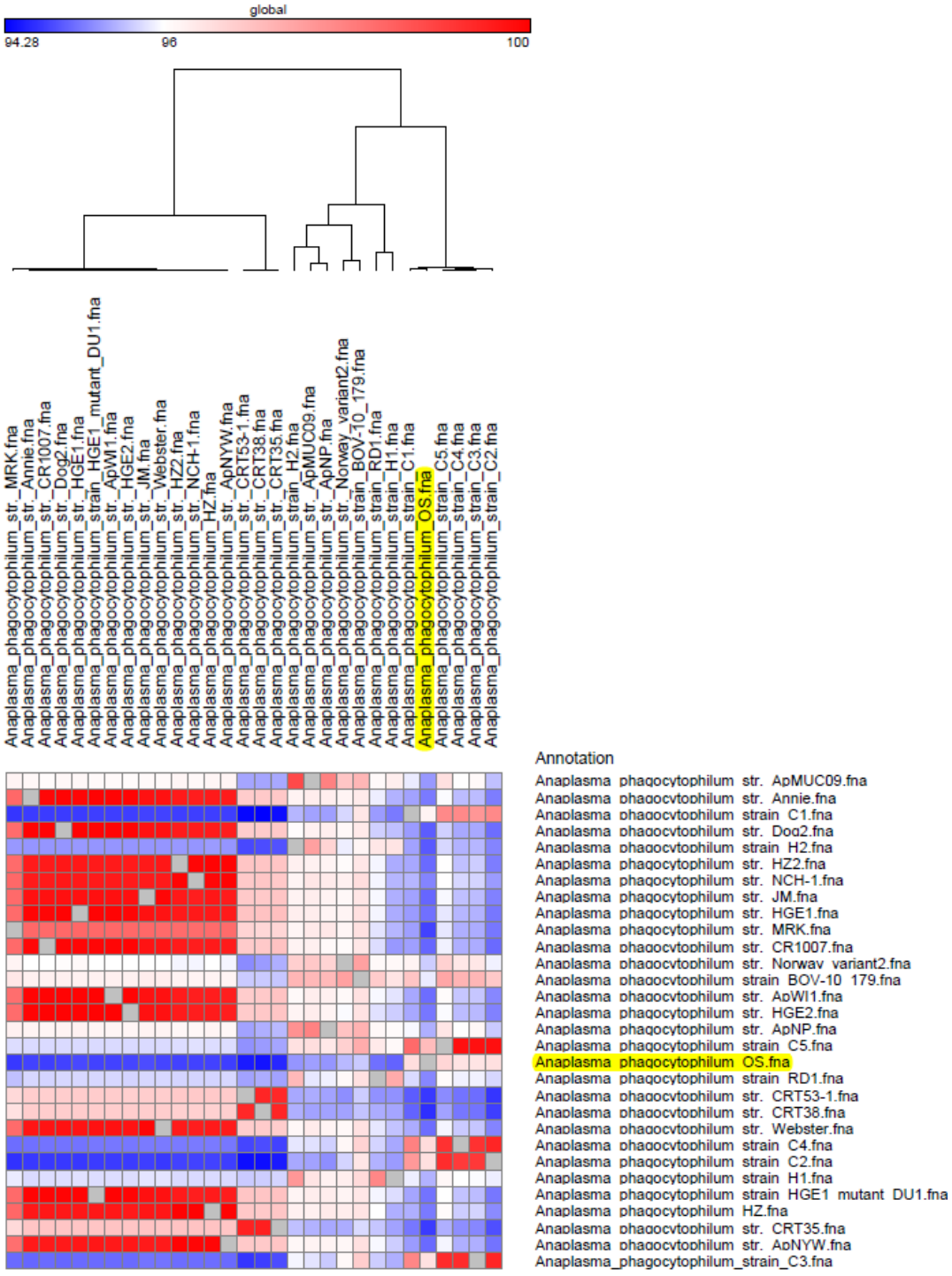


Figure 4. 18 Heatmap of *A. phagocytophilum* genome relationships. Dendrograms were generated by hierarchical clustering based on pair-wise distance. Strain OS is highlighted.

4.3.7 Targeted proteomics

Shotgun proteomics (section 4.3.3) was effective for evaluating regulation of tick cell proteins during *Anaplasma* infection, but failed to quantify sufficient bacterial proteins for any meaningful analysis of the pathogen. Therefore, we applied a targeted proteomics approach (PRM), as it provides high sensitivity and selectivity, together with high-throughput quantification with confident targeted peptide confirmation.

The Msp2/p44 outer membrane protein family is antigenically variable and known to interact with host cells as it is expressed on the pathogen's surface. The particular Msp2 proteins were chosen for PRM because there was some evidence for them in previous shotgun MS runs; individual peptides were chosen if there was prior evidence for them and they were unique to that protein. In addition, some peptides were included that were shared with many Msp2 proteins (labelled "universal" or Prokka_00004 and 00074) as a general positive control.

Overall, there was no real change in abundance in Msp2 proteins between the two time points. Indeed, the only significant (t-test, $p < 0.05$) change observed was for protein Prokka_00835, which decreased approximately twofold in the second timepoint (T2). Notably, only one protein was detected in the IRE/CTVM19 experiment (Prokka_00790) (**Figure 4. 19**) ,and this protein also constituted >90% of the total Msp2 protein signal in the ISE6 samples [considering both shotgun (Progenesis) and PRM analyses] (**Figure 4. 20**). The 'universal' peptides showed an overall increase in abundance over time, although this was not statistically significant

(**Figure 4. 19**). Surprisingly, PROKKA_00071 (**Figure 4. 20**) was detected by shotgun analysis in a small quantity, but was not recognised in the PRM analysis.

Orthologous clusters of strain OS Msp2 protein sequences with homologous sequences from the other 29 *A. phagocytophilum* strains, together with 15 *A. marginale* strains, was performed to elucidate if the dominant MSP2 (PROKKA_00790) (**Figure 4. 20**) detected by PRM had orthologues in other strains. Orthologues of this Msp2 protein were found in both *Anaplasma* spp., but the *A. marginale* sequences showed distinct differences compared with those from *A. phagocytophilum* (**Figure 4. 21**). Interestingly, most of the *A. phagocytophilum* strains that contain an orthologue of PROKKA_00790 were isolated from ruminant hosts, except strain Webster (from human) and a tick isolate (CRT53-1) with an unknown definitive host (**Figure 4. 21**).

In addition, a comparison of p44 expression between the present *in vitro* study in tick cells and a previous *in vivo* study in sheep using strain OS (Thomas et al. 2013) revealed virtually no overlap (**Table 4. 7)**. Although PROKKA_00004 was homologous to a p44 expressed at the RNA level in sheep blood, this sequence was targeted only by "universal" peptides in our PRM experiment, so cannot be assigned a non-redundant gene identity (**Table 4. 7**).

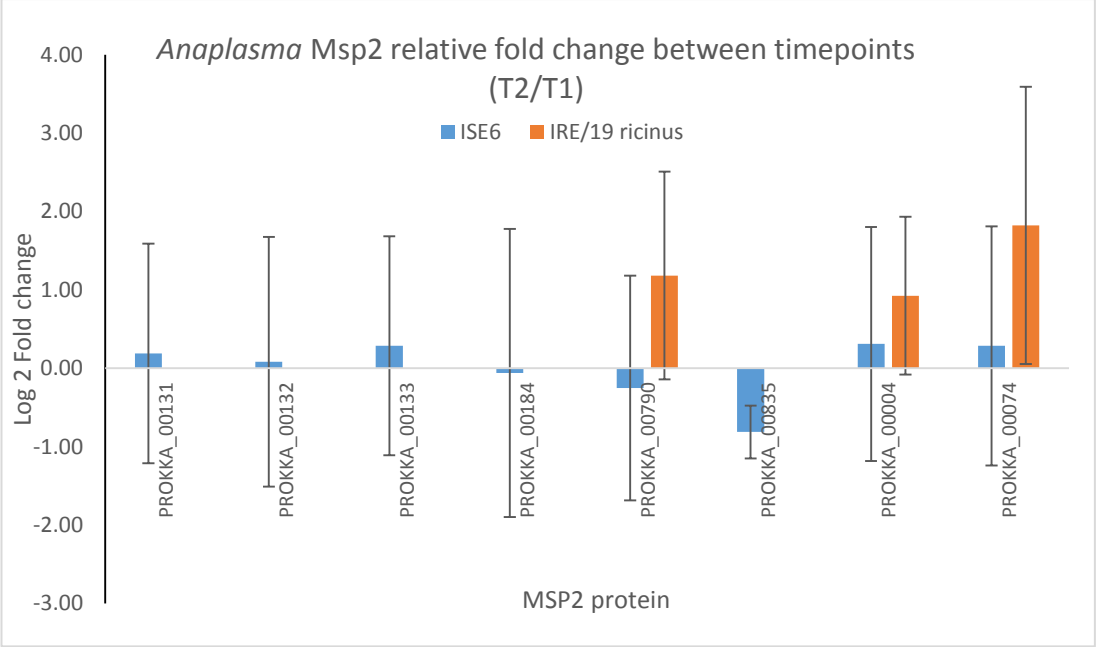


Figure 4. 19 Changes in Msp2 protein abundance between two time-points in ISE6 and IRE/CTVM19 cells infected with *A. phagocytophilum* strain OS.

Error bars are the confidence intervals (95%). Peptides were included that were shared with many Msp2 proteins (labelled universal or PROKKA_00004 and 00074) as a general positive control. The decrease in PROKKA_00835 expression is statistically significant (t-test, $p < 0.05$).

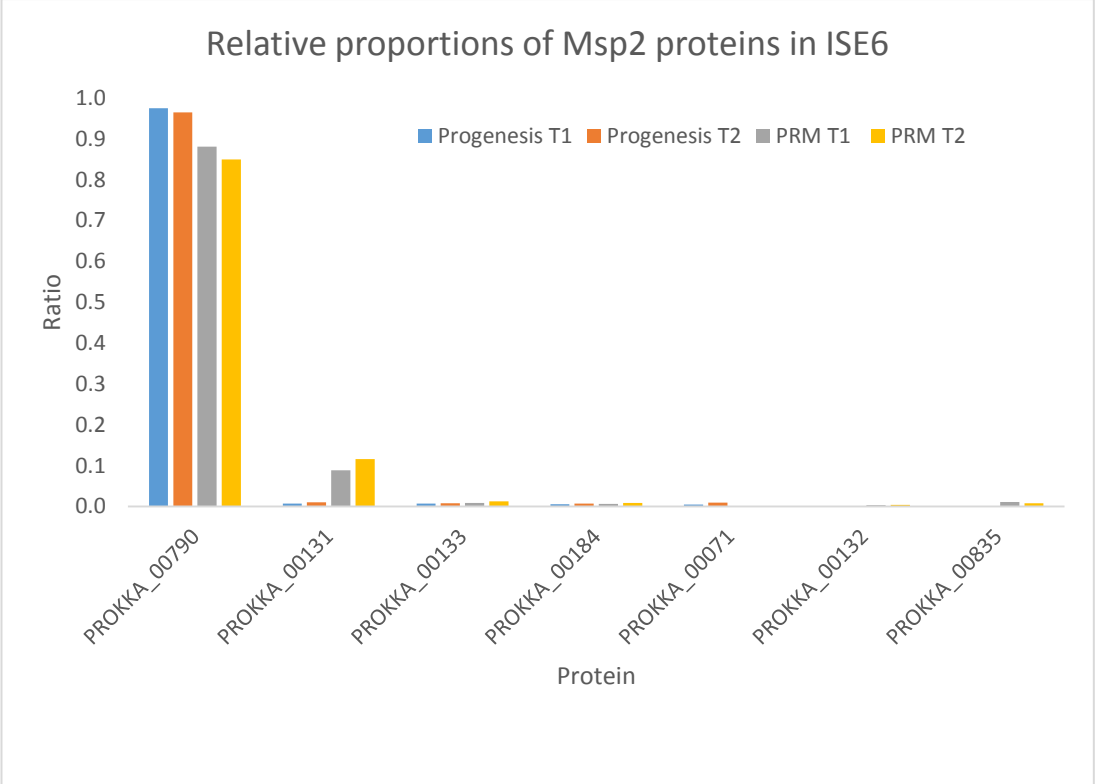


Figure 4. 20 Relative proportion of MSP2 in ISE6 only.

Analysis of quantitative data by both Progenesis (shotgun) and PRM is displayed

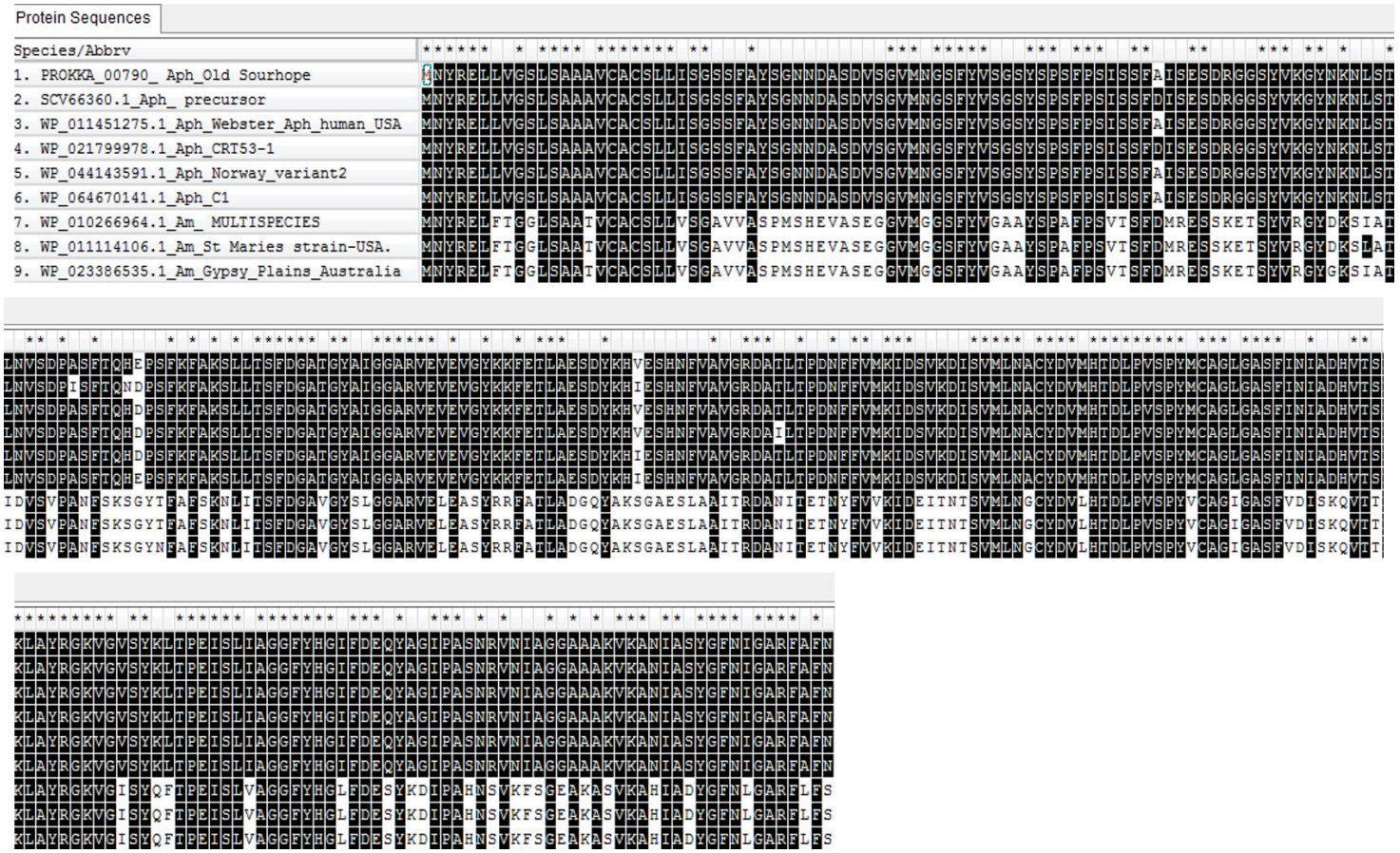


Figure 4. 21 Alignment of an orthologous cluster for the Msp2 family containing the dominantly expressed Msp2 from strain OS, PROKKA_00790. Sequences were searched from among 30 *A*. *phagocytophilum* and 15 *A. marginale* genomes.

Table 4. 7 A comparison between Msp2/p44 proteins expressed *in vitro* in the present study with a previous *in vivo* study in sheep (Thomas et al. 2013).

Msp2/P44 accession number of the present study	Msp2/P44 accession number (NCBI) from (Thomas et al. 2013)	Protein name
PROKKA_00603	KF481646.1_prot_AGV04702.1_1	*Anaplasma phagocytophilum* p44-31 (p44-31) gene
PROKKA_00261	KF481672.1_prot_AGV04727.1_1	*Anaplasma phagocytophilum* p44-57 (p44-57) gene
PROKKA_01212	KF481633.1_prot_AGV04689.1_1	*Anaplasma phagocytophilum* p44-18 (p44-18) gene
PROKKA_01265	KF481655.1_prot_AGV04711.1_1	*Anaplasma phagocytophilum* p44-40 (p44-40) gene
PROKKA_01224	KF481669.1_prot_AGV04724.1_1	*Anaplasma phagocytophilum* p44-54 (p44-54) gene
PROKKA_00672	KF481627.1_prot_AGV04683.1_1	*Anaplasma phagocytophilum* p44-12 (p44-12) gene
PROKKA_00496	KF481625.1_prot_AGV04681.1_1	*Anaplasma phagocytophilum* p44-10 (p44-10) gene
PROKKA_01143	KF481653.1_prot_AGV04709.1_1	*Anaplasma phagocytophilum* p44-38 (p44-38) gene
PROKKA_01172	KF481637.1_prot_AGV04693.1_1	*Anaplasma phagocytophilum* p44-22 (p44-22) gene
PROKKA_01095	KF481671.1_prot_AGV04726.1_1	*Anaplasma phagocytophilum* p44-56 (p44-56) gene
PROKKA_00497	KF481623.1_prot_AGV04679.1_1	*Anaplasma phagocytophilum* p44-8 (p44-8) gene
PROKKA_00553	KF481662.1_prot_AGV04718.1_1	*Anaplasma phagocytophilum* p44-47 (p44-47) gene
PROKKA_01173	KF481645.1_prot_AGV04701.1_1	*Anaplasma phagocytophilum* p44-30 (p44-30) gene
PROKKA_01240	KF481634.1_prot_AGV04690.1_1	*Anaplasma phagocytophilum* p44-19 (p44-19) gene
PROKKA_00260	KF481644.1_prot_AGV04700.1_1	*Anaplasma phagocytophilum* p44-29 (p44-29) gene

PROKKA_01108	KF481641.1_prot_AGV04697.1_1	*Anaplasma phagocytophilum* p44-26 (p44-26) gene
PROKKA_01151	KF481674.1_prot_AGV04729.1_1	*Anaplasma phagocytophilum* p44-59 (p44-59) gene
PROKKA_00004*	KF481629.1_prot_AGV04685.1_1	*Anaplasma phagocytophilum* p44-14 (p44-14) gene
PROKKA_00813	KF481666.1_prot_AGV04722.1_1	*Anaplasma phagocytophilum* p44-51 (p44-51) gene
PROKKA_01112	KF481673.1_prot_AGV04728.1_1	*Anaplasma phagocytophilum* p44-58 (p44-58) gene
PROKKA_00240	KF481679.1_prot_AGV04734.1_1	Anaplasma phagocytophilum p44-64 (p44-64) gene
PROKKA_00934	KF481647.1_prot_AGV04703.1_1	*Anaplasma phagocytophilum* p44-32 (p44-32) gene
PROKKA_00287	KF481621.1_prot_AGV04677.1_1	*Anaplasma phagocytophilum* p44-6 (p44-6) gene
PROKKA_00375	KF481618.1_prot_AGV04675.1_1	*Anaplasma phagocytophilum* p44-3 (p44-3) gene
PROKKA_00566	KF481626.1_prot_AGV04682.1_1	*Anaplasma phagocytophilum* p44-11 (p44-11) gene
PROKKA_01094	KF481643.1_prot_AGV04699.1_1	*Anaplasma phagocytophilum* p44-28 (p44-28) gene
PROKKA_00474	KF481638.1_prot_AGV04694.1_1	*Anaplasma phagocytophilum* p44-23 (p44-23) gene

*Observed in vitro in the present study, but peptides are shared with other p44 proteins.

4.4 Discussion

The mechanisms used by *A. phagocytophilum* to infect and multiply in both tick and vertebrate host cells are still poorly defined. Exploiting "–omics" technologies can help elucidate these host-pathogen interactions and identify key processes involved in the establishment of infection. Within this chapter, I demonstrated that "-omics" technologies can be used on ISE6 and IRE/CTVM19 cell lines and report host cell responses to infection detected by both label-free shotgun and PRM mass spectrometry.

Interestingly, *A. phagocytophilum* infection rates were variable between the two species of tick cell lines, with low replication of OS strain in *I. ricinus* (IRE/CTVM19) cells, even though this is the main natural vector in the UK. However, as *A. phagocytophilum* OS strain is routinely propagated in ISE6 cell lines (Woldehiwet et al. 2002), this finding is not surprising, as some level of laboratory adaptation of this strain to ISE6 cells has likely occurred. However, due to the variation between host cell susceptibility and strain infectivity, proteomics experiments had to be designed to achieve equivalent levels of infection in both cell lines and were not focused on a specific biological event.

Analysis for enriched protein family domains in regulated tick proteins identified several protein families modulated in response to infection, some of which have previously been reported (Rikihisa 2010; Truchan et al. 2013; Ayllón et al. 2013; Alberdi et al. 2016b). Previous studies have identified several key mechanisms that are crucial for successful bacterial invasion. The host response to *A. phagocytophilum* infection involves the modulation of the actin cytoskeleton (for bacterial internalisation and formation of the bacterial vacuole), followed by inhibition of

apoptosis, host cell lipid scavenging (hijacking of cholesterol), and manipulation of immune responses such as antimicrobial peptides (de la Fuente et al. 2016; Alberdi et al. 2016b; Ayllón et al. 2015a).

After *Anaplasma* infection of ISE6 cells, five protein families were upregulated; heat shock proteins (HSP), UBA domain proteins, calreticulin protein family, mannose 6-phosphate receptor and PB1 domain, while in IRE/CTVM19, two protein families were upregulated; SPFH domain/band 7 family and emp24/gp25L/p24 family/GOLD. However, downregulation of spectrin, spectrin repeat and calponin homolog domain were reported in IRE/CTVM19; and RNA recognition motif in ISE6. Several regulated responses where associated with exploiting the host cell cholesterol and host immune response.

Evaluating IRE/CTVM19 cells response to *A. phagocytophilum* infection revealed a downregulation of spectrin, spectrin repeat and calponin homology domains. Previously, a binary role of spectrin alpha chain or alpha-fodrin (an actin cross-linking protein that links the plasma membrane to the actin cytoskeleton) during *A. phagocytophilum* infection was reported. On one hand, activation of this protein promotes host cell cytoskeleton rearrangement that is required for successful pathogen infection, but on the other hand, it plays a role in host cell defence mechanism (host cell apoptosis induction) to limit *A. phagocytophilum* infection (Pascual et al. 1997; Ayllón et al. 2013). These results were further corroborated by GO analysis of the IRE/CTVM19 responses, which revealed downregulation of actin and cytoskeleton binding, suggesting a host defence response to control pathogen infection. In addition, downregulation of calponin homology domains correlated with spectrin downregulation, as spectrin is a multi-

domain protein involving two calponin homology domains (Pascual et al. 1997). Remodelling of actin cytoskeleton even after 20 dpi in IRE/CTM19 suggests early stage modulation related to spectrin and the cytoskeleton, reflecting the tick-pathogen interaction.

Interestingly multiple proteins associated with host immune response were reported in this study. In ISE6, upregulation of ubiquitin-associated domains (UBA domains) was detected. The ubiquitin system is a critical regulator of host-pathogen interactions. In response to infection, host cells initiate antimicrobial mechanisms including inflammatory signalling, phagosomal maturation, autophagy and apoptosis. Each of these responses is regulated by the ubiquitin system to minimise host cell damage. To establish and sustain infection, some pathogens, such as *Shigella flexneri* and *Listeria monocytogenes* have developed strategies to evade mammalian immune responses by exploiting the host ubiquitin system (Skaug et al. 2009). However, very little is understood about how microbes alter host ubiquitination processes in arthropod cells. Recent evidence suggests that *I. scapularis* ticks carry appropriate ubiquitination machinery and that an E3 ubiquitin ligase named X-linked inhibitor of apoptosis protein (XIAP) can restrict *Anaplasma* colonisation in ticks *in vivo* (Naranjo et al. 2013; Severo et al. 2013). As host cell ubiquitination is involved with many processes, further work is required to determine the intricacies of *Anaplasma*-modulated ubiquitin host cell responses.

This study also reports an up-regulation of NADPH phagocyte oxidase (Phox and Bem 1 domain) (PB1) in *A. phagocytophilum* infected ISE6 cells. Professional phagocytes are key components of the innate immune response which engulf and

destroy invading pathogens. A characteristic of phagocytosis is the production of superoxide (O_2^-) by activated NADPH oxidase in a tightly-regulated process (Babior, 1999). In *Anaplasma*-infected human neutrophil HL-60 cells, infection results in an induction of NADPH oxidase assembly and plasma membrane redistribution (Carlyon et al. 2004). In mice, the control of *Anaplasma* appears to be independent of phagocyte oxidase, which is activated during infection and leads to inflammation and hepatic injury; this is thought to be more damaging to the host than the *Anaplasma* itself (Scorpio et al. 2006). As arthropod immune systems are more simplistic than their mammalian counterparts, the role of phagocyte oxidase proteins are likely not equivalent in tick cell lines; however an up-regulation of PB1 hints that *Anaplasma* may modulate tick cell oxidation pathways.

