Evolution and adaptation of tick-borne pathogens across Eurasia

Robert Ethan Rollins

München 2021

Dissertation
der Fakultät für Biologie
der Ludwig-Maximilians-Universität München

vorgelegt von
ROBERT ETHAN ROLLINS
aus Salt Lake City, Utah, USA

München, den 09.08.2021

1. Gutachter: Dr. Noémie Becker
2. Gutachter: Prof. Dr. Silke Robatzek

Dissertation eingereicht am: August 9, 2021
Datum der Disputation: Oktober 28, 2021

Erklärung:

Diese Dissertation wurde im Sinne von § 12 der Promotionsordnung von Dr. Noémie Becker betreut. Ich erkläre hiermit, dass die Dissertation nicht einer anderen Prüfungskommission vorgelegt worden ist und dass ich mich nicht anderweitig einer Doktorprüfung ohne Erfolg unterzogen habe.

Eidesstattliche Erklärung:

Ich versichere hiermit an Eid statt, dass die vorgelegte Dissertation von mir selbständig und ohne unerlaubte Hilfe angefertigt wurde.

München, 09.08.2021

_____

Robert Ethan Rollins

Table of Contents

Declaration of Authors' Contributions

In this dissertation, I present work that was performed and realized with other scientific collaborators during my doctoral studies from October 2018 to July 2021. This work is separated into six individual Papers which represent four published manuscripts (Papers 1-4) and two unpublished manuscripts (Papers 5-6).

Paper 1: Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry

Robert E. Rollins, Sabine Schaper, Claudia Kahlhofer, Dimitrios Frangoulidis, Aurelia F. T. Strauß, Massimiliano Cardinale, Andrea Springer, Christina Strube, Deon K. Bakkes, Noémie S. Becker, Lidia Chitimia-Dobler

I conceived the project with Aurelia F.T. Strauß under the supervision of Noémie S. Becker. Bird collection, identification, and tick collection were performed by Aurelia F.T. Strauß and Massimilianao Cardinale. Lidia Chitimia-Dobler identified collected ticks to species level and organized all further testing. Molecular and in-silico testing were done by me, Lidia Chitimia-Dobler, Sabine Schaper, Claudia Kahlhofer, Dimitrios Frangoulidis, Andrea Springer, Christina Strube, Deon K. Bakkes. I wrote the manuscript with Noémie S. Becker and Lidia Chitimia-Dobler with input from all co-authors. I and Noémie S. Becker read and approved the final manuscript which was published in:

ROLLINS, R.E., et. al. (2021). Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry. Ticks and Tick-borne Diseases. 12(1): 101590

Paper 2: Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010

Mercy Okeyo, Sabrina Hepner, **Robert E. Rollins**, Christina Hartberger, Reinhard K. Straubinger, Durdica Marosevic, Stephanie A. Bannister, Antra Bormane, Michael Donaghy, Andreas Sing, Volker Fingerle and Gabriele Margos

The project was conceived by Gabriele Margos and Volker Fingerle as part of the doctoral project of Mercy Okeyo. Stephanie A. Bannister, Antra Bormane, and Michael Donaghy oversaw and performed all tick collections and tick identification. Mercy Okeyo, Sabrina Hepner, and Christina Hartberger performed all DNA extractions, *Borrelia* screening, and sequencing all multiple locus sequence typing loci for *Borrelia* positive samples. I designed and performed all statistical analyses for the project, outside of the GeoBURST analysis which was performed by Mercy Okeyo under the guidance of Gabriele Margos and Durdica Marosevic. Mercy Okeyo, Gabriele Margos, and Sabrina Hepner wrote the manuscript with input from me, Volker Fingerle, Reinhard K. Straubinger, and Andreas Sing. All co-authors read and approved the final manuscript which was published:

OKEYO, M., et. al. (2020). Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999-2010. Environmental Microbiology. 22:5033-5047.

Paper 3: Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany

Robert E. Rollins, Zehra Yeyin, Maja Wyczanska, Nikolas Alig, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Noémie S. Becker

Noémie S. Becker and I conceived the project with guidance from Volker Fingerle and Gabriele Margos. Zehra Yeyin, Maya Wyczanska, and I collected ticks and performed DNA extractions and screening ticks for *Borrelia* bacteria, including molecular genospecies identification with guidance from Sabrina Hepner. Niko Alig sequenced all 16S sequences for identifying tick species under guidance from me, for which I performed the median-joining network analysis. I performed all statistical analysis for the project. This work was the basis of Zehra Yeyin's Bachelor's thesis who I supervised with Noémie S. Becker. I wrote the manuscript with Noémie S. Becker and the input of all co-authors. Noémie S. Becker and I approved the final manuscript which was published:

ROLLINS, R.E., etal. (2021). Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany. Ticks and Tick-borne Diseases. 12 (1): 101589

Paper 4: High conservation combined with high plasticity: genomics and evolution of *Borrelia bavariensis*

Noémie S. Becker, **Robert E. Rollins**, Kateryna Nosenko, Alexander Paulus, Samantha Martin, Stefan Krebs, Ai Takano, Kozue Sato, Sergey Y. Kovalev, Hiroki Kawabata, Volker Fingerle and Gabriele Margos

Noémie S. Becker, Gabriele Margos, and Volker Fingerle designed the study. Ai Takano, Sergey Y. Kovalev, Hiroki Kawabata, Volker Fingerle and Gabriele Margos provided *Borrelia* samples which were cultured and processed by me, Kozue Sato, Hiroki Kawabata, Ai Takano, and Kateryna Nosenko. Sequencing was performed by Stefan Krebs, Gabriele Margos, and Volker Fingerle. Kateryna Nosenko developed the novel qPCR protocol which I then optimized and used to produce the data for the manuscript. Noémie S. Becker, Kateryna Nosenko, Alexander Paulus, and Samantha Martin analyzed all data. Noémie S. Becker wrote the manuscript with input from me, Kateryna Nosenko, and Gabriele Margos, which was then reviewed by all co-authors. Noémie S. Becker approved the final manuscript which was published:

BECKER, N.S., et. al. (2020) High conservation combined with high plasticity: Genomics and evolution of *Borrelia bavariensis*. BMC Genomics. 21:702.

Paper 5: Out of Asia? Recurrent vector switches leading to the expansion of Eurasian Lyme disease bacteria

Robert E. Rollins, Kozue Sato, Minoru Nakao, Mohammed T. Tawfeeq, Fernanda Herrera-Mesías, Ricardo J. Pereira, Sergey Kovalev, Gabriele Margos, Volker Fingerle, Hiroki Kawabata, and Noémie S. Becker

The study idea was designed by Noémie S. Becker, Gabriele Margos, and Volker Fingerle. I, along with Noémie S. Becker, Kozue Sato, Hiroki Kawabata, and Minoru Nakao performed all tick collections in Japan and then Noémie S. Becker, Kozue Sato, Hiroki Kawabata, and myself performed all bacterial isolations for Japanese samples. Sergey Kovalev performed all tick collection, *Borrelia* isolation, and DNA extraction for the Russian isolates. Ticks in Germany were collected by Mohammed T. Tawfeeq, Fernanda Herrera-Mesías and me who also all worked on isolating *Borrelia* isolates. Tick identification was performed by Hiroki Kawabata (Japan), Sergey Kovalev (Russia), and me (Japan and Germany). Sequence data and DNA for additional samples from Japan, Russia, and Europe were provided by Hiroki Kawabata, Sergey

Kovalev, Gabriele Margos, and Volker Fingerle. I performed all Illumina MiSeq sequencing of novel *Borrelia* isolates from Japan, Russia, and Germany. I also assembled all sequence data from our novel isolates and the additional *Borrelia afzelii* and *Borrelia garinii* isolates provided by collaborators. I performed all phylogenetic, population genetic, and statistical analyses with the guidance of Noémie S. Becker and Ricardo J. Pereira. I wrote the manuscript with input from Noémie S. Becker, Ricardo J. Pereira, Hiroki Kawabata, and Gabriele Margos. The final version was then read and approved by all co-authors. The manuscript is currently in review at Molecular Ecology but remains unpublished.

Paper 6: Disentangling the role of *Borrelia bavariensis* PFam54 proteins *in vitro* and *in vivo* using two strains naturally lacking the PFam54 gene array

Robert E. Rollins, Janna Wülbern, Florian Röttgerding, Tristan Nowak, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Yi-Pin Lin, Peter Kraiczy, Noémie S. Becker

Noémie S. Becker, Janna Wülbern, Sabrina Hepner, Peter Kraiczy, and I conceptualized the project which was then coordinated by Noémie S. Becker and Peter Kraiczy. The isolates under study were provided by Gabriele Margos and Volker Fingerle which were cultivated (depending on lab location) by myself, Florian Röttgerding, and Tristan Nowak. Sabrina Hepner, Gabriele Margos, and Volker Fingerle organized all whole genome sequencing of samples for which the data was processed by Noémie S. Becker (Illumina), Sabrina Hepner (PacBio), and me (Illumina). Janna Wülbern and I characterized the PFam54 gene family in all samples including developing the PCRs for confirming absence or presence of the genes. Florian Röttgerding performed the serum bactericidal assays and immunofluorescence analyses under the guidance of Peter Kraiczy. Tristan Nowak performed all mouse infection studies under the guidance of Yi-Pin Lin. Throughout the study, expert opinions and advice were provided by Peter Kraiczy, Yi-Pin Lin, Noémie S. Becker, Gabriele Margos, and Volker Fingerle. I wrote the manuscript with Noémie S. Becker, Janna Wülbern Peter Kraiczy, and Yi-Pin Lin. The final draft was approved by all co-authors. The current manuscript is under review at Applied and Environmental Microbiology but remains unpublished.

_____ _____

Robert Ethan Rollins Dr. Noémie S. Becker

Abbreviations

DFM – Dark field microscopy

CI – Credible interval

LB – Lyme borreliosis

LD – Lyme disease

RF – Relapsing fever

CCHF – Crimean-Congo hemorrhagic fever

TBE – tick-borne encephalitis

***Bb*sl** – *Borrelia burgorferi* sensu lato

lp – linear plasmid

cp – circular plasmid

CRASP – complement-regulator acquiring surface protein

MAC – membrane attack complex

MLST – multiple locus sequence typing

OspA – outer surface protein A

OspC – outer surface protein C

TROSPA – tick receptor of outer surface protein A

AMOVA – analysis of molecular variance

GLMM – generalized linear mixed effects model

List of Figures

List of Tables

Abstract

Ticks are obligate, blood feeding parasites of vertebrate animals and further act as vectors for many vector-borne pathogens including some causing human and animal disease. One such pathogen is *Borrelia burgdorferi* sensu lato (*Bb*sl), which describes a diverse species complex of spirochete bacteria that exist in an obligate transmission cycle between ixodid ticks and various vertebrate reservoir hosts. Some of these *Bb*sl genospecies are of epidemiological concern as they are the causative agents of human Lyme borreliosis (LB); the most common vector-borne disease in the Northern hemisphere. Due to their ecology, these genospecies undergo complex evolutionary and selective pressures driving adaptation to new host and vector species, but also maintain a high level of within-species diversity. These processes are thought to further drive speciation events and expansion of *Bb*sl genospecies into new geographic ranges. Evolutionary studies are currently impeded by a lack of isolates and associated genomic data including populations, such as those in Asia, having been neglected in the past. Combining ecological and genomic data, would open the way to studying the genetic underpinnings of both host and vector adaptation, variability in the severity of human LB, and how these bacteria may adapt to changing environments. The Eurasian *Bb*sl system offers a unique opportunity to utilize this evolutionary ecology approach to study *Bb*sl evolution. Currently three LB-causing *Bb*sl genospecies (*Borrelia afzelii*, *Borrelia bavariensis*, *Borrelia garinii*) share a Eurasian distribution and are transmitted by different generalist tick vectors: *Ixodes ricinus* (Europe) or *Ixodes persulcatus* (Asia). These genospecies further differ in their host associations and human LB manifestations, making them prime candidates for comparative genomics studies to understand the genetic underpinnings of these factors. The main aim of this dissertation was to study *Bb*sl from an evolutionary ecology approach and to produce isolates of pathogenic, Eurasian *Bb*sl genospecies with associated genomic data to fill current gaps in our knowledge of these genospecies. We utilized ecological (Papers 1-3), genomic (Papers 4-5), and molecular based (Paper 6) studies to answer questions related to the evolution of these genospecies.

In Paper 1, we screened migratory birds transiting through the Italian island of Ponza for exotic tick species and tick-borne pathogens, including bacteria and viruses. Through this we were able to confirm the role of migratory birds in the movement of exotic ticks and associated pathogens into new geographic areas. We then collected

Ixodes ricinus ticks longitudinally over a 10-year period in Latvia (Paper 2) and in a single year (2019) in multiple plots in southern-Germany (Paper 3). These studies allowed us to hypothesize that both along spatial and temporal scales that *Bb*sl diversity and prevalence is influenced predominantly by host community structure, leading to both highly stable and dynamic *Bb*sl communities. Utilizing then existing genomic data and recently sequenced Russian isolates of *B. bavariensis*, we characterized the full genomes of 33 isolates in Paper 4. This allowed us to see that, even in the almost clonal European population, we observe a high level of within-species diversity including novel findings such as higher plasmid copy number than expected from previous literature. From this we further characterized the evolutionary history of 142 isolates belonging to *B. afzelii*, *B. bavariensis*, and *B. garinii* based on assembled chromosomes corrected for recombining regions. Through this we were able to support that all three genospecies share an Asian origin (already shown for *B. bavariensis* in Paper 4 and previous work) and that colonization of Europe resulted through adaptation to a novel tick vector (*I. ricinus*). This also was paired with the fact that post-colonization gene flow for the different genospecies appeared to correlate with the mobility of proposed reservoir hosts. This allowed us to make testable hypotheses regarding the evolution and ecology of these three genospecies.

Both Paper 4 and 5, characterized a high level of within-species diversity even in our samples which we hypothesize could influence *Bb*sl spirochete interactions with both host and vector species. Paper 6 characterizes one instance of within-species variation where two European *B. bavariensis* isolates (PBN and PNi) were found to naturally lack the entire PFam54 gene array. This gene array is known to encode protein products important for evading the host's innate immune system, cell adhesion, and survival in the tick midgut. Their natural absence did increases susceptibility to human complement, an important pillar of innate immunity. Even so, these isolates remained infectious to mice post intradermal inoculation but did differ in how efficiently they can colonize certain mouse tissues. This highlighted that these genes are not required for mouse infectivity but potentially play a larger role in human pathogenicity. There is just one instance of within-species variability and is related to gene loss instead of sequence variation. Our results show that a high level of diversity does exist across the Eurasian range of these three genospecies which requires further research to understand how variability relates to the evolution of human pathogenicity, vector adaptation, and host adaptation.

Taken together, our results allowed us to better understand how these genospecies evolved across their geographic range. We were able to integrate ecological and genomics-based studies to show that geographic expansion appears to relate to vector adaptation while divergence, and potentially speciation, appears to be driven by host-adaptation. Utilizing an evolutionary ecology perspective, we were able to also identify unexpected results, such as the lack of geographic structure in the Asian *B. bavariensis* isolates, and through this create testable hypotheses regarding *Bb*sl ecology. Our analysis further was able to show that natural within-species variation can influence transmission cycles and produced an isolate library (especially for Asian populations of LB-causing genospecies) which can be utilized in future lab-based transmission studies to determine how further variability can lead to adaptation or human pathogenicity in these bacteria.

General Introduction

"It is interesting to contemplate an entangled bank, clothed with many plants of many kinds, with birds singing on the bushes, with various insects flitting about, and with worms crawling through the damp earth, and to reflect that these elaborately constructed forms, so different from each other, and dependent on each other in so complex a manner, have all been produced by laws acting around us. These laws, taken in the largest sense, being Growth with Reproduction; Inheritance which is almost implied by reproduction; Variability from the indirect and direct action of the external conditions of life, and from use and disuse; a Ratio of Increase so high as to lead to a Struggle for Life, and as a consequence to Natural Selection, entailing Divergence of Character and the Extinction of less-improved forms. Thus, from the war of nature, from famine and death, the most exalted object which we are capable of conceiving, namely, the production of the higher animals, directly follows. There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved."

Charles Darwin "On the Origin of Species", 1st ed. 1859

Ecological theory describes six major interspecific relationships which can exist in the natural world: competition, predation, symbiosis, mutualism, commensalism, and parasitism (1). Each relationship describes a potential mode of interaction between organisms where each party can be influenced positively, negatively, or neutrally (1). These relationships influence the fitness of organisms and can impose selective forces resulting in adaptation (1). Parasitism which describes a relationship where one organism (i.e. the parasite) benefits from another (i.e. the host) who is negatively affected by the relationship, is one such example (1). Many parasites currently are undergoing geographic range expansion in relation to changing environments leading to new opportunities for biological interactions and, through this, novel opportunities for adaptation (2–7). As many parasites further act as vectors for pathogenic organisms (2, 8), geographic expansion can not only lead to geographic spread of parasitic organisms but their corresponding pathogens potentially leading to emergent diseases.

Many vector-borne diseases appear to have increased in prevalence and geographic range over the past few decades (2, 9, 10). This brings forward questions whether pathogens will successfully colonize new locations and how will they potentially adapt to these new environments? Adaptation and successful establishment would have wide reaching impacts to human and animal health as contact rates with pathogenic vector-borne pathogens could increase. Vector-borne pathogens face steep

selective barriers not only to infect their respective vectors but further to infect their host species and evade elaborate immune responses (11–15). These selective pressures have shaped the evolutionary history of these pathogens suggesting that studying them from an evolutionary genomics and ecological perspective could answer questions regarding past adaptative trajectories; providing valuable information in predicting future adaptation (11, 16, 17). This dissertation will focus on one vector system, namely ticks and tick-borne pathogens, as a model for understanding how these systems evolve to new environments and what evolutionary and ecological factors result in adaptation or even geographic expansion.

Ticks and the pathogens they carry

Ticks are parasitic arachnids (Family: Arachnida) which belong to the order Acari along with other species of mites and can be found on basically all continents besides Antarctica (18, 19). They are currently split into two major families: the Argasidae (i.e. soft ticks) and the Ixodidae (i.e. hard ticks) (8, 18, 19). The hard ticks comprise the most species diverse family with over 650 recognized species while the soft ticks comprise approximately 175 species (18). A third family does also exist, the Nuttalliellidae, which consists of only a single species *Nuttalliella namaqua* which is found in certain areas of Africa (18, 20). For the rest of this dissertation when ticks are discussed, this will refer predominantly to hard ticks (i.e. Ixodidae).

Hard ticks contains six major genera of medical and veterinarian importance (*Ixodes*, *Hyalomma*, *Dermacentor*, *Rhipicephalus*, *Haemaphysalis*, and *Amblyomma*) as well as eight minor genera containing fewer species which are generally not of medical importance (18, 19, 21). All ticks are blood-sucking ectoparasites that feed on various vertebrate hosts (18, 19). Hard ticks have three distinct life-stages (larvae, nymph, and adult) between which they must consume a blood meal (18, 19). Hard ticks can be further sub-divided based on their host-seeking behavior (i.e. questing behavior). Many species are exophilic meaning they quest for hosts in the open environment either by waiting on vegetation to ambush passing hosts (*Ixodes, Rhipicephalus*, *Haemaphysalis*, *Dermacentor*) or by actively "hunting" hosts by running across the ground towards nearby hosts (*Amblyomma*, *Hyalomma*) (19). As these exophilic species spend much of their time in the open environment, their survival is not only influenced by the presence of suitable hosts but also by many environmental factors. Depending on tick species, factors such as vegetation, relative

humidity, precipitation, and temperature can all play a role in tick survival (18, 19, 22–24). Due to this, habitat modification, such as urbanization, or climate change can further influence tick survival (25).

Ticks are the most important vector of pathogens to domestic animals worldwide and rival mosquitos in their medical importance to humans (18). Even so, not all tick species can vector pathogens, with only a few hundred of the over 800 recognized species acting as competent vectors to pathogenic organisms (18). These pathogens include bacterial species from various genera (e.g., *Borrelia*, *Anaplasma*, *Ehrlichia*, *Rickettsia*), hemoparasites (e.g., *Babesia, Theleria*), and a diverse group of viruses representing multiple viral families (e.g., *Flaviviridae*, *Bunyavirales*, *Orthomyxoviridae*, *Reoviridae*) (8, 18, 26, 27). Most of these pathogens are obligate parasites and do not have free-living life stages; only existing in transmission cycles between vertebrate reservoir hosts and their respective tick vectors (8, 18). Most tick-borne pathogens are maintained by stable trans-stadial (i.e., between molts) transmission between tick life stages (8). However in rare cases, ticks can act both as vector and reservoir in the presence of stable trans-ovarial transmission (many viral infections), where adult females transmit pathogens to their eggs (8, 28). Many tick-borne pathogens are of epidemiological importance causing various diseases in both live-stock animals and humans (8, 18). These include widespread viral diseases such as tick-borne encephalitis (TBE) or Crimean-Congo hemorrhagic fever (CCHF) but also rarer viral diseases many of which are increasing in incidence across their geographic ranges (8, 27). Besides viral infections ticks transmit many bacterial caused diseases such as spotted fever rickettsioses (e.g., Rocky Mountain spotted fever, African tick bite fever), human anaplasmosis, human ehrlichiosis, tularemia, human relapsing fever borreliosis, and Lyme borreliosis (LB, also termed Lyme disease (LD) in North America) which is the most common vector-borne zoonosis in the northern hemisphere (8, 29, 30). Lyme borreliosis is caused by specific genospecies of spirochete bacteria belonging to the genus *Borrelia* (29, 30).

Ecology of *Borrelia burgdorferi* sensu lato spirochetes

Borrelia are spiral shaped bacteria (i.e. spirochetes), which belong to the family Spirochaetales, along with other medically relevant bacterial species such as *Treponema pallidum*; the causative agent of syphilis (31). *Borrelia* bacteria are currently separated into three clades described as the LB spirochetes (containing the

genospecies causing LB in humans), the relapsing-fever (RF) spirochetes (containing the causative agents of RF in humans), and reptile-associated *Borrelia* which also include monotreme associated genospecies (31, 32). In 2014, a split was proposed to create two genera, the historic *Borrelia* to contain the RF spirochetes and *Borreliella* gen. nov. to contain the LB spirochetes (33). This sparked a large debate regarding the taxonomy of *Borrelia* with currently no scientific consensus regarding the correct taxonomy (31, 32, 34–38). For the ease of comparison to previous work, this dissertation will utilize the nomenclature proposed by Margos et al., (2018), which does not split *Borrelia* into two genera and maintains the clade descriptions of RF clade, reptile-associated *Borrelia*, and the LB clade. The LB clade is normally referred to as the *Borrelia burgdorferi* sensu lato (hereafter *Bb*sl) species complex

The *Bb*sl species complex currently contains 20 accepted and 3 proposed genospecies which differ in their geographic distribution, tick vector as well as reservoir host associations (39–42). Of these genospecies, only six are thought to cause LB in humans (*Borrelia afzelii*, *Borrelia burgdorferi* sensu stricto (s.s.), *Borrelia bavariensis*, *Borrelia garinii*, *Borrelia mayonii*, and *Borrelia spielmanii*) (29, 30, 43, 44). As with many tick-borne pathogens, *Bb*sl spirochetes are obligate parasites with no free living life stage and exist in a transmission cycle between competent, hard tick vectors (*Ixodes* spp.) and various vertebrate hosts (39, 42, 43). Genospecies belonging to the *Bb*sl complex are found across the northern Hemisphere including marine-based, avian transmission cycles (42, 43, 45). Across this large geographic range these spirochetes are predominately vectored by very few, competent tick species, namely: *Ixodes ricinus* (Europe), *Ixodes persulcatus* (Asia), *Ixodes uriae* (marine), *Ixodes scapularis* (eastern North America), *Ixodes pacificus* (western North America) (42, 43, 46). Many *Bb*sl genospecies have highly specific host associations with most genospecies surviving only in a single host class (i.e., passerine birds, rodents, marine birds, etc.) (42, 47, 48). Even so, genospecies such as *B. burgdorferi* s.s. can act as host generalists, where spirochetes belonging to this genospecies are able to infect various vertebrate host classes (43, 46). Although, it does seem that specific genotypes of *B. burgdorferi* s.s. display variable fitness in the different host classes it can infect (11, 49). Some vertebrate species such as deer or large vertebrates interestingly do not act as reservoirs for *Bb*sl, even though they play an integral role in maintaining tick populations (43, 46, 50, 51). These types of vertebrates are termed dead-end hosts as they do not play a role in further transmission of *Bb*sl (42). Some vertebrates can be

infected by certain *Bb*sl genospecies but do not transmit these spirochetes further, therefore halting the transmission cycle (43, 46). These hosts are termed incidental hosts and include humans (43, 46). For humans, infection with *Bb*sl spirochetes belonging to specific genospecies can manifest as human LB (29, 30). Early human LB generally manifests with a characteristic erythema migrans or "bulls-eye shaped" rash along with non-specific flu-like symptoms (29, 30). Left untreated *Bb*sl spirochetes can migrate to distant tissues such as the central nervous system and joints causing various symptoms such as neuroborreliosis or Lyme arthritis (29, 30).

As the successful transmission of *Bb*sl spirochetes is dependent on the presence and viability of both a competent tick vector and reservoir host population, ecological factors which influence either of these will have a pronounced impact on *Bb*sl community structure. Indeed, factors such as humidity, temperature, and vegetation cover all can influence the prevalence of *Bb*sl in the environment (52–54) through the proposed effects these have on the tick vector (22, 51, 53). Reservoir host community structure will also influence if *Bb*sl spirochetes can survive in an environment and, if so, which genospecies are maintained (43, 55, 56). Even so, a high diversity of hosts can also have negative effects on *Bb*sl transmission in certain situations (57, 58). This phenomenon is known as the dilution effect, where vectors are spread across a high diversity of hosts, of which some may not be competent reservoirs for *Bb*sl spirochetes (57, 58). This reduces the probability of an infected tick to attach to a competent host resulting in an overall reduction in *Bb*sl prevalence (57, 58). Although, this effect has been shown to not hold true in every situation (59). Taken together, the successful transmission and therefore evolution of *Bb*sl spirochetes is intimately linked to the reservoir host and vector species to which these bacteria are adapted.

Molecular mechanism of *Borrelia burgdorferi* sensu lato transmission

The concept of ecological niche or all the factors required for an organism to successfully complete its life cycle (11, 60), can be a valuable tool in understanding the evolution of *Bb*sl genospecies. For these spirochetes, we can define the ecological niche as the reservoir host and competent vector species which are required for *Bb*sl spirochetes to complete their life cycle (13, 43, 46). On the side of the host, *Bb*sl spirochetes need to overcome the host's innate and adaptive immune system to establish viable infections (12, 61–63). In terms of tick colonization, the spirochetes need to successfully migrate into a feeding tick, overcome tick immunity, survive in a

nutrient poor environment until the tick feeds again, and then again successfully infect a naïve host (13, 43, 46, 64).

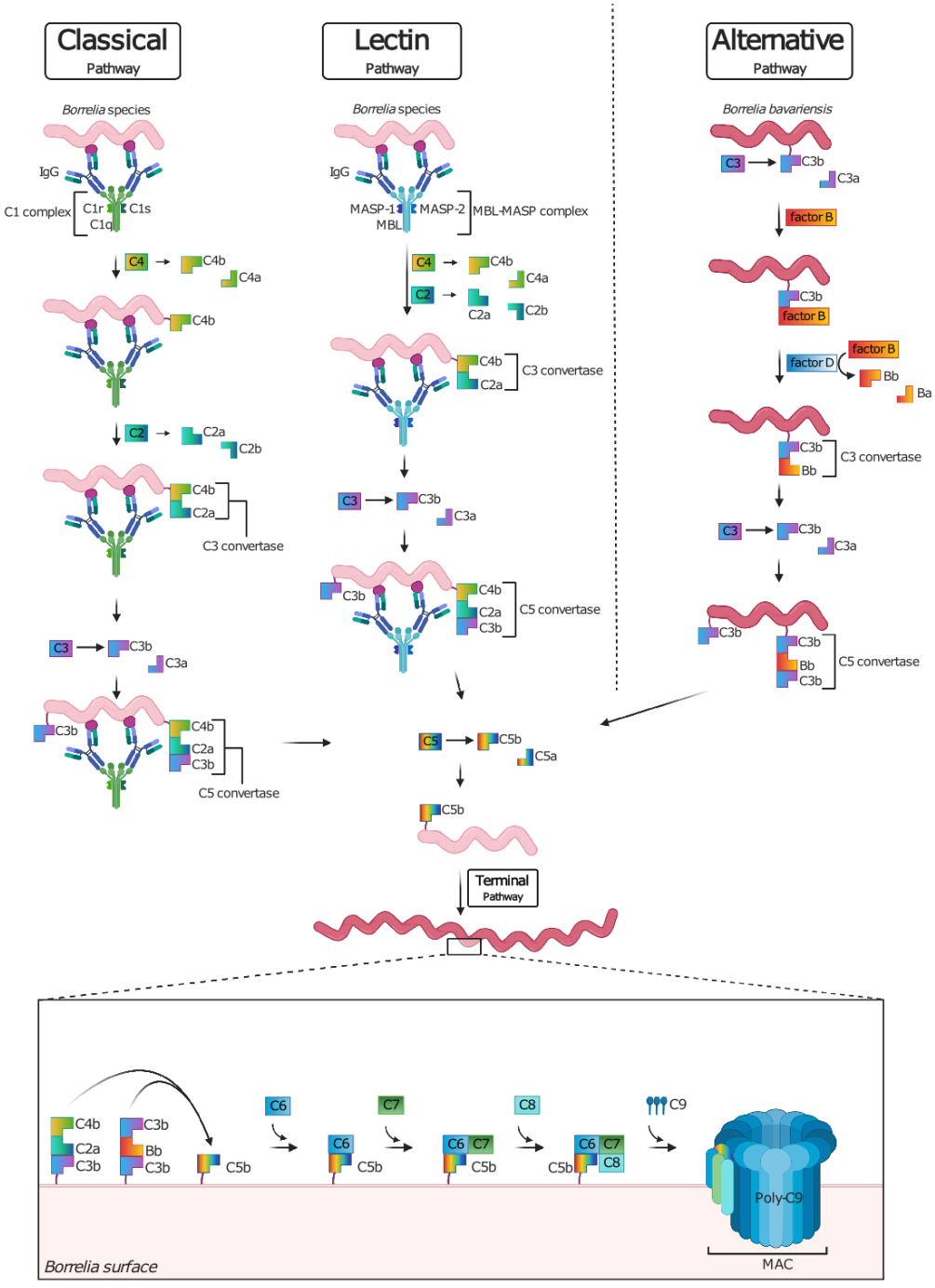


Figure 1. Overview of the three activation pathways of the human complement system: classical, lectin, and alternative pathways. All three pathways recognize foreign antigens in unique ways which all lead to the cleavage of complement component C3 into activated C3b and co-component C3a. This initiates a protein cascade beginning with the cleavage of C5 leading to the recruitment of late stage complement proteins (C6, C7, C8, C9) which ends in the formation of the membrane attack complex (MAC) and cell lysis. See (14, 15) for a complete overview of the complement system and the various proteins involved in all three pathways. Complement system figure adapted from the Master thesis of Janna Wülbern who has given permission for the inclusion of the figure in this dissertation.

To establish host infection, *Bb*sl spirochetes need to be able to evade the host's immune system, which consists of the innate and adaptive immune systems (14, 15). *Borrelia burgdorferi* s.l. spirochetes have evolved various mechanisms to achieve this such as overcoming complement, an integral part of innate immunity, either indirectly through the acquisition of complement regulators or directly through interactions with complement proteins (61, 62, 65–67). The complement system consists of three distinct pathways (classical, lectin, and alternative) all leading to the cleavage of C3 to form activated C3b (14, 15) (Figure 1). This initiates the activation of other complement components ending in the recruitment of late stage complement proteins (C6, C7, C8, and C9) to form the membrane attack complex (MAC) which ultimately leads to bacterial cell lysis (14, 15) (Figure 1). Additionally, cleavage of C3 and C5 by the C3 and C5 convertases, respectively, leads to the release of anaphylatoxins C3a and C5a, which can recruit additional immune cells and, therefore, are integral in mounting further host immune responses to infection (14, 15) (Figure 1). Host cells regulate complement by utilizing membrane-bound regulators or by binding fluid-phase regulatory proteins, such as factor H (14). All complement regulators can terminate the complement cascade at specific activation levels to protect self-cells from complement-mediated damage (14). *Borrelia burgdorferi* s.l. spirochetes produce various proteins, known as CRASPs (**c**omplement **r**egulator **a**cquiring **s**urface **p**roteins), which can recruit factor H or factor H-like proteins and confer resistance to complement mediated killing (65, 66). Further mechanisms have been described as well, where through direct interactions with specific complement proteins such as C1 (Figure 1) or late stage complement proteins C7-C9 (Figure 1) *Bb*sl spirochetes are also able to protect themselves from complement (68, 69). Most of this work though has been done utilizing human complement with less focus on how these interactions function in the natural reservoir hosts. It has been shown that susceptibility to complement active serum is host specific and appears to mirror proposed *Bb*sl host associations (42, 47, 48, 70, 71). This, paired with recent work that showed that some *Bb*sl genospecies display host specific factor H binding (71, 72), suggests that complement evasion is most likely a major component of host adaptation, although more work is needed. Overcoming complement is just one mechanisms *Bb*sl spirochetes utilize to successfully infect their hosts. Other mechanisms include the ability to produce variable antigenic surface proteins through the *vslE* locus to protect spirochetes from

the host's adaptive immune system (63) and producing proteins with cell adhesin properties to disseminate to specific tissues within a host (73–75).

Although highly important, successful host infection and dissemination is just one half of the *Bb*sl lifecycle. These spirochetes additionally need to colonize the tick vector. The current transmission mechanism has only been thoroughly studied in the North American system focusing on the main LD-causing *Bb*sl genospecies, *Borrelia burgdorferi* s. s., and its vector *Ixodes scapularis* (13, 64). In this case, *B. burgdorferi* s. s. appears to first colonize the midgut of the tick where it attaches to the epithelial wall through binding of the borrelial outer surface protein A (OspA) and the tick receptor of outer surface protein A (TROSPA) (13, 64, 76). Here the bacteria enter a dormant state during the tick's molt to the next life stage (13, 43, 64). Once the next bloodmeal commences, the spirochetes begin to replicate and undergo transcriptional modifications to produce outer surface proteins which are thought to protect from immune factors still present in the host blood meal (77). The spirochetes then migrate from the midgut, through the hemolymph, into the salivary glands through an interaction with the tick protein TRE31 and the borrelial protein BBE31 (13, 64, 78). The spirochetes then finally migrate out of the salivary glands partially due to interactions between an additional outer surface protein, OspC, which can bind and interact with the tick salivary component Salp15 (13, 64, 79, 80). The tick saliva itself contains many immunomodulatory factors which facilitates *Bb*sl spirochetes colonization at the tick feeding site (13, 64). Although this mechanism is thought to hold true for other *Bb*sl genospecies, recent work on *Borrelia afzelii*, one of the main causative agents of human LB in Eurasia, has shown that *B. afzelii* spirochetes do not appear to migrate through the salivary glands but instead migrate directly into the host from the tick midgut (81). This opens the viable possibility that different *Bb*sl genospecies may use different transmission mechanisms within the tick vector requiring further research.

The evolution of the *Borrelia burgdorferi* sensu lato species complex

Borrelia burgdorferi s.l. spirochetes have complex and fragmented genomes comprising a linear chromosome of approximately 900kb and up to or over 20 unique linear and circular plasmids (82–85). The chromosome mostly encodes for proteins related to metabolism and general housekeeping functions (32, 82, 83, 85) with many adaptation genes related to host and vector colonization being plasmid located (83,

85). In general, the chromosome has been found to be highly conserved among and within genospecies while plasmids are known to be diverse not only in their presence or absence in specific isolates but also gene content (84–87).

Ecological interactions between *Bb*sl spirochetes and reservoir hosts will select for variation in genes and encoded products, such as CRASPs, which facilitate successful infection of the vertebrate host (11, 16, 71, 88). It can further be observed that the proposed strength of these selective forces should be proportional to the rate of interaction between a specific vertebrate host and the *Bb*sl spirochetes (11, 43). Namely, rarer vertebrate species which *Bb*sl spirochetes do not encounter that often will not exert selective forces that could drive host adaptation or association (11, 43). This can also mean that certain *Bb*sl genospecies can be excluded from an environment if a suitable vertebrate host is not present; modifying therefore the ecological distribution and prevalence of *Bb*sl genospecies (11, 55, 89). Host associations can then influence other factors such as gene flow (i.e. migration) of *Bb*sl spirochetes due to host mobility and further influencing population structure (11, 16, 56, 90, 91). This can affect passive divergence due to stochastic processes such as genetic drift, making less mobile populations more prone to divergence due to drift (11). Ticks also can display host preferences, suggesting that *Bb*sl vector association could also lead to geographic structuring (11, 16). Although, whether structuring is driven more by vector or host association is challenging to disentangle. Many have argued that these processes could drive speciation events (11, 42, 44, 92), such as in the case of *B. garinii* and its sister species *B. bavariensis* (39, 93, 94). These two genospecies most likely diverged due to specializing on different vertebrate reservoir hosts: *B. garinii* to birds, and *B. bavariensis* to rodents (39, 93, 94). Examples such as these are prime opportunities to employee comparative genomics to study the influences of host adaptation to genospecies structure but also to research the genetic underpinnings of host switching in a vector-borne bacterium.

In a similar way to the host, recurrent interactions with vectors will impose selective pressure on genes promoting successful vector colonization (11, 13, 43). Vector adaptation genes studied so far display low within-genospecies diversity while between-genospecies diversity is higher (11, 16). Lower diversity in vector-related genes is thought to arise from selection for specific mutations which maximize the colonization efficiency of a whole *Bb*sl genospecies to specific vectors (11, 16). This would lead to higher variability between *Bb*sl genospecies which utilize different tick

vectors while leading to lower within-genospecies diversity due to potential purifying selection (11, 16). Interactions with novel vectors and consequent adaptation, can introduce *Bb*sl spirochetes to novel ecological conditions and further lead to geographic expansion (11, 16). This can be observed with the example of *B. bavariensis* which consists of two populations: a high diversity, ancestral Asian population and a low diversity, almost clonal European population (39, 93). It is hypothesized that this population split resulted through *B. bavariensis* being able to invade a novel vector, *Ixodes ricinus* (39, 93). This resulted in *B. bavariensis* expanding geographically but was accompanied by a population bottleneck which greatly reduced overall diversity in the European population (39, 93). *Borrelia bavariensis* could therefore provide an opportunity to utilize comparative genomics to study the genetic underpinnings of vector adaptation. Additional genospecies, such as *B. garinii*, *B. afzelii*, and *B. turdi* also are found in European and Asian transmission cycles but if they as well have undergone a selective vector shift is not well studied (39).

Motivation and research objectives

Since the description of *B. burgdorferi* s. s. in 1982 (95), many aspects of *Bb*sl transmission have been clarified. Even so, almost 40 years later many genes, and their encoded products, still do not have a known function or are still described as hypothetical (96). This leaves many open questions regarding infection mechanisms. This includes the role of within-species diversity as most functional studies to date have focused on very few type-strains (12, 13, 46, 66). *Borrelia burgdorferi* sensu lato spirochetes offer a unique opportunity to study both the genetic underpinnings of host and vector adaptation plus the evolutionary constraints experienced by vector-borne bacteria. As these spirochetes further differ in their capacity to cause human disease (29, 30), studies related to evolutionary ecology would be able to inform on future geographic range expansions and potential impacts to human disease risk (11).

A major impediment to the study of many *Bb*sl genospecies from an eco-evolutionary perspective, is a lack of viable isolates and samples from the field. This includes a bias in samples towards patient isolates and a focus on type strains which most likely do not capture the global diversity of the genospecies. Without these samples, it is not possible to study *Bb*sl spirochetes from an evolutionary ecology perspective (11, 16, 97). The Eurasian *Bb*sl system allows for a fascinating opportunity to study both host and vector adaptation. Three human pathogenic *Bb*sl genospecies

(*Borrelia afzelii*, *Borrelia bavariensis*, and *Borrelia garinii*) all share a Eurasian distribution and currently exist in separate transmission cycles vectored predominately by different generalist ticks in Asia (*Ixodes persulcatus*) and Europe (*Ixodes ricinus*) (43). This means that each genospecies has at least once in its evolutionary history successfully adapted to a novel tick vector. Previous work in *B. bavariensis* has supported an Asian origin suggesting that the adaptation was from *I. persulcatus* into *I. ricinus* (39, 93), but whether this is also true for the other genospecies is not known. This alone sets up this system to be an invaluable opportunity to study vector adaptation in three independent *Bb*sl genospecies, all of which can cause LB in humans. Additionally, though, these three genospecies differ in their host associations with both *B. bavariensis* and *B. afzelii* utilizing rodents as reservoir hosts while *B. garinii* is associated with birds (39, 43, 45, 98–100). *Borrelia garinii* and *B. bavariensis* are closely related sister taxa (39, 101), which suggests that at the split between these two genospecies there was also a switch in host association (39, 101). This means that this system can also be used to study the underlaying factors related to host adaptation in addition to vector adaptation.

The overarching aim of this dissertation was to answer open questions regarding the evolution of Eurasian *Bb*sl genospecies in relation to host and vector adaptation through the collection and genomic characterization of novel tick isolates of LB causing genospecies from across their Eurasian range. Paper 1, 2, and 3 concern themselves with the ecological spread and prevalence of *Bb*sl (Papers 2 & 3) along with other tick-borne pathogens (Paper 1). Utilizing then whole genome data from existing isolates and novel *Bb*sl isolates, we aimed to further understanding the basic evolutionary history and full genomes of *B. bavariensis* across Eurasia (Paper 4) before expanding to describe how all three Eurasian genospecies (*B. afzelii*, *B. bavariensis*, and *B. garinii*) have evolved across their geographic range (Paper 5). Paper 6 then characterizes the phenotypic effects related to the genomic variability observed in two *B. bavariensis* isolates described in Papers 4 & 5.

Results

Paper 1: Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry

Robert E. Rollins, Sabine Schaper, Claudia Kahlhofer, Dimitrios Frangoulidis, Aurelia F. T. Strauß, Massimiliano Cardinale, Andrea Springer, Christina Strube, Deon K. Bakkes, Noémie S. Becker, Lidia Chitimia-Dobler

Ticks and Tick-borne Diseases (2021), 12(1): 101590
https://doi.org/10.1016/j.ttbdis.2020.101590

Ticks and Tick-borne Diseases 12 (2021) 101590

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

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

Short communication

Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry

Robert E. Rollins [a], Sabine Schaper [b], Claudia Kahlhofer [b], Dimitrios Frangoulidis [b,c], Aurelia F. T. Strauß [d], Massimiliano Cardinale [e], Andrea Springer [f], Christina Strube [f], Deon K. Bakkes [g,h], Noémie S. Becker [a,*], Lidia Chitimia-Dobler [b]

[a] Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Großhaderner Str. 2, 82512 Planegg-Martinsried, Germany
[b] Bundeswehr Institute of Microbiology, Munich, Germany
[c] Bundeswehr Medical Service Headquarters VI-2, Medical Intelligence & Information (MI²), Munich, Germany
[d] Research Group Evolution of Sensory Systems, Max Plank Institute for Ornithology, Eberhard-Gwinner-Strasse 5, 82319 Seewiesen, Germany
[e] Swedish University of Agricultural Sciences, Department of Aquatic Resources, Institute of Marine Research, Lysekil 45330, Turistgatan 5, Sweden
[f] Institute for Parasitology, Centre for Infection Medicine, University of Veterinary Medicine Hannover, Buenteweg 17, 30559 Hanover, Germany
[g] Gertrud Theiler Tick Museum, Onderstepoort Veterinary Research, Pretoria, South Africa
[h] Evolutionary Genomics Group, Department of Botany and Zoology, Stellenbosch University, South Africa

ARTICLE INFO

Keywords:
Ticks
Migratory birds
Tick-borne pathogens
Africa
Ponza

ABSTRACT

Seasonal migration of birds between breeding and wintering areas can facilitate the spread of tick species and tick-borne diseases. In this study, 151 birds representing 10 different bird species were captured on Ponza Island, an important migratory stopover off the western coast of Italy and screened for tick infestation. Ticks were collected and identified morphologically. Morphological identification was supported through sequencing a fragment of the 16S mitochondrial gene. In total, 16 captured birds carried ticks from four tick species: *Hyalomma rufipes* (n = 14), *Amblyomma variegatum* (n = 1), *Amblyomma* sp. (n = 1), and *Ixodes ventalloi* (n = 2). All specimens were either larvae (n = 2) or nymphs (n = 16). All ticks were investigated for tick-borne pathogens using published molecular methods. *Rickettsia aeschlimannii* was detected in six of the 14 collected *H. rufipes* ticks. Additionally, the singular *A. variegatum* nymph tested positive for *R. africae*. In all 14 *H. rufipes* specimens (2 larvae and 12 nymphs), *Francisella*-like endosymbionts were detected. Four *H. rufipes* ticks tested positive for *Borrelia burgdorferi* sensu lato in a screening PCR but did not produce sufficient amplicon amounts for species identification. All ticks tested negative for tick-borne encephalitis virus, Crimean-Congo hemorrhagic fever virus, *Coxiella burnetii*, *Coxiella*-like organisms, *Babesia* spp., and *Theileria* spp. This study confirms the role of migratory birds in the spread and establishment of both exotic tick species and tick-borne pathogens outside their endemic range.