In agreement with the observed up-regulation of host heat-shock protein 20 (HSP20) in ISE6 cells in this study, a modulation of host HSP20 has also previously been reported both *in vitro* and *in vivo* in response to *Anaplasma* infection and blood feeding (Mulenga et al. 2007; Busby et al. 2012). In particular, this response is characteristic of tick responses to blood-feeding and heat-shock stress. As *A. phagocytophilum* is transmitted via blood feeding, it would not be surprising if tick cells have developed a mechanism to protect the susceptible host against *Anaplasma* transmission during blood feeding (Busby et al. 2012). A previous transcriptomic and proteomic study comprising ISE6 cells and *I. scapularis* whole ticks revealed that both *A. marginale* and *A. phagocytophilum* infection at high temperatures leads to upregulation of HSPs expression (Villar et al. 2010). However, the transcriptional results did not demonstrate significant differences in mRNA levels of HSP20 and HSP70 expression in *A. phagocytophilum*-infected ISE6, suggesting that a

posttranscriptional mechanism prompted by *A. phagocytophilum* limits tick responses to infection (Villar et al. 2010).

Upregulation of host calreticulin in ISE6 cells in response to infection fits in well with known host-pathogen interactions previously reported (Truchan et al. 2016). During invasion of host cells, *Anaplasma* exploits the host cell resources by forming a pathogen-occupied vacuole (POV) where it can evade host immunity allowing it to persist and replicate. Within this vacuole, *Anaplasma* exploits the host by recruiting host cell cholesterol and calreticulin for bacterial growth and replication (Truchan et al. 2013; Huang et al. 2011). To date, this has been demonstrated in both HL-60 mammalian and ISE6 tick cells *in vitro* (Truchan et al. 2016). In addition, mannose-6-phosphate receptor (located on trans-Golgi vesicles), has been identified in the *A. phagocytophilum*-occupied vacuole as well (Huang et al. 2011); thus, up-regulation of calreticulin and mannose-6-phosphate together may indicate the formation of the POV (Truchan et al. 2016).

In IRE/CTVM19 cells, two protein families were upregulated following infection for 20 days; SPFH domain/Band 7 family and emp24/gp25L/p24 family/GOLD. The SPFH (stomatin, prohibitin, flotillin, and HflK/C) domain proteins are membrane proteins and one of the lipid raft membrane micro-domain components in eukaryotic and prokaryotic cells (Kuwahara et al. 2009). Many of the cellular activities like intracellular transportation and cell signalling are related to lipid raft membrane micro-domain that also contains considerable amounts of glycosphingolipid and cholesterol; however, the exact function of this family is still unclear (Dermine et al. 2001). A recent study demonstrated a role of SPFH in regulating the formation and function of large protein–cholesterol super-complexes

in the plasma membrane (Kuwahara et al. 2009). *Anaplasma phagocytophilum* lacks genes for cholesterol biosynthesis, and acquires cholesterol from host cells (Rikihisa 2010). Generation of protein–cholesterol super-complexes may explain the observed upregulation of (SPFH) domain proteins in IRE/CTVM19 cells reported here.

Another physiological role of SPFH domain family involves somatin, a cytoskeleton protein that plays an important role in maintaining the cell membrane structure in red blood cells (Kuwahara et al. 2009). In addition, interaction with F-actin is mediated by the SPFH domain (Langhorst et al. 2007), and reggie-1/flotillin-2 micro-domains at the plasma membrane were organized along actin filaments of the membrane cytoskeleton. Actin polymerisation resulted in reggie-1/flotillin-2 micro-domains immobilization, while actin reorganisation promoted exchange of reggie-1/flotillin-2 molecules between micro-domains (Langhorst et al. 2007). Therefore, rearrangement of the actin cytoskeleton by *A. phagocytophilum* identified in ticks and mammals *in vitro* (de la Fuente et al. 2017) could be another explanation for SPFH domain upregulation.

Another upregulated protein reported in IRE/CTVM19 was emp24/gp25L/p24 family/GOLD, which belongs to GOLD domain proteins (one of the eukaryotic domains) involved with transportation of molecules from the ER to the Golgi complex (Anantharaman & Aravind 2002). The P24 proteins of the GOLD domain appear to play an important role in vesicular or Golgi-membrane-associated protein complexes assembly via specific protein-protein interactions and lipid-binding of proteins to membranes (Anantharaman & Aravind 2002).

The GOLD-domain proteins comprises three different types: p24 domains that bind the membrane via a membrane-spanning helix; GCP60 domains that are located

peripherally of Golgi apparatus membranes; and GOLD domains that are located on the extreme amino or carboxyl terminus. Generation of protein-protein and protein-lipid complexes suggests a selective interaction role with specific proteins that might affect their packing into vesicles (Anantharaman & Aravind 2002). Highlighting the role of emp24 family/GOLD suggests that upregulation of emp24 family/GOLD observed in IRE/CTVM19 could be related again with the cholesterol hijacking mechanism induced by *A. phagocytophilum* infection.

Focusing on the most upregulated and downregulated proteins in response to *A. phagocytophilum* infection, similarities and differences were observed in host cell responses between IRE/CTVM19 and ISE6 cells by GO analysis. A shared host response to *Anaplasma* infection could be observed by a downregulation in proteins associated with cellular processes in both cell lines. However, both ISE6 and IRE/CTVM19 also exhibited species-specific host cell responses. This may be explained by the heterogenous nature of arthropod cell lines *in vitro* that has been illustrated previously (Alberdi et al. 2015; Villar et al. 2015b; Bell-Sakyi et al. 2007). Analysis of Molecular Function (MF) in ISE6 shows an upregulation of proteins involved with protein binding and calcium ion binding and a downregulation of RNA recognition motifs (RRM); and in IRE/CTVM19, a downregulation of proteins involved with actin and cytoskeleton binding was detected. Analysis of biological processes (BP) demonstrates a shared host cell response, with cell metabolic and catabolic processes downregulated in both ISE6 and IRE/CTVM19 cells, which agreed with the transcriptional response of *A. phagcocytophiulm* infection in another study (Alberdi et al 2016b).

One of the main molecular findings of current study is an upregulation of protein binding and calcium ion binding in ISE6. As previously reported, *E. chaffeensis* and *A. phagocytophilum* induce sequential signalling events involving cross-linking through transglutaminase 3, phospholipase C (PLC) and Tec protein tyrosine kinase (PTK), which consequently leads to upregulation of intracellular [Ca^{+}] required for host cell invasion. *Ehrlichia chaffeensis* and *A. phagocytophilum* infection depends on transglutaminase activity that necessitates receptor-mediated endocytosis to enter host monocytes instead of microfilament activity; thus cytosolic free calcium [Ca^{2+}] upregulation is required for *A. phagocytophilum* invasion (Rikihisa 2003). Other researchers documented manipulation of tick voltage-gated calcium (Ca^{2+} ion channels) by *A. phagocytophilum* to counteract tick immune responses or protein exocytosis to facilitate pathogen transmission (Alberdi et al. 2016b). In addition, upregulation of [Ca^{+}] has been confirmed in the human monocytic cell line THP-1 as a result of *Ehrlichia chaffeensis* infection (Lin et al. 2002).

Interestingly, GO analysis corroborated with Pfam in indicating downregulation of RRM and RNA binding in both cell lines, named as RNA-binding domain (RBD) or ribonucleoprotein domain (RNP); which could be explained by the inhibition of nutrient uptake that occurs as a result of *A. phagocytophilum* infection. In addition, *A. phagocytophilum* has previously been shown to regulate host cell gene transcription (which would involve RNA-binding proteins) in both IRE/CTVM19 cells and whole *I. scapularis* ticks (Sultana et al. 2010).

GO analysis of *Anaplasma* infected ISE6 cells revealed downregulation of biological processes involving metabolism, biosynthesis processes and translation. Upregulation of metabolic processes in the adult female midgut has been reported

previously using iTRAQ-based quantitative proteomics, which probably reflects the pathogen lifecycle (Villar et al. 2015a). In addition to these, other cellular processes related to the immune response, protein processing and glucose metabolism upon *A. phagocytophilum* infection have been previously reported (Villar et al. 2015a). Recent research suggests that ISE6 cells respond in the same way as hemocytes, and this could be because ISE6 was derived from embryonated eggs and contains cells with similar morphology and behaviour as haemocytes. These might attempt to counteract *A. phagocytophilum* infection, supporting a tissue-specific signature (Alberdi et al. 2016a).

Furthermore, GO analysis of IRE/CTVM19 cells in the present study again correlated with previous transcriptional analysis (Alberdi et al. 2016a), and integrated metabolomic, transcriptomic, and proteomic data (Villar et al. 2015a). Our findings illustrated downregulation of responses in IRE/CTVM19 cells against *A. phagocytophilum* infection associated with catabolic processes and cellular protein-protein assembly; and also correlated with tick midgut responses identified previously, consequently boosting the tissue-specific signature suggested before (Ayllón et al. 2015b; Alberdi et al. 2016a).

The current results provide evidence to support the use of tick *I. scapularis* ISE6 cells and *I. ricinus* IRE/CTVM19 cells to uncover the complex nature of tick-pathogen interactions, despite some differences within and between different cell lines. We have investigated the host response to *A. phagocytophilum* infection, which uncovered that ISE6 cells are more susceptible to infection than IRE/CTVM19. In turn, proteins were identified that might reflect the differences in *Anaplasma* growth rate in ISE6 versus IRE/CTVM19. Interestingly, our study identified a shared

response to infection: that is, cholesterol hijacking. Thus, the current study identified cell line-specific responses involving actin cytoskeleton remodelling that may reflect the early invasion stage in IRE/CTVM19; in contrast, several of the ISE6 responses were related with immune responses, which highlighted the host response towards the establishment phase of *A. phagocytophilum*. Finally, downregulation of metabolic and catabolic processes in both cell lines may be linked with the bacterial lifecycle associated with transcription and replication (correlated with midgut responses in whole ticks) as suggested previously.

Remarkably, *A. phagocytophilum* is able to invade and replicate inside mammalian leukocytes as well as multiple tick cell types, surviving significant environmental pressures successfully. This points to the importance of further exploring *A. phagocytophilum* genomic diversity. Here, we wanted to achieve greater insights into the British ovine strain OS by whole genome sequencing, followed by identification and characterization of the Msp2/p44 gene family. Consequently, this would facilitate understanding of the striking role that p44/Msp2 proteins play in *A. phgocytophilum* infection, pathogenesis, host adaptation, and evolution of this pathogen (Lin et al. 2004; Caspersen et al. 2002; Lin et al. 2003).

Notably, in spite of considerable economic loses as a result of tick-borne fever in sheep in the UK and Europe (de la Funte wt al.2005), no genome sequence was available for a British isolate of *A. phagocytophilum* prior to our study. To the best of our knowledge, the strain OS draft genome that we present here is only the second ovine strain sequenced worldwide, even though sheep are the major host presenting with clinical disease in Europe.

Interestingly, the current study revealed by phylogenetic analysis that the OS strain is closely related to European *A. phagocytophilum* isolates from cattle, and is more distant from human, equine, canine and other USA strains which cluster together. This finding corroborates previous studies which documented that European ruminant strains are segregated for human, canine and equine strains from the USA (Fuente et al. 2005; Barbet et al. 2013). This suggests a possible differentiating factor using DNA sequence-based diagnostic inspection to differentiate human-infective strains, which would help in tracking the existence of human-virulent strains in tick vector populations (Barbet et al. 2013). A recent study highlights a potential marker of human-virulent *A. phagocytophilum* using the *drhm* gene (for distantly related to human marker); sequencing of Ap-variant 1 (ruminant), MRK (horse), European sheep, and Californian woodrat strains confirmed the existence of this gene, in contrast with consistent deletion of this gene in strains that infect human cells *in vivo* and *in vitro.* However, canine strains from Minnesota or Wisconsin also exhibit a deletion of *drhm* (Al-Khedery & Barbet 2014), which could be applied in epidemiological studies in order to discriminate human and canine virulent strains from other clades; but at the same time, this highlights the need for further research of diverse *A. phagocytophilum* strains.

The clustering analysis in the current study using 30 genomes altogether revealed 123 genes unique to the OS strain. This number was high compared to most of the other strains, but it could be due to a low minimum protein length in the OS annotated genome. A total of 341 gene clusters comprised the core genome of *A. phagocytophilum*. Notably, 138 proteins in strain OS corresponded to the Msp2/p44

gene family, which is of a similar magnitude to other sequenced strains (Dunning Hotopp et al. 2006; Thomas et al. 2013).

Successful propagation and persistence mechanisms exhibited by *A. phagocytophilum* promote and secure transmission and self-propagation beyond the life span of a single host, so the host-bacterial interaction must lead to manipulation of the host immune response by *A. phagocytophilum*. Antigenic variation is an outstanding molecular mechanism that allows *A. phagocytophilum* to evade host immune responses and assure persistency (Caspersen et al. 2002). Availability of Msp2/p44 assures a combinational strategy that *A. phagocytophilum* uses to generate new variants during chronic infection (Caspersen et al. 2002; Lin et al. 2003; Rikihisa 2010; Barbet et al. 2003).

The present study revealed that several Msp2/p44 proteins were expressed in ISE6 cells infected with *A. phagocytophilum* OS strain. Msp2 (classical *A. marginale* nomenclature) or p44 (classical *A. phagocytophilum* nomenclature) have a similar organization of conserved flanking 5' and 3' sequences and central pseudogenes containing a hypervariable region, and are also encoded by a multi-gene family containing some truncated gene copies, suggesting a possible source for a recombination mechanism (Zhi et al. 1999). Antigenic switching between *A. marginale* (Msp3 and Msp2) proteins is responsible for continuous generation of new epitopes that allows evasion of host immune responses in much the same way as the VSG in the monomolecular coat of trypanosomes (Brayton et al. 2003), thus assuring persistency through antigenic variation. Diminishment of antibody towards Msp2/P44 expressed during acute infection as persistent infection develops in lambs

corroborates the antigenic variation mechanism and guarantees persistency of *A. phagocytophilum* (Granquist et al. 2010).

Parallel reaction monitoring (PRM) was used to look for specific peptides from individual Msp2/p44 family members, as it is more sensitive than conventional proteomics and can obtain sufficient statistical power to establish association with proteins or biomarkers in a disease. However, the PRM analysis did not detect real changes in abundance between the two time-points, although protein Prokka_00835 decreased approximately twofold by the second time-point. Whereas 7 Msp2/p44 proteins were identified in *Anaplasma*-infected ISE6 cells, the Prokka_00790 protein was clearly dominant. Interestingly, phylogenetic mapping of the dominant protein with orthologous Msp2/p44 proteins from *A. phagocytophilum*, as well as from some *A. marginale* strains, revealed the conservation of this Msp2/p44 orthologue primarily in strains that infect ruminants. Thus, these strains could share some common characteristics, including reservoirs and pathogenicity, which may be different from strains that infect humans.

A previous study that targeted expression patterns of *A. marginale* Msp2 variants in response to growth in tick and mammals *in vitro* showed the predominance of a specific Msp2 sequence structure during infection of a host type (tick versus mammalian), highlighting the mosaic antigenic variation which is seen in *A. marginale* but not in *A. phagocytophilum* (Chávez et al. 2012). Comparing between *A. phagocytophilum* Msp2 variant expression patterns during infection of tick versus mammalian cells could be an area for future studies, using reactivity with specific antibodies in IFAT and Western blot approaches as applied with *A. marginale* before (Chávez et al. 2012). Remarkably, elegant use of monoclonal antibodies specific to

distinct Msp2 varaints was able to identify differences between Msp2 expressions in individual tick cells in the same culture by IFAT (Chávez et al. 2012). This might also occur with *A. phagocytophilum*, albeit rarely, judging by the levels of several minor p44 members observed in our PRM study.

CHAPTER FIVE

Investigation of *Rickettsia buchneri* presence in tick salivary glands

5.1 Introduction

The blacklegged tick *I. scapularis* (deer tick) constitutes a considerable threat to human health, as it is the predominant vector of Lyme disease, human graulocytic anaplasmosis and babesiosis in North America (Adelson et al. 2004; Magnarelli & Swihart 1991; Piesman & Spielman 1980). While ticks may harbour both pathogenic and non-pathogenic microflora, little is known about how the microflora of *I. scapularis* within the tick may influence the transmission of pathogens (Moreno et al. 2006). This is true even of its primary endosymbiont, a *Rickettsia* spp. originally named simply as the "rickettisal endosymbiont of *Ixodes scapularis*" (REIS) (Felsheim et al. 2009).

Recently, REIS has been named formally as *Rickettsia buchneri* sp. nov. (strain ISO-7^{T}) (Kurtti et al. 2016). This obligate intracellular bacterium was thought to be restricted to the ovaries of *I. scapularis* females on the basis of light and electron microscopic evidence (Munderloh et al. 2005). It was isolated in the tick embryonic cell lines IRE11 (*I. ricinus*) and ISE6 (*I. scapularis*), and this was confirmed by transmission electron microscopy (Kurtti et al. 2016). Ultrastructural identification in cultured cells revealed a similar appearance to *Rickettsia* in the ovaries of the blacklegged tick (coccobacillary rickettsiae identified in the cytoplasm, resulting in hyperatrophy), and to the *Rickettsia* genus in general (Kurtti et al. 2016).

Rickettsia buchneri belongs to the SFG rickettsiae (Gillespie et al. 2012); nevertheless, this bacterium has been detected in *I. scapularis* only (Kurtti et al.

2016). Phylogenetic analysis of *R. buchneri* genomes compared with other rickettsial species illustrated that it forms a distinct clade from rickettsiae of European tick species, although this also indicated that *R. monacensis* from *I. ricinus* is its closest relative (Kurtti et al. 2016).

Understanding the relationship between *I. scapularis* and *R. buchneri* could provide insights into the distribution of tick-borne diseases in the Eastern USA, as there is a strong potential for this symbiont to interact with pathogens within the tick. Moreover, if *R. buchneri* is found in tick tissues outside the ovary, especially the salivary glands, it may be exposed to the mammalian host, generating immune reponses of potential diagnostic significance.

5.2 Material and Methods

5.2.1 Tissue dissection

Partially-fed, adult *Ixodes scapularis* ticks were collected by deer hunters from various locations in New Hampshire, fixed in ethanol, and collated by Dr Clotilde Carlow and Dr Zhiru Li (New England Biolabs, Masschusetts). Dissection of salivary glands and ovaries was done by Stefano Gaiarsa and Prof. Sara Epis at the University of Milan. Ten ovaries (OV) and 10 salivary gland pairs (SG) from individual ticks were preserved in ethanol in 1.5 ml tubes.

5.2.2 Protein and DNA extraction

In Liverpool, protein extraction was performed by Dr. Stuart Armstrong on the pooled SG and OV samples. The protein concentration was measured using the Pierce Coomassie Plus (Bradford) protein assay (Bio-Rad, Hercules, CA, USA). A DNA

extraction from the remaining material was conducted using a Qiagen Blood and Tissue DNeasy kit according to the manufacturer's protocol.

5.2.3 GeLC-MS

To get an optimal separation of a wide range of proteins under denaturing conditions, we used 4 - 12% NuPAGE Bis-Tris protein gels (Invitrogen). Approx. 30 µg of the pooled OV and SG lysates were loaded onto the gel, and electrophoresis was conducted at 200 V (constant) for 35 min. The gel was stained overnight with GelCode Blue Stain Reagent (ThermoFisher Scientific). The entire gel lane was excised into 7 equal slices (~1 mm wide), which were then cut into smaller pieces (~1 mm^3). Gel pieces were destained in 50% acetonitrile/50 mM ammonium bicarbonate (pH ~8), reduced for 30 min at 37°C with 10 mM dithiothreitol (Sigma) in 50 mM ammonium bicarbonate, and alkylated with 55 mM iodoacetamide (Sigma) in 50 mM ammonium bicarbonate for 30 min in the dark at room temperature. Gel pieces were washed for 15 min in 50 mM ammonium bicarbonate and dehydrated with absolute acetonitrile. Acetonitrile was removed, and the gel plugs rehydrated with 0.01 µg/µL proteomic grade trypsin (Sigma) in 50 mM ammonium bicarbonate. Digestion was performed overnight at 37°C. Peptides were extracted from the gel plugs using successive 15 min incubations in 3% (v/v) acetonitrile in 0.1% (v/v) TFA. Peptide extracts were pooled and reduced to dryness using a centrifugal evaporator (Eppendorf Concentrator Plus), then re-suspended in 3% (v/v) acetonitrile, 0.1% (v/v) TFA for analysis by mass spectrometry.

5.2.4 Conventional PCR and qPCR

Both conventional PCR and qPCR (TaqMan) targeting the *Rickettsia* spp. *gltA* gene, a highly conserved gene for the majority of SFG rickettsiae, was used on the OV and SG tissue extracts as described in section 2.2.4. Normalisation of the qPCR against a tick single-copy nuclear gene, ribosomal protein L6 (*ISRPL6*), was done as described in section 2.2.3. The conventional PCR was performed as described in section 2.2.4, and PCR products from SG and OV were purified using a Qiaquick PCR purification kit (Qiagen, Germany) according to the manufacturer's instructions. Sequences were obtained by the Sanger method from Eurofins MWG Operon (Ebersberg, Germany), identified using BLAST (www.ncbi.nlm.nih.gov/BLAST), and compared with the reference gene *gltA* sequence (WP_014409697.1; *R. buchneri* strain ISO7 contig076, whole genome shotgun sequence) available in GenBank. The alignment in ClustalW was used to align and compare DNA sequences to the nucleotide collection database.

5.2.5 Proteomic analysis

Label-free mass spectrometry and geLC-MS followed by proteomics analysis were performed similarly to the *Anaplasma* experiments *in vitro* (section 4.2.11.1). Peptides were analysed by LC-MS/MS using a Q-Exactive mass spectrometer platform (Thermo Fisher Scientific). Each Thermo RAW file was imported into Progenesis QI (version 4.1, Nonlinear Dynamics) individually. Peak peaking and Mascot parameters were as described elsewhere (section 4.2.11.4). Tandem MS data were searched against a combined database from *I. scapularis* (Uniprot, June 2017) and REIS (NCBI BioProject PRJNA33979); a total of 22,584 combined sequences;

6,279,968 residues. Separate sample fractions were combined using Progenesis QI to create one output file. Quantitative analysis for proteins detected in both ovaries OV and SG was done in Progenesis QI for proteomics by Dr Stuart Armstrong. Protein intensity values were used to create heat-maps using the Morpheus online tool (Morpheus, https://software.broadinstitute.org/morpheus).

5.3 Results

The qPCR assay detected *R. buchneri* DNA in both OV and SG of *I. scapularis*. However, the bacterial density was significantly higher in OV (the primary location (**Figure 5.2**). Conventional PCR (**Figure 5. 1**), fallowed by Sanger sequencing, detected DNA that was 100% identical to *R. buchneri* in both tissues (**Figure 5. 3**). Moreover, geLC-MS demonstrated the presence of *R. buchneri* proteins in both OV and SG (**Figure 5. 4**). A comparison of expression levels of 29 proteins (with ≥2 unique peptides) between the two tissues showed that most proteins could only be detected in OV; although two chaperonins, heat-shock protein 20, and peptidoglycan-associated lipoprotein were robustly detected in SG (**Figure 5. 4**). Interestingly, several proteins encoded by the plasmids of *R. buchneri*, including a biotin synthase expressed in OV, were detected by the geLC-MS analysis (**Figure 5. 4**).

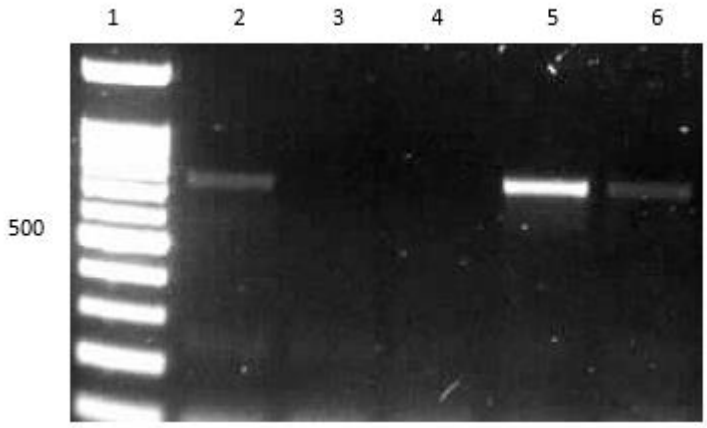


Figure 5. 1 Gel electrophoresis of *R. buchneri* in ovaries (OV) and salivary glands (SG) visualised on a 1.5% agarose gel. Lane 1= DNA ladder; lane 2, SG; lane 5, OV; and lane 6 (1/10 dilution of OV sample).

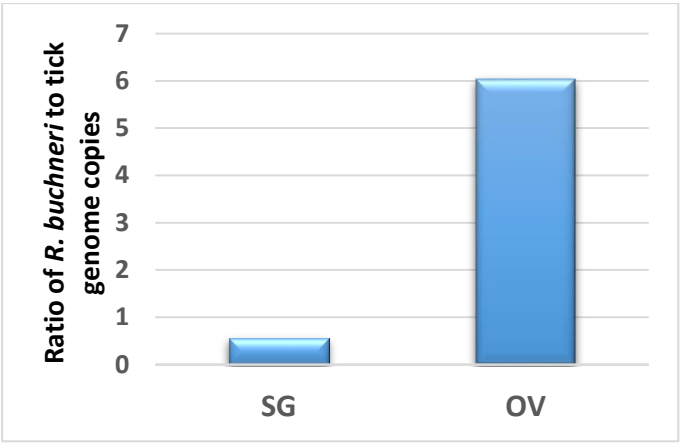


Figure 5.2 The ratio of *R. buchneri* to tick genome copies in different tissues (ovaries, OV; salivary glands, SG) using qPCR.