1. Introduction

Seasonal migration of birds between breeding and overwintering areas can facilitate the spread of tick species and tick-borne diseases (Buczek et al., 2020; Comstedt et al., 2006), as birds easily cross geographical and ecological barriers. Migratory birds are of interest for tick-borne diseases as they carry immature ticks (Buczek et al., 2020) and can potentially spread pathogens far beyond their original home ranges.

Migratory bird species can be classified into: 1) short-distance migrants tending to move between European breeding areas and the non-breeding, overwintering regions of Europe and Northern Africa and 2) long-distance or trans-Saharan migrants which spend the non-breeding period in sub-Saharan African countries. Both migratory types can play a role in the dispersal of exotic tick species and various tick-borne pathogens from their natural habitats (Georgopoulou and Tsiouris, 2008; Hasle, 2013).

In the present study, ticks were collected from birds (both long- and

* Corresponding author at: Division of Evolutionary Biology, Faculty of Biology Ludwig Maxilian Universität München, Großhadernerstr. 2, 82152 Planegg-Martinsried, Germany.
E-mail address: nbecker@bio.lmu.de (N.S. Becker).

https://doi.org/10.1016/j.ttbdis.2020.101590
Received 8 May 2020; Received in revised form 23 September 2020; Accepted 28 September 2020
Available online 9 October 2020

short-distance migrants) captured in spring on the island of Ponza off the western coast of Italy. This island is recognized as an important stopover site for migratory birds returning from Africa in spring (Cecere et al., 2010; Maggini et al., 2020). The study aims to investigate the role of migratory birds for introduction of exotic tick species and tick-borne pathogens into Europe.

2. Materials and methods

2.1. Tick collection and identification

Ticks were collected from migratory birds trapped in mist nets on Ponza Island during spring migration from March until May 2019 (see Supplemental Methods for details). All ticks from each bird were stored in 80 % ethanol at room temperature until DNA/RNA extraction. Ticks were identified to the taxonomic level of species using published phenotypic keys (Apanaskevich and Horak, 2008; Pérez-Eid, 2007; Voltzit and Keirans, 2003).

2.2. Nucleic acid extraction and PCR

RNA and DNA were extracted from individual ticks using a the QIAamp mini Viral RNA extraction kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions and stored at −80 °C until use (tests for viral RNA were performed prior to freezing the samples). For molecular identification of tick specimens, the 16S rRNA gene was amplified using the polymerase chain reaction protocol as described by Mangold et al. (1998) and analysed using a phylogenetic method (see Suppl. Meth. for details).

Ticks were tested for tick-borne encephalitis virus (TBEV) RNA and for CCHF virus using previously described real time (RT) PCR methods (Atkinson et al., 2012; Schwaiger and Cassinotti, 2003). Detection of *Rickettsia* spp. was performed using a previously published RT-PCR assay targeting a conserved 70-bp part of the *gltA* gene (Wölfel et al., 2008). For species identification of *Rickettsia*, a PCR assay targeting the *23S-5S* intergenic spacer region was utilized using primers described by Chitimia-Dobler et al. (2018) and the thermoprofile described by Jado et al. (2006). *Rickettsia* sequences were aligned and analysed together with GenBank references for *R. aeschlimannii*, *R. africae*, *R. conorii* and *R. sibirica* (see Suppl. Meth.). Pairwise distances were calculated using maximum composite likelihood (G = 4) in MEGA 10.1.7 (Kumar et al., 2016). Furthermore, the ticks were tested for piroplasms using a conventional PCR amplifying part of the *18S* rRNA gene (Casati et al., 2006). This primer set amplifies *Babesia* spp. as well as *Theileria* spp. (Lempereur et al., 2017). Additionally, ticks were tested for *Coxiella burnetii* by real-time PCR targeting the IS1111 gene element (Frangoulidis et al., 2012) and conventional 16S rRNA PCR to test for further endosymbionts as previously described (Al-Deeb et al., 2016). 16S amplicons were identified through an NCBI-Blastn analysis. To test for *B. burgdorferi* s.l. we used an RT-PCR targeting the 23S-5S intergenic spacer as previously described (Strube et al., 2010). To determine *Borrelia* species of positive samples, we used a semi-nested PCR amplifying the housekeeping gene *recG* (Margos et al., 2008).

All primers used are to be found in Supplementary Table 1. Appropriate positive and negative controls were included in all PCR reactions and products were sequenced with the Sanger method. All RT-PCRs were run once per sample.

3. Results

Out of the 151 screened birds, 16 birds, belonging to 10 different species (Table 1), were carrying one or two ticks. Most infested birds (n = 14) belonged to long distance migratory species, the most common being *Sylvia communis*. Two short distance migratory species also carried ticks (Table 1).

In total, 18 ticks belonging to three genera and four tick species, were identified: *H. rufipes*, *A. variegatum*, *Amblyomma* spp., and *I. ventalloi* (Table 2).

All ticks tested negative for TBE virus, CCHF virus, as well as *Babesia* and *Theileria* species. Seven tick specimens were positive for *Rickettsia* spp., with *R. aeschlimannii* identified in six *H. rufipes* ticks and *R. africae* identified in one *A. variegatum* (Table 2). Using appropriate reference samples from GenBank (see Methods) average genetic pairwise distances within *Rickettsia* species ranged from 0.15−0.5%, while distance between species ranged from 1.72 to 5.59%. All *Hyalomma* ticks (n = 14) tested positive for *Francisella*-like endosymbionts but not for *C. burnetti* and *Coxiella*-like organisms (Table 2). Furthermore, four *H. rufipes* ticks tested positive for *B. burgdorferi* s.l. in a screening PCR but did not produce sufficient amplicon amounts for species identification (Table 2).

4. Discussion

Migratory birds can influence the spread of tick species and tick-borne diseases across geographic barriers even leading to introduction of exotic species to new environments (Hasle, 2013). This can also apply to tick endosymbionts (Bonnet et al., 2017), but to what extent is not fully understood. Ticks carry many endosymbiotic bacteria, which play varied roles in tick biology, most notably in diet supplementation (Bonnet et al., 2017).

4.1. The role of migratory birds in tick dispersal

In the present study, ticks belonging to three genera (i.e. *Hyalomma*, *Amblyomma*, and *Ixodes*) were collected from 10 migratory bird species, most of which were long distance migrants.

Amblyomma ticks were only found on the tree pipit (*Anthus trivialis*), which is a long distance migrant. Only one of these *Amblyomma* ticks could be identified to species level (confirmed as *A. variegatum*). The *Amblyomma* genus contains 130 species, none of which is distributed in Europe (Guglielmone et al., 2014). There have been sporadic reports of

Table 1
Information regarding bird species carrying ticks during this study.

Bird common name	Bird scientific name	Migration Distance	Migration Classification	Partial Resident	Breeding	Overwintering
Common whitethroat	*Sylvia communis*	Long	Full Migrant	–	Europe	sub-Saharan
Spotted flycatcher	*Muscicapa striata*	Long	Full Migrant	–	Europe/Western Asia	sub-Saharan
European robin	*Erithacus rubecula*	Short	Full Migrant	Western Europe	Europe/Western Asia	North Africa/Mediterranean
Tree pipit	*Anthus trivialis*	Long	Full Migrant	–	Europe/Western Asia	sub-Saharan/India
Northern wheatear	*Oenanthe oenanthe*	Long	Full Migrant	–	Europe/Asia	sub-Saharan
Eurasian golden oriole	*Oriolus oriolus*	Long	Full Migrant	Spain and Italy	Europe/Western Asia	sub-Saharan
Common redstart	*Phoenicurus phoenicurus*	Long	Full Migrant	–	Europe/Western Asia	sub-Saharan
Black redstart	*Phoenicurus ochruros*	Short	Full Migrant	–	Europe/Central Asia	North Africa/India
Whinchat	*Saxicola rubetra*	Long	Full Migrant	–	Europe/Western Asia	sub-Saharan
Collared flycatcher	*Ficedula albicollis*	Long	Full Migrant	–	Europe/Western Asia	sub-Saharan

Note: All information on the migratory behaviour and residential status of bird species in this project was taken from the BirdLife International (http://www.birdlife.org/) database.

Table 2
Information on the ticks collected from migratory birds and which pathogens they carried.

Bird common name	Tick species (ID)	Tick stage	*Rickettsia* spp.	*Francisella*-like endosymbionts	*Borrelia* spp.	*Babesia/ Theileria* spp.	TBE virus	CCHF virus	*Coxiella* spp./*C.*-like organisms
Common whitethroat	*Hyalomma rufipes* (G12)	nymph	–	+	–	–	–	–	–
Common whitethroat	*Hyalomma rufipes* (H1)	nymph	*R. aeschlimannii*	+	–	–	–	–	–
Common whitethroat	*Hyalomma rufipes* (H2)	nymph	*R. aeschlimannii*	+	–	–	–	–	–
Common whitethroat	*Hyalomma rufipes* (H3)	nymph	*R. aeschlimannii*	+	–	–	–	–	–
Common whitethroat	*Hyalomma rufipes* (H4)	nymph	–	+	+	–	–	–	–
Common whitethroat	*Hyalomma rufipes* (H5)	nymph	–	+	+	–	–	–	–
Spotted flycatcher	*Hyalomma rufipes* (H6)	larva	–	+	–	–	–	–	–
European robin	*Ixodes ventalloi* (H8)	nymph	–	–	–	–	–	–	–
Tree pipit	*Amblyomma* sp. (H9)	nymph	–	–	+	–	–	–	–
Tree pipit	*Hyalomma rufipes* (H10)	larva	*R. aeschlimannii*	+	–	–	–	–	–
	Amblyomma variegatum (H11)	nymph	*R. africae*	–	–	–	–	–	–
Northern wheatear	*Hyalomma rufipes* (H12)	nymph	*R. aeschlimannii*	+	+	–	–	–	–
Eurasian golden oriole	*Hyalomma rufipes* (A1)	nymph	–	+	+	–	–	–	–
Common redstart	*Hyalomma rufipes* (A2)	nymph	–	+	–	–	–	–	–
Black redstart	*Ixodes ventalloi* (A3)	nymph	–	–	–	–	–	–	–
Whinchat	*Hyalomma rufipes* (A4)	nymph	*R. aeschlimannii*	+	–	–	–	–	–
Black redstart	*Hyalomma rufipes* (A5)	nymph	–	+	–	–	–	–	–
Collared flycatcher	*Hyalomma rufipes* (A6)	nymph	–	+	–	–	–	–	–

All tick morphological identification was confirmed molecularly by sequencing a segment of the 16S mitochondrial gene (Mangold et al., 1998). Tick samples were screened for pathogens and endosymbionts using standard methods: *Rickettsia* spp. (Chitimia-Dobler et al., 2018; Wölfel et al., 2008), *Coxiella* spp. and *Coxiella*-like organisms (Frangoulidis et al., 2014; Al-Deeb et al., 2016), *Borrelia* spp. (Margos et al., 2008; Strube et al., 2010), *Babesia/Theileria* spp. (Casati et al., 2006), TBE virus (Schwaiger and Cassinotti, 2003), and CCHF virus (Atkinson et al., 2012).

A. variegatum and other *Amblyomma* spp. nymphs on long-distance migratory birds captured in Italy and Greece (Albanese et al., 1971; Papadopoulos et al., 1996; Toma et al., 2014; Wallménius et al., 2014) and adult *A. variegatum* specimens were sampled from cattle on the island of Corsica, suggesting a possible risk for the establishment of *Amblyomma* populations in Europe (Cicculli et al., 2019).

Hyalomma rufipes was the most common tick species found on the birds collected in this study (n = 14). All samples came from long-distance migrant species. Five individual *S. communis* carried *H. rufipes* suggesting that these birds play a substantial role as a host for *H. rufipes* and shuttle for its expansion. *Sylvia communis* birds captured at Ventotene island, located 30 km south-east from Ponza Island, also yielded the highest number of *H. rufipes* ticks (Pascucci et al., 2019). Both *H. rufipes* immature stages have been frequently found on birds migrating from Africa (Hoffman et al., 2018; Jaenson et al., 1994; Jameson et al., 2012), including birds migrating through Italy (Pascucci et al., 2019). *Hyalomma marginatum* and *H. rufipes* have been observed in multiple European countries (Chitimia-Dobler et al., 2019; Duscher et al., 2018; Grandi et al., 2020; Hansford et al., 2019).

Two *I. ventalloi* nymphs were found on the two short-distance migrant bird species (Table 2). This tick species can be found in multiple European countries and in northern Africa (Hillyard, 1996). Thus, there is a high probability that these ticks were locally acquired (Hillyard, 1996). Studies to verify which local tick populations are present on Ponza, through flagging or drag sampling, would be needed to determine if tick species are established.

4.2. The role of migratory birds in the potential establishment of exotic tick-borne pathogens in Europe

It has been suggested that migratory birds can play a role in the establishment of new tick-borne pathogens (Hasle, 2013). In our study, we identified several tick-borne pathogens that might be in the process of establishing populations in Europe.

First, the singular *A. variegatum* nymph tested positive for *R. africae*, the causative agent of the African tick bite fever, the most widespread tick-borne rickettsiosis in sub-Saharan Africa (Pascucci et al., 2019). *Rickettsia africae* has previously been detected in *A. variegatum* ticks collected on Mediterranean islands (Cicculli et al., 2019; Pascucci et al., 2019), a *Hyalomma* spp. and an *I. ricinus* feeding on migratory birds captured in Central Italy, including Ponza Island (Toma et al., 2014). This repeated detection of this pathogen along the same migratory route from sub-Saharan Africa to Europe emphasizes the potential risk of introduction of *R. africae* due to avian migratory behaviour.

Additionally, *R. aeschlimannii* was detected in 6 of the 14 *H. rufipes* ticks collected in this study. *Rickettsia aeschlimannii* is a common *Rickettsia* in *Hyalomma* ticks, especially in *H. rufipes* and *H. marginatum*, originating from Africa and Europe (Chitimia-Dobler et al., 2019; Duscher et al., 2018). It has also been responsible for several cases of human spotted fever in the Mediterranean (Tosoni et al., 2016). The prevalence reported here for *R. aeschlimannii* (33.3 % of collected ticks) is comparable to results from previous studies on Mediterranean ticks (ranging from 20 to 48%) (Azagi et al., 2017; Toma et al., 2014; Wallménius et al., 2014). In our study, all ticks which tested positive for *R. aeschlimannii* were feeding on migratory birds overwintering in

sub-Saharan Africa (*S. communis, A trivialis, Oenanthe oenanthe,* and *Saxicola rubetra*). We can hypothesize here that long-distance migratory birds could contribute to the spreading of *R. aeschlimannii*-infected *Hyalomma* ticks, which have been reported as far north as Germany feeding on various mammals (Chitimia-Dobler et al., 2019).

Another rare finding in the present study were the four *H. rufipes* ticks which tested positive for *B. burgdorferi* s.l.. Unfortunately, none of these samples produced a sufficient PCR product for species identification due, most likely, to low concentrations of *Borrelia* DNA in the tick sample. Various *Borrelia* species occurring in Europe are adapted to birds (Margos et al., 2011). *Borrelia burgdorferi* s.l. is usually vectored by species of the genus *Ixodes* (Gern, 2008) and studies reporting *B. burgdorferi* s.l. in *Hyalomma* species are rare and were always from Mediterranean countries (Güner et al., 2004; Orkun et al., 2014; Toma et al., 2014) and sub-Saharan Africa (Diarra et al., 2017; Ehounoud et al., 2016). The only other report of *B. burgdorferi* s.l. in *H. rufipes* was from migratory birds in Central Italy in several collection sites, including the island of Ponza (Toma et al., 2014).

Two birds in our sample were found to carry two ticks each, both positive for *Rickettsia* which opens the way for further research on the influence of co-feeding transmission on the spread of these pathogens.

4.3. The distribution of Francisella-like tick endosymbionts

All *H. rufipes* specimens collected in our study tested positive for *Francisella*-like endosymbionts (FLEs). FLEs were originally detected in ticks of the genus *Dermacentor* and are related to *F. tularensis*, the agent of tularemia in humans (Ahantarig et al., 2013). Their role as endosymbionts is thought to be related to diet complementation (Gerhart et al., 2016). The presence of FLEs has already been described in *Hyalomma* spp. from different locations including Sardinia island (Brinkmann et al., 2019; Chisu et al., 2019; Ivanov et al., 2011; Montagna et al., 2012; Szigeti et al., 2014; Wang et al., 2018) as well as in one *H. rufipes* collected in Ethiopia (Szigeti et al., 2014). The incidence observed in our study (100 %) is in accordance with Azagi et al. (2017) (Israel 90 %) but not Montagna et al. (2012) (Yemen 10 %) and suggests that FLE are obligate endosymbionts of *H. rufipes*.

5. Conclusions

In this study, we confirm the role of migratory bird in the spread of exotic tick species and tick-borne pathogens and find them along the same migratory route as in previous studies. The Mediterranean islands are confirmed as a major entry point of these pathogens into Europe. We suggest continuous monitoring of these regions will help evaluate whether this should raise public health concerns.

Ethics statement

Ringing was conducted under permit from the Regione Lazio (Determinazione Dirigenziale B0332/06; B0084/09; A12042/11; G00575/15; and G00668/18) and further experimental use (Permissions No. G17371) following European and National standards for animal welfare.

Author contributions

A.F.T.S., R.E.R., & N.S.B. conceived the project. A.F.T.S. & M.C. captured the birds and collected the tick samples. L.C.D. identified the ticks and organized all further testing. L.C.D, S.S., C.K., D.F., A.S., C.S., D.B., & R.E.R. worked on molecular and in silico analysis of tick samples. L.C.D, R.E.R., and N.S.B wrote the manuscript with final input and approval from all co-authors.

Funding

Bird capture was part of the "Small Island Project" and funded in part by the Max Plank Research Society, Germany and by the University of Veterinary Medicine Vienna. *Borrelia* molecular testing was funded through the German Research Foundation (DFG Grant No. BE 5791/2-1). Internal Bundeswehr funding was used for tick identification, TBE virus screening, CCHF virus, *Rickettsia* spp.

Research data

Tick 16S sequences were uploaded to the GenBank database (accession numbers: MT374742-MT374759). All *Rickettsia* 23S-5S intergenic spacer sequences were uploaded to the GenBank database (accession numbers: MT374762-MT374768).

CRediT authorship contribution statement

Robert E. Rollins: Conceptualization, Investigation, Formal analysis, Writing - original draft. **Sabine Schaper:** Investigation. **Claudia Kahlhofer:** Investigation. **Dimitrios Frangoulidis:** Investigation, Formal analysis. **Aurelia F.T. Strauß:** Conceptualization, Investigation. **Massimiliano Cardinale:** Investigation. **Andrea Springer:** Investigation, Formal analysis. **Christina Strube:** Investigation. **Deon K. Bakkes:** Investigation, Formal analysis. **Noémie S. Becker:** Conceptualization, Supervision, Writing - review & editing. **Lidia Chitimia-Dobler:** Supervision, Investigation, Formal analysis, Writing - original draft.

Declaration of Competing Interest

The authors do not have any conflicts of interests to declare.

Acknowledgements

We are thankful to the many ringers and volunteers of the CISCA Team, and to the personnel of Castel Porziano, Presidenza della Repubblica Italia, who helped with bird capture and data collection on Ponza during the year of the study. This is publication N. 71 of the Piccole Isole Project. The study complies with current Italian laws and was performed under permit from the Regione Lazio. We would like to thank the National Reference Center for Borrelia of the Bavarian Food and Health Safety Authority in Munich, Germany, for use of lab space.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2020.101590.

References

Ahantarig, A., Trinachartvanit, W., Baimai, V., Grubhoffer, L., 2013. Hard ticks and their bacterial endosymbionts (or would be pathogens). Folia Microbiol. (Praha). 58, 419–428.

Albanese, M., Bruno-Smiraglia, C., Lavagnino, A., 1971. Notizie sulle zecche di Sicilia con segnalazione di *Hyalomma detritum* e *Amblyomma variegatum*. Riv. Parassitol. 32, 273–276.

Al-Deeb, M.A., Frangoulidis, D., Walter, M.C., Kömpf, D., Fischer, S.F., Petney, T., Muzaffar, S.Bin, 2016. *Coxiella*-like endosymbiont in argasid ticks (*Ornithodoros muesebecki*) from a Socotra Cormorant colony in Umm Al Quwain, United Arab Emirates. Ticks Tick. Dis. 7, 166–171.

Apanaskevich, D.A., Horak, I.G., 2008. The genus *Hyalomma* Koch, 1844: V. Re-evaluation of the taxonomic rank of taxa comprising the *H. (Euhyalomma) marginatum* Koch complex of species (Acari: ixodidae) with redescription of all parasitic stages and notes on biology. Int. J. Acarol. 34, 13–42.

Atkinson, B., Chamberlain, J., Logue, C.H., Cook, N., Bruce, C., Dowall, S.D., Hewson, R., 2012. Development of a real-time RT-PCR assay for the detection of Crimean-Congo hemorrhagic fever virus. Vector-Borne Zoonotic Dis. 12, 786–793.

Azagi, T., Klement, E., Perlman, G., Lustig, Y., Mumcuoglu, K.Y., 2017. *Francisella*-Like Endosymbionts and *Rickettsia* Species in Local and Imported *Hyalomma* Ticks, 83, pp. 1–14.

Bonnet, S.I., Binetruy, F., Hernández-Jarguín, A.M., Duron, O., 2017. The tick microbiome: why non-pathogenic microorganisms matter in tick biology and pathogen transmission. Front. Cell. Infect. Microbiol. 7, 1–14.
Brinkmann, A., Hekimoğlu, O., Dinçer, E., Hagedorn, P., Nitsche, A., Ergünay, K., 2019. A cross-sectional screening by next-generation sequencing reveals Rickettsia, Coxiella, Francisella, Borrelia, babesia, Theileria and Hemolivia species in ticks from Anatolia. Parasit. Vectors 12, 1–13.
Buczek, A.M., Buczek, W., Buczek, A., Bartosik, K., 2020. The potential role of migratory birds in the rapid spread of ticks and tick-borne pathogens in the changing climatic and environmental conditions in Europe. Int. J. Environ. Res. Public Health 17, 1–12.
Casati, S., Sager, H., Gern, L., Piffaretti, J.-C., 2006. Presence of potentially pathogenic *Babesia* sp. For human in *Ixodes ricinus* in Switzerland. Ann. Agric. Environ. Med. 13, 65–70.
Cecere, J.G., Matricardi, C., Frank, B., Imperio, S., Spina, F., Gargallo, G., Barboutis, C., Boitani, L., 2010. Nectar exploitation by songbirds at Mediterranean stopover sites. Ardeola 57, 143–147.
Chisu, V., Foxi, C., Masala, G., 2019. First molecular detection of *Francisella*-like endosymbionts in *Hyalomma* and *Rhipicephalus* tick species collected from vertebrate hosts from Sardinia island, Italy. Exp. Appl. Acarol. 79, 245–254.
Chitimia-Dobler, L., Rieß, R., Kahl, O., Wölfel, S., Dobler, G., Nava, S., Estrada-Peña, A., 2018. Ixodes inopinatus – Occurring also outside the Mediterranean region. Ticks Tick. Dis. 9, 196–200.
Chitimia-Dobler, L., Schaper, S., Rieß, R., Bitterwolf, K., Frangoulidis, D., Bestehorn, M., Springer, A., Oehme, R., Drehmann, M., Lindau, A., Mackenstedt, U., Strube, C., Dobler, G., 2019. Imported *Hyalomma* ticks in Germany in 2018. Parasit. Vectors 12, 1–9.
Cicculli, V., de Lamballerie, X., Charrel, R., Falchi, A., 2019. First molecular detection of *Rickettsia africae* in a tropical bont tick, *Amblyomma variegatum*, collected in Corsica, France. Exp. Appl. Acarol. 77, 207–214.
Comstedt, P., Bergstrom, S., Olsen, B., Garpmo, U., Marjavaara, L., Mejlon, H., Barbour, A.G., Bunikis, J., 2006. Migratory passerine birds as reserviors of Lyme borreliosis in Europe. Emerg. Infect. Dis. 12, 1087–1095.
Diarra, A.Z., Almeras, L., Laroche, M., Berenger, J.M., Koné, A.K., Bocoum, Z., Dabo, A., Doumbo, O., Raoult, D., Parola, P., 2017. Molecular and MALDI-TOF identification of ticks and tick-associated bacteria in Mali. PLoS Negl. Trop. Dis. 11, 1–25.
Duscher, G.G., Hodžić, A., Hufnagl, P., Wille-Piazzai, W., Schötta, A.M., Markowicz, M. A., Estrada-Peña, A., Stanek, G., Allerberger, F., 2018. Adult *Hyalomma marginatum* tick positive for *Rickettsia aeschlimannii* in Austria, October 2018. Eurosurveillance 23, 1800595.
Ehounoud, C.B., Yao, K.P., Dahmani, M., Achi, Y.L., Amanzougaghene, N., Kacou N'Douba, A., N'Guessan, J.D., Raoult, D., Fenollar, F., Mediannikov, O., 2016. Multiple pathogens including potential new species in tick vectors in Côte d'Ivoire. PLoS Negl. Trop. Dis. 10, 1–18.
Frangoulidis, D., Meyer, H., Kahlhofer, C., Splettstoesser, W.D., 2012. Real-time" PCR-based detection of *Coxiella burnetii* using conventional techniques. FEMS Immunol. Med. Microbiol. 64, 134–136.
Georgopoulou, I., Tsiouris, V., 2008. The potential role of migratory birds in the transmission of zoonoses. Vet. Ital. 44, 671–677.
Gerhart, J.G., Moses, A.S., Raghavan, R., 2016. A *Francisella*-like endosymbiont in the Gulf Coast tick evolved from a mammalian pathogen. Sci. Rep. 6, 1–6.
Grandi, G., Chitimia-Dobler, L., Choklikitumnuey, P., Strube, C., Springer, A., Albihn, A., Jaenson, T.G.T., Omazic, A., 2020. First records of adult Hyalomma marginatum and H. Rufipes ticks (Acari: ixodidae) in Sweden. Ticks Tick. Dis. 11, 101403.
Guglielmone, A.A., Apanaskevich, D.A., Estrada-Peña, A., Robbins, R.G., Petney, T.N., Horak, I.G., 2014. The Hard Ticks of the World: (Acari: Ixodida: Ixodidae). Springer, Berlin.
Güner, E.S., Watanabe, M., Hashimoto, N., Kadosaka, T., Kawamura, Y., Ezaki, T., Kawabata, H., Imai, Y., Kaneda, K., Masuzawa, T., 2004. *Borrelia turcica* sp. nov., isolated from the hard tick *Hyalomma aegyptium* in Turkey. Int. J. Syst. Evol. Microbiol. 54, 1649–1652.
Hansford, K.M., Carter, D., Gillingham, E.L., Hernandez-Triana, L.M., Chamberlain, J., Cull, B., McGinley, L., Paul Phipps, L., Medlock, J.M., 2019. *Hyalomma rufipes* on an untraveled horse: is this the first evidence of *Hyalomma* nymphs successfully moulting in the United Kingdom? Ticks Tick. Borne. Dis. 10, 704–708.
Hasle, G., 2013. Transport of ixodid ticks and tick-borne pathogens by migratory birds. Front. Cell. Infect. Microbiol. 3, 1–6.
Hillyard, P.D., 1996. Ticks of North-West Europe. Feild Studies Council, Schrewsbury, London.
Hoffman, T., Lindeborg, M., Barboutis, C., Erciyas-Yavuz, K., Evander, M., Fransson, T., Figuerola, J., Jaenson, T.G.T., Kiat, Y., Lindgren, P.E., Lundkvist, Å., Mohamed, N., Moutailler, S., Nyström, F., Olsen, B., Salaneck, E., 2018. Alkhurma hemorrhagic fever virus RNA in *Hyalomma rufipes* ticks infesting migratory birds, Europe and Asia Minor. Emerg. Infect. Dis. 24, 879–882.
Ivanov, I.N., Mitkova, N., Reye, A.L., Hübschen, J.M., Vatcheva-Dobrevska, R.S., Dobreva, E.G., Kantardjiev, T.V., Muller, C.P., 2011. Detection of new *Francisella*-like tick endosymbionts in *Hyalomma* spp. and *Rhipicephalus*spp. (Acari: Ixodidae) from Bulgaria. Appl. Environ. Microbiol. 77, 5562–5565.
Jado, I., Escudero, R., Gil, H., Jiménez-Alonso, M.I., Sousa, R., García-Pérez, A.L., Rodríguez-Vargas, M., Lobo, B., Anda, P., 2006. Molecular method for identification of Rickettsia species in clinical and environmental samples. J. Clin. Microbiol. 44, 4572–4576.
Jaenson, T.G.T., Tälleklint, L., Lundqvist, L., Olsen, B., Chirico, J., Mejlon, H., 1994. Geographical distribution, host associations, and vector roles of ticks (Acari: ixodidae, argasidae) in Sweden. J. Med. Entomol. 31, 240–256.
Jameson, L.J., Morgan, P.J., Medlock, J.M., Watola, G., Vaux, A.G.C., 2012. Importation of *Hyalomma marginatum*, vector of Crimean-Congo haemorrhagic fever virus, into the United Kingdom by migratory birds. Ticks Tick. Dis. 3, 95–99.
Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.
Lempereur, L., Beck, R., Fonseca, I., Marques, C., Duarte, A., Santos, M., Zúquete, S., Gomes, J., Walder, G., Domingos, A., Antunes, S., Baneth, G., Silaghi, C., Holman, P., Zintl, A., 2017. Guidelines for the detection of *Babesia* and *Theileria* parasites. Vector-Borne Zoonotic Dis. 17, 51–65.
Maggini, I., Trez, M., Cardinale, M., Fusani, L., 2020. Stopover dynamics of 12 passerine migrant species in a small Mediterranean island during spring migration. J. Ornithol. 161, 793–802.
Mangold, A.J., Bargues, M.D., Mas-Coma, S., 1998. Mitochondrial 16S rDNA sequences and phylogenetic relationships of species of *Rhipicephalus* and other tick genera among Metastriata (Acari: ixodidae). Parasitol. Res. 84, 478–484.
Margos, G., Gatewood, A.G., Aanensen, D.M., Hanincová, K., Terekhova, D., Vollmer, S. A., Cornet, M., Piesman, J., Donaghy, M., Bormane, A., Hurn, M.A., Feil, E.J., Fish, D., Casjens, S., Wormser, G.P., Schwartz, I., Kurtenbach, K., 2008. MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. Proc. Natl. Acad. Sci. U. S. A. 105, 8730–8735.
Margos, G., Vollmer, S.A., Ogden, N.H., Fish, D., 2011. Population genetics, taxonomy, phylogeny and evolution of *Borrelia burgdorferi* sensu lato. Infect. Genet. Evol.
Montagna, M., Chouaia, B., Pella, F., Mariconti, M., Pistone, D., Fasola, M., Epis, S., 2012. Screening for bacterial DNA in the hard tick Hyalomma marginatum (Ixodidae) from Socotra Island (Yemen): detection of Francisella-like endosymbiont. J. Entomol. Acarol. Res. 44, 60–63.
Orkun, Ö., Karaer, Z., Çakmak, A., Nalbantoğlu, S., 2014. Identification of tick-borne pathogens in ticks feeding on humans in Turkey. PLoS Negl. Trop. Dis. 8, e3067.
Papadopoulos, B., Morel, P.C., Aeschlimann, A., 1996. Ticks of domestic animals in the Macedonia region of Greece. Vet. Parasitol. 63, 25–40.
Pascucci, I., Di Domenico, M., Capobianco Dondona, G., Di Gennaro, A., Polci, A., Capobianco Dondona, A., Mancuso, E., Cammà, C., Savini, G., Cecere, J.G., Spina, F., Monaco, F., 2019. Assessing the role of migratory birds in the introduction of ticks and tick-borne pathogens from African countries: an Italian experience. Ticks Tick. Dis. 10, 101272.
Pérez-Eid, C., 2007. Les Tiques: Identification, Biologie, Importance Médicalvétérinaire. Lavoisier, Paris.
Schwaiger, M., Cassinotti, P., 2003. Development of a quantitative real-time RT-PCR assay with internal control for the laboratory detection of tick borne encephalitis virus (TBEV) RNA. J. Clin. Virol. 27, 136–145.
Strube, C., Montenegro, V.M., Epe, C., Eckelt, E., Schnieder, T., 2010. Establishment of a minor groove binder-probe based quantitative real time PCR to detect *Borrelia burgdorferi* sensu lato and differentiation of *Borrelia spielmanii* by ospA-specific conventional PCR. Parasit. Vectors 3, 1–10.
Szigeti, A., Kreizinger, Z., Hornok, S., Abichu, G., Gyuranecz, M., 2014. Detection of *Francisella*-like endosymbiont in *Hyalomma rufipes* from Ethiopia. Ticks Tick. Dis. 5, 818–820.
Toma, L., Mancini, F., Di Luca, M., Cecere, J.G., Bianchi, R., Khoury, C., Quarchioni, E., Manzia, F., Rezza, G., Ciervo, A., 2014. Detection of microbial agents in ticks collected from migratory birds in central Italy. Vector-Borne Zoonotic Dis. 14, 199–205.
Tosoni, A., Mirijello, A., Ciervo, A., Mancini, F., Rezza, G., Damiano, F., Cauda, R., Gasbarrini, A., 2016. Human *Rickettsia aeschlimannii* infection: first case with acute hepatitis and review of the literature. Eur. Rev. Med. Pharmacol. Sci. 20, 2630–2633.
Voltzit, O.V., Keirans, J.E., 2003. A review of African *Amblyomma* species (Acari, Ixodida, Ixodidae). Acarina 11, 135–214.
Wallménius, K., Barboutis, C., Fransson, T., Jaenson, T.G., Lindgren, P.E., Nyström, F., Olsen, B., Salaneck, E., Nilsson, K., 2014. Spotted fever Rickettsia species in Hyalomma and Ixodes ticks infesting migratory birds in the European Mediterranean area. Parasit. Vectors 7, 1–20.
Wang, Y., Mao, L., Sun, Y., Wang, Z., Zhang, Jiayong, Zhang, Jibo, Peng, Y., Xia, L., 2018. A novel *Francisella*-like endosymbiont in *Haemaphysalis longicornis* and *Hyalomma asiaticum*, China. Vector-Borne Zoonotic Dis. 18, 669–676.
Wölfel, R., Essbauer, S., Dobler, G., 2008. Diagnostics of tick-borne rickettsioses in Germany: a modern concept for a neglected disease. Int. J. Med. Microbiol. 298, 368–374.

Paper 1: Supplementary Material

Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry

Robert E. Rollins, Sabine Schaper, Claudia Kahlhofer, Dimitrios Frangoulidis, Aurelia F. T. Straußß, Massimiliano Cardinale, Andrea Springer, Christina Strube, Deon K. Bakkes, Noémie S. Becker, Lidia Chitimia-Dobler

Ticks and Tick-borne Diseases (2021), 12(1): 101590
https://doi.org/10.1016/j.ttbdis.2020.101590

Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry

Ticks and Tick-borne Disease

Robert E. Rollins, Sabine Schaper, Claudia Kahlhofer, Dimitrios Frangoulidis, Aurelia F. T. Strauß, Massimiliano Cardinale, Andrea Springer, Christina Strube, Deon K. Bakkes, Noémie S. Becker, Lidia Chitimia-Dobler

--Online Supplementary Material—

Supplemental Methods

Tick collection and identification

Ticks were collected from migratory birds trapped in mist nets on Ponza Island (40°55′ N, 12°58′ E) during spring migration from March until May 2019. Ponza is the largest island of the Italian Pontine Islands archipelago, located 33 km (21 nautical miles) south of Cape Circeo in the Tyrrhenian Sea. A ringing station (www.inanellamentoponza.it) operates here throughout the spring season (March-May) since 2002 in connection with the Italian "Small Island Project" (102). Birds were captured with mist-nets spanning 340m that were checked every hour from sunrise to sunset each day. Nets were not checked on days with rain or strong winds. After being captured, birds were ringed and measured using standard procedures (103). If a tick was noticed at this stage it was collected immediately (n = 9). Additionally, 142 birds were systematically checked for ticks on their heads (104) and all ticks were collected from infested birds using fine tweezers (n = 7).

Nucleic acid extraction and PCR

Identification of tick species is difficult and is complicated by misidentified sequences on GenBank. To remedy this, a phylogenetic analysis of a species comprehensive dataset was done to ensure thorough comparison between species and confidence in identifications. Tick sequences were compiled into a dataset of 57 sequences with GenBank references for *H. dromedarii*, *H. somalicum*, *H. impeltatum*, *H. isaaci*, *H. rufipes, H. marginatum, H. anatolicum*, *H. truncatum*, *H.*

albiparmatum, *H. impressum*, *H. scupense*, *H. schulzei*, *R. appendiculatus*, *I. ventalloi*, *I. festai*, *A. hebraeum* and *A. variegatum* and aligned in MAFFT (Q-INS-i, 200PAM / k=2, Gap opening penalty: 1.53) (105). Alignments were analyzed under the maximum likelihood criterion in MEGA v7.0.14 (106) with 1000 bootstrap replicates, using an K3Pu+F+G4 nucleotide substitution model determined using W-IQ-TREE (107).

Rickettsia sequences were compiled into a dataset of 24 sequences with GenBank references for *R. aeschlimanni*, *R. africae*, *R. conorii* and *R. sibirica*. These data were aligned and analyzed using the same procedure as above but using an HKY+F+G4 nucleotide substitution model. Pairwise distances were calculated using maximum composite likelihood (G=4) in MEGA 10.1.7 (106).

Supplemental Table 1. Overview of screening methods used for all tick-borne pathogens in this study.

Pathogen	Method	Positive control	Primer	Sequence (5' to 3')	Citation
***Babesia*/ *Theileria* spp.**	PCR	*Babesia divergens* (isolate 564)	*BJ1* *BN2*	GTCTTGTAATTGGAATGATGG TAGTTTATGGTTAGGACTACG	**Casati et al. (2006)**
***Rickettsia* spp.**	RT PCR	*Rickettsia helvetica* (AS 819)	*PanRick_2_for* *PanRick_2_rev* *PanRick_3_taq*	ATAGGACAACACCGTTTATTT CAAACATCATATGCAGAAA FAM-CCTGATAATTCGTTAGATTTTACCG–TMR3	**Wölfel et al. (2008)**
	PCR	*Rickettsia helvetica* (AS 819)	*23S_for* *23S_rev*	GATAGGTCGGGTGTGGAAGCAC GGGA132TGGGATCGTGTGTTTCAC	**Chitimia-Dobler et al. (2018) & Jado et al. (2006)**
***Borrelia burgdorferi* sensu lato**	RT-PCR	*Borrelia mayonii* (DSM 102811.MN14-1420)	*IGS-MGB for* *IGS-MGB rev* *IGS-MGB probe*	TCC TAG GCA TTC ACC ATA GAC T TGG CAA AAT AGA GAT GGA AGA T 6-FAM-ATT ACT TTG ACC ATA TTT-MGBNFQ	**Strube et al. (2010)**
	PCR	*Borrelia mayonii* (DSM 102811.MN14-1420)	*recF917* *recR1658* *recF890* *recR1694*	CTTTAATTGAAGCTGGATATC CAAGTTGCATTTGGACAATC CCCTTGTTGCCTTGCTTTC GAAAGTCCAAAACGCTCAG	**Margos et al. (2008)**
CCHF virus	RT-PCR	CCHF IVT Hewson	*CCHF S1* *CCHF S122* *CCHF probe*	TCTCAAAGAAACACGTGCC CCTTTTTGAACTCTTCAAACC FAM-ACTCAAGGKAACACTGTGGGCGTAAG-BHQ1	**Atkinson et al. (2012)**
TBE virus	RT-PCR	Langat Virus	*TBEV_for* *TBEV_rev*	GGGCGGTTCTTGTTCTCC ACACATCACCTCCTTGTCAGACT	**Schwaiger & Cassinotti (2003)**
***Coxiella* spp./*Coxiella*-like endosymbionts**	RT-PCR	*Coxiella burnetii* (Nine Mile Strain RSA 493)	*CB_S4k* *CB_A2k* *CB_LCR* *CB_Flu*	GAAACGGGTGTTTGAATTGTTTG TCACA TTGCCGCGTTTACT LC Red640-TAATCACCAATCGCTTCGTCCCGGT GCCACCGCTTTTAATTCCTCCTC–FL	**Frangoulidis et al. (2012)**
***Francisella*-like endosymbionts**	PCR	-	*16S_P1L* *16S_P1R*	CTAGCGGCATGCTTAACACA CCCAGGCGGAGAACTTAAC	**Al-Deeb et al. (2016)**

Paper 2: Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010

Mercy Okeyo, Sabrina Hepner, **Robert E. Rollins**, Christina Hartberger, Reinhard K. Straubinger, Durdica Marosevic, Stephanie A. Bannister, Antra Bormane, Michael Donaghy, Andreas Sing, Volker Fingerle and Gabriele Margos

Environmental Microbiology (2020), 22:5033-5047
https://doi.org/10.1111/1462-2920.15100

environmental microbiology

Environmental Microbiology (2020) **00**(00), 00–00 doi:10.1111/1462-2920.15100

Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010

Mercy Okeyo,[1*] Sabrina Hepner,[1] Robert E. Rollins,[2] Christina Hartberger,[1] Reinhard K. Straubinger,[3] Durdica Marosevic,[1] Stephanie A. Bannister,[4] Antra Bormane,[5] Michael Donaghy,[6] Andreas Sing,[1] Volker Fingerle[1] and Gabriele Margos[1]

[1]*Bavarian Health and Food Safety Authority, German National Reference Centre for Borrelia, Veterinärstr. 2, 85764 Oberschleissheim, Germany.*

[2]*Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany.*

[3]*Bacteriology and Mycology, Institute for Infectious Diseases and Zoonoses, Faculty of Veterinary Medicine, LMU Munich, Veterinärstraße 13, 80539 Munich, Germany.*

[4]*Binx Health, Trowbridge, WILTS BA14 0XG, UK.*

[5]*Center for Disease Prevention and Control, Infectious Diseases Surveillance and Immunization Unit, Duntes iela 22-4, Riga LV-1005, Latvia.*

[6]*Department of Clinical Neurology, Green Templeton College, The University of Oxford, Oxford, UK.*

Summary

Members of the *Borrelia burgdorferi* sensu lato (s.l.) species complex are known to cause human Lyme borreliosis. Because of longevity of some reservoir hosts and the *Ixodes* tick vectors' life cycle, long-term studies are required to better understand species and population dynamics of these bacteria in their natural habitats. Ticks were collected between 1999 and 2010 in three ecologically different habitats in Latvia. We used multilocus sequence typing utilizing eight chromosomally located housekeeping genes to obtain information about species and population fluctuations and/or stability of *B. burgdorferi* s.l. in these habitats. The average prevalence over all years was 18.9%. From initial high-infection prevalences of 25.5%, 33.1% and 31.8%, from 2002 onwards the infection rates steadily decreased to 7.3%. *Borrelia afzelii* and *Borrelia garinii* were the most commonly found genospecies but striking local differences were obvious. In one habitat, a significant shift from rodent-associated to bird-associated *Borrelia* species was noted whilst in the other habitats, *Borrelia* species composition was relatively stable over time. Sequence types (STs) showed a random spatial and temporal distribution. These results demonstrated that there are temporal regional changes and extrapolations from one habitat to the next are not possible.

Introduction

Lyme borreliosis (LB) is the most common tick-borne disease in temperate regions of the northern hemisphere (Lindgren and Jaenson, 2006). The disease is endemic roughly between 40° and 60° latitude (Steere, 2001; Hubalek, 2009) including some regions in North America, for example, the Northeast, the Midwest, California and Canada (Lane and Lavoie, 1988; Fritz and Kjemtrup, 2003; Bacon *et al*., 2008; Ogden *et al*., 2008; Hoen *et al*., 2009a,b; Mechai *et al*., 2015; Schwartz *et al*., 2017), in Europe (Lindgren and Jaenson, 2006), Eastern Europe and Asia (Platonov *et al*., 2011; Takano *et al*., 2011). According to the World Health Organization, the highest incidence of LB is reported in Central European countries such as Czech Republic, Estonia, Lithuania, Slovenia, Austria, Germany and some northern countries bordering the Baltic Sea. The disease is caused by bacteria that belong to the *Borrelia burgdorferi* sensu lato (s.l.) species complex (Wang *et al*., 1999; Kurtenbach *et al*., 2006; Stanek *et al*., 2011; Hanincova *et al*., 2013). In Europe, known human pathogenic species of these bacteria include *Borrelia afzelii*, *Borrelia bavariensis*, *B. burgdorferi* sensu stricto (s.s.), *Borrelia garinii* and *Borrelia spielmanii*; with *B. afzelii* and *B. garinii* being the most common in questing ticks.