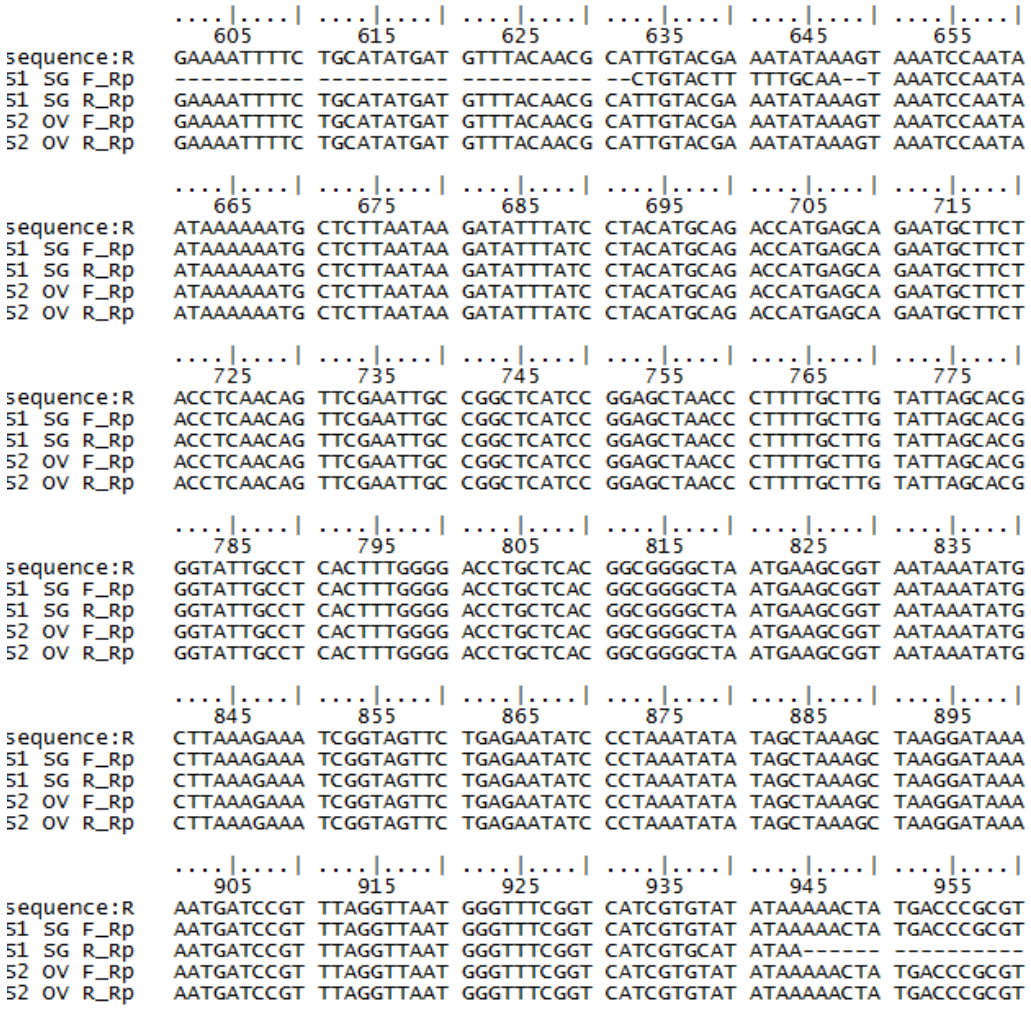


Figure 5. 3 Multiple alignment (clustalW) of *R. buchneri* (ovaries, OV; salivary gland, SG) *gltA* sequences with the corresponding sequence from the *R. buchneri* draft genome (top).

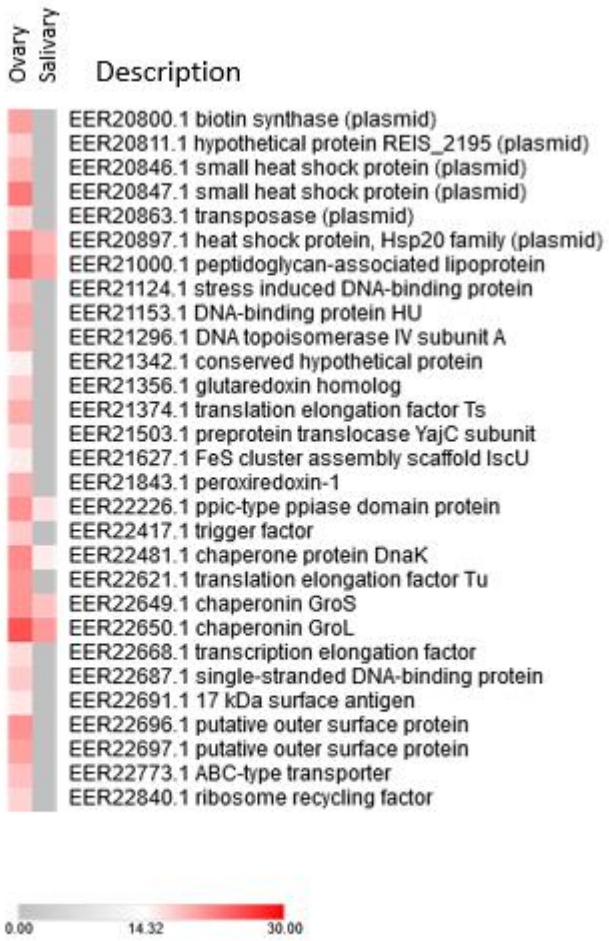


Figure 5. 4 A heatmap representing relative expression levels for 29 *R. buchneri* proteins (detected with ≥2 unique peptides) between SG and OV.

5.4 Discussion

According to the findings of the present study, TaqMan real-time PCR detected *R. buchneri* in the salivary gland of the blacklegged tick (*I. scapularis*); moreover, geLC-MS analysis illustrated bacterial proteins in the salivary glands as well. To date, this is the first time that *R. buchneri* has been detected in *I. scapularis* salivary glands. Notably, this contradicts the current view that *R. buchneri* resides exclusively in the ovaries (Munderloh et al. 2005) . An important implication of this

finding is that symbiotic bacteria could be a potential biomarker for tick bite, as the existence of *R. buchneri* in the salivary gland suggests potential transmission to mammals via saliva during tick feeding. However, the density of *R. buchneri* was variable between both tissues; bacterial density was significantly higher in ovaries (the primary location), and it might be assumed to increase with blood engorgement as for *Ca*. M. mitochondrii, since the blood provides nutrients for bacteria to multiply effectively (Sassera et al. 2008).

Although *I. scapularis* is not known to be a vector of pathogenic rickettsiae (unlike *I. ricinus*) (Mariconti et al. 2012), colonisation of salivary glands by bacterial symbionts may be a common phenomenon, and suggests similarities in the biology of the symbionts of *I. scapularis* and *I. ricinus.* Moreover, *Ca*. M. mitichondrii proteins are known to be immunogenic in animals parasitized by ticks (Bazzocchi et al. 2013; Mariconti et al. 2012), providing a potential diagnostic biomarker for tick bite; although complications for the diagnosis of pathogenic Rickettsiales infections in humans are also possible implications (Vaughn et al. 2014). Future studies using *in situ* hybridization and electron microscopy would be particularly useful to confirm bacterial existence in both tissues, in the unlikely event that our findings were caused by cross-contamination of tissues during dissection.

Interestingly, proteomic analysis highlighted the expression in the ovary of a biotin synthase. *Rickettsia buchneri* is unusual among rickettsiae in containing four plasmids (pREIS1 to pREIS4), and pREIS2 has acquired a complete biotin synthesis operon by lateral gene transfer (Gillespie et al. 2012). This suggests a possible role of the symbiont in providing biotin for the tick vector, perhaps in a similar manner as for other tick endosymbionts that have been revealed recently. These include the

Coxiella-like endosymbiont in *Rhipicephalus microplus* and the *Francisella*-type endosymbiont in soft ticks.

The maternally-transmitted *Francisella* symbiont, "F-Om", is an obligate nutritional mutualist bacterium in *Ornithodoros moubata* (a soft tick) that synthesizes and provides their host tick with B vitamins. This allows the tick to subsist on a single dietary source, blood, which is deficient in B vitamins (Duron et al. 2018). Physical abnormalities in *O. moubata* were detected as a result of experimental antibiotic elimination of F-Om, suggesting that F-Omis is a B-vitamin providing symbiont in line with its genomic features, in which the relevant synthesis pathways are intact (Duron et al. 2018). Similarly, elimination of a *Coxiella* endosymbiont (vertically transmitted mutualist symbiont in *Rhipicephalus microplus*) by tetracycline treatment resulted in negative impacts on tick fitness. Specifically, elimination of the *Coxiella* symbiont by anitibiotic treatment impaired the development and maturation of the tick metanymph stage to the adult stage (Guizzo et al. 2017). The existence of biosynthetic pathways for several vitamins and key metabolic cofactors in the *Coxiella* symbiont genome suggests that it may provide a nutritional complement to the tick host (Gottlieb et al. 2015). Thus, the vital role that endosymbiotic bacterial play in their tick host points to a new objective to control tick-borne disease by targeted disruption of the tick-symbiont interface (Guizzo et al. 2017).

6. Final discussion and conclusion

Based on the findings of the second chapter, our results represent the first survey of *Ca.* Midichloria mitochondrii prevalence and density in the British Isles. Although we were unable to screen nymphs or male adults from Scotland for *Ca*. Midichloria in this study, there was a very high prevalence of the symbiont in nymphs and male adults from Wales (100% and 94% infected, respectively) in contrast with a prevalence of only ~75% in English nymphs. This supports the hypothesis that evolutionary forces experienced by *I. ricinus* in the UK may differ significantly to those in its heartland in mainland Europe. Thus, the prevalence in the English nymphs tallies perfectly with what would be anticipated if 100% of females and ~50% of male nymphs were infected (assuming a 1:1 sex ratio), as observed in continental Europe (Lo et al., 2006); whereas the Welsh data do not conform to these expected frequencies. In addition, the current study corroborated previous findings in which the highest bacterial densities were observed in ffs females, suggesting that bacterial growth rate is linked with the blood meal and engorgement level (Sassera et al 2006). This study was intended to address the hypothesis of potential interactions between some rickettsial pathogens *(A. phagocytophilum*, *Rickettsia* spp. and *Ca.* Neoehrlichia mikurensis), and symbiotic bacteria *Ca*. M. mitochondrii within *I. ricinus* by quantifying their levels in different parts of the UK. In general, the prevalence of *A. phagocytophilum* within *I. ricinus* ticks was low across different locations of Great Britain; these findings correlated with the *A. phagocytophilum* prevalence reported in continental Europe (Henningsson et al. 2015; Silaghi et al. 2008). Transovarial transmission of *A. phagocytophilum* probably does not occur in *I. ricinus*, so the

reservoir animal is required to complete the life cycle (Rizzoli et al. 2014). Interestingly, our results suggest diversity between different strains of *A. phagocytophilum* based on geographical location within the UK according to mapping of 16S rRNA gene sequences (a highly conserved gene) that showed one nucleotide difference between infected English nymphs and ticks collected from Welsh fallow deer. This finding might be related to a previous UK investigation that observed a higher infection rate of *A. phagocytophilum* in adult Welsh *I. ricinus* compared with southern English ticks (Ogden et al. 1998). The same study highlighted divergence between *A. phagocytophilum* sequences obtained from *I. ricinus* feeding on Welsh sheep and those collected from Scottish sheep. This suggests that Scottish sheep might act as competent reservoirs for several genotypes of *A. phagocytophilum* (Ogden et al. 1998).

Deer play an important role as the predominant natural mammalian reservoir for *A. phagocytophilum* in the UK and continental Europe involving roe deer (Alberdi et al. 2000), red deer (*Cervus elaphus*) (Stuen et al. 2010), and fallow deer (*Dama dama*) (Veronesi et al. 2011), which consequently affects *A. phagocytophilum* epidemiology and pathogenesis across its geographical distribution. On the other hand, other large animals like cattle, horses and sheep (Hulínská et al. 2004); as well as small mammals such as *Apodemus flavicollis* rodents, have been reported as reservoir animals. Importantly, the strains circulating between large and small mammals may be distinct from one another, occurring in non-overlapping ecological cycles (Hildebrandt et al. 2002; Bown et al. 2009).

Based on real-time PCR results, and in agreement with previous studies (Jahfari et al. 2012), no *Ca.* N. mikurensis could be detected in the Great Britain yet, even though

it has been detected in some of the European countries such as the Netherlands, Belgium and Sweden (Jahfari et al. 2012; Andersson et al. 2013; Derdáková et al. 2014). However, global warming may suggest a possible expansion of this pathogen toward the British Isles, which could be an area for future interest.

The co-existence of pathogenic Rickettsiales bacteria has been reported in Europe before; for instance, tick coinfections with *A. phagocytophilum* and *Ca.* N. mikurensis (Glatz et al. 2014), and *A. phagocytophilum* and pathogenic *Rickettsia* species (Hildebrandt et al. 2010). Therefore, a similar situation might be expected in the *I. ricinus* vector within the UK (Tijsse-klasen et al. 2011). However, although real-time PCR revealed a number of *Rickettsia*-positive samples, no *Rickettsia* infections could be confirmed by sequencing. Previously, *R. raoultii* was reported only in *I. ricinus* ticks across Welsh and English areas, while *R. helvetica* were widespread across the British Isles, but generally with a low infection rate (Tijsse-klasen et al. 2011).

Although the current research was intended to measure potential interactions between pathogenic rickettsial bacteria and symbiotic bacteria (*Ca*. M.mitochondrii) in wild ticks by quantifying their levels in different parts of the UK, due to low infection rates, the statistical power was not sufficient to identify if the symbiont and pathogen densities are positively or negatively correlated. However, future work should be done using a larger sample size, or different life stages, and could target additional pathogens like *B. burgdorferi* (the causative agent of Lyme disease) to help answer this question. The existence of symbiont-pathogen interactions would suggest a new area for possible microbiome manipulation of the tick vector, which could lead to amelioration of tick-borne diseases as previously suggested (Gall et al. 2016; Macaluso et al. 2002). Indeed, new strategies for

biological control of tick-borne diseases may be on the horizon, similar to the recent approach of using endosymbiotic bacteria such as *Wolbachia pipientis* to control mosquito-borne diseases (LePage & Bordenstein 2013). Future *in vitro* studies could also target symbiont-pathogen interactions. For instance, *R. buchneri* can already be cultured in tick cells, so experiments could be done with *R. buchneri* and pathogenic rickettsiae or *A. phagocytophilum* in the same cultures, to determine if there is an inhibitory or facilitating effect. However, in the case of mixed rickettsial infections, the specificity of molecular assays would be critical to ensure accurate differentiation between the two *Rickettsia* spp. In addition, ongoing attempts to culture *Ca*. Midichloria in the Tick Cell Biobank (unfortunately unsuccessful to date) would allow similar experiments to investigate endosymbiont *Ca*. Midichloria-pathogen interactions *in vitro*.

Along a similar line of potential bacterial interactions, the tropical bont tick (*A. variegatum*) from Cameroon, the main vector of heartwater disease in ruminants (*E. ruminantium*), also transmits *R. africae* to humans (Robinson et al. 2009; Kelly et al. 2010; Dumler et al. 2001), which has been investigated in this study. As *R. africae* is very common in *A. variegatum* and is vertically transmitted, it could be mutualistic. The hypothesis is that the density of *E. ruminantium* may be inversely correlated with that of *R. africae* due to displacement or innate immune stimulation.

Real-time PCR revealed a high prevalence of *R. africae* in *A. variegatum* ticks across the Adamawa Region in Cameroon. This highlights two suggestions: vertical transmission of *R. africae* may be very efficient, as demonstrated with the closely related *A. hebraeum* (Robinson et al. 2009); also, as the ovaries are considered the primary location for *R. africae*, this explains the higher density in females.

Alternatively, horizontal transmission via an infected blood meal would be another possible source of infection (Robinson et al. 2009). The current findings agreed with previous serological studies that highlighted a high *R. africae* prevalence associated with both humans and ticks feeding on cattle in the southern regions of Cameroon (Ndip et al. 2004). Moreover, our results concurred with a very recent paper that reported a high prevalence of *R. africae* (57%) in *A. variegatum* ticks in the Adamawa Region using conventional PCR only (Vanegas et al. 2018).

It should be noted that the densities of *R. africae* in *A. variegatum* were often quite low, although the prevalence was very high. An important consideration for the epidemiology of ATBF is whether all PCR-positive ticks are actually infected with rickettsial bacteria or are false-positives due to insertion of rickettsial genes into the nuclear genome of the tick. Such lateral gene tranfers are common in other symbiont systems, particularly *Wolbachia* (Kondo et al. 2002; Hotopp et al. 2007; Koutsovoulos et al. 2014). This question could be answered by tick genome sequencing studies and imaging analysis of *Rickettsia* bacteria in the ticks, thus clarifying the public health risk of *A. variegatum* to local people and tourists.

Despite the low infection rate of *E. ruminantium* in our *A. vareigatum* specimens, our results corroborated previous findings suggesting local transmission of the heartwater agent in the north of Cameroon (Awa 1997) (the Adamawa Region, where our survey was based, has similar biogeography to the North and Far North Regions sampled in the latter study). However, our study to investigate the potential interactions of *R. africae* and *E. ruminantium* was designed assuming a much higher prevalence than we actually observed in the Adamawa Region. This was because previous surveys of *E. ruminantium* infection in *A. variegatum* on cattle in Cameroon

reported prevalence rates of 24.7 % (in the Southwest Region) and 40.9 % in the (Northwest Region) (Esemu et al. 2013). Unfortunately, these two Anglophone regions of Cameroon have experienced civil unrest in recent years, making surveys in those areas very difficult to conduct. It is possible that environmental or seasonal factors (northern Cameroon has a harsh dry season from October – April) are responsible for the lower prevalence we observed in the Adamawa Region. In agreement with previous studies, detection of a higher prevalence of *E. ruminantium* in male ticks compared with females reflects the feeding lifestyle that allows males to be attached to their hosts for a longer time, maximising the chance for infection (Jordaan & Baker 1981). Again, because of low infection rate, identification a positive or negative relationship between *R. africae* and *E. ruminantium* was not possible due to insufficient statistical power. In addition, we should mention that our findings were based on a limited number of *A. variegatum,* and the majority were ffs females, suggesting a potential PCR inhibition effect of blood (although template DNA was diluted heavily to overcome this). We propose that further research should be undertaken using a larger sample size and expanding to a different geographic area might be required to address the question of possible positive or negative interactions between the two pathogens.

Ca. M. mitochondrii is fixed in *I. ricinus* female ticks, suggesting an obligate mutualism. However, genetic variation in *Ca*. M. mitochondrii or possible co-cladogenesis with its host across its geographic range had not been investigated before our study commenced. In chapter 3, we presented the first combined population genetic analysis of *I. ricinus* and *Ca*. M. mitochondrii using multi-locus sequence typing (MLST) of tick mitochondrial markers and bacterial housekeeping

genes in ticks from Great Britain and continental Europe. These data showed that although *Ca*. M. mitochondrii exhibits a marked low level of genetic diversity, tick and symbiont populations in Great Britain, and particularly in Scotland, have a distinct signature that provides evidence of local coevolution in isolation from continental Europe. Horizontal transmission of the symbiont, possibly by co-feeding, would explain the general low level of co-cladogensis between *Ca*. Midichloria and tick hosts across multiple genera, which is corroborated by the presence of the bacterial symbiont in tick salivary glands and transmission to the host via the blood meal (Cafiso et al. 2016; Bazzocchi et al. 2013; Mariconti et al. 2012).

The results of our study were consistent with the previous clear differentiation in tick genotypes observed between western Norway and the British Isles (Roed et al. 2016), and between the British Isles and Latvia [(Dinnis et al. 2014); data incorporated into our study]. Thus, in agreement with previous studies, our results observed a marine barrier around the British Isles as a significant impediment to gene flow from the European landmass. The limited gene flow that does occur is probably associated with spring migrations of birds from the British Isles seeding continental Europe with sfs immature stages, whereas the autumn migrations in the opposite direction are less likely to lead to successful tick establishment due to winter attrition. Importantly, the success rate of moulting to the next stage in the lifecycle has been estimated to be only 10% for *I. ricinus* (Randolph et al. 2002).

Moreover, our data indicate that the ticks of Scotland form a separate clade that contains STs from other parts of Britain, but very little contribution from continental Europe. The only clear evidence for structure in our *M. mitochondrii* dataset was observed in STs from Scotland, and to a lesser extent Wales, relative to the locations

on the continent. It is possible that differences in the most abundant larger hosts (red deer) between Scotland and other parts of Europe may have acted to select a "race" of *I. ricinus* that is better adapted for feeding on red deer (Kempf et al. 2011). Physical factors (longer winter) at higher latitudes increasing attrition rates when ticks are dormant and reduced questing periods may have also played a role in selecting for certain tick-symbiont genotype combinations. Thus, the evolutionary history of the tick-symbiont relationship in these parts of the British Isles seems to have been subject to either genetic drift caused by a population bottleneck, or selection on ticks and their symbionts for traits that are important for reproductive success in certain locales. This has potential implications for the natural history of *I. ricinus* in the United Kingdom and its role in the spread of disease.

Future work should consider that focusing only on *Ca*. M. mitochondrii housekeeping genes under strong purifying selection may have significantly underestimated the genomic diversity of the symbiont. This could be remedied by whole genome sequencing of the symbiont where feasible. Moreover, the *Ca.* Midichloria MLST study needs to be expanded into other tick species to address the question of tick: symbiont co-speciation on a longer evolutionary timescale. While our MLST study was being completed, a study of genetic diversity in *Ca*. Midichloria using an entirely different MLST scheme was performed across two tick families, three genera and 10 species (Buysse & Duron 2018). The authors concluded from their work that there are three lineages of *Ca*. Midichloria, two of which are generalists that spread by horizontal transmission, and a third lineage in *Ixodes* spp. that appears to be more host specific. However, the total number of specimens examined in the whole study was only 17 (usually one per species); this is approximately one-quarter of our sample

size for our single tick species of interest. Sequencing the genome of *Ca*. Midichloria from Scottish ticks will help determine if this is a significantly different symbiont strain. Furthermore, analysis of mitochondria and *Ca*. Midichloria genomes together would address the question of whether *Ca.* Midichloria is often horizontally transmitted. Attaining such an objective could radically overhaul our understanding of the biology of this resilient vector species.

Since our main aim was to identify potential interactions between multiple Rickettsiales and the tick host, *A. phagocytophilum* was chosen as an exemplar for investigation to identify proteins involved in host-bacteria interactions *in vitro*. Also, there is a growing interest to identify the common and specific strategies that *A. phagocytophilum* uses to establish infection in ticks and vertebrate hosts, which has important implications for the development of vaccines for the control of pathogen infection and transmission.

Based on the findings of the fourth chapter, use of *A. phagocytophilum*-infected *I. scapularis* and *I. ricinus* embryonic cells confirms the reliability of these cells for "omics" analysis (in this case, involving genomics and proteomics), and this could play a crucial part in furthering our understanding of tick-pathogen interactions. However, our results highlighted a higher susceptibility of ISE6 to *A. phagocytophilum* infection compared with IRE/CTVM19, suggesting differences in the cell phenotypes or populations represented in different cell lines, which might affect their permissiveness to *Anaplasma* infection (Alberdi et al. 2016a). However, to confirm this suggestion, more experiments comparing infected ISE6 and IRE/CTVM19 cells at different infection time points and with various *A. phagocytophilum* strains are required to determine the possible variation between

different cell lines. Moreover, *in vivo* studies in ticks could compare the susceptibility of *I. ricinus* and *I. scapularis* to distinct *A. phagocytophilum* strains affecting different mammalian hosts, although such experiments would be very complex and expensive to perform.

The current study supports the previous suggestion of evolutionary adaptation that *A. phagocytophilum* uses to infect different host species. There appear to be common molecular strategies that *A. phagocytophilum* uses to establish the infection in both the tick vector and vertebrate host (de la Fuente et al. 2017), involving cytoskeleton rearrangement, inhibition of cell apoptosis, manipulation of the immune response, and the use of *A. phagocytophilum* proteins to modulate the intracellular environment. The current study revealed the manipulation of IRE/CTVM19 cytoskeleton, which reflects the initial stages of infection; while the majority of ISE6 responses were associated with immune responses, which might be associated with the establishment phase of infection. However, both ISE6 and IRE/CTVM19 seemed to be subjected to the cholesterol hijacking mechanism induced by *A. phagocytophilum* infection. On the other hand, the biological responses related to downregulation of metabolic and catabolic processes might be connected to the bacterial life cycle, leading to reduced transcription and replication by the host cell as a means to try to control the infection. In future studies, silencing of genes *in vitro* in tick cell lines could be conducted, which may indicate the functions of the tick proteins at the cellular level, advancing our understanding of these interactions.

Exploring *A. phagocytophilum* genomic diversity in the current study provided insights into the distinctiveness of the British *A. phagocytophilum* sheep strain, OS. Improving our understanding of the molecular mechanisms underlying regulation of

Msp2/p44 expression represents an important first step towards developing new therapies and vaccines against anaplasmosis. Our study confirmed the expression of several Msp2/p44 proteins in ISE6 infected with the OS strain, underlining the key role of the antigenic variation mechanism. In the definitive host, continuous generation of new epitopes enables evasion of host immune responses and allows a persistent phase to develop in the host animal reservoir, consequently assuring successful bacterial transmission to the tick vector. Further work on Msp/p44 diversity and expression in *A. phagocytophilum* is urgently required, as extrapolation from the closely related *A. marginale* used in most Msp2 studies may not be appropriate (Bowie et al. 2002; de la Funte et al. 2005). However, future research on *A. phagocytophilum in vitro* could borrow from some of the technical approaches used successfully with *A. marignale*, which have previously highlighted the role of Msp2 in host adaptation and evolution of this pathogen in different host cell environments (Chávez et al. 2012). Although our use of advanced proteomics analysis (PRM) targeting specific peptides of Msp2/p44 family members revealed few changes in abundance between the two time-points, optimising the time course would help determine of the existence of certain variants is correlated with specific host cell types of the tick vector, which could be new area for future studies.