Received 22 March, 2020; revised 18 May, 2020; accepted 23 May, 2020. *For correspondence. E-mail mercyakinyi2012@gmail.com Tel.: +49 9131/6808-5615, Fax +49 9131-6808-5183, +49 1521 6384 275.

The bacteria are maintained in natural transmission cycles between reservoir hosts and ticks of the genus *Ixodes*. The four main vectors of human pathogenic LB genospecies are *Ixodes ricinus* in Europe, *Ixodes persulcatus* in Eastern Europe and Asia and *Ixodes scapularis* and *Ixodes pacificus* in North America (Rauter and Hartung, 2005; Geller *et al*., 2013; Schillberg *et al*., 2018; Gasmi *et al*., 2019). Over 100 vertebrate species are speculated to be reservoir competent hosts to *Borrelia* (Gern *et al*., 1998; Eisen and Lane, 2002; Schrauber and Ostfeld, 2002; Brisson and Dykhuizen, 2006; Hanincova *et al*., 2006; Norte *et al*., 2013; Heylen *et al*., 2017). In Europe, competent reservoir hosts include several species of rats, mice, dormouse, voles, squirrels, hedgehogs, shrews and birds including passerines and seabirds (Matuschka *et al*., 1994; Gern *et al*., 1998; Piesman and Gern, 2004; Pfaffle *et al*., 2011; Richter *et al*., 2011; Norte *et al*., 2013; Heylen *et al*., 2017).

In the past decades much progress has been made in understanding the diversity and geographic distribution of *Borrelia* species (Kurtenbach *et al*., 2006). However, many studies investigating the distribution of *Borrelia* species and strains, accumulated data for short-time periods (Etti *et al*., 2003; Ferquel *et al*., 2006; Rosef *et al*., 2009; Halos *et al*., 2010; Geller *et al*., 2013). Due to the complexity of the ecology of both, pathogen, reservoir host and vector, long-term studies are required to understand changing epidemiological patterns relating to the structure and dynamics of species and populations (Killilea *et al*., 2008).

Multilocus sequence typing (MLST) and multilocus sequence analysis (MLSA) (Postic *et al*., 2007; Margos *et al*., 2011) have become valuable tools for bacterial epidemiological and taxonomic studies (Vitorino *et al*., 2008; Ogden *et al*., 2011; Vollmer *et al*., 2013; Mechai *et al*., 2015; Coipan *et al*., 2016). In this study, we used MLST to determine the prevalence of *B. burgdorferi* s.l. species in ticks from three different Latvian habitats over a period of 11 years. The aim was to examine spatial and temporal variation and population fluctuation in different habitats.

Results

Infection prevalence of B. burgdorferi *s.l. in questing ticks from Latvia*

All ticks (n = 3165) in this study were collected by drag sampling in the three defined habitats in Latvia (Table 1, Figs S1 and S2). The highest numbers of ticks were acquired from Babite (n = 1695), followed by Kemeri (n = 802) and finally Jaunciems (n = 668; Table 2). An additional 1319 ticks collected in the years 2002, 2006 and 2007 in the same habitats and processed in a previous study (Vollmer *et al*., 2011; 2013) were included in our analyses. There were no samples available for the years 2004, 2005, 2008 and 2009.

Table 1. Number of tick DNA isolates processed at each step in the current study.

Step	Number of tick DNA isolates
Total number of ticks processed from 1999, 2000, 2001, 2003, 2010	3165
Screening PCR negative	2470
Screening PCR positive	698
Mixed infections in screening PCR	30
Samples with Ct value > 35 (not subjected to MLST PCR)	304
Samples with Ct value ≤ 35 (subjected to MLST PCR)	364
No PCR product in MLST PCR (untypeable)	82
At least one MLST gene amplified and sequenced	282
Mixed infections after MLST PCR and sequencing	55
<8 MLST genes amplified (spatio-temporal analysis)	156
All 8 MLST genes amplified and sequenced (goeBURST and spatio-temporal analysis)	71

Out of a total of 4484 screened ticks, 1931 were adults, 2516 were nymphs and 37 larvae (Table 2). In total, 848/4484 ticks (18.9%) were identified as *Borrelia* positive in the screening real-time polymerase chain reaction (PCR) targeting the flagellin B (*flaB*) locus. Of the 1931 screened adults, 522 (27%) were positive and of 2516 nymphs, 326 (13.0%) were positive. Overall, adults were more likely to be *Borrelia* positive when compared to nymphs (life stage, Table 3). None of the 37 screened larvae were positive (Table 2). Babite, Jaunciems and Kemeri had mean prevalences for *B. burgdorferi* s.l. in ticks of 19.5%, 16.7% and 19.8% respectively. Ticks coming from the three sites were equally likely to be *Borrelia* positive (Table 3). However, ticks coming from different years did differ in how likely they were to be *Borrelia* positive (Table 3). Remarkably, infection prevalences declined with time in all habitats (Fig. 1).

For MLST only samples with a cycle threshold (Ct) value ≤ 35 were considered to contain sufficient *Borrelia* DNA for conventional MLST PCR. This cut-off was chosen because previous studies had shown that samples with high Ct values (35–40) contained low amounts of *Borrelia* DNA making amplification of the MLST genes unfeasible (Okeyo *et al*., 2019). Following this, 315 (Table S2) out of the 848 samples positive in real-time PCR had at least two MLST genes sequenced and could be used for *Borrelia* species determination and to test for host adaptation types of genospecies per year,

Table 2. Total number of positive ticks per year and region.

Sampling years	1999	2000	2001	2002[a]	2003	2006[a]	2007[a]	2010	Total
Total number of collected ticks									
Total per year	271	236	1133	368	883	492	459	642	4484
Total adults	271	200	678	nd	410	nd	nd	372	1931
Total nymphs		36	418	368	473	492	459	270	2516
Total larvae			37						37
Total number of positives per year (%)	69 (25.5)	78 (33.1)	360 (31.8)	53 (14.4)	141 (16.0)	48 (9.8)	52 (11.3)	47 (7.3)	848 (18.9)
Positive ticks per stage									
Adult, number positive (% of total adults)	69 (25.5)	60 (30.0)	280 (41.3)	nd	77 (18.8)	nd	nd	36 (9.7)	522 (27)
Nymph, number positives (% of total nymphs)	0 (0.0)	18 (50.0)	80 (19.1)	53 (14.4)	64 (13.5)	48 (9.8)	52 (11.3)	11 (4.1)	326 (13)
Positive ticks per region									
Babite, number of positive ticks (%)	46 (25.6)	19 (22.4)	169 (26.9)	31 (20.8)	83 (18.9)	35 (13.5)	36 (16.4)	32 (9.1)	453 (19.5)
Babite, number of positive adult ticks (%)	46 (25.6)	18 (22.4)	127 (42.8)	nd	35 (17.1)	nd	nd	23 (12.8)	250 (26.4)
Babite, number of positive nymphs (%)	0	0	42 (13.7)	31 (20.8)	48 (20.6)	35 (13.5)	36 (16.4)	9 (5.2)	201 (15.0)
Jaunciems, number of positive ticks (%)	7 (15.6)	23 (26.1)	73 (34.6)	16 (12.6)	30 (14.6)	7 (6.6)	4 (4.0)	7 (5.9)	167 (16.7)
Jaunciems, number of positive adults (%)	7 (15.6)	23 (26.1)	62 (36.7)	nd	15 (14.7)	nd	nd	6 (9.8)	113 (24.3)
Jaunciems, number of positive nymphs (%)	0	0	11 (26.2)	16 (12.6)	15 (14.4)	7 (6.6)	4 (4.0)	1 (1.8)	54 (10.1)
Kemeri, number of positive ticks (%)	16 (34.8)	36 (57.1)	118 (41.8)	6 (6.5)	28 (11.7)	6 (4.7)	12 (8.6)	8 (4.7)	230 (19.8)
Kemeri, number of positive adults (%)	16 (34.8)	18 (66.7)	91 (42.9)	nd	27 (26.2)	nd	nd	7 (5.3)	159 (30.6)
Kemeri, number of positive nymphs (%)	0	18 (50.0)	27 (38.6)	6 (6.5)	1 (0.7)	6 (4.7)	12 (8.6)	1 (2.5)	71 (11.1)

nd = no data.
[a] Data from these years were analysed by Vollmer (2010) and Vollmer *et al*. (2011).

Table 3. Model outputs of *Borrelia* presence analysis. *Borrelia* presence models a binary factor of a tick being either infected (1) or non-infected (0) with *Borrelia*.

Fixed effects	β (95% Credible Interval [CI])
Intercept	−1.17 (−1.59, −0.75)
Riga District Babite[a]	0.00 (−0.19, 0.18)
Riga Jaunciems[a]	−0.15 (−0.37, 0.08)
Life stage[b]	−0.75 (−0.95, −0.54)

Random effects	σ^2 (95% CI)
Year	0.34 (0.24, 0.46)
Residual	$0.32\pi^2$ ($0.32\pi^2$, $0.29\pi^2$)

[a] Reference category: Kemeri.
[b] Reference category: Adult.

as well as the potential carry over year effects on *Borrelia* genospecies occurrence. For 71 samples, good sequence data were obtained for all 8 MLST genes in this study, and 88 samples with all MLST sequences were obtained from a previously published study (Vollmer *et al*., 2011)

Spatio-temporal distribution of B. burgdorferi *s.l. in Latvia*

To gain insights in the geographical distribution and fluctuations of *B. burgdorferi* s.l. species in Latvia we examined the spatio-temporal distribution of a total of 315 isolates. Of these 175 were from Babite (Table S3-A), 57 from Jaunciems (Table S3-B) and 83 from Kemeri (Table S3-C). Due to insufficient data for comparison from some year-habitat combinations (i.e. from Kemeri 2002 [single sample available], Kemeri 2010 [no samples] and Jaunciems 2010 [no samples]), these were excluded from this comparison. Overall, the presence of specific *Borrelia* genospecies varied with site and year (Table 4). Spatial distribution analysis showed that *B. garinii* was very abundant in Babite making 43% of the total species analysed. *Borrelia afzelii* followed with 29%, *Borrelia valaisiana* with 17%, *B. burgdorferi* s.s. with 9%, *Borrelia lusitaniae* and *B. bavariensis* at 2% each (Fig. 2A). *Borrelia afzelii* accounted for more than three quarters (82%) of the species in Jaunciems (Fig. 2B). The remaining species were present at 7% for *B. garinii*, 9% *B. valaisiana* and *B. lusitaniae* at 2%. Even though *B. afzelii* dominated also in Kemeri (40%), *B. garinii*, *B. valaisiana* and *B. burgdorferi* s.s. were also well

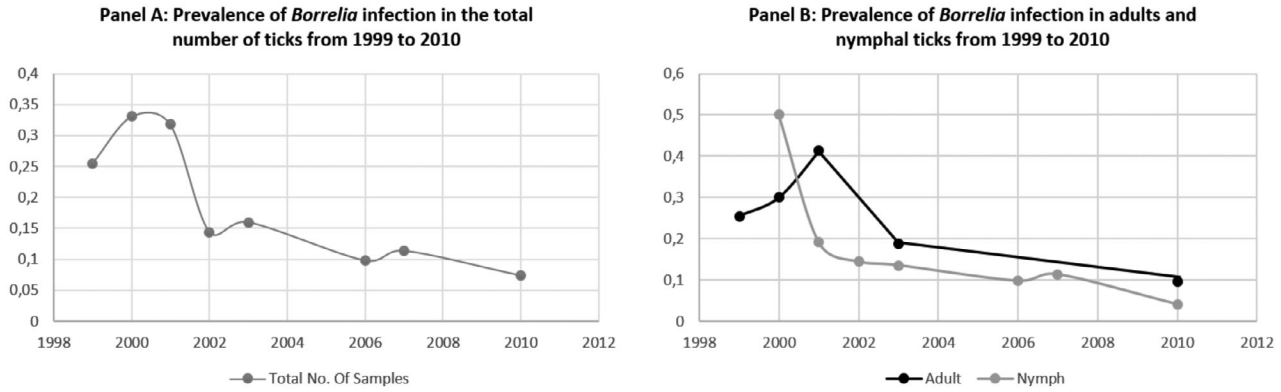


Fig 1. Prevalence of infected ticks (total number) versus adults and nymphal ticks. *Y*-axis shows proportion of all ticks that were *Borrelia* positive and *X*-axis shows the years.
A. Panel A shows the general progression of both adults and nymphal ticks from 1999 to 2010. The highest infection rate was in the year 2000, from then onwards the infection rate gradually decreased with slight increases in the years 2003 and 2007.
B. Panel B shows the infection rate of adults and nymphal ticks, the highest infection rates were in the years 2000 and 2001 for nymphs and adults respectively. In addition, a constant decrease from the peak infection years to 2010 is noticeable.

Table 4. Results of GLMM exploring impacts to the absolute number of ticks infected with specific *Borrelia* genospecies in a given year.

Fixed effects	β (95% CI)
Intercept	−0.64 (−1.44, 0.15)
Genospecies prevalence in previous sampling[a]	1.75 (1.16, 2.33)
Genospecies found in previous sampling[b]	1.22 (0.80, 1.64)

Random effects	σ^2 (95% CI)
Year	0.64 (0.47, 0.84)
Site	0.14 (0.07, 0.24)
Residual	2.63 (3.20, 2.20)

[a] Calculated per genospecies per year.
[b] Binary factor given per genospecies per year.

represented at 32%, 16%, 11% respectively (Fig. 2C). The bird-associated species *B. garinii* and *B. valaisiana* represented almost half of the total species. *Borrelia lusitaniae* was identified in a single tick in this habitat. Overall, Babite was the most species rich region with six species observed (Fig. 2A–C). Kemeri and Jaunciems had five and four species respectively.

Temporal distribution analyses showed that in Babite (Fig. S3) *B. afzelii* dominated in the years 1999 (41%), 2000 (72%) and 2002 (72%). In 2006, 2007 and 2010, *B. garinii* dominated with 61%, 48% and 57% respectively. *Borrelia valaisiana* was also well represented in the years 2006 (35%) and 2007 (36%). *Borrelia burgdorferi* s.s. was observed in the years 1999 (18%), 2000 (14%), 2001 (8%), 2003 (17%) and 2007 (12%). It appears that there was a significant shift in the distribution of host-adaptation types in Babite from rodent-adapted species to bird-adapted species over the study period (Fisher's exact test; $p = 0.002$). *Borrelia afzelii* was the dominating species in Jaunciems (Fig. S4). It was the only species identified in the years 2002, 2003, 2006 and 2007. Whilst in the years 1999, 2000 and 2001 it was dominating with 67%, 86% and 73% respectively. Other species (*B. lusitaniae*, *B. valaisiana* and *B. garinii*) were identified in small percentages (14%—33%) during that time. Correspondingly, rodent-adapted species significantly dominated this site throughout the entire study (Fisher'' exact test; $p = 0.002$). In Kemeri (Fig. S5), although species fluctuations were observed over years they did not differ significantly (Fisher's exact test; $p = 0.83$). *Borrelia afzelii* was well presented or dominated (36%–67%) in the years 2000, 2001, 2006 and 2007 but other species like *B. valaisiana*, *B. garinii* and *B. burgdorferi* s.s. were also well represented.

Combined data over all years and all habitats showed that the most prevalent species detected in Latvian ticks was *B. afzelii* (42%), followed by *B. garinii* (34%), *B. valaisiana* (15%), and *B. burgdorferi* s.s. (5%). *Borrelia bavariensis* and *B. lusitaniae* had low prevalences of about 1%. The three most common genospecies in Europe: *B. afzelii*, *B. garinii* and *B. valaisiana* accounted for 90% of the total genospecies identified over this time period in Latvia. Analysis of species distributions over years showed that prevalence of genospecies in the previous year positively correlated with the frequency of the same genospecies in the subsequent year [prevalence in previous sampling, mean: 1.75; 95% confidence interval (CI): 1.16–2.33; Table 4]. Additionally, the frequency of a genospecies was generally higher if it had been found in the previous year when compared to if it had not been

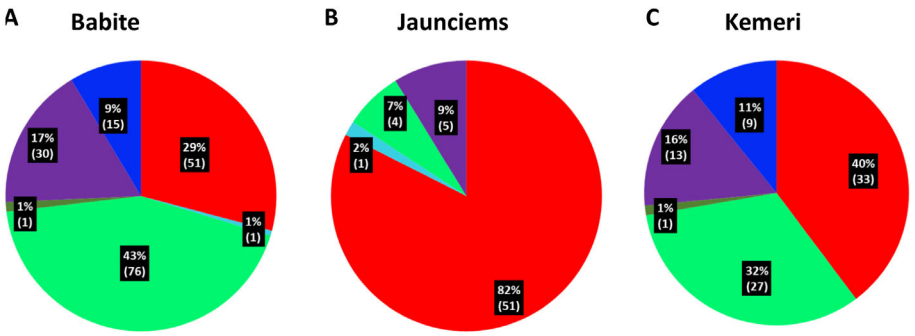


Fig 2. Pie chart showing the overall spatial distribution of *Borrelia burgdorferi* s.l. species in the three sampled habitats. The percentage of species is given; number shown in parenthesis refers to the actual number. Colours were assigned to the species as follows:
● *Borrelia afzelii*, ● *Borrelia bavariensis*, ● *B. burgdorferi* s.s., ● *Borrelia garinii*, ● *Borrelia lusitaniae* and ● *Borrelia valaisiana*
A. In Babite, six different species were identified with bird-associated species [*B. garinii* (43%) and *B. valaisiana* (17%)] dominating.
B. In Jaunciems, four different species were identified with *B. afzelii* clearly dominating this region.
C. In Kemeri, five species were identified with *B. afzelii* and *B. garinii* most commonly found.

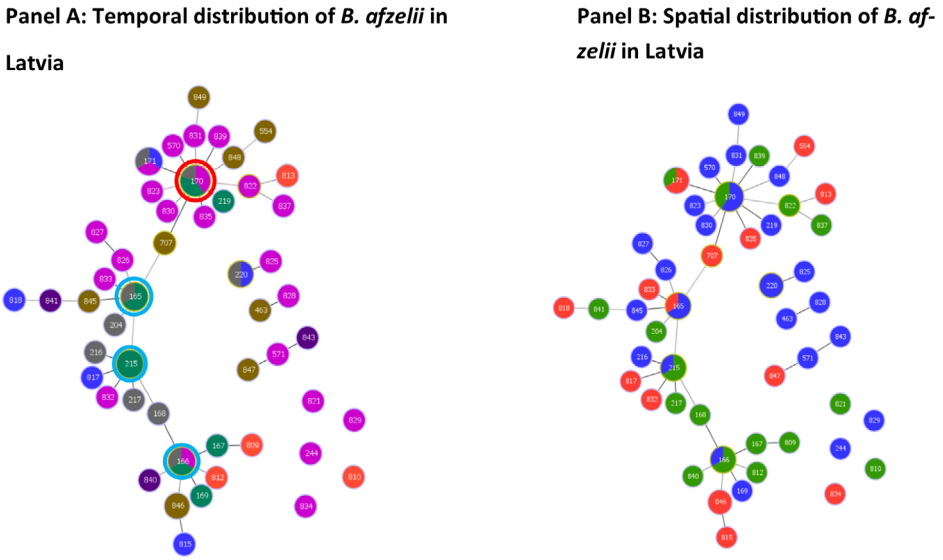


Fig 3. Population structure of *Borrelia afzelii* and the spatio-temporal distribution of STs in Lativa. The diagram shows the relationship of *B. afzelii* STs based on TLV.
A. Groupings per year. The distribution of STs from 1999 to 2007 in Latvia is indicated with different colours as follows:
● 1999, ● 2000, ● 2001, ● 2002, ● 2003, ● 2006 and ● 2007.
B. Groupings per collection region are displayed. For regions, the following colours were assigned: ● Babite, ● Jaunciems, and ● Kemeri.
One major CC was formed with founder ST170 (circled in red in A) and three sub-founders (circled blue in A). A second major CC was identified but no founder ST was recognized. There were two minor CC and five singletons.

found (found in previous sampling, mean: 1.22; 95% CI: 0.80–1.64; Table 4). Even so, genospecies frequencies varied between sites and over years (Table 4).

Identification of clonal complexes and their spatio-temporal distribution

In order to get a detailed view about the relationship of isolates and to determine if temporal fluctuations or changes within the populations occurred, that is, to see whether different sequnce types (STs) dominate at different times, a goeBURST analysis was conducted on 159 isolates (Table S1). The 159 samples were grouped into 109 STs, which means that identical STs were identified. Several major clonal complexes (CCs) and minor CCs (in which only two STs were connected) were formed in the different species (Figs 3–6). There was no evidence of specific STs dominating a habitat at specific times; rather there was arbitrary distribution of STs regardless of the sampling year for the four analysed species.

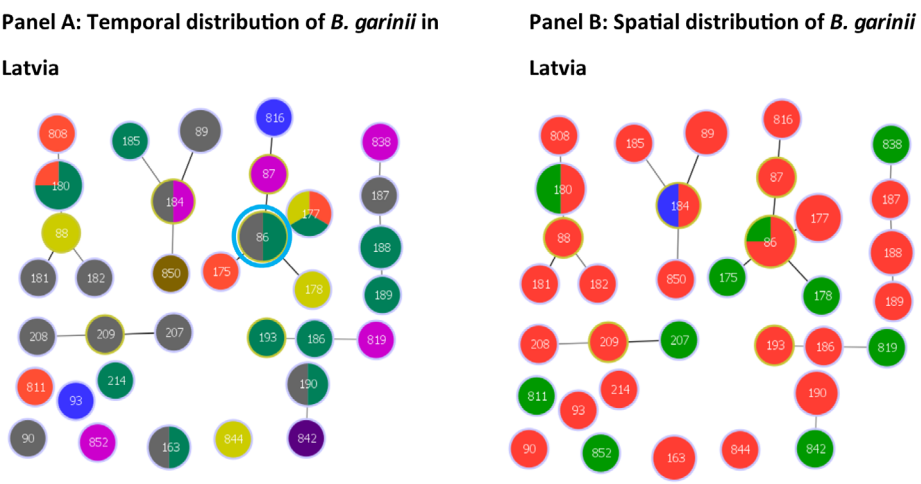


Fig 4. Shows spatio-temporal distribution of *Borrelia garinii*. There are six CC, one minor CC and seven singletons.
ST86 (circled in red in A) was identified as the founder ST to its complex. The colour-coding as in Fig. 3. The largest number of *B. garinii* STs were found in the years 2006 (teal) and 2007 (dark grey) (A).
ST177 appeared in the years 1999, 2007 and 2010 whilst ST86, ST163, ST180, ST184 and ST190 appeared in two different years each.