In order to determine if the presence of certain variants is correlated with specific mammalian hosts, more diverse sampling of *A. phagocytophilum* genomes is required. Clustering of Msp2/p44 proteins from sequenced *A. phagocytophilum* and *A. marginale* genomes in this study revealed that a dominantly expressed *A. phagocytophilum* OS strain p44 protein is found almost exclusively in strains infecting ruminants. This suggests possible sharing of several characteristics such as

epidemiology, pathogenesis and animal reservoir that may differentiate animal strains from human strains. However, it is less clear how the Msp2/p44 repertoire of an individual strain is utilised between the tick vector and the mammalian host. In our study, there was almost no overlap between the Msp2/p44 variants detected in tick cells infected with *A. phagocytophilum* OS compared with those reportedly expressed in sheep infected with the same strain (Thomas et al. 2013). This is consistent with previous work in *A. marginale*, in which it was shown that there is a predominance of a specific Msp2 structure during infection of a particular host type (tick versus blood sample) (Chávez et al. 2012).

Finally, identifying *R. buchneri* in *I. scapularis* salivary glands in the present study highlights possible future studies that should consider serological cross-reactions between pathogenic and symbiotic rickettsiae (Lindblom et al. 2016). This may be complicating diagnosis of tick-borne diseases in humans, as people bitten by ticks might have antibodies to *R. buchneri* even if a different tick-borne agent caused their disease. Detecting expression of biotin synthase in tick ovaries in the present proteomics analysis of *R. buchneri* suggests a role in providing the tick vector with the nutrients that are essential to adapt to a blood diet; a key function for symbionts that has been highlighted in two very recent studies [characterisation of the *Coxiella*-like symbiont in *R. microplus* (Guizzo et al. 2017) and the *Francisella*-like symbiont of soft ticks (Duron et al. 2018)]. These findings demonstrate that bacterial symbionts act as a source of essential B-vitamins for their tick hosts and have a crucial role in tick fitness and adaptation. Thus, disruption of the tick-symbiont relationship might constitute a new target for control of tick-borne diseases.

Conclusions: Many tick species are potentially infected with multiple species of Rickettsiales. To date, the main research focus has been on those species that cause disease in vertebrates. However, non-pathogenic Rickettsiales are likely to interact with their pathogenic relatives and possibly other tick-borne pathogens in the tick host. Here, we have contributed to our understanding of the roles of three vertically-transmitted symbionts in three of the most important tick vectors worldwide: *Ca*. M. mitochondrii in *I. ricinus, R. africae* in *A. variegatum*, and *R. buchneri* in *I. scapularis*. The use of tick cells to examine Rickettsiales:tick interactions in vitro, as shown here with *A. phagocytophilum*, could be expanded in future to include pathogen-symbiont coinfections.

References

Adelson, M.E. et al., 2004. Prevalence of Borrelia burgdorferi, Bartonella spp., Babesia microti,Anaplasma phagocytophila in Ixodes scapularis Ticks Collected in Northern New Jersey. *Society*, 42(6), pp.2799–2801.

Ahantarig, A. et al., 2013. Hard ticks and their bacterial endosymbionts (or would be pathogens). *Jurnal of*, 50(5), pp.419–428.

Akuffo, R. et al., 2016. Crimean-Congo hemorrhagic fever virus in livestock ticks and animal handler seroprevalence at an abattoir in Ghana. *BMC Infectious Diseases*, 16(1), p.324. Available at: http://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-016-1660-6.

Al-Khedery, B. & Barbet, A., 2014. Comparative Genomics Identifies a Potential Marker of Human-Virulent Anaplasma phagocytophilum. *Pathogens*, 3(1), pp.25–35. Available at: http://www.mdpi.com/2076-0817/3/1/25/.

Alberdi, M.P. et al., 2012. Detection and identification of putative bacterial endosymbionts and endogenous viruses in tick cell lines. *Ticks and Tick-borne Diseases*, 3(3), pp.137–146. Available at: http://dx.doi.org/10.1016/j.ttbdis.2012.05.002.

Alberdi, M.P. et al., 1998. Natural prevalence of infection with Ehrlichia (Cytoecetes) phagocytophila of Ixodes ricinus ticks in Scotland. *Veterinary Parasitology*, 78, pp.203–213.

Alberdi, M.P., Walker, A.R. & Urquhart, K.A., 2000. Field evidence that roe deer (Capreolus capveolus) are a natural host for Ehrlichia phagocytophila. *Epidemiology and Infection*, 124(2), pp.315–323.

Alberdi, P., et al., 2016a. Anaplasma phagocytophilum Manipulates Host Cell Apoptosis by Different Mechanisms to Establish Infection. *Veterinary Sciences*, 3(3), p.15. Available at: http://www.mdpi.com/2306-7381/3/3/15.

Alberdi, P. et al., 2015. Infection of Ixodes spp. tick cells with different Anaplasma phagocytophilum isolates induces the inhibition of apoptotic cell death. *Ticks and Tick-borne Diseases*, 6(6), pp.758–767. Available at: http://dx.doi.org/10.1016/j.ttbdis.2015.07.001.

Alberdi, P., et al., 2016b. Tissue-Specific Signatures in the Transcriptional Response to Anaplasma phagocytophilum Infection of Ixodes scapularis and Ixodes ricinus Tick Cell Lines. *Frontiers in cellular and infection microbiology*, 6(February), p.20. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26904518%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4748044.

Allsopp, B.A., 2015. Heartwater – Ehrlichia ruminantium infection Aetiology Ante-mortem diagnosis Pathology and post-mortem diagnosis. *Revue scientifique et technique (International Office of Epizootics)*, 34(2), pp.557–568.

Ammazzalorso, A.D. et al., 2015. To beat or not to beat a tick: comparison of DNA extraction methods for ticks (Ixodes scapularis). *PeerJ*, 3, p.e1147. Available at: https://peerj.com/articles/1147.

Anantharaman, V. & Aravind, L., 2002. The GOLD domain, a novel protein module involved in Golgi function and secretion. *Genome Biology*, 3(5), p.research0023.1–0023.7. Available at: http://genomebiology.com/2002/3/5/research/0023.1%5Cnhttp://genomebiology.com/2002/3/5/research/0023.

Anderson, B.E. et al., 1991. Ehrlichia chaffeensis, a new species associated with human ehrlichiosis. *Journal of Clinical Microbiology*, 29(12), pp.2838–2842.

Andersson, M. et al., 2013. Co-Infection with ' Candidatus Neoehrlichia mikurensis .' , 13(7), pp.438–442.

Armstrong, S.D. et al., 2014. Comparative Analysis of the Secretome from a Model Filarial Nematode (*Litomosoides sigmodontis*) Reveals Maximal Diversity in Gravid Female Parasites. *Molecular & Cellular Proteomics*, 13(10), pp.2527–2544. Available at: http://www.mcponline.org/lookup/doi/10.1074/mcp.M114.038539.

Ashida, H. et al., 2011. Cell death and infection: A double-edged sword for host and pathogen survival. *Journal of Cell Biology*, 195(6), pp.931–942.

Awa, D.N., 1997. Serological survey of heartwater relative to the distribution of the vector Amblyomma variegatum and other tick species in north Cameroon. *Veterinary Parasitology*, 68(1–2), pp.165–173.

Ayllón, N. et al., 2013. Anaplasma phagocytophilum Inhibits Apoptosis and Promotes Cytoskeleton Rearrangement for Infection of Tick Cells. *Infection and Immunity*, 81(7), pp.2415–2425.

Ayllón, N., et al., 2015a. Nuclease Tudor-SN is involved in tick dsRNAMediated RNA interference and feeding but not in defense against flaviviral or anaplasma phagocytophilum rickettsial infection. *PLoS ONE*, 10(7), pp.1–18.

Ayllón, N., et al., 2015b. Systems Biology of Tissue-Specific Response to Anaplasma phagocytophilum Reveals Differentiated Apoptosis in the Tick Vector Ixodes scapularis. *PLOS Genetics*, 11(3), p.e1005120. Available at: http://dx.plos.org/10.1371/journal.pgen.1005120.

Bandyopadhyay, S. et al., 2006. Mechanism of apoptosis induced by the inhibition of fatty acid synthase in breast cancer cells. *Cancer Research*, 66(11), pp.5934–5940.

Barbet, A. et al., 2013. An Emerging Tick-Borne Disease of Humans Is Caused by a Subset of Strains with Conserved Genome Structure. *Pathogens*, 2(3), pp.544–555. Available at: http://www.mdpi.com/2076-0817/2/3/544/.

Barbet, A.F. et al., 2000. Antigenic variation of Anaplasma marginale by expression of MSP2 mosaics. *Infection and Immunity*, 68(11), pp.6133–6138.

Barbet, A.F. et al., 2003. Expression of multiple outer membrane protein sequence variants from a single genomic locus of Anaplasma phagocytophilum. *Infection and immunity*, 71(4), pp.1706–18. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=152091&tool=pmcentrez&rendertype=abstract.

Barker, S. & Murrell, A., 2004. Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology*, 129, pp.15–36.

Barré, N., Garris, G. & Camus, E., 1995. Propagation of the tick Amblyomma variegatum in the Caribbean. *Revue scientifique et technique (International Office of Epizootics)*, 14(3), pp.841–855.

Barre, N. & Garris, G.I., 1989. Biological and ecology of Amblyomma variegatum (Acare:Ixodidae) in the Carribean: implications fir regional eradication program. *Agricultural Entomology*, 3(June 1987).

Bazzocchi, C. et al., 2013. Molecular and serological evidence for the circulation of the tick symbiont Midichloria (Rickettsiales : Midichloriaceae) in different mammalian species. , pp.1–7.

Beati, L. et al., 1993. Confirmation that Rickettsia helvetica sp. nov. Is a Distinct Species of the Spotted Fever Group of Rickettsiae. *International Journal of Systematic Bacteriology*, 43(3), pp.521–526.

Bell-sakyi, L. et al., 2018. Ticks and Tick-borne Diseases The Tick Cell Biobank : A global resource for in vitro research on ticks , other arthropods and the pathogens they transmit. *Ticks and Tick-borne Diseases*, (March), pp.0–1. Available at: https://doi.org/10.1016/j.ttbdis.2018.05.015.

Bell-Sakyi, L. et al., 2007. Tick cell lines: tools for tick and tick-borne disease research. *Trends in*

Parasitology, 23(9), pp.450–457.

Bell-Sakyi, L. et al., 2012. Tick cell lines for study of Crimean-Congo hemorrhagic fever virus and other arboviruses. , 12(9). Available at: http://eprints.gla.ac.uk/72113/.

Belongia, E.A. et al., 2001. Population-Based Incidence of Human Granulocytic Ehrlichiosis in Northwestern Wisconsin, 1997–1999. *The Journal of Infectious Diseases*, 184(11), pp.1470–1474. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11709792.

Beninati, T. et al., 2004. A novel alpha-proteobacterium resides in the mitochondria of ovarian cells of the tick Ixodes ricinus. *Applied and Environmental Microbiology*, 70(5), pp.2596–2602.

Beninati, T. et al., 2002. First detection of spotted fever group rickettsiae in Ixodes ricinus from Italy. *Emerging Infectious Diseases*, 8(9), pp.983–986.

Benjamini, Y. & Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society*, 57(1), pp.289–300. Available at: http://www.jstor.org/stable/2346101%5Cnhttp://about.jstor.org/terms.

Bergey, D.H., Buchanan, R. E., & & Gibbons, N.E., 1974. Bergey's manual of determinative bacteriology. In *Baltimore, Williams & Wilkins Co*.

Beugnet, F. & Marié, J.L., 2009. Emerging arthropod-borne diseases of companion animals in Europe. *Veterinary Parasitology*, 163(4), pp.298–305.

Bezuidenhout, J.D., 1987. Natural Transmission of Heartwater. *Onderstepoort J. vet Res*, 54, pp.349–351.

Black, W.C. & Piesman, J., 1994. Phylogeny of hard- and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences. *Proceedings of the National Academy of Sciences*, 91(21), pp.10034–10038. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.91.21.10034.

Bonnet, S.I. et al., 2017. The Tick Microbiome: Why Non-pathogenic Microorganisms Matter in Tick Biology and Pathogen Transmission. *Frontiers in cellular and infection microbiology*, 7(June), p.236. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28642842%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5462901.

Bown, K.J. et al., 2009. Delineating Anaplasma phagocytophilum ecotypes in coexisting, discrete enzootic cycles. *Emerging Infectious Diseases*, 15(12), pp.1948–1954.

Bown, K.J. et al., 2003. Seasonal dynamics of Anaplasma phagocytophila in a rodent-tick (Ixodes trianguliceps) system, United Kingdom. *Emerging Infectious Diseases*, 9(1), pp.63–70.

Braendle, C. et al., 2003. Developmental Origin and Evolution of Bacteriocytes in the Aphid–Buchnera Symbiosis Benjamin Normark, ed. *PLoS Biology*, 1(1), p.e21. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14551917.

Bram, R.A. et al., 2002. Threat of Foreign Arthropod-Borne Pathogens to Livestock in the United States. *Medical and Veterinary Entomology*, 39(3), pp.405–416.

Brayton, K. a et al., 2003. Variation of the Immunodominant Outer Membrane Anaplasma marginale. , 71(11), pp.6627–6632.

Brayton, K.A. et al., 2001. Efficient use of a small genome to generate antigenic diversity in tick-borne ehrlichial pathogens. *Proceedings of the National Academy of Sciences*, 98(7), pp.4130–4135. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.071056298.

Brayton, K.A., 2012. Persistence and Antigenic Variation. *Intracellular Pathogens II: Rickettsiales*, pp.366–390.

Brouqui, P. et al., 2003. A New Ehrlichia Detected in Ixodes ricinus Tick Collected from Asymptomatic Humans in Northern Italy. *Annals of the New York academy of science*, 140, pp.134–140.

Burgdorfer, W. et al., 1982. Lyme Disease--A Tick-Borne Spirochetosis ? , 216(June), pp.1317–1319.

Burkot, T.R. et al., 2001. Use of a sentinel host system to study the questing behavior of Ixodes spinipalpis and its role in the transmission of Borrelia bissettii, human granulocytic ehrlichiosis, and Babesia microti. *American Journal of Tropical Medicine and Hygiene*, 65(4), pp.293–299.

Busby, A.T. et al., 2012. Expression of heat shock proteins and subolesin affects stress responses, Anaplasma phagocytophilum infection and questing behaviour in the tick, Ixodes scapularis. *Medical and Veterinary Entomology*, 26(1), pp.92–102.

Buysse, M. & Duron, O., 2018. Multi-locus phylogenetics of the Midichloria endosymbionts reveals variable specificity of association with ticks.

Cabezas-Cruz, A. et al., 2016. Anaplasma phagocytophilum increases the levels of histone modifying enzymes to inhibit cell apoptosis and facilitate pathogen infection in the tick vector Ixodes scapularis. *Epigenetics*, 11(4), pp.303–319.

Cafiso, A. et al., 2016. Molecular screening for Midichloria in hard and soft ticks reveals variable prevalence levels and bacterial loads in different tick species. *Ticks and Tick-borne Diseases*, 7(6), pp.1186–1192. Available at: http://dx.doi.org/10.1016/j.ttbdis.2016.07.017.

Cariou, M., Duret, L. & Charlat, S., 2017. The global impact of Wolbachia on mitochondrial diversity and evolution. *Journal of Evolutionary Biology*, 30(12), pp.2204–2210.

Carlyon, J.A. et al., 2004. Anaplasma phagocytophilum utilizes multiple host evasion mechanisms to thwart NADPH oxidase-mediated killing during neutrophil infection. *Infection and Immunity*, 72(8), pp.4772–4783.

Carpi, G. et al., 2011. Metagenomic profile of the bacterial communities associated with Ixodes ricinus ticks. *PLoS ONE*, 6(10).

Carpi, G. et al., 2016. Mitogenomes reveal diversity of the European Lyme borreliosis vector Ixodes ricinus in Italy. *Molecular Phylogenetics and Evolution*, 101, pp.194–202. Available at: http://dx.doi.org/10.1016/j.ympev.2016.05.009.

Carver, T. et al., 2012. Artemis: An integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics*, 28(4), pp.464–469.

Casati, S. et al., 2008. Assessment of intraspecific mtDNA variability of European Ixodes ricinus sensu stricto (Acari: Ixodidae). *Infection, Genetics and Evolution*, 8(2), pp.152–158.

Caspersen, K. et al., 2002. Genetic Variability and Stability of Anaplasma. , 70(3), pp.1230–1234.

Cen-Aguilar, J.F. et al., 1998. Studies on the effect of infection by Babesia sp. on oviposition of Boophilus microplus engorged females naturally infected in the Mexican tropics. *Veterinary Parasitology*, 78(4), pp.253–257.

Chaudhari, N.M., Gupta, V.K. & Dutta, C., 2016. BPGA-an ultra-fast pan-genome analysis pipeline. *Scientific Reports*, 6(April), pp.1–10. Available at: http://dx.doi.org/10.1038/srep24373.

Chávez, A.S.O. et al., 2012. Expression patterns of Anaplasma marginale Msp2 variants change in response to growth in cattle, and tick cells versus mammalian cells. *PloS one*, 7(4), p.e36012. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3338850&tool=pmcentrez&rendertype=abstract.

Chen, S.M. et al., 1994. Identification of a granulocytotropic *Ehrlichia* species as the etiologic agent of human disease. *Journal of clinical microbiology*, 32(3), pp.589–95. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=263091&tool=pmcentrez&rendertype=abstract.

Chen, X., Li, S. & Aksoy, S., 1999. Concordant Evolution of a Symbiont with Its Host Insect Species: Molecular Phylogeny of Genus Glossina and Its Bacteriome-Associated Endosymbiont,

Wigglesworthia glossinidia. *Journal of Molecular Evolution*, 48(1), pp.49–58. Available at: http://link.springer.com/10.1007/PL00006444.

Choi, M. et al., 2014. MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics*, 30(17), pp.2524–2526. Available at: https://academic.oup.com/bioinformatics/bioinformatics/article/2748156/MSstats:

Clark, C.D. et al., 2012. Pattern and timing of retreat of the last British-Irish Ice Sheet. *Quaternary Science Reviews*, 44, pp.112–146. Available at: http://dx.doi.org/10.1016/j.quascirev.2010.07.019.

Clay, K. et al., 2008. Microbial communities and interactions in the lone star tick, Amblyomma americanum. *Molecular Ecology*, 17(19), pp.4371–4381.

Comstedt, P. et al., 2006. Migratory passerine birds as reserviors of Lyme borreliosis in Europe. *Emerging Infectious Diseases*, 12(7), pp.1087–1095. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3291064/.

Cowdry, E. V., 1925. A GROUP OF MICROORGANISMS TRANSMITTED HEREDITARILY IN TICKS AND APPARENTLY UNASSOCIATED WITH DISEASE .* In the course of a study of the relation of Rickettsia ruminantiuml to found in them must have been inherited through the eggs , since the Soon somewha. *J Exp Med*, 41, pp.817–830.

Dautel, H. et al., 2008. Winter activity of Ixodes ricinus in a Berlin forest. *International Journal of Medical Microbiology*, 298(SUPPL. 1), pp.50–54.

Derdáková, M. et al., 2014. Candidatus Neoehrlichia mikurensis and its co-circulation with Anaplasma phagocytophilum in Ixodes ricinus ticks across ecologically different habitats of Central Europe. *Parasites & vectors*, 7, p.160. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3984398&tool=pmcentrez&rendertype=abstract.

Dermine, J.F. et al., 2001. Flotillin-1-enriched Lipid Raft Domains Accumulate on Maturing Phagosomes. *Journal of Biological Chemistry*, 276(21), pp.18507–18512.

Dinnis, R.E. et al., 2014. Multilocus sequence typing using mitochondrial genes (mtMLST) reveals geographic population structure of Ixodes ricinus ticks. *Ticks and Tick-borne Diseases*, 5(2), pp.152–160. Available at: http://dx.doi.org/10.1016/j.ttbdis.2013.10.001.

Du, Z. et al., 2010. agriGO: A GO analysis toolkit for the agricultural community. *Nucleic Acids Research*, 38(SUPPL. 2), pp.64–70.

Dugat, T. et al., 2014. Comparative genomics of first available bovine Anaplasma phagocytophilum genome obtained with targeted sequence capture. *BMC Genomics*, 15(1), p.973. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-15-973.

Dumler, J.S. et al., 2001. Reorganization of gene in families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of Ehrlichia with Anaplasm, Cowdria with Ehriichia with neorickettsia, description of six new species combinations and designatio. *International Journal of systematic and evolutionary microbiology*, 51(2001), pp.2145–2165.

Dunning Hotopp, J.C. et al., 2006. Comparative genomics of emerging human ehrlichiosis agents. *PLoS Genetics*, 2(2), pp.208–223.

Duron, O. et al., 2017. Evolutionary changes in symbiont community structure in ticks. *Molecular Ecology*, 26(11), pp.2905–2921.

Duron, O. et al., 2015. The Recent Evolution of a Maternally-Inherited Endosymbiont of Ticks Led to the Emergence of the Q Fever Pathogen, Coxiella burnetii. *PLoS Pathogens*, 11(5), pp.1–23.

Duron, O. et al., 2018. Tick-Bacteria Mutualism Depends on B Vitamin Synthesis Pathways. *Current Biology*, 0(0), pp.1–7. Available at: https://www.cell.com/current-biology/fulltext/S0960-

9822(18)30466-4?rss=yes.

Eddy, S.R., 2011. Accelerated profile HMM searches. *PLoS Computational Biology*, 7(10).

Edwards, T. & Kenyon, W., 2013. Briefing: Wild Deer in Scotland. *Scottish Parliment Information Centre (SPICe)*, 13/74(November), pp.1–30. Available at: http://www.parliament.scot/parliamentarybusiness/69742.aspx.

Enright, M.C. et al., 2000. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. *Journal of clinical microbiology*, 38(3), pp.1008–15. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10698988.

Epis, S. et al., 2013. Localization of the bacterial symbiont Candidatus Midichloria mitochondrii within the hard tick Ixodes ricinus by whole-mount FISH staining. *Ticks and Tick-borne Diseases*, 4(1–2), pp.39–45. Available at: http://dx.doi.org/10.1016/j.ttbdis.2012.06.005.

Epis, S. et al., 2008. Midichloria mitochondrii is widespread in hard ticks (Ixodidae) and resides in the mitochondria of phylogenetically diverse species. *Parasitology*, 135(4), pp.485–494.

Eremeeva, M.E. et al., 2006. Isolation and Identification of Rickettsia massiliae from Rhipicephalus sanguineus Ticks Collected in Arizona. *Applied and Environmental Microbiology*, 72(8), pp.5569–5577. Available at: http://aem.asm.org/cgi/doi/10.1128/AEM.00122-06.

Esemu, S.N. et al., 2013. Prevalence of Ehrlichia ruminantium in adult Amblyomma variegatum collected from cattle in Cameroon. *Experimental and Applied Acarology*, 59(3), pp.377–387.

Esemu, S.N., Ndip, L.M.& & Ndip, R.N., 2011. Ehrlichia species, probable emerging human pathogens in sub-Saharan Africa. *Environmental Health*, 26, pp.269–279.

Eshoo, M.W. et al., 2015. Survey of ixodes pacificus ticks in California reveals a diversity of microorganisms and a novel and widespread anaplasmataceae species. *PLoS ONE*, 10(9), pp.1–14.

Estrada-Peña, A. et al., 2013. Association of environmental traits with the geographic ranges of ticks (Acari: Ixodidae) of medical and veterinary importance in the western Palearctic. A digital data set. *Experimental and Applied Acarology*, 59(3), pp.351–366.

Estrada-Peña, A., 2001. Forecasting habitat suitability for ticks and prevention of tick-borne diseases. *Veterinary Parasitology*, 98(1–3), pp.111–132.

Estrada-Peña, A. et al., 2006. The known distribution and ecological preferences of the tick subgenus Boophilus (Acari: Ixodidae) in Africa and Latin America. *Experimental and Applied Acarology*, 38(2–3), pp.219–235.

Estrada-Peña, A., 2015. Ticks as vectors: taxonomy, biology and ecology. *Revue scientifique et technique (International Office of Epizootics)*, 34(1), pp.53–65. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26470449.

Estrada-Peña, A. et al., 2007. Using invaded range data to model the climate suitability for Amblyomma variegatum (Acari : Ixodidae) in the New World. *Experimental & Applied Acarology*, 41, pp.203–214.

Estrada-Peña, A., Ayllón, N. & de la Fuente, J., 2012. Impact of climate trends on tick-borne pathogen transmission. *Frontiers in Physiology*, 3 MAR(March), pp.1–12.

Estrada-Peña, A., Horak, I.G. & Petney, T., 2008. Climate changes and suitability for the ticks Amblyomma hebraeum and Amblyomma variegatum (Ixodidae) in Zimbabwe (1974 – 1999). *Veterinary Parasitology*, 151, pp.256–267.

Estrada-Peña, A., Nava, S. & Petney, T., 2014. Description of all the stages of Ixodes inopinatus n. sp. (Acari: Ixodidae). *Ticks and Tick-borne Diseases*, 5(6), pp.734–743.

Feil, E.J. et al., 2004. eBURST : Inferring Patterns of Evolutionary Descent among Clusters of Related

Bacterial Genotypes from Multilocus Sequence Typing Data. , 186(5), pp.1518–1530.

Felsheim, R.F., Kurtti, T.J. & Munderloh, U.G., 2009. Genome sequence of the endosymbiont Rickettsia peacockii and comparison with virulent Rickettsia rickettsii: Identification of virulence factors. *PLoS ONE*, 4(12).

Filippova, N.A., 2017. History of the species range of ixodid ticks, vectors of pathogens with natural nidality (Acari, Ixodidae), as a prerequisite of their intraspecific biodiversity. *Entomological Review*, 97(2), pp.255–275. Available at: http://link.springer.com/10.1134/S0013873817020117.

Finn, R.D. et al., 2014. Pfam: The protein families database. *Nucleic Acids Research*, 42(D1), pp.222–230.

Foley, J. et al., 2008. Possible differential host tropism in Anaplasma phagocytophilum strains in the western United States. *Annals of the New York Academy of Sciences*, 1149, pp.94–97.

Foster, W.N.M. & Cameron, A.E., 1970. OBSERVATIONS TICK-BORNE OVINE STRAINS OF In the course of field investigations into the epidemiology of enzootic staphy- lococcal infection (tick pyaemia) in lambs , tick infested localities have been encountered in which pyaemia is not a problem , alth.