Borrelia afzelii samples formed two major CCs, two minor CCs and five singletons. The largest major CC (CC170) had 38 contributing STs (Table 5, Fig. 3A–B) indicating a close relatedness amongst *B. afzelii* STs. The CC founder, ST170, was identified with 99% probability of being the founder of this complex. ST165, ST166 and ST215 were classified as sub-founders with probabilities of 94%, 98% and 99% respectively. Another CC consisted of three STs: ST571 (94% probability of being the group founder), ST843 and ST847. The spatial distribution of *B. afzelii* showed that identical ST were found in Kemeri and Jaunciems (n = 3) whilst between Kemeri and Babite and Jaunciems and Babite one ST was shared (Fig. 3B), most of them occurred in different years (Fig. 3A).

Borrelia garinii formed six major CCs, one minor CC and seven singletons (Table 5, Fig. 4A–B). This may mirror the diversity of this genospecies (Vollmer *et al*., 2011; Jacquot *et al*., 2014). There were 47 sampled assigned to 34 STs (Fig. 4). One founder ST (ST86) was identified with 76% probability (Table 5). The spatial distribution of *B. garinii* STs (Fig. 4B) revealed that identical STs were found in Jaunciems and Babite (n = 1) and in Kemeri and Babite (n = 2), mostly in different years (Fig. 4B).

Isolates of *B. valaisiana* and *B. burgdorferi* s.s. were well connected (Table 5, Figs. 5 and 6). There were 17 STs of 32 *B. valaisiana* samples available which were resolved in one major CC and one minor CC (Fig. 5). One founder (probability 65%) was identified and two sub-founders (Table 5). Five STs, ST96, ST97, ST100, ST199 and ST201 were found in several years (Fig. 5A) and five in the different habitats (Fig. 5B). ST199 was found in Babite in 2002, in Kemeri in 2001 and 2007, and in Jaunciems in 2007.

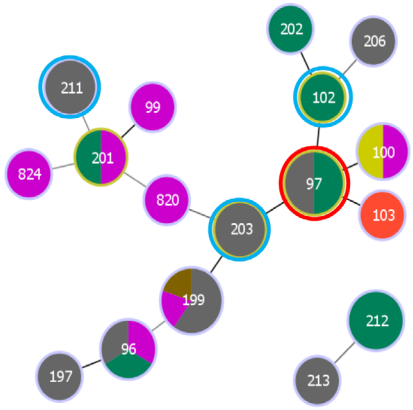
There were eight *B. burgdorferi* s.s. STs (Fig. 6). One major CC, with unrecognized sub-founder, and one minor CC were identified. The random distribution of *B. burgdorferi* s.s. STs is comparable to other species with the difference that here we had a much smaller sample size. ST1 was noticed in the years 1999 and 2000. ST20 and ST21 were noted twice in the years 1999 and 2003 respectively. ST21 was also noted in the year 2000 but just once and ST161 was noted twice in the year 2007 (Fig. 6A). *B. burgdorferi* s.s. was found in Babite and Kemeri but no identical ST was found in the two habitats (Fig. 6B).

Discussion

Prevalence of I. ricinus *ticks and* Borrelia *in Latvia*

Ixodes ticks have a long-life cycle that – depending on environmental conditions – may take up to 6 years to complete (Gray *et al*., 2016). This means that pathogens that are transmitted by ticks will remain in the environment for extended periods of time (or as long as the tick stays alive). Ticks represent one important part in transmission cycles of tick-borne pathogens (TBP). Thus, in order to gain insights into changes in species distribution and fluctuations of TBP, short-term studies are insufficient. They can produce a snapshot of the prevalence of organisms or their population structure (Etti *et al*., 2003; Ferquel *et al*., 2006; Rosef *et al*., 2009; Halos *et al*., 2010; Geller *et al*., 2013) but do not describe the stability of species/strains or their fluctuations over the years. Therefore, the aim of this study was to investigate the temporal and spatial distribution of *Borrelia* species and populations, their fluctuations and/ or stability in defined habitats in Latvia over several years. MLST was used to characterize *B. burgdorferi* s.l. in questing *I. ricinus* ticks. This is the first longitudinal study covering a period of 11 years of tick collection in Latvia in defined

Panel A: Temporal distribution of *B. valaisiana* in Latvia



Panel B: Spatial distribution of *B. valaisiana* in Latvia

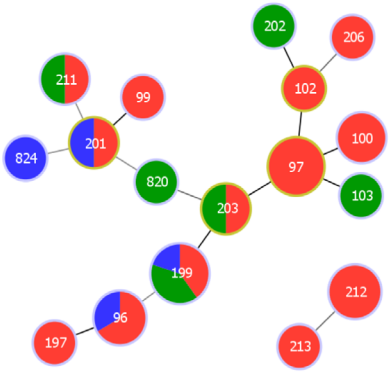
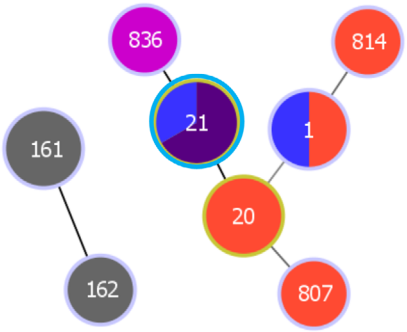


Fig 5. Shows spatio-temporal distribution of *B. valaisiana*.
One major CC with founder ST (circled red in A) and two sub-founders (blue circled in A) were formed. There was one minor CC and no singletons. The colour-coding as in Fig. 3.

Panel A: Temporal distribution of *B. burgdorferi* s.s. STs in Latvia



Panel B: Spatial distribution of *B. burgdorferi* s.s. STs in Latvia

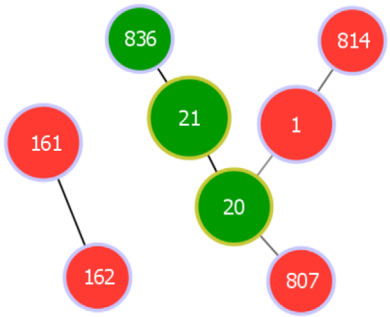


Fig 6. Shows temporal distribution of *B. burgdorferi* s.s.
One major CC with a sub-founder (circled blue) and one minor CC was identified. The colour-coding as in Fig. 3.

Table 5. CC and their respective founder and sub-founder STs.

Species	Major clonal complex	Minor clonal complex	Founder ST
Borrelia afzelii	CC170		ST170
	CC571		Not identified
		CC220	Not identified
		CC463	Not identified
Borrelia garinii	CC86		ST86
	CC180		Not identified
	CC185		Not identified
	CC187		Not identified
	CC207		Not identified
	CC193		Not identified
		CC190	Not identified
Borrelia valaisiana	CC97		ST97
		CC203	Not identified
		CC102	Not identified
Borrelia burgdorferi s.s.	CC21		ST21 (sub-founder)
		CC161	Not identified

Not identified = No founder ST was identified.

habitats and using MLST/MLSA as instrument for strain typing.

The overall average *Borrelia* infection prevalence of 18.9% determined in our study in adults and nymphs was comparable to previously published tick prevalences in Europe (Bormane *et al*., 2004; Rauter and Hartung, 2005; Strnad *et al*., 2017). Remarkably, over the years mean prevalences declined from > 30% in 2000 and 2001 to 7.3% in 2010. A prevalence of 28% for *B. burgdorferi* s.l. in the years 1999 and 2000 in the same habitats (Etti *et al*., 2003) is consistent with our data. It may be worth mentioning that the decline in tick infection prevalence with *Borrelia* was not due to the decrease of infection in one particular habitat but was registered in all three investigated habitats. Prevalence fluctuations reported in previous studies usually covered shorter periods of time (3–6 years) (Wielinga *et al*., 2006; Hoen *et al*., 2009a,b; Herrmann *et al*., 2013). One longitudinal study carried out in the Netherlands found that at one particular collection site, where ticks were collected between 2000 and 2009, the prevalence of *B. burgdorferi* s.l. was relatively stable (prevalence of 7% for *B. burgdorferi* s.l.). Except for the years 2004 and 2005, in which the highest nymphal tick densities and peak prevalences of *B. burgdorferi* s.l. were observed (20%–25%). The authors speculated that the variation in tick density may have been due fluctuations in the availability of reservoir hosts, so-called mast years. Mast years refer to when there is a high acorn production and a subsequent increase in hosts (Ostfeld *et al*., 2006). Although in these particular years, peak densities of ticks coincided with high *Borrelia* prevalence, in general, that is, when all collection sites (sampled from 2000 to 2006) were included, the authors found a negative correlation between tick densities and *Borrelia* infection prevalence (Coipan *et al*., 2013). We do not know what the underlying causes for the steady decline of *Borrelia* tick infections we observed during our study are. There are suggestions that a decline of reservoir host populations may lower the risk of tick infection (Paul *et al*., 2016). This would be consistent with studies conducted by Ostfeld *et al*. (2006). These authors found that high availability of hosts due to mast years may influence the risk for acquiring LB in humans, where the abundance of suitable hosts (mice and chipmunks) in a particular year led to increased risk for LB in the following year due to an increase in numbers of infected nymphs (Ostfeld *et al*., 2006). A steady decline of infection prevalences as observed in our study was surprising in view of reports of increasing tick infection rates due to climate and other environmental changes (Rosa *et al*., 2018). We can only speculate here that this may be due to changes in reservoir host populations as reservoir hosts may be a strong factor impacting the presence and structure of *Borrelia* populations (Etti *et al*., 2003; Kurtenbach *et al*., 2006; Ostfeld *et al*., 2006; Randolph, 2008).

Generally, the infection rate in adult ticks was higher than that in nymphs. This is consistent with previous studies on the prevalence of *Borrelia* in ticks in Europe (Rauter and Hartung, 2005; Strnad *et al*., 2017). This has been ascribed to the fact that adult ticks had an opportunity to feed on at least two hosts; assuming that both hosts were infected and that the feeding was not interrupted. In contrast, nymphs usually have only fed on one host (Kurtenbach *et al*., 2010; Estrada-Pena *et al*., 2011; Rizzoli *et al*., 2011; Strnad *et al*., 2017). Also, an efficient trans-stadial transmission of the spirochete purportedly contributes to this (Mejlon and Jaenson, 1993). In our study for the year 2000 an opposite trend was observed (Fig. 1B). Similar observations have been made in Italy, France, Slovakia and Sweden where higher infection rates in nymphs than in adults were described (Aureli *et al*., 2015; Chvostáč *et al*., 2018; Ehrmann *et al*., 2018; Akl *et al*., 2019). Since the interactions between tick, host and pathogen are highly complex and many factors could be involved (some of which may work in synergy) several explanations are possible (Randolph, 2008; Randolph, 2009; van Duijvendijk *et al*., 2015). Amongst them the number of incompetent hosts on which nymphs of the previous year may have fed leading to an overall reduction in prevalence in the adult population in the subsequent year (dilution effect) (Jouda *et al*., 2004; Nahimana *et al*., 2004; Kurtenbach *et al*., 2006), strong temporal variation in tick population (Randolph, 2008) or unusual high transovarial transmission of *Borrelia* (van Duijvendijk *et al*., 2016). We also cannot discount the possibility that the observed high infection rate in nymphs may have been due to sampling bias as in the particular year the number of collected nymphs was very low.

Borrelia species depend on vector and host for survival (Kurtenbach *et al*., 2006; Margos *et al*., 2011; Radolf *et al*., 2012). Accordingly, fluctuation/stability of *Borrelia* species may depend on many factors although these may not directly impact the bacteria themselves. It is well described that numerous abiotic and biotic conditions directly influence ticks and reservoir hosts in their natural habitats and this will – indirectly – affect TBP and contribute to fluctuation that can be observed (Randolph, 2004; 2008; Kurtenbach *et al*., 2006; Pfaffle *et al*., 2013; Paul *et al*., 2016). Due to the physiological requirements of ticks, their abundance in any given habitat may be very sensitive to climatic or microclimatic changes (Kurtenbach *et al*., 2006; Randolph, 2008; 2009; Margos *et al*., 2011) but host availability may have a stronger impact on fluctuations of tick and TBP populations than climatic changes (Ostfeld *et al*., 2006; Randolph, 2008). Apart from the year 2001 where numbers of collected ticks were almost twice or three times

that of other years, the observed decline in *Borrelia* infection rates in our study was not accompanied by dramatic shifts or decline in tick abundance over the years which are known to show strong temporal variability (Randolph, 2008). It can also be assumed that in a given habitat the effect of climate or other environmental changes on tick abundance similarly affect the tick population as a whole (Randolph, 2008; 2009). Thus, we propose that the pattern of presences of *Borrelia* species in the three habitats investigated here and that we observed in our study were mainly due to the impact of fluctuations/stability of *Borrelia* reservoir hosts (Randolph, 2008).

In Europe, the pronounced host associations of *Borrelia* species have been attributed to the ability of species to resist host complement. *Borrelia garinii* and *B. valaisiana* are adapted to birds as reservoir hosts whilst *B. afzelii* is associated with rodent hosts (Hu *et al*., 1997; Kurtenbach *et al*., 1998; Taragel'ova *et al*., 2008; Dubska *et al*., 2009; Norte *et al*., 2013; Heylen *et al*., 2017). These are also the most commonly identified species in Europe which are geographically unequally distributed (Kurtenbach *et al*., 2006; Strnad *et al*., 2017; Estrada-Pena *et al*., 2018). Expectedly, these were also the most common species identified in the studied habitats in Latvia.

Species distribution in habitats

The species composition differed markedly between the investigated habitats. In Babite, the sylvatic habitat, six species were detected: *B. afzelii*, *B. bavariensis*, *B. burgdorferi* s.s., *B. garinii*, *B. lusitaniae* and *B. valaisiana*. Interestingly, in this habitat, a significant shift in *Borrelia* species was observed with time. The rodent-adapted species *B. afzelii* was the most frequently detected species in 1999, 2000 and 2002. Afterwards the bird-adapted species *B. garinii* became the dominant species. Particularly, in sampling years 2006 and 2007 the bird-adapted species *B. garinii* and *B. valaisiana* dominated suggesting a shift in the reservoir host composition in this particular habitat from competent rodent reservoir hosts of *B. afzelii* to bird populations that are reservoir hosts for *B. garinii* and *B. valaisiana*. This striking and significant shift was contrasted by a more stable *Borrelia* species composition in the other habitats. In Kemeri four *Borrelia* species were found and – although some variation occurred in species abundance from year to year – there was no significant shift in *Borrelia* species composition suggesting relative stability of bird and rodent reservoir host populations. In addition, prevalence and presence of genospecies in a previous year positively correlated to the matching genospecies frequency in the following year, supporting that carry over of genospecies in general is plausible within a given site. However, we do observe, over long time periods, that species can go extinct or re-emerge locally, which would explain the site and year fluctuations in genospecies frequency. Stability was also observed in Jaunciems, where *B. afzelii* was the dominant species throughout the study, a result that was consistent with previous studies in this habitat (Etti *et al*., 2003). Many rodent hosts including *Apodemus* sp., *Myodes* sp. and squirrels can be reservoir for *B. afzelii*. An interesting aspect that we do currently not know would be to investigate which rodent species occur in this habitat.

Population analysis

Using MLST as typing system, our data also provided a temporal and spatial view on populations of individual *Borrelia* species. Whilst host association of *Borrelia* is governed by the innate immune system of the host (Marcinkiewicz *et al*., 2017) and may result in shifts in species composition/abundance in response to host shifts in the habitat, the host adaptive immune response impinges on populations of *Borrelia* species (Kurtenbach *et al*., 2006). Previous work suggested that adaptive immune responses directed towards outer surface proteins, for example, outer surface protein C (OspC) of *Borrelia* (Qiu *et al*., 2002; Brisson and Dykhuizen, 2004), may lead to fluctuations of genotypes (Kurtenbach *et al*., 2006). In these scenarios, abundant genotypes would decrease due to frequency-dependent selection of adaptive immune responses, whilst rare genotypes would become more frequent. However, such fluctuations of OspC genotypes were not noticed in a longitudinal study conducted in Switzerland on *B. afzelii* and *B. garinii* (Durand *et al*., 2017). Instead, dominance of a *B. afzelii* OspC-type termed A10 was observed over the whole time of the study (11 years). As the samples in our study had been acquired in the different habitats over several years, we expected to see some kind of pattern and clustering of closely related STs according to the collection time or region. However, no recognizable trend concerning collection year or site was observed relating to population structure of species. Generally, we did not observe the dominance of specific STs in certain years and their replacement with less frequent STs (Qiu *et al*., 2002; Brisson and Dykhuizen, 2004). In none of the *Borrelia* species, not in *B. afzelii* or in any of the other species, did we find a dominant ST. In fact, there were no STs that dominated any of the habitats at any time, similar to studies conducted in the United States (Hoen *et al*., 2009a,b). In all the species, a minority of STs were found in several years, and if so, often not in the same habitat. STs that were found in consecutive years were closely related often belonging to the same CC, in

particular in *B. afzelii* whose population consisted of two major CCs and only few singletons. Therefore, our data are not consistent with either of the scenarios described above. A confounding factor may have been that many STs were found only once and that for many samples all eight MLST genes could not be sequenced and thus, could not be identified to ST.

STs were mostly specific to a region whilst few were shared between regions. Only few shared STs were found in the regions, for instance out of a total of 109 STs, four shared STs were found between Jaunciems and Kemeri, five shared STs between Babite and Kemeri and five shared STs between Babite and Jaunciems suggesting very little exchange of STs between habitats, even for the bird-associated species. As these habitats are approximately 40 km apart, one might have expected that birds migrate back and forth between these three regions. However, in such a scenario it would be expected to find identical STs of bird-associates species, *B. garinii* and *B. valaisiana*, more frequently in all the three regions (Vollmer *et al*., 2011; 2013) than *B. afzelii* which was not the case, neither in our study nor in a previous study conducted in the same habitats (Etti *et al*., 2003).

Taken together, our data suggest that the population dynamics of *Borrelia* species is far more complex than previously suggested. A limitation of our study was that we did not obtain the full array of MLST sequences for all of the samples. Thus, improved methods for sequence acquisition are required. It has been shown in the United States that enrichment of bacterial DNA and whole genome sequencing may give additional and better data but that may be more complicated to accomplish in Europe with a many more species to consider than just *B. burgdorferi* s.s. in the United States (Carpi *et al*., 2015; Walter *et al*., 2016).

Conclusions

In conclusion, *Borrelia* is a vector obligate bacterium, and the vectors in turn are host-reliant, which makes the ecology of these bacteria highly complex. Generally, host abundance is a determining factor for vector survival hence spirochete survival. The distribution of *B. burgdorferi* s.l. species is dependent on the host distribution and movement (Kurtenbach *et al*., 2002; Comstedt *et al*., 2011; James *et al*., 2014) but species variation may also be the result of strong bottlenecks within sites (Bruyndonckx *et al*., 2009; Lemoine *et al*., 2018).

Our data reflect the complexity of *Borrelia* ecology. Contrary to common opinion that climate change will increase vector populations and the risk for acquiring TBP for humans, we did neither observe an increase in tick abundance nor an increase in *Borrelia* infection in ticks. In fact, after 2001 tick infection rates with *Borrelia* steadily declined in all three habitats. The three investigated habitats, Babite, Kemeri and Jaunciems, differed markedly in *Borrelia* species structure suggesting differences in reservoir host composition. Temporal changes in species structure were observed in one habitat (Babite) whilst the other two habitats (Kemeri and Jaunciems) showed relative stability. Furthermore, our results suggested carryover of genospecies from year to year is possible. However, frequencies did vary between sites and over years, not excluding the possibility of local extinction or re-emergence over longer time periods.

Our long-term data show that the ecology of different habitats may respond differently to ecological changes over time and extrapolation of data relating to prevalence of TBP, to fluctuations or stability of species or populations is not possible. Certainly, further long-term studies are needed, which focus on the same sites to have a comparable data over years that include data on reservoir hosts, prevalences and population structure.

Materials and methods

Tick sampling

Ticks were collected by drag sampling using a white cotton cloth from Spring to Autumn in three different habitats in Latvia between the years 1999 and 2010 (Rulison *et al*., 2013). Sampling habitats were sylvatic (Babite), peridomestic (Jaunciems), and peridomestic-sylvatic (Kemeri). Collected ticks were morphologically identified as *I. ricinus* (Sonenshine and Roe, 2014) and were stored in 70% ethanol (EtOH) until further use. MLST sequences from ticks sampled in habitats from the years 2002, 2006 and 2007 were already processed and published in previous studies (Vollmer *et al*., 2011; 2013) and were included in our study. See Table 1 for the summary of processed ticks in the current study.

DNA extraction and PCR

Whole ticks were washed shortly using approximately 1 ml distilled water prior to DNA purification. DNA extraction was done using two different methods: (i) alkaline hydrolysis using 1.25% ammonium hydroxide and (ii) a commercially available DNA extraction kit as per manufactures' instructions. The methods have been extensively described elsewhere (Okeyo *et al*., 2019).

A screening real-time PCR targeting the *flaB* gene encoding the flagellin B protein (P41) was conducted (Venczel *et al*., 2016). Briefly, QuantiTectMultilpex PCR (NoROX Master Mix, Qiagen, Hilden, Germany) was used with the following thermo-profile: Taq polymerase

activation at 95°C for 10 min, followed by 45 cycles of DNA annealing, extension and elongation at 95°C for 10 s, 56°C for 40 s and 72°C for 30 s respectively. Because of differences in sensitivity between real-time PCR and conventional PCR, only samples with Ct values ≤ 35 were used for further MLST analysis (Okeyo *et al*., 2019).

A MLST scheme with eight housekeeping loci (Margos *et al*., 2008; 2009) (*clpA*, *clpX*, *nifS*, *pepX*, *pyrG*, *recG*, *rplB*, *uvrA*; see also http://pubmlst.org/borrelia) was employed as described in detail by Wang and colleagues (2014). Briefly, a two-step (nested) PCR with a touchdown profile in the first round was conducted for seven genes, to minimize non-specific binding of the primers. For *recG* a regular two-step PCR was used (Margos *et al*., 2009; Vollmer *et al*., 2011; Jungnick *et al*., 2015). The PCR products were purified using polyethylenglycol (PEG) as described at https://openwetware.org/wiki/PEG_purification_of_PCR_products. The PEG solution was prepared by mixing 10.0 g PEG with 7.3 g sodium chloride (NaCl) and the volume adjusted to 50 ml with distilled water. The DNA was eluted with 30 μl nuclease free water (Promega, Germany).

Sequencing, sequence analysis and species assignment

Twenty samples from this study were sent to a commercial sequencing company for sequencing (GATC, Koblenz, Germany; www.gatc.de) and the remaining samples were sequenced in-house with Illumina MiSeq technique as described (Kingry *et al*., 2018).