Fourie, L.J., Belozerov, V.N. & Needham, G.R., 2001. Ixodes rubicundus nymphs are short-day diapause-induced ticks with thermolabile sensitivity and desiccation resistance. *Medical and Veterinary Entomology*, 15(3), pp.335–341.

Francisco, A.P. & Monteiro, P.T., 2012. PHYLOViZ : phylogenetic inference and data visualization for sequence based typing methods.

Funk, D.J. et al., 2000. Intraspecific phylogenetic congruence among multiple symbiont genomes. *Proceedings of the Royal Society B: Biological Sciences*, 267(1461), pp.2517–2521. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2000.1314.

Galardini, M. et al., 2011. CONTIGuator: A bacterial genomes finishing tool for structural insights on draft genomes. *Source Code for Biology and Medicine*, 6(1), p.11. Available at: http://www.scfbm.org/content/6/1/11.

Galindo, R.C. & la Fuente, J. de, 2012. Transcriptomics Data Integration Reveals Jak-STAT as a Common Pathway Affected by Pathogenic Intracellular Bacteria in Natural Reservoir Hosts. *Journal of Proteomics & Bioinformatics*, 05(04), pp.108–115. Available at: http://www.omicsonline.org/0974-276X/JPB-05-108.digital/JPB-05-108.html.

Gall, C.A. et al., 2016. The bacterial microbiome of Dermacentor andersoni ticks influences pathogen susceptibility. *ISME Journal*, 10(8), pp.1846–1855. Available at: http://dx.doi.org/10.1038/ismej.2015.266.

Gassner, F. et al., 2011. Geographic and temporal variations in population dynamics of Ixodes ricinus and associated Borrelia infections in the Netherlands. *Vector-Borne and Zoonotic Diseases*, 11(5), pp.523–532. Available at: http://online.liebertpub.com/doi/abs/10.1089/vbz.2010.0026.

Ge, Y. et al., 2005. Anaplasma phagocytophilum inhibits human neutrophil apoptosis via upregulation of bfl-1, maintenance of mitochondrial membrane potential and prevention of caspase 3 activation. *Cellular Microbiology*, 7(1), pp.29–38.

Ghosh, M. et al., 2017. Sexual alteration in antioxidant response and esterase profile in Hyalomma anatolicum anatolicum (Acari: Ixodidae) ticks. *Journal of Parasitic Diseases*, 41(1), pp.106–111. Available at: http://link.springer.com/10.1007/s12639-016-0758-5.

Gillespie, J.J. et al., 2012. A rickettsia genome overrun by mobile genetic elements provides insight into the acquisition of genes characteristic of an obligate intracellular lifestyle. *Journal of Bacteriology*, 194(2), pp.376–394.

Glatz, M. et al., 2014. Detection of Candidatus Neoehrlichia mikurensis, Borrelia burgdorferi sensu lato genospecies and Anaplasma phagocytophilum in a tick population from Austria. *Ticks and tick-borne diseases*, 5(2), pp.139–44. Available at: http://dx.doi.org/10.1016/j.ttbdis.2013.10.006.

Gottlieb, Y., Lalzar, I. & Klasson, L., 2015. Distinctive genome reduction rates revealed by genomic analyses of two Coxiella-like endosymbionts in ticks. *Genome Biology and Evolution*, 7(6), pp.1779–1796.

Graham, R.I. et al., 2012. Wolbachia in a major African crop pest increases susceptibility to viral disease rather than protects. *Ecology Letters*, 15(9), pp.993–1000.

Graham, R.I., Mainwaring, M.C. & Du Feu, R., 2010. Detection of spotted fever group Rickettsia spp. from bird ticks in the U.K. *Medical and Veterinary Entomology*, 24(3), pp.340–343.

Granquist, E.G. et al., 2014. Evaluation of microbial communities and symbionts in Ixodes ricinus and ungulate hosts (Cervus elaphus and Ovis aries) from shared habitats on the west coast of Norway. *Ticks and Tick-borne Diseases*, 5(6), pp.780–784. Available at: http://dx.doi.org/10.1016/j.ttbdis.2014.05.005.

Granquist, E.G. et al., 2010. Variant-specific and diminishing immune responses towards the highly variable MSP2(P44) outer membrane protein of Anaplasma phagocytophilum during persistent infection in lambs. *Veterinary Immunology and Immunopathology*, 133(2–4), pp.117–124.

Gray, J.S., 2002. Biology of Ixodes species ticks in relation to tick-borne zoonoses. *European Journal of Epidemiology*, 16, pp.101–103.

Grindle, N. et al., 2003. Identification of Arsenophonus-type bacteria from the dog tick Dermacentor variabilis. *Journal of Invertebrate Pathology*, 83(3), pp.264–266.

Grøva, L. et al., 2011. Prevalence of Anaplasma phagocytophilum infection and effect on lamb growth. *Acta Veterinaria Scandinavica*, 53(30), pp.1–7.

Guglielmone, A.A. et al., 2010. The argasidae, Ixodidae and Nuttalliellidae (Acari: Ixodida) of the world: A list of valid species names. *Zootaxa*, (2528), pp.1–28.

Guizzo, M.G. et al., 2017. A Coxiella mutualist symbiont is essential to the development of Rhipicephalus microplus. *Scientific Reports*, 7(1), pp.1–10.

Gulia-Nuss, M. et al., 2016. Genomic insights into the Ixodes scapularis tick vector of Lyme disease. *Nature Communications*, 7(May 2015).

Hagedorn, P. et al., 2014. Human granulocytic anaplasmosis acquired in Scotland, 2013. *Emerging Infectious Diseases*, 20(6), pp.1079–1081.

Haig, D.A., 1952. Note on the use of the white mouse for the transport of strains of heartwater. *Veterinarni Medicina*, 23, pp.167–170.

Hajdušek, O. et al., 2013. Interaction of the tick immune system with transmitted pathogens. *Frontiers in cellular and infection microbiology*, 3(July), p.26. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3712896&tool=pmcentrez&rendertype=abstract.

Hall, T.A., 1999. BioEdit:a user-friendly biological sequence alingment editor and analysis program for windows 95/98/nt. *Nucleic Acids Symposium Series*, 41, pp.95–98.

Harrus, S. et al., 2011. Molecular detection of Ehrlichia canis, Anaplasma bovis, Anaplasma platys, Candidatus Midichloria mitochondrii and Babesia canis vogeli in ticks from Israel. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases*, 17(3), pp.459–63. Available at: http://dx.doi.org/10.1111/j.1469-0691.2010.03316.x.

Hartelt, K. et al., 2004. Pathogens and symbionts in ticks: prevalence of Anaplasma phagocytophilum

(Ehrlichia sp.), Wolbachia sp., Rickettsia sp., and Babesia sp. in Southern Germany. *International journal of medical microbiology : IJMM*, 293 Suppl, pp.86–92.

Hedges, L.M. et al., 2008. Wolbachia and virus protection in insects. *Science*, 322(5902), p.702.

Henningsson, A.J. et al., 2015. Detection of Anaplasma phagocytophilum in Ixodes ricinus ticks from Norway using a realtime PCR assay targeting the Anaplasma citrate synthase gene gltA. *BMC microbiology*, 15, p.153. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4521461&tool=pmcentrez&rendertype=abstract.

Herren, J.K. & Lemaitre, B., 2011. Spiroplasma and host immunity: Activation of humoral immune responses increases endosymbiont load and susceptibility to certain Gram-negative bacterial pathogens in Drosophila melanogaster. *Cellular Microbiology*, 13(9), pp.1385–1396.

Hildebrandt, A. et al., 2010. Detection of Rickettsia spp. and Anaplasma phagocytophilum in Ixodes ricinus ticks in a region of Middle Germany (Thuringia). *Ticks and Tick-borne Diseases*, 1(1), pp.52–56.

Hildebrandt, A. et al., 2002. Prevalence of granulocytic Ehrlichiae in Ixodes ricinus ticks in Middle Germany (Thuringia) detected by PCR and sequencing of a 16S ribosomal DNA fragment. *FEMS Microbiology Letters*, 211(2), pp.225–230.

Hillyard, P.D., 1996. *Ticks of North-West Europe: keys and notes for identification of the species. Field Studies Council, Shrewsbury*,

Hodzic, E. et al., 1998. Acquisition and transmission of the agent of human granulocytic ehrlichiosis by Ixodes scapularis ticks. *Journal of Clinical Microbiology*, 36(12), pp.3574–3578.

Hornok, S. et al., 2013. First evidence of Candidatus Neoehrlichia mikurensis in Hungary. *Parasites & Vectors*, 6(1), p.267. Available at: http://www.parasitesandvectors.com/content/6/1/267.

Hornok, S. et al., 2011. First Molecular Evidence of *Anaplasma ovis* and *Rickettsia* spp. in Keds (Diptera: Hippoboscidae) of Sheep and Wild Ruminants. *Vector-Borne and Zoonotic Diseases*, 11(10), pp.1319–1321. Available at: http://www.liebertonline.com/doi/abs/10.1089/vbz.2011.0649.

Hotopp, J.C.D. et al., 2007. Widespread Lateral Gene Transfer from Intracellular Bacteria to Multicellular Eukaryotes\r10.1126/science.1142490. *Science*, 317(5845), pp.1753–1756. Available at: http://www.sciencemag.org/cgi/content/abstract/317/5845/1753.

Huang, B. et al., 2011. The Anaplasma phagocytophilum-occupied vacuole selectively recruits Rab-GTPases that are predominantly associated with recycling endosomes. *Cellular Microbiology*, 12(9), pp.1292–1307.

Hudson, J.R., 1950. The recognition off tick-borne fever as a disease of cattle. *British Veterinary Journal*, 106(1), pp.3–17. Available at: http://dx.doi.org/10.1016/S0007-1935(17)53027-4.

Hulínská, D. et al., 2004. Detection of Anaplasma phagocytophilum in animals by real-time polymerase chain reaction. *APMIS : acta pathologica, microbiologica, et immunologica Scandinavica*, 112(4–5), pp.239–247.

Hurst, G.D.D., Jiggins, F.M. & Pomiankowski, A., 2002. Which Way to Manipulate Host Reproduction? Wolbachia That Cause Cytoplasmic Incompatibility Are Easily Invaded by Sex Ratio–Distorting Mutants. *The American Naturalist*, 160(3), pp.360–373. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18707445.

Huson, D.H. & Bryant, D., 2006. Application of Phylogenetic Networks in Evolutionary Studies. *Molecular Biology and Evolution*, 23(2), pp.254–267.

Ishida, S. & Taylor, D.J., 2007. Mature habitats associated with genetic divergence despite strong dispersal ability in an arthropod. *BMC evolutionary biology*, 7, p.52.

Jaenike, J. et al., 2013. Adaptation via Symbiosis : Recent Spread of a Drosophila Defensive Symbiont. *Scientific Reports*, 212(2010), pp.212–215.

Jahfari, S. et al., 2012. Prevalence of Neoehrlichia mikurensis in ticks and rodents from North-west Europe. *Parasites & vectors*, 5(1), p.74. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3395572&tool=pmcentrez&rendertype=abstract.

James, M.C. et al., 2014. The Heterogeneity, Distribution, and Environmental Associations of Borrelia burgdorferi Sensu Lato, the Agent of Lyme Borreliosis, in Scotland. *Frontiers in Public Health*, 2(August), pp.1–10. Available at: http://journal.frontiersin.org/article/10.3389/fpubh.2014.00129/abstract.

James, M.C. et al., 2011. The importance of passerine birds as tick hosts and in the transmission of Borrelia burgdorferi, the agent of Lyme disease: A case study from Scotland. *Ibis*, 153(2), pp.293–302.

Jasinskas, A., Zhong, J. & Barbour, A.G., 2007. Highly prevalent Coxiella sp. bacterium in the tick vector Amblyomma americanum. *Applied and Environmental Microbiology*, 73(1), pp.334–336.

Jayaprakash, A.D. et al., 2015. Stable heteroplasmy at the single-cell level is facilitated by intercellular exchange of mtDNA. *Nucleic Acids Research*, 43(4), pp.2177–2187.

Jongejan, F. & Uillenberg, G., 2004. The global importance of ticks. *Parasitology*, 129(5), pp.643–645.

Jordaan, J. & Baker, J., 1981. Survival rate on the host and mating capacity of Amblyomma hebreaum (Koch) male ticks. *Tick biology and control*, pp.115–117.

Jore, S. et al., 2014. Climate and environmental change drives Ixodes ricinus geographical expansion at the northern range margin. *Parasites & Vectors*, 7(1), p.11. Available at: http://www.parasitesandvectors.com/content/7/1/11/abstract%5Cnhttp://www.parasitesandvectors.com/content/7/1/11%5Cnhttp://www.parasitesandvectors.com/content/pdf/1756-3305-7-11.pdf.

Kagemann, J. & Clay, K., 2013. Effects of Infection by Arsenophonus and Rickettsia Bacteria on the Locomotive Ability of the Ticks Amblyomma americanum, Dermacentor variabilis, and Ixodes scapularis. *J. Med. Entomol*, 50(1), pp.155–162. Available at: http://dx.doi.org/10.1603/ME12086.

Kambris, Z. et al., 2009. Immune activation by life-shortening wolbachia and reduced filarial competence in mosquitoes. *Science*, 326(5949), pp.134–136.

Kawahara, M., 2004. Ultrastructure and phylogenetic analysis of “Candidatus Neoehrlichia mikurensis” in the family Anaplasmataceae, isolated from wild rats and found in Ixodes ovatus ticks. *INTERNATIONAL JOURNAL OF SYSTEMATIC AND EVOLUTIONARY MICROBIOLOGY*, 54(5), pp.1837–1843. Available at: http://ijs.microbiologyresearch.org/content/journal/ijsem/10.1099/ijs.0.63260-0.

Kelly, P. et al., 2010. RICKETTSIA AFRICAE IN AMBLYOMMA VARIEGATUM AND DOMESTIC RUMINANTS ON EIGHT CARIBBEAN ISLANDS. *Journal of Parasitology*, 96(6), pp.1086–1088.

Kelly, P.J. et al., 1991. Role of cattle in the epidemiology of tick-bite fever in Zimbabwe. *Journal of Clinical Microbiology*, 29(2), pp.256–259.

Kempf, F. et al., 2011. Host races in Ixodes ricinus, the European vector of Lyme borreliosis. *Infection, Genetics and Evolution*, 11(8), pp.2043–2048. Available at: http://dx.doi.org/10.1016/j.meegid.2011.09.016.

Klyachko, O. et al., 2007. Localization and visualization of a Coxiella-type symbiont within the lone star tick, Amblyomma americanum. *Applied and Environmental Microbiology*, 73(20), pp.6584–6594.

Kondo, N. et al., 2002. Genome fragment of Wolbachia endosymbiont transferred to X chromosome of host insect. *Proceedings of the National Academy of Sciences*, 99(22), pp.14280–14285. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.222228199.

Kotál, J. et al., 2015. Modulation of host immunity by tick saliva. *Journal of Proteomics*, 128(December 2017), pp.58–68.

Koutsovoulos, G. et al., 2014. Palaeosymbiosis Revealed by Genomic Fossils of Wolbachia in a Strongyloidean Nematode. *PLoS Genetics*, 10(6).

Kumar, S., Stecher, G. & Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular biology and evolution*, 33(7), pp.1870–1874.

Kurtti & al, 1996. NOTE Tick Cell Culture Isolation of an Intracellular Prokaryote from the Tick Ixodes scapularis. *Journal of Invertebrate Pathology*, 5158(67), pp.318–321. Available at: https://ac.els-cdn.com/S0022201196900506/1-s2.0-S0022201196900506-main.pdf?_tid=bb005d15-21cd-4db7-ab98-ab8f4d262e84&acdnat=1524829201_e54342fd19a847aa51a8910bfcfc97ca.

Kurtti, T.J. et al., 2016. Rickettsia buchneri sp . nov ., a rickettsial endosymbiont of the blacklegged tick Ixodes scapularis. , (2015), pp.37–64.

Kuwahara, Y. et al., 2009. Unusual thermal disassembly of the SPFH domain oligomer from Pyrococcus horikoshii. *Biophysical journal*, 97(7), pp.2034–43. Available at: http://dx.doi.org/10.1016/j.bpj.2009.07.034.

de la Fuente, J. et al., 2015. Anaplasma phagocytophilum Uses Common Strategies for Infection of Ticks and Vertebrate Hosts. *Trends in Microbiology*, xx, pp.1–8. Available at: http://dx.doi.org/10.1016/j.tim.2015.12.001.

de la Fuente, J. et al., 2016. Anaplasma phagocytophilum Uses Common Strategies for Infection of Ticks and Vertebrate Hosts. *Trends in Microbiology*, 24(3), pp.173–180. Available at: http://dx.doi.org/10.1016/j.tim.2015.12.001.

de la Fuente, J. et al., 2001. Differential adhesion of major surface proteins 1a and 1b of the ehrlichial cattle pathogen Anaplasma marginale to bovine erythrocytes and tick cells. *International Journal for Parasitology*, 31(2), pp.145–153.

de la Fuente, J. et al., 2005. Gene expression profiling of human promyelocytic cells in response to infection with Anaplasma phagocytophilum. *Cellular Microbiology*, 7(4), pp.549–559.

de la Fuente, J. et al., 2005. Sequence Analysis of the msp4 Gene of Anaplasma phagocytophilum Strains. *Clinical Microbiology*, 43(3), pp.1309–1317.

de la Fuente, J. et al., 2017. Tick-Pathogen Interactions and Vector Competence: Identification of Molecular Drivers for Tick-Borne Diseases. *Frontiers in Cellular and Infection Microbiology*, 7(April), pp.1–13. Available at: http://journal.frontiersin.org/article/10.3389/fcimb.2017.00114/full.

Ladbury, G.A.F. et al., 2008. Dynamic transmission of numerous Anaplasma phagocytophilum genotypes among lambs in an infected sheep flock in an area of anaplasmosis endemicity. *Journal of Clinical Microbiology*, 46(5), pp.1686–1691.

Laetsch, D.R. & Blaxter, M.L., 2017. BlobTools: Interrogation of genome assemblies. *F1000Research*, 6, p.1287. Available at: https://f1000research.com/articles/6-1287/v1.

Langhorst, M.F. et al., 2007. Linking membrane microdomains to the cytoskeleton: Regulation of the lateral mobility of reggie-1/flotillin-2 by interaction with actin. *FEBS Letters*, 581(24), pp.4697–4703.

Lee, H.C. et al., 2008. Anaplasma phagocytophilum-induced gene expression in both human neutrophils and HL-60 cells. *Genomics*, 92(3), pp.144–151.

Lee, H.C. & Goodman, J.L., 2006. Anaplasma phagocytophilum causes global induction of antiapoptosis in human neutrophils. *Genomics*, 88(4), pp.496–503.

LePage, D. & Bordenstein, S.R., 2013. Wolbachia: Can we save lives with a great pandemic? *Trends in Parasitology*, 29(8), pp.385–393. Available at: http://dx.doi.org/10.1016/j.pt.2013.06.003.

Lewis, D., 1979. The Detection of Rickettsia-Like Microorganisms Within the Ovaries of Female Ixodes ricinus Ticks David. *Parasitology Research*, 298, pp.295–298.

Li, L., Stoeckert, C.J.J. & Roos, D.S., 2003. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes -- Li et al. 13 (9): 2178 -- Genome Research. *Genome Research*, 13(9), pp.2178–2189. Available at: http://genome.cshlp.org/cgi/content/full/13/9/2178.

Lin, M. et al., 2007. Anaplasma phagocytophilum AnkA secreted by type IV secretion system is tyrosine phosphorylated by Abl-1 to facilitate infection. *Cellular Microbiology*, 9(11), pp.2644–2657.

Lin, M. et al., 2011. Global proteomic analysis of two tick-borne emerging zoonotic agents: Anaplasma phagocytophilum and Ehrlichia chaffeensis. *Frontiers in Microbiology*, 2(FEB), pp.1–19.

Lin, M. et al., 2002. Rapid Activation of Protein Tyrosine Kinase and Phospholipase C- γ 2 and Increase in Cytosolic Free Calcium Are Required by Ehrlichia chaffeensis for Internalization and Growth in THP-1 Cells Rapid Activation of Protein Tyrosine Kinase and Phospholipase C. *Infection and Immunity*, 70(2), pp.889–898.

Lin, M. & Rikihisa, Y., 2003. Ehrlichia chaffeensis and Anaplasma phagocytophilum lack genes for lipid A biosynthesis and incorporate cholesterol for their survival. *Infection and Immunity*, 71(9), pp.5324–5331.

Lin, Q. et al., 2004. Anaplasma phagocytophilum has a functional msp2 gene that is distinct from p44. *Infection and Immunity*, 72(7), pp.3883–3889.

Lin, Q. et al., 2003. Mechanisms of variable p44 expression by Anaplasma phagocytophilum. *Infection and Immunity*, 71(10), pp.5650–5661.

Lin, Q., Zhang, C. & Rikihisa, Y., 2006. Analysis of involvement of the RecF pathway in p44 recombination in Anaplasma phagocytophilum and in Escherichia coli by using a plasmid carrying the p44 expression and p44 donor loci. *Infection and Immunity*, 74(4), pp.2052–2062.

Lindblom, A. et al., 2016. Prevalence of rickettsia spp. in ticks and serological and clinical outcomes in tick-bitten individuals in Sweden and on the Åland Islands. *PLoS ONE*, 11(11), pp.1–14.

Lindquist, E., Krantz, G.W. & Walter, D.E., 2009. *A Manual of Acarology*, Lubbok,Texas:Texas: Tech University Press.

Lo, N. et al., 2006. Widespread distribution and high prevalence of an alpha-proteobacterial symbiont in the tick Ixodes ricinus. *Environmental Microbiology*, 8(7), pp.1280–1287. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16817936.

Lockwood, S. et al., 2016. Comparative genomics reveals multiple pathways to mutualism for tick-borne pathogens. *BMC Genomics*, 17(1), p.481. Available at: http://bmcgenomics.biomedcentral.com/articles/10.1186/s12864-016-2744-9.

Lommano, E. et al., 2012. Infections and coinfections of questing Ixodes ricinus ticks by emerging zoonotic pathogens in Western Switzerland. *Applied and Environmental Microbiology*, 78(13), pp.4606–4612.

Lorusso, V. et al., 2013. Rickettsia africae in Amblyomma variegatum Ticks, Uganda and Nigeria. *Emerging Infectious Diseases*, 19(10), pp.1705–1707. Available at: http://wwwnc.cdc.gov/eid/article/19/10/13-0389_article.htm.

Love, R.R. et al., 2016. Evaluation of DISCOVAR de novo using a mosquito sample for cost-effective

short-read genome assembly. *BMC Genomics*, 17(1), pp.1–10. Available at: http://dx.doi.org/10.1186/s12864-016-2531-7.

Macaluso, K.R. et al., 2002. Rickettsial Infection in *Dermacentor variabilis* (Acari: Ixodidae) Inhibits Transovarial Transmission of a Second *Rickettsia*. *Journal of Medical Entomology*, 39(6), pp.809–813. Available at: https://academic.oup.com/jme/article-lookup/doi/10.1603/0022-2585-39.6.809.

Magnarelli, L.A. & Swihart, R.K., 1991. Spotted fever group rickettsiae or Borrelia burgdorferi in Ixodes cookei (Ixodidae) in Connecticut. *Journal of Clinical Microbiology*, 29(7), pp.1520–1522.

Maiden, M.C. et al., 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America*, 95(6), pp.3140–5. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=19708&tool=pmcentrez&rendertype=abstract.

Maina, A.N. et al., 2014. High Prevalence of Rickettsia africae Variants in Amblyomma variegatum Ticks from Domestic Mammals in Rural Western Kenya : Implications for Human Health. *VECTOR-BORNE AND ZOONOTIC DISEASES*, 14(10), pp.693–702.

Makepeace, B.L., Rodgers, L. & Trees, A.J., 2006. Rate of elimination of Wolbachia pipientis by doxycycline in vitro increases following drug withdrawal. *Antimicrobial Agents and Chemotherapy*, 50(3), pp.922–927.

Mariconti, M. et al., 2012. Humans parasitized by the hard tick Ixodes ricinus are seropositive to Midichloria mitochondrii: is Midichloria a novel pathogen, or just a marker of tick bite? *Pathogens and global health*, 106(7), pp.391–6. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3589662&tool=pmcentrez&rendertype=abstract.

Massung, R.F. et al., 2005. Anaplasma phagocytophilum in white-tailed deer. *Emerging Infectious Diseases*, 11(10), pp.1604–1606.

Massung, R.F. et al., 2003. Inability of a variant strain of Anaplasma phagocytophilum to infect mice. *The Journal of infectious diseases*, 188(11), pp.1757–63. Available at: http://www.ncbi.nlm.nih.gov/pubmed/14639548.

Massung, R.F. et al., 2007. Isolation and propagation of the Ap-variant 1 strain of Anaplasma phagocytophilum in a tick cell line. *Journal of Clinical Microbiology*, 45(7), pp.2138–2143.

Massung, R.F. et al., 1998. Nested PCR assay for detection of granulocytic ehrlichiae. *Journal of clinical microbiology*, 36(4), pp.1090–1095.

Massung, R.F., Mather, T.N. & Levin, M.L., 2006. Reservoir competency of goats for Anaplasma phagocytophilum. *Annals of the New York Academy of Sciences*, 1078(2), pp.476–478.

McCarthy, F.M. et al., 2006. AgBase: A functional genomics resource for agriculture. *BMC Genomics*, 7, pp.1–13.

Mediannikov, O. et al., 2012. Isolation of Arsenophonus nasoniae from Ixodes ricinus ticks in Slovakia. *Ticks and Tick-borne Diseases*, 3(5–6), pp.367–370. Available at: http://dx.doi.org/10.1016/j.ttbdis.2012.10.016.

Mediannikov, O. & Fenollar, F., 2014. Looking in ticks for human bacterial pathogens. *Microbial Pathogenesis*, 77, pp.142–148. Available at: http://dx.doi.org/10.1016/j.micpath.2014.09.008.

Medlock, J.M. et al., 2013. Driving forces for changes in geographical distribution of Ixodes ricinus ticks in Europe. *Parasites & Vectors*, 6(1), pp.31–50. Available at: http://www.karger.com/doi/10.1159/000213069.

Meeus, T. de et al., 2002. SEX-BIASED GENETIC STRUCTURE IN THE VECTOR OF LYME DISEASE,

IXODES RICINUS. *Evolution*, 56(9), pp.1802–1807. Available at: http://doi.wiley.com/10.1111/j.0014-3820.2002.tb00194.x.

Milne, I. et al., 2009. TOPALi v2: A rich graphical interface for evolutionary analyses of multiple alignments on HPC clusters and multi-core desktops. *Bioinformatics*, 25(1), pp.126–127.

Moniuszko, A. et al., 2014. Coinfection of tick cell lines has variable effects on replication of intracellular bacterial and viral pathogens. *Ticks and Tick-borne Diseases*, 5(4), pp.415–422. Available at: http://dx.doi.org/10.1016/j.ttbdis.2014.01.010.

Moran, N.A., McCutcheon, J.P. & Nakabachi, A., 2008. Genomics and Evolution of Heritable Bacterial Symbionts. *Annual Review of Genetics*, 42(1), pp.165–190. Available at: http://www.annualreviews.org/doi/10.1146/annurev.genet.41.110306.130119.

Moreno, C.X. et al., 2006. Molecular analysis of microbial communities identified in different developmental stages of Ixodes scapularis ticks from Westchester and Dutchess Counties, New York. *Environmental Microbiology*, 8(5), pp.761–772. Available at: http://doi.wiley.com/10.1111/j.1462-2920.2005.00955.x.

Mott, J., Barnewall, R.E. & Rikihisa, Y., 1999. Human granulocytic Ehrlichiosis agent and Ehrlichia chaffeensis reside in different cytoplasmic compartments in HL-60 cells. *Infection and Immunity*, 67(3), pp.1368–1378.

Moutailler, S. et al., 2016. Co-infection of Ticks: The Rule Rather Than the Exception. *PLoS Neglected Tropical Diseases*, 10(3), pp.1–17.

Mulenga, A., Blandon, M. & Khumthong, R., 2007. The molecular basis of the Amblyomma americanum tick attachment phase. *Experimental and Applied Acarology*, 41(4), pp.267–287.

Munderloh, U.G. et al., 1999. Invasion and Intracellular Development of the Human Granulocytic Ehrlichiosis Agent in Tick Cell Culture Invasion and Intracellular Development of the Human Granulocytic Ehrlichiosis Agent in Tick Cell Culture. , 37(8), pp.2518–2524.

Munderloh, U.G. et al., 1996. Isolation of the equine granulocytic ehrlichiosis agent, Ehrlichia equi, in tick cell culture. *Journal of Clinical Microbiology*, 34(3), pp.664–670.

Munderloh, U.G., Jauron, S.D. & and Kurtti, T.J., 2005. The Tick: A Different Kind of Host for Human Pathogens. *Tick-Borne Diseases of Humans*, pp.37–64.

Munderloh, U.G. & Kurtti, T.J., 1989. Formulation of medium for tick cell culture. *Experimental & Applied Acarology*, 7(3), pp.219–229.

N, E.T.I.J.S.S.E.L.A.S.E. & N, L.J.J.A.M.E.S.O., 2011. First detection of spotted fever group rickettsiae in Ixodes ricinus and Dermacentor reticulatus ticks in the UK. , (2011), pp.524–529.

Nakao, R. et al., 2013. Ticks and Tick-borne Diseases High prevalence of spotted fever group rickettsiae in Amblyomma variegatum from Uganda and their identification using sizes of intergenic spacers. *Ticks and Tick-borne Diseases*, 4(6), pp.506–512.

Naranjo, V. et al., 2013. Reciprocal Regulation of NF-kB (Relish) and Subolesin in the Tick Vector, Ixodes scapularis. *PLoS ONE*, 8(6), pp.1–12.

Ndi, C. et al., 1998. Preliminary observations on ticks and tick-borne diseases in the North West Province of Cameroon II. Bovine heartwater. , 44(3), pp.263–265.

Ndip, L.M. et al., 2004. Detection of Rickettsia africae in patients and ticks along the coastal region of Cameroon. *American Journal of Tropical Medicine and Hygiene*, 71(3), pp.363–366.

Ndip, L.M., 2013. Prevalence of Ehrlichia ruminantium in adult Amblyomma variegatum collected from cattle. *Experimental & Applied Acarology*, 59, pp.377–387.

Neelakanta, G. et al., 2010. Anaplasma phagocytophilum induces Ixodes scapularis ticks to express an antifreeze glycoprotein gene that enhances their survival in the cold. *Journal of Clinical*

Investigation, 120(9).

Niebylski, M.L. et al., 1997. Rickettsia peacockii sp. nov., a new species infecting wood ticks, Dermacentor andersoni, in western Montana. *International Journal of Systematic Bacteriology*, 47(2), pp.446–452.

Nilsson, K., Elfving, K. & Påhlson, C., 2006. Rickettsia helvetica in patient with meningitis, Sweden, 2006. *Emerging Infectious Diseases*, 16(3), pp.490–492.

Ninio, C. et al., 2015. Antibiotic treatment of the hard tick Ixodes ricinus: influence on Midichloria mitochondrii load following blood meal. *Ticks and Tick-borne Diseases*, 6(5), pp.653–657. Available at: http://linkinghub.elsevier.com/retrieve/pii/S1877959X1500103X.

Niu, H. et al., 2010. Anaplasma phagocytophilum Ats-1 Is imported into host cell mitochondria and interferes with apoptosis induction. *PLoS Pathogens*, 6(2).

Niu, H., Yamaguchi, M. & Rikihisa, Y., 2008. Subversion of cellular autophagy by Anaplasma phagocytophilum. *Cellular Microbiology*, 10(3), pp.593–605.

Noda, H., Munderloh, U.G. & Kurtti, T.J., 1997. Endosymbionts of ticks and their relationship to Wolbachia spp. and tick-borne pathogens of humans and animals. *Applied and environmental microbiology*, 63(10), pp.3926–32. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9327557.

Norris, D.E. et al., 1996. Population genetics of Ixodes scapularis (Acari: Ixodidae) based on mitochondrial 16S and 12S genes. *Journal of medical entomology*, 33(1), pp.78–89. Available at: https://academic.oup.com/jme/article-lookup/doi/10.1093/jmedent/33.1.78.

Noureddine, R., Chauvin, a. & Plantard, O., 2011. Lack of genetic structure among Eurasian populations of the tick Ixodes ricinus contrasts with marked divergence from north-African populations. *International Journal for Parasitology*, 41(2), pp.183–192. Available at: http://dx.doi.org/10.1016/j.ijpara.2010.08.010.

Ogden, N.H. et al., 1998. Granulocytic Ehrlichia infection in Ixodid ticks and mammals in woodlands and uplands of the U.K. *Medical and Veterinary Entomology*, 12(4), pp.423–429.

Ogden, N.H. et al., 2008. Role of migratory birds in introduction and range expansion of Ixodes scapularis ticks and of Borrelia burgdorferi and Anaplasma phagocytophilum in Canada. *Applied and Environmental Microbiology*, 74(6), pp.1780–1790.

Oksi, J. et al., 1993. Fatal encephalitis caused by concomitant infection with tick-borne encephalitis virus and borrelia burgdorferi. *Clinical Infectious Diseases*, 16(3), pp.392–396.

Oliver, K.M. et al., 2008. Population dynamics of defensive symbionts in aphids. *Proceedings of the Royal Society B: Biological Sciences*, 275(1632), pp.293–299. Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2007.1192.

Oliver, K.M., Moran, N.A. & Hunter, M.S., 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proceedings of the National Academy of Sciences*, 102(36), pp.12795–12800. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0506131102.

Oteo, J.A. & Portillo, A., 2012. Tick-borne rickettsioses in Europe. *Ticks and Tick-borne Diseases*, 3(5–6), pp.271–278. Available at: http://dx.doi.org/10.1016/j.ttbdis.2012.10.035.

Palmer, G.U.Y.H., Futse, J.E. & Knowles, D.P., 2006. Insights into Mechanisms of Bacterial Antigenic Variation Derived from the Complete Genome Sequence of Anaplasma marginale. *Annals of the New York Academy of Sciences*, 25, pp.15–25.

Parola, P. et al., 2001. Detection and identification of spotted fever group Rickettsiae and Ehrlichiae in African ticks. *Emerging Infectious Diseases*, 7(6), pp.1014–1017.

Parola, P. et al., 1998. First isolation of Rickettsia helvetica from Ixodes ricinus ticks in France. *European Journal of Clinical Microbiology and Infectious Diseases*, 17, pp.95–100. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/9629973.

Pascual, J., Castresana, J. & Saraste, M., 1997. Evolution of the spectrin repeat. *BioEssays*, 19(9), pp.811–817.

Passos, L.M.F., 2012. In vitro cultivation of Anaplasma marginale and A. phagocytophilum in tick cell lines: a review. *Revista Brasileira de Parasitologia Veterinária*, 21(2), pp.81–86. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22832744%5Cnhttp://www.scielo.br/scielo.php?script=sci_arttext&pid=S1984-29612012000200002&lng=en&nrm=iso&tlng=en.

Patrican, L.A., 1997. Acquisition of lyme disease spirochetes by cofeeding Ixodes scapularis ticks. *American Journal of Tropical Medicine and Hygiene*, 57(5), pp.589–593.

Pedra, J.H.F. et al., 2010. Fucosylation enhances colonization of ticks by Anaplasma phagocytophilum. *Cellular Microbiology*, 12(9), pp.1222–1234.

Pedra, J.H.F. et al., 2005. Modulation of NB4 promyelocytic leukemic cell machinery by Anaplasma phagocytophilum. *Genomics*, 86(3), pp.365–377.

Pegram, R.G. & Banda, D.S., 1990. Ecology and Phenology of Cattle Ticks in Zambia : Development and Survival of Free-Living Stages. *Experimental & Applied Acarology*, 8, pp.291–301.

Pekova, S. et al., 2011. Candidatus Neoehrlichia mikurensis infection identified in 2 hematooncologic patients : benefit of molecular techniques for rare pathogen detection. *Diagnostic Microbiology and Infectious Disease*, 69(3), pp.266–270. Available at: http://dx.doi.org/10.1016/j.diagmicrobio.2010.10.004.

Piesman, J. & Spielman, A., 1980. Human babesiosis on Nantucket island: Prevalence of Babesia microti in ticks. *American Journal of Tropical Medicine and Hygiene*, 29(5), pp.742–746.

Pietzsch, M.E. et al., 2005. Distribution of Ixodes ricinus in the British Isles: Investigation of historical records. *Medical and Veterinary Entomology*, 19(3), pp.306–314.

Pietzsch, M.E. et al., 2008. Preliminary evaluation of exotic tick species and exotic pathogens imported on migratory birds into the British Isles. *Veterinary Parasitology*, 155(3–4), pp.328–332.

Pilgrim, J. et al., 2017. Torix group Rickettsia are widespread in Culicoides biting midges (Diptera: Ceratopogonidae), reach high frequency and carry unique genomic features. *Environmental Microbiology*, 19(10), pp.4238–4255.

Pond, S.L.K. & Frost, S.D.W., 2005. Datamonkey : rapid detection of selective pressure on individual sites of codon alignments. , 21(10), pp.2531–2533.

Portillo, A. et al., 2015. Rickettsioses in Europe. *Microbes and Infection*, 17(11–12), pp.834–838.

Pusterla, N., Pusterla, JB., Braun, U., Lutz, H., 1999. Experimental cross-infections with Ehrlichia phagocytophila and human granulocytic ehrlichia-like agent in cows and horses. *Veterinary Record*, 145(11), pp.311–314.

Putman, R. et al., 2011. Identifying threshold densities for wild deer in the UK above which negative impacts may occur. *Mammal Review*, 41(3), pp.175–196.

Qviller, L. et al., 2016. The influence of red deer space use on the distribution of Ixodes ricinus ticks in the landscape. *Parasites & vectors*, 9(1), p.545. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27737695%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5064927.

Rajput, Z.I. et al., 2006. Importance of ticks and their chemical and immunological control in livestock. *Journal of Zhejiang University SCIENCE B*, 7(11), pp.912–921. Available at: http://www.springerlink.com/index/10.1631/jzus.2006.B0912.

Randolph, S.E. et al., 2002. An empirical quantitative framework for the seasonal population

dynamics of the tick Ixodes ricinus. *International Journal for Parasitology*, 32(8), pp.979–989.

Raoult, L.B.P.J.K.L.A.M.P.R.M.D., 1995. Prevalence of Rickettsia-Like Organisms and Spotted Fever Group Rickettsiae in Ticks (Acari: Ixodidae) from Zimbabwe. *ournal of Medical Entomology*, 32(6), pp.787–792.

Rich, S.M. et al., 1995. Distribution of the Ixodes ricinus-like ticks of eastern North America. *Proceedings of the National Academy of Sciences of the United States of America*, 92(14), pp.6284–8. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=41502&tool=pmcentrez&rendertype=abstract.

Richter, D. & Matuschka, F.R., 2012. "Candidatus Neoehrlichia mikurensis," Anaplasma phagocytophilum, and Lyme Disease spirochetes in questing European vector ticks and in feeding ticks removed from people. *Journal of Clinical Microbiology*, 50(3), pp.943–947.

Richter, M. & Rossello-Mora, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106(45), pp.19126–19131. Available at: http://www.pnas.org/cgi/doi/10.1073/pnas.0906412106.

Rikihisa, Y., 2010. Anaplasma phagocytophilum and Ehrlichia chaffeensis: Subversive manipulators of host cells. *Nature Reviews Microbiology*, 8(5), pp.328–339. Available at: http://dx.doi.org/10.1038/nrmicro2318.

Rikihisa, Y., 2006. Ehrlichia subversion of host innate responses. *Current Opinion in Microbiology*, 9(1), pp.95–101.

Rikihisa, Y., 2011. Mechanisms of obligatory intracellular infection with Anaplasma phagocytophilum. *Clinical Microbiology Reviews*, 24(3), pp.469–489.

Rikihisa, Y., 2003. Mechanisms to Create a Safe Haven by Members of the Family Anaplasmataceae. *Annals of the New York Academy of Sciences*, 990(1), pp.548–555. Available at: http://doi.wiley.com/10.1111/j.1749-6632.2003.tb07425.x.

Rikihisa, Y., 1991. The tribe Ehrlichieae and ehrlichial diseases. *Clinical Microbiology Reviews*, 4(3), pp.286–308.

Rikihisa, Y., Lin, M. & Niu, H., 2010. Type IV secretion in the obligatory intracellular bacterium Anaplasma phagocytophilum. *Cellular Microbiology*, 12(9), pp.1213–1221.

Rio, R.V.M., Attardo, G.M. & Weiss, B.L., 2016. Grandeur Alliances: Symbiont Metabolic Integration and Obligate Arthropod Hematophagy. *Trends in Parasitology*, 32(9), pp.739–749. Available at: http://dx.doi.org/10.1016/j.pt.2016.05.002.

Rizzoli, A. et al., 2014. Ixodes ricinus and Its Transmitted Pathogens in Urban and Peri-Urban Areas in Europe: New Hazards and Relevance for Public Health. *Frontiers in Public Health*, 2(December), p.251. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4248671&tool=pmcentrez&rendertype=abstract.

Rizzoli, A. et al., 2011. Lyme borreliosis in Europe. *Eurosurveillance*, 16(27), pp.1–8.

Robinson, J.B. et al., 2009. New approaches to detection and identification of Rickettsia africae and Ehrlichia ruminantium in Amblyomma variegatum (Acari: Ixodidae) ticks from the Caribbean. *Journal of medical entomology*, 46(4), pp.942–951.

Roed, K.H. et al., 2016. Phylogenetic Lineages and Postglacial Dispersal Dynamics Characterize the Genetic Structure of the Tick, Ixodes ricinus, in Northwest Europe U. G. Munderloh, ed. *PLOS ONE*, 11(12), p.e0167450. Available at: http://dx.plos.org/10.1371/journal.pone.0167450.

Roux, V. et al., 1997. Citrate Synthase Gene Comparison, a New Tool for Phylogenetic Analysis, and Its Application for the Rickettsiae. *International Journal of Systematic Bacteriology*, 47(2),

pp.252–261.

Ruiz-Fons, F. & Gilbert, L., 2010. The role of deer as vehicles to move ticks, Ixodes ricinus, between contrasting habitats. *International Journal for Parasitology*, 40(9), pp.1013–1020. Available at: http://dx.doi.org/10.1016/j.ijpara.2010.02.006.

Růžek, D. et al., 2008. Growth of tick-borne encephalitis virus (European subtype) in cell lines from vector and non-vector ticks. *Virus Research*, 137(1), pp.142–146.

Sacchi, L. et al., 2004. A symbiont of the thick Ixodes ricinus invades and consumes mitochondria in a mode similar to that of the parasitic bacterium Bdellovibrio bacteriovorus. *Tissue and Cell*, 36(1), pp.43–53.

Sassera, D. et al., 2008. "Candidatus midichloria" endosymbionts bloom after the blood meal of the host, the hard tick Ixodes ricinus. *Applied and Environmental Microbiology*, 74(19), pp.6138–6140.

Sassera, D. et al., 2006. Candidatus Midichloria mitochondrii', an endosymbiont of the Ixodes ricinus with a unique intramitochondrial lifestyle. *International Journal of Systematic and Evolutionary Microbiology*, 56(11), pp.2535–2540.

Scarpulla, M. et al., 2016. Molecular detection and characterization of spotted fever group rickettsiae in ticks from Central Italy. *Ticks and Tick-borne Diseases*, 7(5), pp.1052–1056. Available at: http://dx.doi.org/10.1016/j.ttbdis.2016.06.003.

Schouls, L.M. et al., 1999. Detection and Identification of Ehrlichia, Borrelia burgdorferi Sensu Lato, and Bartonella Species in Dutch Ixodes ricinus Ticks. *Journal of Clinical Microbiology*, 37(7), pp.2215–2222.

Scoles, G.A., 2004. Phylogenetic Analysis of the Francisella -like Endosymbionts of Dermacentor Ticks. *Journal of Medical Entomology*, 41(3), pp.277–286. Available at: https://academic.oup.com/jme/article-lookup/doi/10.1603/0022-2585-41.3.277.

Scorpio, D.G. et al., 2006. Innate immune response to Anaplasma phagocytophilum contributes to hepatic injury. *Clinical and Vaccine Immunology*, 13(7), pp.806–809.

Scott, G.R.& & Horsburgh, D., 1983. New rickettsial isolates. , 36, p.17.

Seemann, T., 2014. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, 30(14), pp.2068–2069.

Seidman, D. et al., 2015. Essential Domains of Anaplasma phagocytophilum Invasins Utilized to Infect Mammalian Host Cells. *PLOS Pathogens*, 11(2), p.e1004669. Available at: http://dx.plos.org/10.1371/journal.ppat.1004669.

Severinsson, K. et al., 2010. Detection and prevalence of Anaplasma phagocytophilum and Rickettsia helvetica in Ixodes ricinus ticks in seven study areas in Sweden. *Parasites and Vectors*, 3, p.66. Available at: http://www.embase.com/search/results?subaction=viewrecord&from=export&id=L51021826%5Cnhttp://dx.doi.org/10.1186/1756-3305-3-66%5Cnhttp://rc5hy6jd4f.search.serialssolutions.com?sid=EMBASE&issn=17563305&id=doi:10.1186/1756-3305-3-66&atitle=Detection+and+pre.

Severo, M.S. et al., 2012. Anaplasma phagocytophilum : deceptively simple or simply deceptive? *Future Microbiology*, 7(6), pp.719–731. Available at: http://www.futuremedicine.com/doi/10.2217/fmb.12.45.

Severo, M.S. et al., 2013. The E3 ubiquitin ligase XIAP restricts anaplasma phagocytophilum colonization of ixodes scapularis ticks. *Journal of Infectious Diseases*, 208(11), pp.1830–1840.

Silaghi, C. et al., 2008. Anaplasma phagocytophilum infection in ixodes ricinus, Bavaria, Germany. *Emerging Infectious Diseases*, 14(6), pp.972–974.

Simser, J. a, Palmer, a N.N.T. & Munderloh, U.G., 2001. Isolation of a Spotted Fever Group Rickettsia,. *Society*, 67(2), pp.546–552.

Simser, J.A. et al., 2002. Rickettsia monacensis sp . nov ., a Spotted Fever Group Rickettsia , from Ticks (Ixodes ricinus) Collected in a European City Park Rickettsia monacensis sp . nov ., a Spotted Fever Group Rickettsia , from Ticks (Ixodes ricinus) Collected in a European. *Applied and environmental microbiology*, 68(9), pp.4559–4566.

Sinclair, S.H.G., Yegnasubramanian, S. & Dumler, J.S., 2015. Global DNA methylation changes and differential gene expression in Anaplasma phagocytophilum-infected human neutrophils. *Clinical Epigenetics*, 7(1), p.77. Available at: http://www.clinicalepigeneticsjournal.com/content/7/1/77.

Skaug, B., Jiang, X. & Chen, Z.J., 2009. The Role of Ubiquitin in NF-κB Regulatory Pathways. *Annual Review of Biochemistry*, 78(1), pp.769–796. Available at: http://www.annualreviews.org/doi/10.1146/annurev.biochem.78.070907.102750.

Skog, A. et al., 2009. Phylogeography of red deer Cervus elaphus in Europe. *Journal of Biogeography*, 36(1), pp.66–77. Available at: http://doi.wiley.com/10.1111/j.1365-2699.2008.01986.x.

Smith, F.D., Wall, L. & Richard, E., 2013. Prevalence of Babesia and Anaplasma in ticks infesting dogs in Great Britain. *Veterinary Parasitology*, 198(1–2), pp.18–23. Available at: http://dx.doi.org/10.1016/j.vetpar.2013.08.026.

Sonenshine, D.E., 1992. Biology of Ticks (Vol . I) Oxford University Press, New York, 3-66. , 8(8), p.1992.

Stannard, A.A., Gribble, D.H. & Smith, R.S., 1969. Equine ehrlichiosis: a disease with similarities to tick-borne fever and bovine petechial fever. *Veterinary Record.*, (6)(84), pp.149–150.

Stenos, J., Graves, S.R. & Unsworth, N.B., 2005. A highly sensitive and specific real-time PCR assay for the detection of spotted fever and typhus group Rickettsiae. *The American journal of tropical medicine and hygiene*, 73(6), pp.1083–5. Available at: http://www.ajtmh.org/cgi/content/long/73/6/1083%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/16354816.

Steyn, H.C. et al., 2008. A quantitative real-time PCR assay for Ehrlichia ruminantium using pCS20. *Veterinary Microbiology*, 131(3–4), pp.258–265. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0378113508001351.

Stuen, S., 2007. Anaplasma Phagocytophilum - The most widespread tick-borne infection in animals in Europe. *Veterinary Research Communications*, 31(SUPPL. 1), pp.79–84.

Stuen, S. et al., 2005. Anaplasma phagocytophilum infection in north Norway. The first laboratory confirmed case. *Acta veterinaria Scandinavica*, 46(3), pp.167–171.

Stuen, S. et al., 2010. Experimental Infection in Lambs With a Red Deer (Cervus Elaphus) Isolate of Anaplasma Phagocytophilum. *Journal of Wildlife Diseases*, 46(3), pp.803–809. Available at: http://www.jwildlifedis.org/doi/10.7589/0090-3558-46.3.803.

Stuen, S., Granquist, E.G. & Silaghi, C., 2013. Anaplasma phagocytophilum—a widespread multi-host pathogen with highly adaptive strategies. *Frontiers in Cellular and Infection Microbiology*, 3(July), pp.1–33. Available at: http://journal.frontiersin.org/article/10.3389/fcimb.2013.00031/abstract.

Sukumaran, B. et al., 2006. An Ixodes scapularis protein required for survival of Anaplasma phagocytophilum in tick salivary glands. *The Journal of experimental medicine*, 203(6), pp.1507–1517.

Sultana, H. et al., 2010. Anaplasma phagocytophilum induces actin phosphorylation to selectively regulate gene transcription in Ixodes scapularis ticks. *The Journal of experimental medicine*, 207(8), pp.1727–1743. Available at:

http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2916137&tool=pmcentrez&rendertype=abstract.

Sumption, K.J. et al., 1995. Human ehrlichiosis in the UK. *the lacet*, 346, pp.1487–1488.

Swanson, S.J. et al., 2006. Coinfections acquired from Ixodes ticks. *Clinical Microbiology Reviews*, 19(4), pp.708–727.

Tamura, K. et al., 2013. MEGA6: Molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, 30(12), pp.2725–2729.

Tappe, J. & Strube, C., 2013. Anaplasma phagocytophilum and Rickettsia spp. infections in hard ticks (Ixodes ricinus) in the city of Hanover (Germany): Revisited. *Ticks and Tick-borne Diseases*, 4(5), pp.432–438. Available at: http://dx.doi.org/10.1016/j.ttbdis.2013.04.009.

Thomas, R.J. et al., 2013. Expression of p44 variant-specific antibodies in sheep persistently infected with Anaplasma phagocytophilum. *Veterinary Microbiology*, 167(3–4), pp.484–493. Available at: http://dx.doi.org/10.1016/j.vetmic.2013.07.032.

Tian, T. et al., 2017. AgriGO v2.0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Research*, 45(W1), pp.W122–W129.

Tijsse-Klasen, E. et al., 2013. Spotted fever group rickettsiae in Dermacentor reticulatus and Haemaphysalis punctata ticks in the UK. *Parasites & vectors*, 6(1), p.212. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3725166&tool=pmcentrez&rendertype=abstract.

Troese, M.J. et al., 2011. Proteomic analysis of anaplasma phagocytophilum during infection of human myeloid cells identifies a protein that is pronouncedly upregulated on the infectious dense-cored cell. *Infection and Immunity*, 79(11), pp.4696–4707.