Quality assessment of sequences obtained by Sanger sequencing was done using the software provided by Smartgene Inc., Lausanne, Switzerland (https://apps.idns-smartgene.com/apps/IDNSPortal.po) or by using the Seqman Pro software v.15 (DNAStar, Lasergene, Madison, WI). Sequences that were generated by MiSeq sequencing were mapped onto MLST reference sequences for B31 extracted from MLST database using the CLC genomic workbench 11 (Qiagen). A cut-off value of 0.3 was used to determine mixed sequences (Hepner *et al*., 2020). All sequences that were identified as mixed infections were removed from further MLST analyses. MEGA 6.0 was used to align and trim sequences to the required size. Species assignment, as well as designation of known allele numbers and sequence type numbers was done through the pubmlst *Borrelia* website (http://pubmlst.org/borrelia). Novel alleles were submitted to the curator of the MLST website and novel allele numbers were assigned. Allele and sequence type numbers allowed determination of *Borrelia* species.

eBURST and goeBURST analyses

goeBURST (Francisco *et al*., 2009; 2012) and eBUSRT (Feil *et al*., 2004) were used to obtain a view on the relatedness of STs, to create a network of CC and to predict the founder of each CC (Feil *et al*., 2004). The analysis is based on allelic profiles determined for each ST (Feil *et al*., 2004). CCs were defined for all samples that had a full complement of MLST alleles using the triple-locus variant parameter in eBURST (Feil *et al*., 2004) and goeBURST (Francisco *et al*., 2009; 2012). STs that differed in no more than three alleles were connected to form a CC. The formed CCs were designated according to the ST number of the potential founder.

Statistical analysis

The probability of a tick to be *Borrelia* positive was modelled using a binomial generalized linear mixed effect model (GLMM). Site and life stage were fitted as fixed effects and a random intercept for year was included. This model utilized all positive tick specimens (Ct value < 40) regardless of complete *Borrelia* species identification. No larvae were positive for *Borrelia* and therefore were excluded from this analysis. For all further statistical analyses, only positive tick specimens with *Borrelia* species identification were used.

Prevalence of host adaptation types per year were calculated and tested using Fisher's exact tests. For this all *Borrelia* spp. found in a site in a given year were combined into four categories: bird-adapted (*B. valaisiana*, *B. garinii*), rodent-adapted (*B. afzelii*, *B. bavariensis*), generalists (*B. burgdorferi* s.s.) or were identified as mixed infections (i.e. potentially more than one *Borrelia* species present).

To test potential carry over year effects on *Borrelia* genospecies prevalence, we modelled the absolute number of ticks infected with a specific *Borrelia* genospecies each year using a GLMM assuming a Poisson error distribution. For this, fixed effects were fitted for the prevalence of the given *Borrelia* genospecies in the previous sampling event and a binary factor if the genospecies was found (1) or not found (0) in the previous sampling event. Random effects were also included for site and year.

All analyses were done in R (version 3.6.1; R Core Team, 2019). Fisher's exact tests were performed using the *fisher.test* command with a simulated *p* value, based on 5000 simulations, from the base R package (R Core Team, 2019). All GLMMs were run with the *glmer* function from the 'lme4' package (Bates *et al*., 2015). The posterior distributions of the model parameters were simulated using the *sim* function from the 'arm' package (Gelman and Su, 2016). Mean estimates and their 95%

CIs, were extracted estimated based on 5000 simulations. Residual errors were calculated according to Nakagawa and Schielzeth (2010).

Acknowledgements

We are grateful to Cecilia Hizo-Teufel, Silvia Stockmeier and Wiltrud Strehle in the NRZ *Borrelia* laboratory for their assistance during the course of this research. This work was supported by Robert Koch Institute through German National Reference Center for *Borrelia* and by the Deutschen Akademischen Austausch Dienst (DAAD).

References

Akl, T., Bourgoin, G., Souq, M.L., Appolinaire, J., Poirel, M. T., Gibert, P., *et al*. (2019) Detection of tick-borne pathogens in questing *Ixodes ricinus* in the French Pyrenees and first identification of *Rickettsia monacensis* in France. *Parasite* **26**: 20–20.

Aureli, S., Galuppi, R., Ostanello, F., Foley, J.E., Bonoli, C., Rejmanek, D., *et al*. (2015) Abundance of questing ticks and molecular evidence for pathogens in ticks in three parks of Emilia-Romagna region of northern Italy. *Ann Agric Environ Med* **22**: 459–466.

Bacon, R.M., Kugeler, K.J., and Mead, P.S. (2008) Surveillance for Lyme disease United States, 1992–2006. *MMWR Surveill Summ* **57**: 1–9.

Bates, D., Mächler, M., Bolker, M.B., and Walker, S.C. (2015) Fitting linear mixed-effects models using lme4. *J Stat Softw* **67**: 48.

Bormane, A., Lucenko, I., Duks, A., Mavtchoutko, V., Ranka, R., Salmina, K., and Baumanis, V. (2004) Vectors of tick-borne diseases and epidemiological situation in Latvia in 1993–2002. *Int J Med Microbiol* **293**: 36–47.

Brisson, D., and Dykhuizen, D.E. (2004) OspC diversity in *Borrelia burgdorferi*: different hosts are different niches. *Genetics* **168**: 713–722.

Brisson, D., and Dykhuizen, D.E. (2006) A modest model explains the distribution and abundance of *Borrelia burgdorferi* strains. *Am J Trop Med Hyg* **74**: 615–622.

Bruyndonckx, N., Henry, I., Christe, P., and Kerth, G. (2009) Spatio-temporal population genetic structure of the parasitic mite *Spinturnix bechsteini* is shaped by its own demography and the social system of its bat host. *Mol Ecol* **18**: 3581–3592.

Carpi, G., Walter, K.S., Bent, S.J., Hoen, A.G., Diuk-Wasser, M., and Caccone, A. (2015) Whole genome capture of vector-borne pathogens from mixed DNA samples: a case study of *Borrelia burgdorferi*. *BMC Genomics* **16**: 434.

Chvostáč, M., Špitalská, E., Václav, R., Vaculová, T., Minichová, L., and Derdáková, M. (2018) Seasonal patterns in the prevalence and diversity of tick-borne *Borrelia burgdorferi* Sensu Lato, *Anaplasma phagocytophilum* and *Rickettsia* spp. in an urban temperate forest in South Western Slovakia. *Int J Environ Res Public Health* **15**: 994.

Coipan, E.C., Jahfari, S., Fonville, M., Maassen, C.B., van der Giessen, J., Takken, W., *et al*. (2013) Spatiotemporal dynamics of emerging pathogens in questing *Ixodes ricinus*. *Front Cell Infect Microbiol* **3**: 36–36.

Coipan, E.C., Jahfari, S., Fonville, M., Oei, G.A., Spanjaard, L., Takumi, K., *et al*. (2016) Imbalanced presence of *Borrelia burgdorferi* s.l. multilocus sequence types in clinical manifestations of Lyme borreliosis. *Infect Genet Evol* **42**: 66–76.

Comstedt, P., Jakobsson, T., and Bergström, S. (2011) Global ecology and epidemiology of *Borrelia garinii* spirochetes. *Infect Ecol Epidemiol* **1**: 9545. https://doi.org/10.3402/iee.v1i0.9545.

Dubska, L., Literak, I., Kocianova, E., Taragelova, V., and Sychra, O. (2009) Differential role of passerine birds in distribution of *Borrelia* spirochetes, based on data from ticks collected from birds during the postbreeding migration period in Central Europe. *Appl Environ Microbiol* **75**: 596–602.

Durand, J., Herrmann, C., Genne, D., Sarr, A., Gern, L., and Voordouw, M.J. (2017) Multistrain infections with Lyme borreliosis pathogens in the tick vector. *Appl Environ Microbiol* **83**: 3.

Ehrmann, S., Ruyts, S.C., Scherer-Lorenzen, M., Bauhus, J., Brunet, J., Cousins, S.A.O., *et al*. (2018) Habitat properties are key drivers of *Borrelia burgdorferi* (s.l.) prevalence in *Ixodes ricinus* populations of deciduous forest fragments. *Parasit Vectors* **11**: 23–23.

Eisen, L., and Lane, R.S. (2002) Vectors of *Borrelia burgdorferi* sensu lato. In *Lyme Borreliosis: Biology, Epidemiology, and Control*, Gray, J., Kahl, O., Lane, R.S., and Stanek, G. (eds). Wallingford, UK: CABI.

Estrada-Pena, A., Cutler, S., Potkonjak, A., Vassier-Tussaut, M., Van Bortel, W., Zeller, H., *et al*. (2018) An updated meta-analysis of the distribution and prevalence of *Borrelia burgdorferi* s.l. in ticks in Europe. *Int J Health Geogr* **17**: 41.

Estrada-Pena, A., Ortega, C., Sanchez, N., Desimone, L., Sudre, B., Suk, J.E., and Semenza, J.C. (2011) Correlation of *Borrelia burgdorferi* sensu lato prevalence in questing *Ixodes ricinus* ticks with specific abiotic traits in the western palearctic. *Appl Environ Microbiol* **77**: 3838–3845.

Etti, S., Hails, R., Schafer, S.M., De Michelis, S., Sewell, H. S., Bormane, A., *et al*. (2003) Habitat-specific diversity of *Borrelia burgdorferi* sensu lato in Europe, exemplified by data from Latvia. *Appl Environ Microbiol* **69**: 3008–3010.

Feil, E.J., Li, B.C., Aanensen, D.M., Hanage, W.P., and Spratt, B.G. (2004) eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. *J Bacteriol* **186**: 1518–1530.

Ferquel, E., Garnier, M., Marie, J., Bernède-Bauduin, C., Baranton, G., Pérez-Eid, C., and Postic, D. (2006) Prevalence of *Borrelia burgdorferi* sensu lato and *Anaplasmataceae* members in *Ixodes ricinus* ticks in Alsace, a focus of Lyme borreliosis endemicity in France. *Appl Environ Microbiol* **72**: 3074–3078.

Francisco, A.P., Bugalho, M., Ramirez, M., and Carrico, J.A. (2009) Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* **10**: 152.

Francisco, A.P., Vaz, C., Monteiro, P.T., Melo-Cristino, J., Ramirez, M., and Carriço, J.A. (2012) PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics* **13**: 87–87.

Fritz, L.C., and Kjemtrup, A.M. (2003) Lyme borreliosis. *Vet Med Today* **223**: 1261–1270.

Gasmi, S., Ogden, N.H., Ripoche, M., Leighton, P.A., Lindsay, R.L., Nelder, M.P., *et al*. (2019) Detection of municipalities at-risk of Lyme disease using passive surveillance of *Ixodes scapularis* as an early signal: a province-specific indicator in Canada. *PLoS One* **14**: e0212637.

Geller, J., Nazarova, L., Katargina, O., and Golovljova, I. (2013) *Borrelia burgdorferi* sensu lato prevalence in tick populations in Estonia. *Parasit Vectors* **6**: 202.

Gelman, A. and Su, Y.-S. (2016) Data analysis using regression and multilevel/hierarchical models URL https://cran.r-project.org/package=arm.

Gern, L., Estrada-Pena, A., Frandsen, F., Gray, J.S., Jaenson, T.G., Jongejan, F., *et al*. (1998) European reservoir hosts of *Borrelia burgdorferi* sensu lato. *Zentralbl Bakteriol* **287**: 196–204.

Gray, J.S., Kahl, O., Lane, R.S., Levin, M.L., and Tsao, J.I. (2016) Diapause in ticks of the medically important *Ixodes ricinus* species complex. *Ticks Tick-Borne Dis* **7**: 992–1003.

Halos, L., Bord, S., Cotté, V., Gasqui, P., Abrial, D., Barnouin, J., *et al*. (2010) Ecological factors characterizing the prevalence of bacterial tick-borne pathogens in *Ixodes ricinus* ticks in pastures and woodlands. *Appl Environ Microbiol* **76**: 4413–4420.

Hanincova, K., Kurtenbach, K., Diuk-Wasser, M., Brei, B., and Fish, D. (2006) Epidemic spread of Lyme borreliosis, Northeastern United States. *Emerg Infect Dis* **12**: 604–611.

Hanincova, K., Mukherjee, P., Ogden, N.H., Margos, G., Wormser, G.P., Reed, K.D., *et al*. (2013) Multilocus sequence typing of *Borrelia burgdorferi* suggests existence of lineages with differential pathogenic properties in humans. *PLoS One* **8**: e73066.

Hepner, S., Fingerle, V., Duscher, G.G., Felsberger, G., Marosevic, D., Rollins, R.E., *et al*. (2020) Population structure of *Borrelia turcica* from Greece and Turkey. *Infect Genet Evol* **77**: 104050.

Herrmann, C., Gern, L., and Voordouw, M.J. (2013) Species co-occurrence patterns among Lyme borreliosis pathogens in the tick vector *Ixodes ricinus*. *Appl Environ Microbiol* **79**: 7273–7280.

Heylen, D., Fonville, M., Docters van Leeuwen, A., Stroo, A., Duisterwinkel, M., van Wieren, S., *et al*. (2017) Pathogen communities of songbird-derived ticks in Europe's low countries. *Parasit Vectors* **10**: 497.

Hoen, A.G., Margos, G., Bent, S.J., Diuk-Wasser, M.A., Barbour, A., Kurtenbach, K., and Fish, D. (2009a) Phylogeography of *Borrelia burgdorferi* in the eastern United States reflects multiple independent Lyme disease emergence events. *Proc Natl Acad Sci USA* **106**: 15013–15018.

Hoen, A.G., Rollend, L.G., Papero, M.A., Carroll, J.F., Daniels, T.J., Mather, T.N., *et al*. (2009b) Effects of tick control by acaricide self-treatment of white-tailed deer on host-seeking tick infection prevalence and entomologic risk for *Ixodes scapularis*-borne pathogens. *Vector Borne Zoonotic Dis* **9**: 431–438.

Hu, C.M., Humair, P.F., Wallich, R., and Gern, L. (1997) *Apodemus* sp.Rodents, reservoir hosts for *Borrelia afzelii* in an endemic area in Switzerland. *Zentralbl Bakteriol* **285**: 558–564.

Hubalek, Z. (2009) Epidemiology of Lyme borreliosis. *Curr Probl Dermatol* **37**: 31–50.

Jacquot, M., Gonnet, M., Ferquel, E., Abrial, D., Claude, A., Gasqui, P., *et al*. (2014) Comparative population genomics of the *Borrelia burgdorferi* species complex reveals high degree of genetic isolation among species and underscores benefits and constraints to studying intra-specific epidemiological processes. *PloS One* **9**: e94384–e94384.

James, M.C., Gilbert, L., Bowman, A.S., and Forbes, K.J. (2014) The heterogeneity, distribution, and environmental associations of *Borrelia burgdorferi* sensu lato, the agent of Lyme borreliosis, in Scotland. *Front Public Health* **2**: 129.

Jouda, F., Perret, J.L., and Gern, L. (2004) Density of questing *Ixodes ricinus* nymphs and adults infected by *Borrelia burgdorferi* sensu lato in Switzerland: Spatiotemporal pattern at a regional scale. *Vector Borne Zoonotic Dis* **4**: 23–32.

Jungnick, S., Margos, G., Rieger, M., Dzaferovic, E., Bent, S. J., Overzier, E., *et al*. (2015) *Borrelia burgdorferi* sensu stricto and *Borrelia afzelii*: population structure and differential pathogenicity. *Int J Med Microbiol* **305**: 673–681.

Killilea, M.E., Swei, A., Lane, R.S., Briggs, C.J., and Ostfeld, R.S. (2008) Spatial dynamics of Lyme disease: a review. *Ecohealth* **5**: 167–195.

Kingry, L.C., Anacker, M., Pritt, B., Bjork, J., Respicio-Kingry, L., Liu, G., *et al*. (2018) Surveillance for and discovery of *Borrelias* species in US patients suspected of tickborne illness. *Clin Infect Dis* **66**: 1864–1871.

Kurtenbach, K., De Michelis, S., Etti, S., Schafer, S.M., Sewell, H.S., Brade, V., and Kraiczy, P. (2002) Host association of *Borrelia burgdorferi* sensu lato-the key role of host complement. *Trends Microbiol* **10**: 74–79.

Kurtenbach, K., Hanincova, K., Tsao, J.I., Margos, G., Fish, D., and Ogden, N.H. (2006) Fundamental processes in the evolutionary ecology of Lyme borreliosis. *Nat Rev Microbiol* **4**: 660–669.

Kurtenbach, K., Hoen, A.G., Bent, S.J., Vollmer, S.A., Ogden, N. and Margos, G. (2010) *Population Biology of Lyme Borreliosis Spirochetes. In Bacterial Population Genetics in Infectious Diseases*. Robinson, D. A., Falush, D., Feil, J.E. (eds). New Jersey: John Wiley & Sons, pp. 217–245.

Kurtenbach, K., Peacey, M., Rijpkema, S.G., Hoodless, A. N., Nuttall, P.A., and Randolph, S.E. (1998) Differential transmission of the genospecies of *Borrelia burgdorferi* sensu lato by game birds and small rodents in England. *Appl Environ Microbiol* **64**: 1169–1174.

Lane, R.S., and Lavoie, P.E. (1988) Lyme borreliosis in California. Acarological, clinical, and epidemiological studies. *Ann NY Acad Sci* **539**: 192–203.

Lemoine, M., Cornetti, L., and Tschirren, B. (2018) Does *Borrelia burgdorferi* sensu lato facilitate the colonisation of marginal habitats by *Ixodes ricinus*? A correlative study in the Swiss Alps. *bioRxiv* **273490**: 1.

Lindgren, E., and Jaenson, T.G.T. (2006) *Lyme Borreliosis in Europe: Influences of Climate and Climate Change, Epidemiology, Ecology and Adaptation Measures*, Vol. **1**,

Copenhagen, Denmark: Public Health Panorama: WHO Regional Office for Europe, pp. 1–34.

Marcinkiewicz, A.L., Kraiczy, P., and Lin, Y.P. (2017) There is a method to the madness: strategies to study host complement evasion by Lyme disease and relapsing fever spirochetes. *Front Microbiol* **8**: 328.

Margos, G., Gatewood, A.G., Aanensen, D.M., Hanincova, K., Terekhova, D., Vollmer, S.A., *et al*. (2008) MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. *Proc Natl Acad Sci USA* **105**: 8730–8735.

Margos, G., Vollmer, S.A., Cornet, M., Garnier, M., Fingerle, V., Wilske, B., *et al*. (2009) A new *Borrelia* species defined by multilocus sequence analysis of housekeeping genes. *Appl Environ Microbiol* **75**: 5410–5416.

Margos, G., Vollmer, S.A., Ogden, N.H., and Fish, D. (2011) Population genetics, taxonomy, phylogeny and evolution of *Borrelia burgdorferi* sensu lato. *Infect Genet Evol* **11**: 1545–1563.

Matuschka, F.R., Eiffert, H., Ohlenbusch, A., and Spielman, A. (1994) Amplifying role of edible dormice in Lyme disease transmission in Central Europe. *J Infect Dis* **170**: 122–127.

Mechai, S., Margos, G., Feil, E.J., Lindsay, L.R., and Ogden, N.H. (2015) Complex population structure of *Borrelia burgdorferi* in southeastern and south Central Canada as revealed by phylogeographic analysis. *Appl Environ Microbiol* **81**: 1309–1318.

Mejlon, H.A., and Jaenson, T.G. (1993) Seasonal prevalence of *Borrelia burgdorferi* in *Ixodes ricinus* in different vegetation types in Sweden. *Scand J Infect Dis* **25**: 449–456.

Nahimana, I., Gern, L., Blanc, D., Praz, G., Francioli, P., and Péter, O. (2004) Risk of *Borrelia burgdorferi* infection in western Switzerland following a tick bite. *J Clin Microbiol Infec Dis* **23**: 603–608.

Nakagawa, S., and Schielzeth, H. (2010) Repeatability for Gaussian and non-Gaussian data: a practical guide for biologists. *Biol Rev Camb Philos Soc* **85**: 935–956.

Norte, A.C., Ramos, J.A., Gern, L., Nuncio, M.S., and Lopes de Carvalho, I. (2013) Birds as reservoirs for *Borrelia burgdorferi* s.l. in Western Europe: circulation of *B. turdi* and other genospecies in bird-tick cycles in Portugal. *Environ Microbiol* **15**: 386–397.

Ogden, N.H., Lindsay, L.R., Hanincova, K., Barker, I.K., Bigras-Poulin, M., Charron, D.F., *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. *Appl Environ Microbiol* **74**: 1780–1790.

Ogden, N.H., Margos, G., Aanensen, D.M., Drebot, M.A., Feil, E.J., Hanincová, K., *et al*. (2011) Investigation of genotypes of *Borrelia burgdorferi* in *Ixodes scapularis* ticks collected during surveillance in Canada. *Appl Environ Microbiol* **77**: 3244–3254.

Okeyo, M., Hartberger, C., Margos, G., Straubinger, R.K., Sing, A., and Fingerle, V. (2019) Comparison of methods for economic and efficient tick and *Borrelia* DNA purification. *Ticks Tick-Borne Dis* **10**: 1041–1045.

Ostfeld, R.S., Canham, C.D., Oggenfuss, K., Winchcombe, R.J., and Keesing, F. (2006) Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. *PLoS Biol* **4**: e145–e145.

Paul, R.E., Cote, M., Le Naour, E., and Bonnet, S.I. (2016) Environmental factors influencing tick densities over seven years in a French suburban forest. *Parasit Vectors* **9**: 309.

Pfaffle, M., Littwin, N., Muders, S.V., and Petney, T.N. (2013) The ecology of tick-borne diseases. *Int J Parasitol* **43**: 1059–1077.

Pfaffle, M., Petney, T., Skuballa, J., and Taraschewski, H. (2011) Comparative population dynamics of a generalist (*Ixodes ricinus*) and specialist tick (*I. hexagonus*) species from European hedgehogs. *Exp Appl Acarol* **54**: 151–164.

Piesman, J., and Gern, L. (2004) Lyme borreliosis in Europe and North America. *Parasitology* **129**: S191–S220.

Platonov, A.E., Karan, L.S., Kolyasnikova, N.M., Makhneva, N. A., Toporkova, M.G., Maleev, V.V., *et al*. (2011) Humans infected with relapsing fever spirochete *Borrelia miyamotoi*, Russia. *Emerg Infect Dis* **17**: 1816–1823.

Postic, D., Garnier, M., and Baranton, G. (2007) Multilocus sequence analysis of atypical *Borrelia burgdorferi* sensu lato isolates-description of *Borrelia californiensis* sp. nov., and genomospecies 1 and 2. *Int J Med Microbiol* **297**: 263–271.