Truchan, H.K. et al., 2016. Anaplasma phagocytophilum Rab10-dependent parasitism of the trans-Golgi network is critical for completion of the infection cycle. *Cellular Microbiology*, 18(2), pp.260–281.

Truchan, H.K., Seidman, D. & Carlyon, J.A., 2013. Breaking in and grabbing a meal: Anaplasma phagocytophilum cellular invasion, nutrient acquisition, and promising tools for their study. *Microbes and Infection*, 15(14–15), pp.1017–1025. Available at: http://dx.doi.org/10.1016/j.micinf.2013.10.010.

Uchiyama, T., 2012. Tropism and pathogenicity of rickettsiae. *Frontiers in Microbiology*, 3(JUN), pp.1–11.

Uilenberg, G., CJ, van V. & NM., P., 1979. Blood parasites of sheep in the Netherlands. I. Anaplasma mesaeterum sp.n. (Rickettsiales, Anaplasmataceae). *Veterinary Quarterly*, 2(104), pp.14–22.

Vanegas, A. et al., 2018. Molecular detection of spotted fever group rickettsiae in ticks from Cameroon. *Ticks and Tick-borne Diseases*, 9(4), pp.1049–1056. Available at: https://doi.org/10.1016/j.ttbdis.2018.03.022.

Vaughn, M.F. et al., 2014. Seroepidemiologic study of human infections with spotted fever group rickettsiae in North Carolina. *Journal of Clinical Microbiology*, 52(11), pp.3960–3966.

Vayssier-Taussat, M. et al., 2015. Emerging horizons for tick-borne pathogens: from the 'one pathogen–one disease' vision to the pathobiome paradigm. *Future Microbiology*, 10(12), pp.2033–2043. Available at: http://www.futuremedicine.com/doi/10.2217/fmb.15.114.

Veronesi, F. et al., 2011. Prevalence of Anaplasma phagocytophilum in fallow deer (Dama dama) and feeding ticks from an Italy preserve. *Research in Veterinary Science*, 90(1), pp.40–43. Available at: http://dx.doi.org/10.1016/j.rvsc.2010.05.019.

Villar, M. et al., 2010. Expression of Heat Shock and Other Stress Response Proteins in Ticks and Cultured Tick Cells in Response to *Anaplasma* spp. Infection and Heat Shock. *International*

Journal of Proteomics, 2010, pp.1–11. Available at: http://www.hindawi.com/journals/ijpro/2010/657261/.

Villar, M. et al., 2015a. Identification and Characterization of Anaplasma phagocytophilum Proteins Involved in Infection of the Tick Vector, Ixodes scapularis. *Plos One*, 10(9), p.e0137237. Available at: http://dx.plos.org/10.1371/journal.pone.0137237.

Villar, M. et al., 2015b. Integrated Metabolomics, Transcriptomics and Proteomics Identifies Metabolic Pathways Affected by Anaplasma phagocytophilum Infection in Tick Cells. *MCP Papers in Press. Published on*, pp.3105–3117.

Vollmer, S.A. et al., 2013. Spatial spread and demographic expansion of Lyme borreliosis spirochaetes in Eurasia. *Infection, Genetics and Evolution*, 14(1), pp.147–155. Available at: http://dx.doi.org/10.1016/j.meegid.2012.11.014.

Walker, A.R., 2001. Age structure of a population of Ixodes ricinus (Acari: Ixodidae) in relation to its seasonal questing. *Bulletin of Entomological Research*, 91(MARCH 2001), pp.69–78.

Walker, A.R. et al., 2003. *Ticks of Domestic Animals in Africa: a Guide to Identification of Species*, Available at: http://www.alanrwalker.com/assets/PDF/tickguide-africa.pdf.

Walker, T. et al., 2011. The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations. *Nature*, 476(7361), pp.450–455. Available at: http://dx.doi.org/10.1038/nature10355.

Welinder-Olsson, C. et al., 2010. First Case of Human "Candidatus Neoehrlichia mikurensis" Infection in a Febrile Patient with Chronic Lymphocytic Leukemia. *Journal of Clinical Microbiology*, 48(5), pp.1956–1959. Available at: http://jcm.asm.org/cgi/doi/10.1128/JCM.02423-09.

Weller, S.J. et al., 1998. Phylogenetic placement of rickettsiae from the ticks Amblyomma americanum and ixodes scapularis. *Journal of Clinical Microbiology*, 36(5), pp.1305–1317.

Wenneras, C., 2015. Infections with the tick-borne bacterium Candidatus Neoehrlichia mikurensis. *Clinical Microbiology and Infection*, 21(7), pp.621–630. Available at: http://dx.doi.org/10.1016/j.cmi.2015.02.030.

Woldehiwet, Z., 2006. Anaplasma phagocytophilum in ruminants in Europe. *Annals of the New York Academy of Sciences*, 1078, pp.446–460.

Woldehiwet, Z. et al., 2002. Cultivation of an ovine strain of Ehrlichia phagocytophila in tick cell cultures. *Journal of Comparative Pathology*, 127(2–3), pp.142–149.

Woldehiwet, Z. & & Scott, G.R., 1993. Tick-borne (pasture) fever: Rickettsial and chlamydial diseases of domestic animals. *Oxford: Pergamon Press*, pp.233–254.

Xu, G. et al., 2003. Molecular phylogenetic analyses indicate that the Ixodes ricinus complex is a paraphyletic group. *Journal of Parasitology*, 89(3), pp.452–457. Available at: http://www.jstor.org/stable/3286032.

Yabsley, M.J. et al., 2008. Characterization of "Candidatus Neoehrlichia lotoris" (family Anaplasmataceae) from raccoons (Procyon lotor). *International Journal of Systematic and Evolutionary Microbiology*, 58(12), pp.2794–2798.

YONOW, T., 1995. The Life-cycle of Amblyomma variegatum (Atari : Ixodidae): a Literature Synthesis with a View to Modelling. *Parasitology*, 25(95), pp.1023–1060.

Yssouf, A. et al., 2011. Colonization of Grande Comore Island by a lineage of Rhipicephalus appendiculatus ticks. *Parasites & vectors*, 4(1), p.38.

Zele, F. et al., 2014. Wolbachia increases susceptibility to Plasmodium infection in a natural system. *Proceedings of the Royal Society B: Biological Sciences*, 281(1779), pp.20132837–20132837.

Available at: http://rspb.royalsocietypublishing.org/cgi/doi/10.1098/rspb.2013.2837.

Zhi, N. et al., 2002. Transcript Heterogeneity of the p44 Multigene Family in a Human Granulocytic Ehrlichiosis Agent Transmitted by Ticks. *Infection and Immunity*, 70(3), pp.1175–1184.

Zhi, N., Ohashi, N. & Rikihisa, Y., 1999. Multiple p44 genes encoding major outer membrane proteins are expressed in the human granulocytic ehrlichiosis agent. *The Journal of biological chemistry*, 274(25), pp.17828–36. Available at: http://www.ncbi.nlm.nih.gov/pubmed/10364227.

Zhong, J., Jasinskas, A. & Barbour, A.G., 2007. Antibiotic Treatment of the Tick Vector Amblyomma americanum Reduced Reproductive Fitness F. Romesberg, ed. *PLoS ONE*, 2(5), p.e405. Available at: http://dx.plos.org/10.1371/journal.pone.0000405.

Zhu, Z., Aeschlimann, A. & Gern, L., 1992. Rickettsia-like microorganisms in the ovarian primordia of molting Ixodes ricinus (Acari : Ixodidae) larvae and nymphs. *Annales de parasitologie humaine et comparée*, 67, pp.99–110.

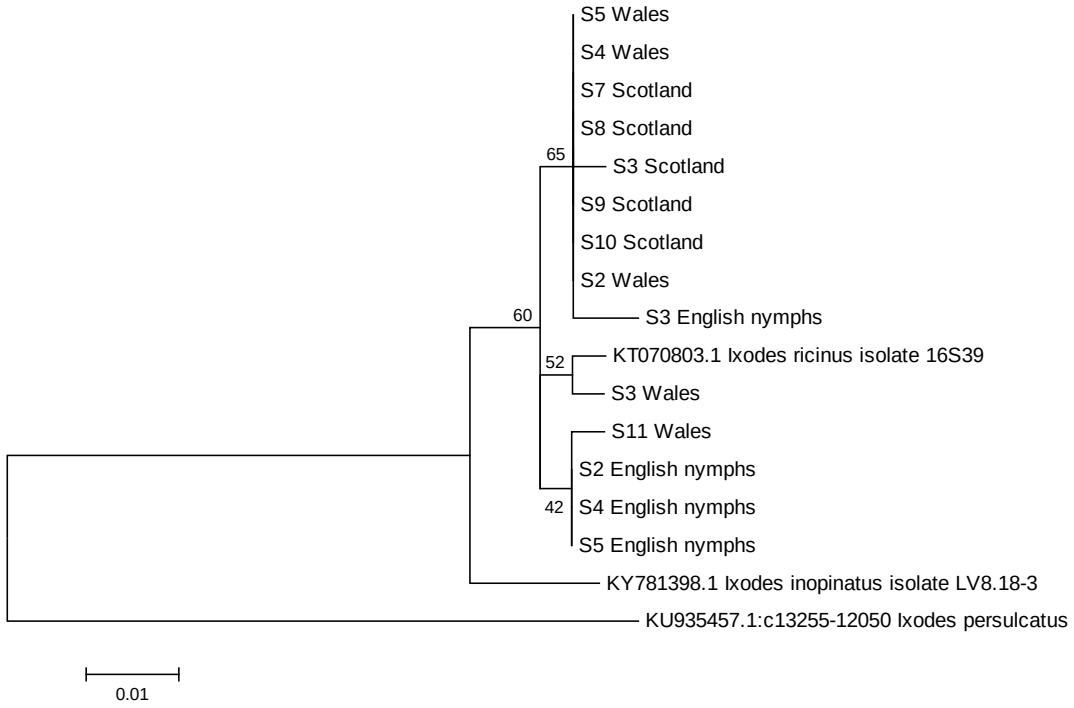
Zimmer, J.S.D. et al., 2006. Advances in proteomics data analysis and display using an accurate mass and time tag approach. *Mass Spectrometry Reviews*, 25(3), pp.450–482.

Zintl, A. et al., 2017. Ticks and Tick-borne diseases in Ireland. *Irish Veterinary Journal*, 70(1), pp.1–10. Available at: http://dx.doi.org/10.1186/s13620-017-0084-y.

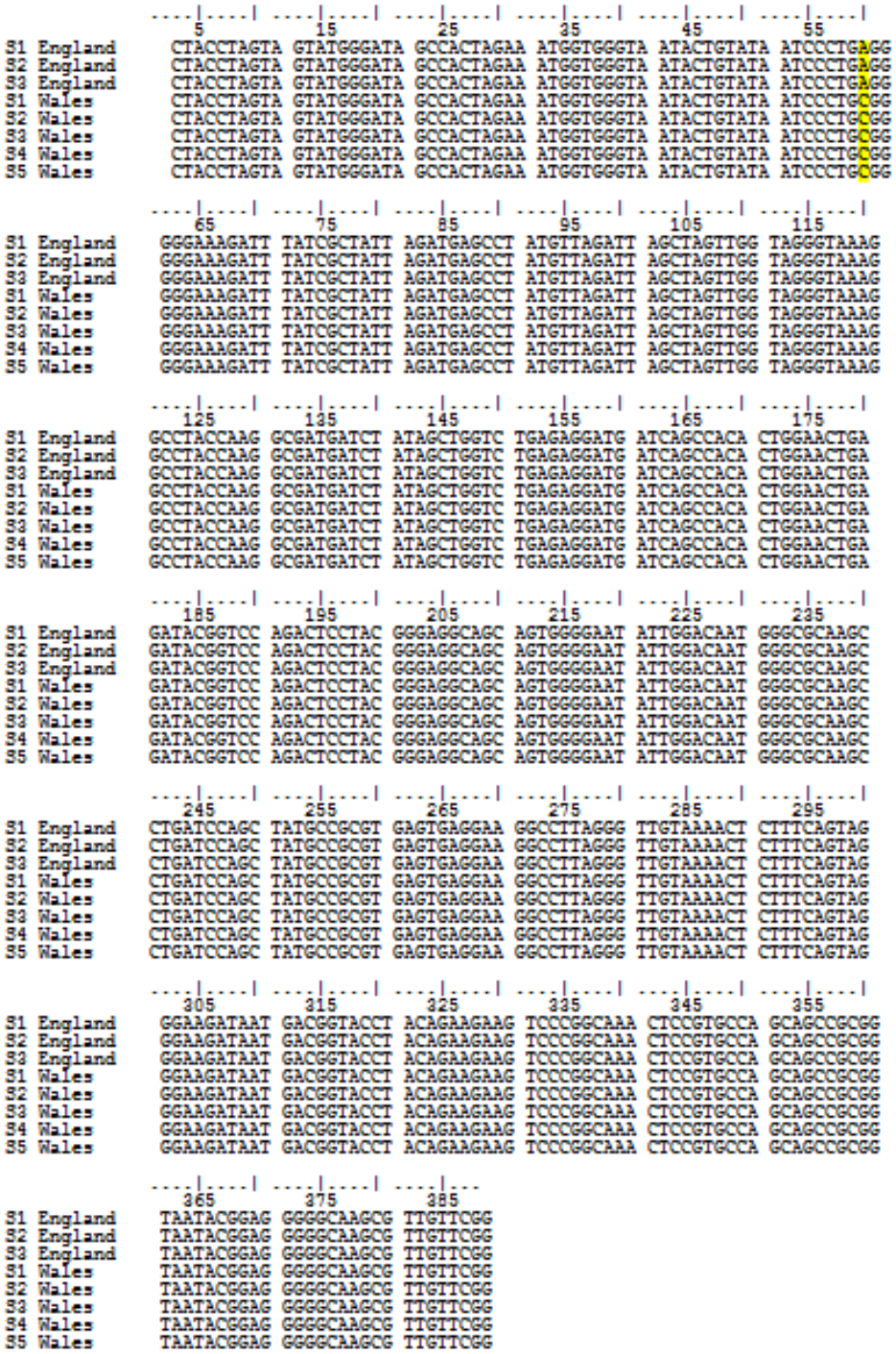
Zivkovic, Z. et al., 2009. Anaplasma phagocytophilum and Anaplasma marginale elicit different gene expression responses in cultured tick cells. *Comparative and Functional Genomics*.

Appendix

Appendix A: Maximum liklehood tree confirm *I. ricinus* species for tick samples used in the present study compared to *I. inopinatus* and *I. persulcatus*. (Bootstrapped 10,000 times) for the 16S rRNA gene.



Appendix B: Alignment figure of *A. phagocytophilum* 16S rRNA gene sequences from Welsh samples used in this study and English nymphs. One nucleotide difference is highlighted in yellow.



Appendix C: samples details, with *Ca*. Midichloria allelic arrangements (DNA), for 64 *Ixodes ricinus* ticks analysed as part of this study.

		MLST allele					
Sample (n.) and country name	Location	ST	ADK	LIPA	NUF2	PPDK	SECY
S10 - Scotland	Aberdeenshire	1	1	1	1	5	1
S9 - Scotland	Aberdeenshire	2	1	1	2	5	1
S1 - Scotland	Aberdeenshire	3	1	1	1	4	1
S10 - France	Chizé	4	1	1	4	4	1
S10 - Wales	powis castle	3	1	1	1	4	1
S11 - Scotland	Aberdeenshire	3	1	1	1	4	1
S12 - Scotland	Aberdeenshire	3	1	1	1	4	1
S2 - Scotland	Aberdeenshire	3	1	1	1	4	2
S3 - Scotland	Aberdeenshire	3	1	1	1	4	1
S4 - Scotland	Aberdeenshire	3	1	1	1	4	1
S5 - Scotland	Aberdeenshire	3	1	1	1	4	1
S6 - Scotland	Aberdeenshire	3	1	1	1	4	1
S7 - Swiss	Zermatt valley and Susten	3	1	1	1	4	1
S7 - Scotland	Aberdeenshire	3	1	1	1	4	1
S8 - Swiss	Zermatt valley and Susten	4	1	1	4	4	1
S8 - Wales	powis castle	3	1	1	1	4	1
S8 - Scotland	Aberdeenshire	3	1	1	1	4	1
S9 - France	Chizé	3	1	1	1	4	1
S4 - Germany	Bonn and le - Lennestadt-Meggen	5	1	1	4	3	1
S2 - Swiss	Zermatt valley and Susten	6	1	1	1	2	1
S1 - Germany	Bonn and le - Lennestadt-Meggen	7	1	1	1	1	1
S1 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S1 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S1 - Wales	powis castle	9	1	1	3	1	1
S10 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S10 - Swiss	Zermatt valley and Susten	9	1	1	3	1	1
S11 - France	Chizé	8	1	1	4	1	1
S11 - Germany	Bonn and le - Lennestadt-Meggen	8	1	1	4	1	1
S11 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S11 - Wales	powis castle	7	1	1	1	1	1
S12 - France	Chizé	7	1	1	1	1	1
S12 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S12 - Wales	powis castle	7	1	1	1	1	1
S13 - France	Chizé	8	1	1	4	1	1

S13 -Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S14 - France	Chizé	8	1	1	4	1	1
S16 - France	Chizé	8	1	1	4	1	1
S2 - France	Chizé	8	1	1	4	1	1
S2 - Germany	Bonn and le - Lennestadt-Meggen	10	2	1	4	1	1
S2 - Wales	powis castle	9	1	1	3	1	1
S3 - Germany	Bonn and le - Lennestadt-Meggen	8	1	1	4	1	1
S3 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S3 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S3 - Wales	powis castle	7	1	1	1	1	1
S4 - Italy	Appennino tosco-emiliano	11	3	1	4	1	1
S4 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S4 - Wales	powis castle	8	1	1	4	1	1
S5 - France	Chizé	8	1	1	4	1	1
S5 - Germany	Bonn and le - Lennestadt-Meggen	8	1	1	4	1	1
S5 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S5 - Wales	powis castle	7	1	1	1	1	1
S6 - France	Chizé	8	1	1	4	1	1
S6 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S6 - Swiss	Zermatt valley and Susten	8	1	1	4	1	1
S6 - Wales	powis castle	8	1	1	4	1	1
S6 - Germany	Bonn	7	1	1	1	1	1
S7 - France	Chizé	7	1	1	1	1	1
S7 - Italy	Appennino tosco-emiliano	7	1	1	1	1	1
S7 - Wales	powis castle	8	1	1	4	1	1
S8 - Germany	Bonn and le - Lennestadt-Meggen	8	1	1	4	1	1
S8 - Italy	Appennino tosco-emiliano	8	1	1	4	1	1
S9 - Germany	Bonn and le - Lennestadt-Meggen	8	1	1	4	1	1
S9 - Swiss	Zermatt valley and Susten	7	1	1	1	1	1
S9 - Wales	powis castle	8	1	1	4	1	1

Appendix D: Sample details, with tick mitochondrial allelic arrangements (DNA), for 64 *Ixodes ricinus* ticks analysed as part of this study. (Red type refers to those that were seen in University of Bath sequences only).

sample(n.) and country name	Location	ST	ATP	COI	COII	COIII	12S	CYTB
S12 - Scotland	Aberdeenshire	447	99	130	9	113	153	145
S11 - Scotland	Aberdeenshire	448	100	132	9	112	154	140
S10 - Scotland	Aberdeenshire	449	101	128	67	111	155	139
S9 - Scotland	Aberdeenshire	450	109	123	100	110	156	141
S8 - Scotland	Aberdeenshire	451	102	129	9	109	157	143
S1 - Scotland	Aberdeenshire	452	106	9	9	9	158	138
S3 - Scotland	Aberdeenshire	453	104	9	101	9	158	138
S4 - Scotland	Aberdeenshire	454	107	127	9	9	158	138
S6 - Scotland	Aberdeenshire	455	103	131	9	9	158	138
S2 - Scotland	Aberdeenshire	456	107	9	69	25	159	138
S5 - Scotland	Aberdeenshire	457	127	124	9	25	160	144
S7 - Scotland	Aberdeenshire	458	108	125	99	25	161	142
S9 - France	Chizé	459	8	126	9	114	162	3
S10 - France	Chizé	460	4	2	2	119	162	3
S2 - Wales	Powis castle	461	110	9	9	9	162	138
S3 - Wales	Powis castle	462	110	118	9	9	162	138
S10 - Wales	Powis castle	463	110	124	98	25	162	138
S11 - Wales	Powis castle	464	8	9	9	9	162	138
S12 - Wales	Powis castle	465	105	9	9	9	162	138
S1 - Germany	Bonn and le - Lennestadt-Meggen	466	13	109	103	131	163	3
S7 - Italy	Appennino tosco-emiliano	467	13	109	105	129	163	118
S7 - France	Chizé	468	13	2	2	131	163	132
S10 - Switzerland	Zermatt valley and Susten	469	111	105	2	11	163	133
S9 - Switzerland	Zermatt valley and Susten	470	124	107	104	126	163	135
S1 - Wales	Powis castle	471	112	110	2	115	163	136
S7 - Switzerland	Zermatt valley and Susten	472	110	124	2	128	163	137
S4 - Wales	Powis castle	473	110	121	92	115	163	138
S5 - Wales	Powis castle	474	125	108	93	115	163	150
S5 - Italy	Appennino tosco-emiliano	475	13	35	102	130	163	152
S6 - Germany	Bonn and le - Lennestadt-Meggen	476	110	109	41	118	163	154
S4 - Switzerland	Zermatt valley and Susten	477	4	113	3	2	164	129

S2 - France	Chizé	478	4	2	2	2	165	3
S11 - Germany	Bonn and le - Lennestadt-Meggen	478	4	2	2	2	165	3
S6 - France	Chizé	479	4	111	3	2	165	3
S11 - France	Chizé	480	4	109	2	2	165	3
S14 - France	Chizé	481	4	15	97	121	165	3
S3 - Germany	Bonn and le - Lennestadt-Meggen	482	4	118	2	2	165	3
S4 - Germany	Bonn and le - Lennestadt-Meggen	483	4	2	39	118	165	3
S7 - Wales	Powis castle	484	126	117	94	116	165	112
S12 - Switzerland	Zermatt valley and Susten	485	113	104	2	123	165	114
S9 - Wales	Powis castle	486	117	119	95	116	165	115
S2 - Switzerland	Zermatt valley and Susten	487	120	106	6	118	165	116
S5 - France	Chizé	488	4	2	65	2	165	120
S13 -Italy	Appennino tosco-emiliano	489	4	2	2	2	165	121
S6 - Italy	Appennino tosco-emiliano	490	4	2	65	2	165	123
S2 - Germany	Bonn and le - Lennestadt-Meggen	491	4	2	2	2	165	124
S8 - Italy	Appennino tosco-emiliano	492	4	2	2	2	165	125
S5 - Germany	Bonn and le - Lennestadt-Meggen	493	4	2	40	2	165	126
S4 - Italy	Appennino tosco-emiliano	494	119	2	65	2	165	127
S12 - France	Chizé	495	110	109	96	122	165	128
S1 - Switzerland	Zermatt valley and Susten	496	122	122	2	2	165	130
S1 - Italy	Appennino tosco-emiliano	497	4	2	2	120	165	131
S11 - Switzerland	Zermatt valley and Susten	498	116	116	2	125	165	134
S16 - France	chize’	499	121	35	2	117	165	146
S8 - Switzerland	Zermatt valley and Susten	500	123	115	2	127	165	147
S6 - Wales	Powis castle	501	114	120	106	124	165	148
S8 - Wales	Powis castle	502	110	9	9	25	165	149
S9 - Germany	Bonn and le - Lennestadt-Meggen	503	4	2	91	2	165	153

S10 - Italy	Appennino tosco-emiliano	504	4	2	105	2	166	119
S13 - France	Chizé	505	4	2	32	2	167	146
S8 - Germany	Bonn and le - Lennestadt-Meggen	506	4	2	2	2	168	122
S3 - Italy	Appennino tosco-emiliano	507	4	2	51	2	169	117
S3 - Switzerland	Zermatt valley and Susten	508	118	112	2	2	170	113
S6 - Switzerland	Zermatt valley and Susten	509	115	114	2	2	171	151

Appendix E: Preparation of basal L-15B medium (Munderloh and Kurtti, 1989):

1) Make up trace mineral stock solutions A, B and C, and use to make up stock solution D. Make up vitamin stock solution. All ingredients obtained from Sigma.Ingredients for the stock solutions should be dissolved in distilled water in the order listed, and the volume of each stock solution brought to 100ml. Store aliquots of stock solution D and the vitamin stock in 1ml amounts at –20°C. Stock solutions A, B and C can be stored in universals at –20°C.

Trace mineral stock solution D

Ingredient	Weight (mg/100ml)
Stock solution A	
$CoCl \cdot 6H_2O$	20
$CuSO_4 \cdot 5H_2O$	20
$MnSO_4 \cdot H_2O$	160
$Zn\ SO_4 \cdot 7H_2O$	200
Stock solution B	
$NaMoO_4 \cdot 2H_2O$	20
Stock solution C	
Na_2SeO_3	20
Stock solution D	
Glutathione (reduced)	1000
Ascorbic acid	1000
$FeSO_4 \cdot 7H_2O$	50
Stock solution A	1ml
Stock solution B	1ml
Stock solution C	1ml

Vitamin stock

p-aminobenzoic acid	100mg
Cyanocobalamine (B_{12})	50
d- Biotin	10

2) Dissolve L-15 powder for 1 litre medium (Invitrogen/Life Technologies Cat No.41300-021) in ~900ml distilled water then add the following components:

Aspartic acid	299mg
Glutamic acid	500mg
Proline	300mg
α-ketoglutaric acid	299mg
D-glucose	2239mg
Mineral stock D	1 ml
Vitamin stock	1 ml

Bring the volume to 1 litre and sterilise by filtration (0.22μm). Store at 4°C for up to 4 months, or at –20°C.

Appendix F: Sequences unique to British *Anaplasma* strain

		Max score	Total score	Query cover	E value	Ident	Accession
>unique/665/1/Org1_Gene1160	**hypothetical protein APHNP_0926 [Anaplasma phagocytophilum str. ApNP]**	182	182	46%	2.00E-60	99%	KJV67757.1
>unique/679/1/Org1_Gene576	**hypothetical protein ANAPRD1_00610 [Anaplasma phagocytophilum]**	111	111	62%	4.00E-32	68%	SCV64364.1
>unique/683/1/Org1_Gene430	**hypothetical protein ANAPH2_00810 [Anaplasma phagocytophilum]**	377	377	100%	3.00E-136	100%	SCV64001.1
>unique/686/1/Org1_Gene121	No significant similarity found						
>unique/721/1/Org1_Gene204	**hypothetical protein ANAPC5_00123 [Anaplasma phagocytophilum]**	365	365	100%	7.00E-132	100%	SCV61940.1
>unique/725/1/Org1_Gene335	**hypothetical protein APHDU1_0110 [Anaplasma phagocytophilum]**	230	230	80%	5.00E-79	82%	KJZ99721.1
>unique/766/1/Org1_Gene694	**hypothetical protein APHCRT_0254 [Anaplasma phagocytophilum str. CRT53-1]**	27.7	27.7	7%	0.5	92%	KJV88063.1
>unique/807/1/Org1_Gene364	No significant similarity found						
>unique/819/1/Org1_Gene1138	**hypothetical protein ANAPC3_01387 [Anaplasma phagocytophilum]**	298	298	100%	5.00E-106	98%	SBO33860.1
>unique/830/1/Org1_Gene888	No significant similarity found						
>unique/843/1/Org1_Gene1038	**hypothetical protein APHNP_1517 [Anaplasma phagocytophilum str. ApNP]**	49.7	49.7	31%	3.00E-08	54%	KJV67406.1
>unique/848/1/Org1_Gene736	No significant similarity found						
>unique/856/1/Org1_Gene1017	**hypothetical protein ANAPC3_00741 [Anaplasma phagocytophilum]**	265	265	100%	2.00E-93	97%	SBO32078.1

>unique/857/1/Org1_Gene837	No significant similarity found						
>unique/859/1/Org1_Gene1118	**hypothetical protein ANAPC4_01400 [Anaplasma phagocytophilum**	259	259	94%	4.00E-91	96%	SBO33963.1
>unique/867/1/Org1_Gene431	**hypothetical protein [Anaplasma phagocytophilum]**	24.6	24.6	22%	6.5	41%	WP_064660561.1
>unique/872/1/Org1_Gene224	**P44/Msp2 family outer membrane protein [Anaplasma phagocytophilum**	200	200	92%	3.00E-65	76%	WP_081250614.
>unique/886/1/Org1_Gene158	**hypothetical protein ANAPC3_00403 [Anaplasma phagocytophilum]**	260	260	100%	6.00E-88	98%	SBO31065.1
>unique/888/1/Org1_Gene1018	**hypothetical protein ANAPC1_00779 [Anaplasma phagocytophilum**	265	265	100%	9.00E-94	99%	SBO14425.1
>unique/897/1/Org1_Gene1213	No significant similarity found						
>unique/907/1/Org1_Gene412	No significant similarity found						
>unique/910/1/Org1_Gene558	No significant similarity found						
>unique/917/1/Org1_Gene723	16S rRNA (cytosine(1402)-N(4))-methyltransferase RsmH [Anaplasma phagocytophilum]	27.3	27.3	25%	1.3	39%	WP_045890236.1
>unique/942/1/Org1_Gene118	**hypothetical protein ANAPH2_01098 [Anaplasma phagocytophilum**	233	233	100%	4.00E-81	98%	SCV64880.1
>unique/948/1/Org1_Gene785	No significant similarity found						
>unique/977/1/Org1_Gene792	No significant similarity found						
>unique/993/1/Org1_Gene225	**p44-29 [Anaplasma phagocytophilum**	75.1	75.1	74%	9.00E-19	52%	AGV04700.1
>unique/1025/1/Org1_Gene844	No significant similarity found						
>unique/1036/1/Org1_Gene399	**hypothetical protein ANAPC1_00970 [Anaplasma phagocytophilum]**	207	207	100%	7.00E-72	96%	SBO14610.1
>unique/1066/1/Org1_Gene1134	**hypothetical protein ANAPH2_01577 [Anaplasma phagocytophilum]**	169	169	100%	3.00E-55	79%	SCV66407.1

>unique/1159/1/Org1_Gene979	No significant similarity found						
>unique/1198/1/Org1_Gene245	**P44 outermembrane protein, silent [Anaplasma phagocytophilum]**	64.3	64.3	63%	2.00E-15	67%	CEH11165.1
>unique/1207/1/Org1_Gene1125	**hypothetical protein [Anaplasma phagocytophilum]**	30.8	30.8	22%	0.008	60%	WP_020848839.1
>unique/1208/1/Org1_Gene768	No significant similarity found						
>unique/1213/1/Org1_Gene1107	**hypothetical protein ANAPH2_00394 [Anaplasma phagocytophilum]**	57.4	57.4	71%	4.00E-13	68%	SCV62708.1
>unique/1281/1/Org1_Gene1149	No significant similarity found						
>unique/1283/1/Org1_Gene141	No significant similarity found						
>unique/1297/1/Org1_Gene334	magnesium transporter [Anaplasma phagocytophilum]	24.6	24.6	38%	4.3	30%	WP_044152415.1
>unique/1320/1/Org1_Gene575	**hypothetical protein ANAPRD1_00610 [Anaplasma phagocytophilum]**	71.6	71.6	93%	4.00E-18	68%	SCV64364.1
>unique/1338/1/Org1_Gene146	**hypothetical protein ANAPH2_01074 [Anaplasma phagocytophilum]**	100	100	63%	2.00E-30	98%	SCV64809.1
>unique/1347/1/Org1_Gene921	No significant similarity found						
>unique/1355/1/Org1_Gene569	**hypothetical protein ANAPH2_01061 [Anaplasma phagocytophilum]**	156	156	100%	2.00E-52	96%	SCV64765.1
>unique/1360/1/Org1_Gene1013	**hypothetical protein [Anaplasma phagocytophilum]**	45.1	45.1	54%	4.00E-04	53%	WP_064103874.1
>unique/1380/1/Org1_Gene134	**hypothetical protein [Anaplasma phagocytophilum]**	95.1	95.1	100%	4.00E-24	64%	WP_060757582.1
>unique/1383/1/Org1_Gene207	**hypothetical protein APHNP_1374 [Anaplasma phagocytophilum str. ApNP]**	93.2	141	94%	1.00E-27	92%	KJV66566.1
>unique/1425/1/Org1_Gene1185	No significant similarity found						

>unique/1453/1/Org1_Gene316	50S ribosomal protein L25 [Anaplasma phagocytophilum]	113	113	78%	2.00E-31	96%	SCV61900.1
>unique/1486/1/Org1_Gene591	No significant similarity found	53.5	53.5	70%	2.00E-07	57%	PAV67337.1
>unique/1511/1/Org1_Gene561	**hypothetical protein APHNP_1279 [Anaplasma phagocytophilum str. ApNP**	111	111	80%	6.00E-35	95%	KJV67961.1
>unique/1554/1/Org1_Gene669	**hypothetical protein P029_02695 [Anaplasma phagocytophilum str. Norway variant2**	87	87	86%	3.00E-25	76%	ANC34276.1
>unique/1566/1/Org1_Gene834	**hypothetical protein ANAPH2_01537 [Anaplasma phagocytophilum]**	133	133	97%	1.00E-39	100%	SCV66264.1
>unique/1567/1/Org1_Gene242	**hypothetical protein ANAPC1_01230 [Anaplasma phagocytophilum]**	112	112	100%	5.00E-35	92%	SBO14858.1
>unique/1568/1/Org1_Gene451	**hypothetical protein ANAPC1_01033 [Anaplasma phagocytophilum]**	127	127	100%	3.00E-41	91%	SBO14671.1
>unique/1570/1/Org1_Gene783	**hypothetical protein [Anaplasma phagocytophilum]**	24.6	24.6	53%	1.7	39%	WP_064670201.1
>unique/1616/1/Org1_Gene155	**hypothetical protein ANAPC1_00717 [Anaplasma phagocytophilum]**	134	134	100%	6.00E-40	100%	SBO14366.1
>unique/1630/1/Org1_Gene111	**hypothetical protein P030_03835 [Anaplasma phagocytophilum str. CRT35]**	81.6	81.6	71%	6.00E-22	83%	KDB57231.1
>unique/1654/1/Org1_Gene1127	No significant similarity found	34.7	34.7	93%	4.7	32%	WP_081892947.1
>unique/1663/1/Org1_Gene993	**hypothetical protein APHNP_0073 [Anaplasma phagocytophilum str. ApNP]**	117	117	93%	3.00E-33	98%	KJV66745.1
>unique/1664/1/Org1_Gene574	**hypothetical protein ANAPRD1_00610 [Anaplasma phagocytophilum]**	33.5	33.5	41%	8	73%	SCV64364.1

>unique/1671/1/Org1_Gene228	**hypothetical protein APHWEB_1397 [Anaplasma phagocytophilum str. Webster]**	60.8	60.8	59%	4.00E-15	81%	KJV59828.1
>unique/1706/1/Org1_Gene935	No significant similarity found						
>unique/1707/1/Org1_Gene677	Protein of unknown function [Anaplasma phagocytophilum	89.4	89.4	75%	2.00E-22	98%	CEG20702.1
>unique/1727/1/Org1_Gene676	No significant similarity found	34.3	34.3	61%	6.2	43%	WP_022869024.1
>unique/1739/1/Org1_Gene96	**hypothetical protein APHNP_1203 [Anaplasma phagocytophilum str. ApNP]**	57.8	57.8	47%	5.00E-10	93%	KJV67110.1
>unique/1746/1/Org1_Gene1053	No significant similarity found						
>unique/1750/1/Org1_Gene982	**hypothetical protein ANAPC1_01240 [Anaplasma phagocytophilum**	123	123	100%	9.00E-40	100%	SBO14868.1
>unique/1766/1/Org1_Gene721	**hypothetical protein ANAPC2_00886 [Anaplasma phagocytophilum]**	115	115	100%	3.00E-32	98%	SBO32029.1
>unique/1773/1/Org1_Gene874	**hypothetical protein ANAPC1_00823 [Anaplasma phagocytophilum]**	114	114	100%	4.00E-32	93%	SBO14467.1
>unique/1780/1/Org1_Gene1090	putative lipoprotein [Anaplasma phagocytophilum str. HZ]	110	110	100%	6.00E-31	96%	ABD43871.1
>unique/1781/1/Org1_Gene953	**hypothetical protein APH_0413 [Anaplasma phagocytophilum str. HZ]**	112	112	100%	1.00E-31	100%	ABD44274.1
>unique/1782/1/Org1_Gene276	**hypothetical protein APHNP_1765 [Anaplasma phagocytophilum str. ApNP]**	114	114	100%	2.00E-32	100%	KJV66890.1
>unique/1785/1/Org1_Gene400	**hypothetical protein APHCRT_0557 [Anaplasma phagocytophilum str. CRT53-1]**	57	57	72%	1.00E-09	70%	KJV86109.1
>unique/1790/1/Org1_Gene458	Protein of unknown function [Anaplasma phagocytophilum]	63.9	63.9	74%	3.00E-16	80%	CEG20606.1

>unique/1797/1/Org1_Gene1215	**hypothetical protein ANAPH2_01037 [Anaplasma phagocytophilum]**	105	105	100%	2.00E-32	100%	SCV64700.1
>unique/1798/1/Org1_Gene484	**hypothetical protein ANAPC1_01270 [Anaplasma phagocytophilum]**	33.1	33.1	55%	2.4	63%	SBO14894.1
>unique/1801/1/Org1_Gene1119	**hypothetical protein CRT38_01757 [Anaplasma phagocytophilum str. CRT38]**	43.9	43.9	86%	3.00E-08	54%	EOA61955.1
>unique/1802/1/Org1_Gene47	putative membrane protein [Anaplasma phagocytophilum str. ApNP]	104	104	100%	2.00E-32	100%	KJV66472.1
>unique/1805/1/Org1_Gene715	**hypothetical protein ANAPC1_01302 [Anaplasma phagocytophilum]**	103	103	100%	4.00E-32	98%	SBO14926.1
>unique/1807/1/Org1_Gene686	No significant similarity found						
>unique/1808/1/Org1_Gene1224	**hypothetical protein APH_0570 [Anaplasma phagocytophilum str. HZ]**	60.5	60.5	50%	4.00E-11	100%	ABD44027.1
>unique/1812/1/Org1_Gene346	**hypothetical protein [Anaplasma phagocytophilum]**	84.7	84.7	82%	2.00E-20	93%	WP_064659998.
>unique/1815/1/Org1_Gene937	**hypothetical protein ANAPC1_00887 [Anaplasma phagocytophilum]**	100	100	100%	3.00E-27	100%	SBO14528.1
>unique/1817/1/Org1_Gene800	**hypothetical protein APH_0670 [Anaplasma phagocytophilum str. HZ]**	105	105	100%	5.00E-29	100%	ABD44145.1
>unique/1821/1/Org1_Gene1073	**putative p44 outer membrane protein, C-terminal [Anaplasma phagocytophilum str. NCH-1]**	102	102	100%	8.00E-28	98%	KJV59448.1
>unique/1823/1/Org1_Gene823	**hypothetical protein APHMUC_0555 [Anaplasma phagocytophilum str. ApMUC09**	105	105	100%	6.00E-29	100%	KJV64704.1
>unique/1826/1/Org1_Gene215	**hypothetical protein APHNP_0067 [Anaplasma phagocytophilum str. ApNP**	98.6	98.6	100%	3.00E-30	98%	KJV67611.1
>unique/1828/1/Org1_Gene722	No significant similarity found						

>unique/1829/1/Org1_Gene735	**hypothetical protein CRT38_01757 [Anaplasma phagocytophilum str. CRT38]**	35.8	35.8	77%	0.35	53%	EOA61955.1
>unique/1831/1/Org1_Gene525	No significant similarity found						
>unique/1832/1/Org1_Gene434	**P44 outermembrane protein, silent [Anaplasma phagocytophilum]**	91.3	91.3	100%	2.00E-27	94%	CEH11201.1
>unique/1837/1/Org1_Gene268	**hypothetical protein ANAPC3_00687 [Anaplasma phagocytophilum**	95.5	95.5	100%	3.00E-29	100%	SBO31921.1
>unique/1845/1/Org1_Gene156	Uncharacterized protein ANAPHAGO_00202 [Anaplasma phagocytophilum]	91.3	91.3	100%	2.00E-27	100%	CEG20354.1
>unique/1848/1/Org1_Gene444	**hypothetical protein APHNP_0850 [Anaplasma phagocytophilum str. ApNP]**	87.8	87.8	100%	3.00E-26	93%	KJV67695.
>unique/1851/1/Org1_Gene101	**hypothetical protein ANAPC3_00851 [Anaplasma phagocytophilum]**	70.9	70.9	86%	2.00E-19	87%	SBO32401.1
>unique/1855/1/Org1_Gene922	**hypothetical protein APH_0992 [Anaplasma phagocytophilum str. HZ]**	82	82	100%	7.00E-24	93%	ABD44084.1
>unique/1857/1/Org1_Gene755	**hypothetical protein ANAPC1_00918 [Anaplasma phagocytophilum]**	91.3	91.3	100%	2.00E-27	100%	SBO14559.1
>unique/1860/1/Org1_Gene456	**hypothetical protein ANAPC1_01334 [Anaplasma phagocytophilum**	85.1	85.1	100%	4.00E-25	90%	SBO14956.1
>unique/1861/1/Org1_Gene349	**hypothetical protein [Anaplasma phagocytophilum]**	37.4	37.4	69%	0.06	63%	WP_020849022.1
>unique/1862/1/Org1_Gene269	**hypothetical protein ANAPC3_00686 [Anaplasma phagocytophilum]**	84.3	84.3	100%	9.00E-25	100%	SBO31919.1
>unique/1864/1/Org1_Gene668	Protein of unknown function [Anaplasma phagocytophilum	88.2	88.2	100%	2.00E-26	100%	CEG20478.1

>unique/1865/1/Org1_Gene699	**hypothetical protein APH_0954 [Anaplasma phagocytophilum str. HZ]**	78.2	78.2	100%	2.00E-22	93%	ABD44219.1
>unique/1870/1/Org1_Gene307	**hypothetical protein APHMUC_0137 [Anaplasma phagocytophilum str. ApMUC09]**	87	87	100%	6.00E-26	100%	KJV64705.1
>unique/1872/1/Org1_Gene939	No significant similarity found						
>unique/1873/1/Org1_Gene429	**hypothetical protein ANAPC1_01127 [Anaplasma phagocytophilum]**	71.2	71.2	100%	9.00E-16	97%	SBO14760.1
>unique/1878/1/Org1_Gene404	**hypothetical protein APHWEB_0022 [Anaplasma phagocytophilum str. Webster]**	69.7	69.7	97%	3.00E-15	94%	KJV60412.1
>unique/1879/1/Org1_Gene1031	**hypothetical protein APHNP_1662 [Anaplasma phagocytophilum str. ApNP]**	76.6	76.6	100%	6.00E-18	100%	KJV67253.1
>unique/1880/1/Org1_Gene363	**hypothetical protein ANAPC1_01124 [Anaplasma phagocytophilum]**	64.3	64.3	100%	4.00E-13	89%	SBO14757.1
>unique/1886/1/Org1_Gene1070	**hypothetical protein APHNP_1556 [Anaplasma phagocytophilum str. ApNP]**	70.9	70.9	100%	1.00E-15	97%	KJV66377.1
>unique/1890/1/Org1_Gene1054	**hypothetical protein ANAPC1_00999 [Anaplasma phagocytophilum]**	67.8	67.8	100%	2.00E-18	100%	SBO14638.1
>unique/1891/1/Org1_Gene531	**hypothetical protein ANAPC3_00697 [Anaplasma phagocytophilum**	68.6	68.6	100%	7.00E-19	100%	SBO31956.1
>unique/1895/1/Org1_Gene742	Protein of unknown function [Anaplasma phagocytophilum]	68.2	68.2	100%	1.00E-14	100%	CEG20448.1
>unique/1896/1/Org1_Gene503	**hypothetical protein ANAPC1_01176 [Anaplasma phagocytophilum]**	60.5	60.5	100%	2.00E-15	94%	SBO14807.1
>unique/1897/1/Org1_Gene897	**hypothetical protein APHWEB_1109 [Anaplasma phagocytophilum str. Webster]**	67.8	67.8	100%	2.00E-18	94%	KJV59606.1

>unique/1898/1/Org1_Gene178	**hypothetical protein ANAPH2_00805 [Anaplasma phagocytophilum**	65.1	65.1	100%	2.00E-13	97%	SCV63985.1
>unique/1899/1/Org1_Gene280	**hypothetical protein ANAPC3_00453 [Anaplasma phagocytophilum]**	67.4	67.4	100%	2.00E-18	97%	SBO31204.1
>unique/1900/1/Org1_Gene1099	**hypothetical protein APH_0967 [Anaplasma phagocytophilum str. HZ]**	98.2	98.2	100%	2.00E-28	100%	ABD43450.1
>unique/1901/1/Org1_Gene119	**hypothetical protein ANAPH2_00667 [Anaplasma phagocytophilum]**	99	99	100%	8.00E-29	100%	SCV63539.1
>unique/1902/1/Org1_Gene333	Protein of unknown function [Anaplasma phagocytophilum]	96.9	96.9	100%	4.00E-28	100%	CEG20738.1
>unique/1903/1/Org1_Gene556	**hypothetical protein ANAPC3_00737 [Anaplasma phagocytophilum]**	96.1	96.1	100%	9.00E-28	100%	SBO32071.1
>unique/1904/1/Org1_Gene577	gram-negative porin family protein [Anaplasma phagocytophilum str. ApNP]	22.7	40.3	62%	10	63%	KJV66605.1
>unique/1905/1/Org1_Gene227	**hypothetical protein ANAPH2_00463 [Anaplasma phagocytophilum]**	98.2	98.2	100%	1.00E-28	100%	SCV62917.1
>unique/1907/1/Org1_Gene436	**hypothetical protein ANAPC3_00788 [Anaplasma phagocytophilum]**	64.7	64.7	79%	1.00E-14	91%	SBO32216.1
>unique/1908/1/Org1_Gene226	sigma-54 interaction domain protein [Anaplasma phagocytophilum str. ApMUC09]	24	24	48%	3.7	52%	KJV64411.1

* (Hypothetical protein [Anaplasma phagocytophilum] in bold; p44 outer membrane proteins highlighted in yellow).

Appendix G: Genes exclusively present in other *A. phagocytophilum* strains - absent in British OSstrain.

	Gene product	*Max score*	*Total score*	*Query cover*	*E value*	*Ident*	*Accession*	*length*
>Org2_Gene539	aconitate hydratase AcnA [Anaplasma phagocytophilum]	1802	1802	100%	0	100%	WP_011451171.1	873
>Org2_Gene412	**hypothetical protein [Anaplasma phagocytophilum**	1748	1748	100%	0	100%	WP_011450904.1	845
>Org2_Gene731	**hypothetical protein [Anaplasma phagocytophilum]**	1372	1372	100%	0	100%	WP_020848823.1	670
	ComEC/Rec2-related protein [Anaplasma phagocytophilum str.HZ	1372	1372	100%	0	100%	ABD43303.1	
>Org2_Gene232	**hypothetical protein YYU_01980 [Anaplasma phagocytophilum str. HZ2]**	1138	1138	100%	0	99%	AGR78778.1	559
>Org2_Gene276	magnesium transporter [Anaplasma phagocytophilum	902	902	100%	0	100%	WP_011450616.1	448
>Org2_Gene438	tRNA (N6-isopentenyl adenosine(37)-C2)-methylthiotransferase MiaB [Anaplasma phagocytophilum	922	922	100%	0	100%	WP_011450970.1	444
>Org2_Gene383	bifunctional folylpolyglutamate synthase/dihydrofolate synthase [Anaplasma phagocytophilum]	887	887	100%	0	100%	WP_011450847.1	429
>Org2_Gene140	AFG1 family ATPase [Anaplasma phagocytophilum str. HGE1]	687	687	100%	0	100%	EOA61840.1	331
>Org2_Gene184	FAD-dependent oxidoreductase [Anaplasma phagocytophilum	721	721	100%	0	100%	WP_011450458.1	352
>Org2_Gene491	quinone-dependent dihydroorotate dehydrogenase [Anaplasma phagocytophilum]	709	709	100%	0	100%	WP_011451087.1	350

>Org2_Gene612	DNA polymerase III subunit delta [Anaplasma phagocytophilum	694	694	100%	0	100%	WP_011451282.1	338
>Org2_Gene108	glycerol-3-phosphate dehydrogenase (NAD(P)(+)) [Anaplasma phagocytophilum]	642	642	100%	0	100%	WP_011450345.1	324
>Org2_Gene278	aspartate carbamoyltransferase [Anaplasma phagocytophilum	615	615	100%	0	100%	WP_011450620.1	300
>Org2_Gene947	RecName: Full=Protoheme IX farnesyltransferase; AltName: Full=Heme B farnesyltransferase; AltName: Full=Heme O synthase	589	589	100%	0	100%	Q2GJ17.1	293
>Org2_Gene476	zinc transporter [Anaplasma phagocytophilum]	592	592	100%	0	100%	WP_011451058.1	286
>Org2_Gene492	type III pantothenate kinase [Anaplasma phagocytophilum]	539	539	100%	0	100%	WP_011451088.1	263
>Org2_Gene390	methyltransferase domain-containing protein [Anaplasma phagocytophilum]	531	531	100%	0	100%	WP_011450866.1	261
>Org2_Gene843	**hypothetical protein [Anaplasma phagocytophilum]**	377	377	100%	3.00E-132	100%	WP_020849043.1	188
>Org2_Gene436	biotin--[acetyl-CoA-carboxylase] ligase [Anaplasma phagocytophilum]	493	493	100%	1.00E-180	100%	WP_011450966.1	241
>Org2_Gene277	tRNA pseudouridine(38-40) synthase TruA [Anaplasma phagocytophilum]	512	512	100%	0	100%	WP_011450619.1	246
>Org2_Gene525	outer membrane protein assembly factor BamD [Anaplasma phagocytophilum]	469	469	100%	4.00E-167	100%	WP_011451146.1	233
>Org2_Gene489	phospholipase/carboxylesterase family protein [Anaplasma phagocytophilum str. HGE1]	458	458	100%	4.00E-167	100%	EOA61251.1	220
>Org2_Gene244	**hypothetical protein [Anaplasma phagocytophilum]**	**420**	420	100%	1.00E-152	100%	WP_011450556.1	207

>Org2_Gene289	orotidine 5'-phosphate decarboxylase [Anaplasma phagocytophilum str. NCH-1]	458	458	100%	7.00E-167	100%	KJV65789.1	228
>Org2_Gene281	Smr domain-containing protein [Anaplasma phagocytophilum]	474	474	100%	1.00E-173	100%	WP_011450630.	228
>Org2_Gene557	cytochrome c oxidase assembly protein [Anaplasma phagocytophilum]	378	378	100%	7.00E-137	100%	WP_011451199.1	183
>Org2_Gene282	DNA-directed RNA polymerase subunit omega [Anaplasma phagocytophilum]	273	273	100%	7.00E-97	100%	WP_011450631.1	135
>Org2_Gene285	NADH:ubiquinone oxidoreductase subunit A [Anaplasma phagocytophilum]	236	236	100%	7.00E-83	100%	WP_011450636.1	120
>Org2_Gene283	Expression regulator ApxR [Anaplasma phagocytophilum]	219	219	100%	2.00E-76	100%	CEG21023.1	111
>Org2_Gene828	acyl carrier protein [Anaplasma phagocytophilum	208	208	100%	3.00E-72	100%	WP_020849015.1	104
>Org2_Gene60	50S ribosomal protein L31 [Anaplasma phagocytophilum]	153	153	100%	3.00E-51	100%	WP_011450244.1	73

* (Hypothetical protein [Anaplasma phagocytophilum] in bold,

Appendix H: The Msp2 sequences of OS strain (contains 115 sequences), the 23 shortest sequence fragments were omitted for clarity.