Qiu, W.-G., Dykhuizen, D.E., Acosta, M.S., and Luft, B.J. (2002) Geographic uniformity of the Lyme disease spirochete (*Borrelia burgdorferi*) and its shared history with tick vector (*Ixodes scapularis*) in the northeastern United States. *Genetics* **160**: 833.

R.Core Team. (2019) *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria. https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing.

Radolf, J.D., Caimano, M.J., Stevenson, B., and Hu, L.T. (2012) Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. *Nat Rev Microbiol* **10**: 87–99.

Randolph, S. (2009) Tick-borne disease systems emerge from the shadows: the beauty lies in molecular detail, the message in epidemiology. *Parasitology* **136**: 1403–1413.

Randolph, S.E. (2004) Tick ecology: processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors. *Parasitology* **129** Suppl: S37–S65.

Randolph, S.E. (2008) The impact of tick ecology on pathogen transmission dynamics. In *Ticks Biology, Disease and Control*, Bowman, A.S., and Nuttall, P.A. (eds). Cambridge, UK: Cambridge University Press.

Rauter, C., and Hartung, T. (2005) Prevalence of *Borrelia burgdorferi* sensu lato genospecies in *Ixodes ricinus* ticks in Europe: a metaanalysis. *Appl Environ Microbiol* **71**: 7203–7216.

Richter, D., Schlee, D.B., and Matuschka, F.R. (2011) Reservoir competence of various rodents for the Lyme disease spirochete *Borrelia spielmanii*. *Appl Environ Microbiol* **77**: 3565–3570.

Rizzoli, A., Hauffe, H., Carpi, G., Vourc, H.G., Neteler, M., and Rosa, R. (2011) Lyme borreliosis in Europe. *Euro Surveill* **16**: 529–533.

Rosa, R., Andreo, V., Tagliapietra, V., Barakova, I., Arnoldi, D., Hauffe, H.C., *et al*. (2018) Effect of climate and land use on the spatio-temporal variability of tick-borne

bacteria in Europe. *Int J Environ Res Public Health* **15**: 732.

Rosef, O., Paulauskas, A., and Radzijevskaja, J. (2009) Prevalence of *Borrelia burgdorferi* sensu lato and *Anaplasma phagocytophilum* in questing *Ixodes ricinus* ticks in relation to the density of wild cervids. *Acta Vet Scand* **51**: 47.

Rulison, E.L., Kuczaj, I., Pang, G., Hickling, G.J., Tsao, J.I., and Ginsberg, H.S. (2013) Flagging versus dragging as sampling methods for nymphal *Ixodes scapularis* (Acari: Ixodidae). *J Vector Ecol* **38**: 163–167.

Schillberg, E., Lunny, D., Lindsay, L.R., Nelder, M.P., Russell, C., Mackie, M., *et al*. (2018) Distribution of *Ixodes scapularis* in Northwestern Ontario: results from active and passive surveillance activities in the northwestern health unit catchment area. *Int J Environ Res Public Health* **15**: 2225.

Schrauber, E.M. and Ostfeld, R.S. (2002) Modeling the effects of reservoir competence decay and demographic turnover in Lyme disease ecology. *Ecological Applications* **12**: 1142–1162.

Schwartz, A.M., Hinckley, A.F., Mead, P.S., Hook, S.A., and Kugeler, K.J. (2017) Surveillance for Lyme disease – United States, 2008–2015. *MMWR Surveill Summ* **66**: 1–12.

Sonenshine, D.E., and Roe, R.M. (2014) *Biology of Ticks*, Vol. **1**. New York, NY: Oxford University Press.

Stanek, G., Fingerle, V., Hunfeld, K.P., Jaulhac, B., Kaiser, R., Krause, A., *et al*. (2011) Lyme borreliosis: clinical case definitions for diagnosis and management in Europe. *Clin Microbiol Infect* **17**: 69–79.

Steere, A.C. (2001) Lyme disease. *N Engl J Med* **345**: 115–125.

Strnad, M., Honig, V., Ruzek, D., Grubhoffer, L., and Rego, R. O.M. (2017) Europe-wide meta-analysis of *Borrelia burgdorferi* sensu lato prevalence in questing *Ixodes ricinus* ticks. *Appl Environ Microbiol* **83**: e00609-17.

Takano, A., Fujita, H., Kadosaka, T., Konnai, S., Tajima, T., Watanabe, H., *et al*. (2011) Characterization of reptile-associated *Borrelia* sp. in the vector tick, *Amblyomma geoemydae*, and its association with Lyme disease and relapsing fever *Borrelia* spp. *Environ Microbiol Rep* **3**: 632–637.

Taragel'ova, V., Koci, J., Hanincova, K., Kurtenbach, K., Derdakova, M., Ogden, N.H., *et al*. (2008) Blackbirds and song thrushes constitute a key reservoir of *Borrelia garinii*, the causative agent of borreliosis in Central Europe. *Appl Environ Microbiol* **74**: 1289–1293.

van Duijvendijk, G., Coipan, C., Wagemakers, A., Fonville, M., Ersöz, J., Oei, A., *et al*. (2016) Larvae of *Ixodes ricinus* transmit *Borrelia afzelii* and *B. miyamotoi* to vertebrate hosts. *Parasit Vectors* **9**: 97.

van Duijvendijk, G., Sprong, H., and Takken, W. (2015) Multi-trophic interactions driving the transmission cycle of *Borrelia afzelii* between *Ixodes ricinus* and rodents: a review. *Parasit Vectors* **8**: 643.

Venczel, R., Knoke, L., Pavlovic, M., Dzaferovic, E., Vaculova, T., Silaghi, C., *et al*. (2016) A novel duplex real-time PCR permits simultaneous detection and differentiation of *Borrelia miyamotoi* and *Borrelia burgdorferi* sensu lato. *Infection* **44**: 47–55.

Vitorino, L.R., Margos, G., Feil, E.J., Collares-Pereira, M., Ze-Ze, L., and Kurtenbach, K. (2008) Fine-scale phylogeographic structure of *Borrelia lusitaniae* revealed by multilocus sequence typing. *PLoS One* **3**: e4002.

Vollmer, S.A. (2010) Towards the origin of Lyme borreliosis. Doctor of Philosophy. Bath, UK: University of Bath. https://purehost.bath.ac.uk/ws/portalfiles/portal/187953846/UnivBath_PhD_2010_S_A_Vollmer.pdf.

Vollmer, S.A., Bormane, A., Dinnis, R.E., Seelig, F., Dobson, A.D.M., Aanensen, D.M., *et al*. (2011) Host migration impacts on the phylogeography of Lyme borreliosis spirochaete species in Europe. *Environ Microbiol* **13**: 184–192.

Vollmer, S.A., Feil, E.J., Chu, C.Y., Raper, S.L., Cao, W.C., Kurtenbach, K., and Margos, G. (2013) Spatial spread and demographic expansion of Lyme borreliosis spirochaetes in Eurasia. *Infect Genet Evol* **14**: 147–155.

Walter, K.S., Carpi, G., Evans, B.R., Caccone, A., and Diuk-Wasser, M.A. (2016) Vectors as epidemiological sentinels: patterns of within-tick *Borrelia burgdorferi* diversity. *PLoS Pathog* **12**: e1005759.

Wang, G., Dankert, J. and van Dam, A.P. (1999) Evidence for frequent OspC gene transfer between *Borrelia valaisiana* sp. nov. and other Lyme disease spirochetes. *FEMS Microbiol Lett*, **177**: 289–296,

Wang, G., Liveris, D., Mukherjee, P., Jungnick, S., Margos, G., and Schwartz, I. (2014) Molecular typing of *Borrelia burgdorferi*. *Curr Protoc Microbiol* **34**: 12C.5.1–12C.5.31.

Wielinga, P.R., Gaasenbeek, C., Fonville, M., de Boer, A., de Vries, A., Dimmers, W., *et al*. (2006) Longitudinal analysis of tick densities and *Borrelia*, *Anaplasma*, and *Ehrlichia* infections of *Ixodes ricinus* ticks in different habitat areas in The Netherlands. *Appl Environ Microbiol* **72**: 7594–7601.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Table S1. Samples used in this study with at least one MLST gene positive

Table S2. Total number of samples used for spatio-temporal distribution analysis

Table S3-A. Samples used for spatio-temporal distribution analysis from Babite

Table S3-B. Samples used for spatio-temporal distribution analysis from Jaunciems

Table S3-C. Samples used for spatio-temporal distribution analysis from Kemeri

Fig. S1. Google map showing tick collection sites in Latvia and Habitat description

Fig. S2. Total number of ticks analysed in the whole study

Fig. S3. Temporal distribution of *B. burgdorferi* s.l. in Babite

Fig. S4. Temporal distribution of *B. burgdorferi* s.l. in Jaunciems

Fig. S5. Temporal distribution of *B. burgdorferi* s.l. in Kemeri

Paper 2: Supplementary Material

Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010

Mercy Okeyo, Sabrina Hepner, **Robert E. Rollins**, Christina Hartberger, Reinhard K. Straubinger, Durdica Marosevic, Stephanie A. Bannister, Antra Bormane, Michael Donaghy, Andreas Sing, Volker Fingerle and Gabriele Margos

Environmental Microbiology (2020), 22:5033-5047
https://doi.org/10.1111/1462-2920.15100

Supplementary information, tables and figures

Habitat description

The three collection regions constitute different ecological habitats. Babite, (longitude, 23° 48'; latitude, 56° 50') is a sylvatic habitat (forested habitation). Jaunciems (longitude, 24° 09'; latitude, 57° 03') is peridomestic (around human habitation) and close to the suburbs of Riga. Kemeri (longitude, 23° 29'; latitude, 56° 56') represents a peridomestic habitat with mixed forest habitats along marshes (Etti et al., 2003).

The tick collection site Babite is far from residential areas with a rural road sectioning it in two parts. The most common tree found in this region is *Picea abies,* also known as *Alnus glutinosa*. In both sections, there is heavy growth of *Picea abies* saplings. Other trees found in this site included: *Populus tremula, Alnus* sp*., Frangula alnus, Padus avium, Betula* sp*., Salix* sp*.,* and *Rubus idaeus. Oxalis acetosella* mainly covers the ground, but there are also areas with dense growth of different common grass species.

Jaunciems is near Riga, the capital city of Latvia. There are two heavily used main roads, private houses and a lake near the tick collection site. The most common tree found in this region was *Pinus sylvestris.* Other tree species found in this region included: *Acer plantanoides, Sorbus aucuparia, Padus avium, Ameanchier spicata, Tilia cordata* amongst others. In contrast to Babite, the underwood and ground are not densely covered. Low plants and grass are found in these regions.

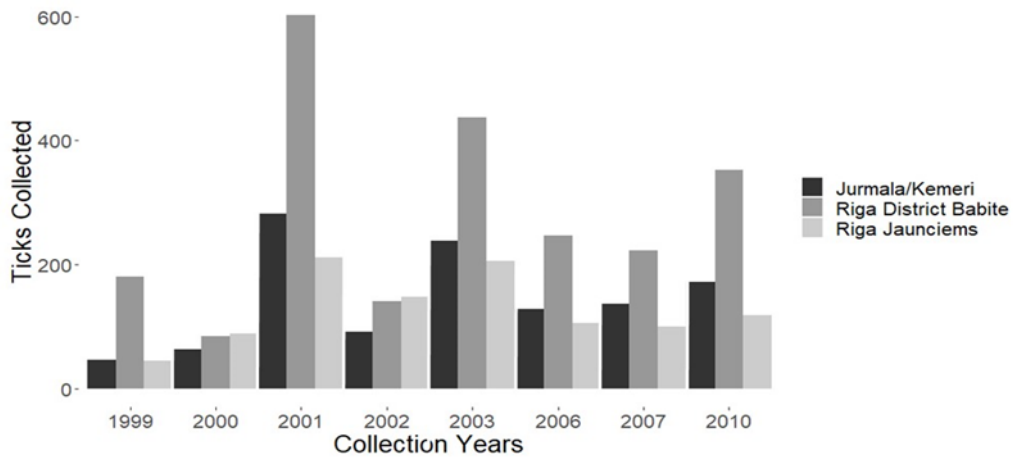
Kemeri is also far away from residential areas and has mixed vegetation. There is a small pedestrians' road. Nearby is woodland and marshes, along which tick collection was conducted. Frequently found trees were *Pinus sylvestris* and *Picea abies,* whilst *Alnus glutinosa* was infrequent at this site. The ground was covered either with *Oxalis acetosella or*

Vaccinium myrtillus or with moss. Different underwood tree species like *Alnus* sp. or *Acer plantanoides* are found. In Kemeri the ground is mostly densely covered.

Fig. S1: Google map showing tick collection sites in Latvia

Babite is a sylvatic habitat (forested habitation), Jaunciems is a peridomestic habitat close to the suburbs of Riga and Kemeri represents a peridomestic habitat with mixed habitats along marshes. Source https://www.google.de/maps.

Fig. S2: Total number of ticks analyzed in the whole study



The highest number of ticks were acquired in the year 2001 in Babite

Table S1: Samples used in this study with at least one MLST gene positive

	Strain	Species	Collection Site	Year	ClpA	ClpX	nifS	pepX	pyrG	recG	rplB	uvrA
1	9-26-55	*B. burgdorferi* s.s.	Babite	1999	14	1	11	1	168	1	8	10
2	9-26-08	*B. garinii*	Babite	1999	43	28	30	90	87	42	28	229
3	9-22-37	*B. afzelii*	Kemeri	1999	36	212	23	30	92	27	23	29
4	9-22-36	*B. afzelii*	Kemeri	1999	43	28	34	90	87	36	28	34
5	9-22-27	*B. garinii*	Kemeri	1999	40	213	26	36	28	34	28	31
6	9-22-21	*B. afzelii*	Kemeri	1999	183	24	23	31	27	27	23	29
7	9-12-26	*B. afzelii*	Babite	1999	109	24	25	31	92	52	23	28
8	9-12-11	*B. burgdorferi* s.s.	Babite	1999	15	1	1	1	168	1	1	1
9	9-26-20	*B. garinii*	Babite	1999	42	27	29	92	29	36	27	33
10	9-25-62	*B. burgdorferi* s.s.	Babite	1999	1	1	1	1	1	1	1	1
11	9-22-34	*B. garinii*	Kemeri	1999	42	27	29	38	29	80	27	33
12	9-22-17	*B. burgdorferi* s.s.	Kemeri	1999	14	1	11	1	1	1	1	10
13	9-22-16	*B. valaisiana*	Kemeri	1999	49	36	36	45	38	44	35	40
14	9-22-09	*B. burgdorferi* s.s.	Kemeri	1999	14	1	11	1	1	1	1	10
15	9-22-04	*B. garinii*	Kemeri	1999	43	28	30	90	87	36	28	34
16	9-20-24	*B. bavariensis*	Jaunciems	1999	41	26	27	37	28	35	26	32
17	0-14-06	*B. afzelii*	Babite	2000	39	24	23	87	92	27	23	78
18	0-14-10b	*B. garinii*	Babite	2000	42	27	29	38	29	39	80	33
19	0-14-26	*B. afzelii*	Babite	2000	51	24	23	86	85	92	23	29
20	0-14-36	*B. garinii*	Babite	2000	45	33	34	36	36	38	30	38
21	0-14-37	*B. afzelii*	Babite	2000	35	24	24	32	21	27	208	28
22	0-14-39	*B. afzelii*	Babite	2000	39	24	24	31	22	92	23	28
23	0-8-03	*B. burgdorferi* s.s.	Kemeri	2000	14	1	11	1	1	10	1	10
24	0-4	*B. afzelii*	Jaunciems	2000	109	24	24	85	90	91	24	29
25	0-13-41	*B. burgdorferi* s.s.	Babite	2000	1	1	1	1	1	1	1	1
26	1-8-29	*B. garinii*	Kemeri	2001	46	214	29	43	98	40	31	37
27	1-8-50	*B. valaisiana*	Kemeri	2001	50	36	37	45	234	44	36	40

28	1-8-12b	*B. afzelii*	Kemeri	2001	172	24	23	85	98	27	24	28
29	1-8-28	*B. afzelii*	Kemeri	2001	109	24	24	31	92	246	23	28
30	1-29-12	*B. afzelii*	Jaunciems	2001	37	215	24	31	22	91	23	29
31	1-29-17	*B. valaisiana*	Jaunciems	2001	257	37	37	86	39	93	36	40
32	1-29-35b	*B. afzelii*	Jaunciems	2001	109	24	24	85	90	91	209	29
33	1-29-60	*B. afzelii*	Jaunciems	2001	36	24	23	89	95	27	23	28
34	1-29-22	*B. afzelii*	Jaunciems	2001	36	24	23	89	118	27	23	28
35	1-29-30a	*B. afzelii*	Jaunciems	2001	38	24	25	32	90	247	24	28
36	1-28	*B. afzelii*	Jaunciems	2001	36	24	24	32	21	31	23	30
37	1-28-18	*B. afzelii*	Jaunciems	2001	37	24	24	31	235	248	23	28
38	1-28-27	*B. afzelii*	Jaunciems	2001	37	24	25	31	236	92	23	28
39	1-41-07	*B. afzelii*	Babite	2001	51	24	23	86	85	27	23	230
40	1-41-18	*B. afzelii*	Babite	2001	36	24	23	86	237	27	23	28
41	1-38-13	*B. afzelii*	Babite	2001	258	24	23	203	26	96	23	156
42	1-37-13	*B. afzelii*	Babite	2001	37	24	24	88	22	92	23	28
43	1-32-30	*B. burgdorferi* s.s.	Kemeri	2001	14	1	11	1	168	10	1	10
44	1-32-49	*B. afzelii*	Kemeri	2001	109	24	24	31	22	246	24	28
45	1-32-53	*B. garinii*	Kemeri	2001	47	73	33	90	91	76	32	33
46	1-31-43	*B. afzelii*	Kemeri	2001	259	24	24	31	22	92	23	28
47	1-31LT	*B. garinii*	Kemeri	2001	40	216	26	36	27	250	25	31
48	1-5-26	*B. valaisiana*	Jaunciems	2001	50	35	37	45	38	44	35	40
49	1-5-15	*B. valaisiana*	Jaunciems	2001	50	37	37	45	39	93	36	40
50	1-5-29	*B. afzelii*	Jaunciems	2001	131	82	24	88	92	27	23	28
51	1-29-03	*B. valaisiana*	Jaunciems	2001	49	35	35	45	38	43	35	40
52	1-29-08	*B. garinii*	Jaunciems	2001	44	29	31	40	31	37	80	35
53	1-29-29	*B. afzelii*	Jaunciems	2001	37	24	24	31	98	92	23	28
54	1-29-43a	*B. afzelii*	Jaunciems	2001	48	34	34	44	37	42	33	39
55	1-4-09	*B. afzelii*	Jaunciems	2001	37	24	24	31	22	92	23	28
56	1-14	*B. valaisiana*	Babite	2001	50	38	36	45	38	44	35	40

57	1-19-17	*B. afzelii*	Babite	2001	39	24	24	31	22	92	23	28
58	1-38-03	*B. garinii*	Babite	2001	42	27	29	38	29	39	27	33
59	1-37-19	*B. valaisiana*	Babite	2001	49	37	37	45	39	45	36	40
60	1-31-25	*B. afzelii*	Kemeri	2001	36	24	23	31	92	27	23	29
61	1-31-42	*B. afzelii*	Kemeri	2001	37	24	24	31	22	92	23	28
62	2-47-03LT	*B. afzelii*	Jaunciems	2002	35	24	23	86	22	27	23	28
63	2-02-11L	*B. afzelii*	Babite	2002	36	24	23	87	92	27	23	78
64	2-16-21L	*B. afzelii*	Babite	2002	36	24	23	87	92	27	23	78
65	2-08-09LT	*B. afzelii*	Babite	2002	36	82	24	88	92	27	23	28
66	2-53LT	*B. afzelii*	Jaunciems	2002	37	24	23	31	22	92	23	30
67	2-58-15L	*B. afzelii*	Jaunciems	2002	37	24	25	31	238	92	23	30
68	2-54-15L	*B. afzelii*	Jaunciems	2002	38	24	25	32	90	29	24	28
69	2-27LT	*B. garinii*	Babite	2002	44	29	76	40	31	37	80	77
70	2-16-17L	*B. lusitaniae*	Kemeri	2002	261	22	94	123	18	118	74	102
71	2-45-9LT	*B. afzelii*	Babite	2002	36	24	24	31	22	92	23	28
72	2-32-12L	*B. afzelii*	Babite	2002	109	24	23	31	23	30	23	30
73	2-18-21L	*B. valaisiana*	Babite	2002	50	35	37	45	38	44	35	40
74	3-5-17	*B. burgdorferi* s.s.	Kemeri	2003	14	1	11	1	1	10	1	10
75	3-8-06	*B. afzelii*	Kemeri	2003	36	24	23	37	92	27	23	231
76	3-8-07	*B. afzelii*	Kemeri	2003	35	24	22	32	98	27	23	28
77	3-17-10	*B. garinii*	Kemeri	2003	48	34	90	44	37	111	33	39
78	3-14-23	*B. afzelii*	Jaunciems	2003	131	217	24	88	92	27	23	28
79	3-8-05	*B. burgdorferi* s.s.	Kemeri	2003	14	1	11	1	1	10	1	10
80	6-44-09L	*B. afzelii*	Babite	2006	36	24	23	86	22	27	23	28
81	6-42-24L	*B. afzelii*	Jaunciems	2006	36	24	23	86	22	27	23	28
82	6-12-18L	*B. afzelii*	Kemeri	2006	36	24	23	30	92	27	23	29
83	6-06-18L	*B. afzelii*	Kemeri	2006	36	24	23	30	92	27	23	29
84	6-07-24L	*B. afzelii*	Jaunciems	2006	36	24	23	31	92	89	23	29
85	6-33-21L	*B. afzelii*	Jaunciems	2006	37	24	24	31	22	92	23	28

86	6-37-21L	*B. afzelii*	Jaunciems	2006	37	24	24	31	22	92	23	28
87	6-35-21L	*B. afzelii*	Jaunciems	2006	51	24	23	86	85	27	23	29
88	6-19-18L	*B. afzelii*	Kemeri	2006	51	24	23	86	85	27	23	29
89	6-22-18L	*B. afzelii*	Kemeri	2006	51	24	23	86	85	27	23	29
90	6-34-24L	*B. afzelii*	Jaunciems	2006	102	24	24	31	22	92	23	28
91	6-10-06L	*B. garinii*	Babite	2006	42	27	29	38	29	36	27	33
92	6-50-06L	*B. garinii*	Babite	2006	42	27	29	38	29	36	27	33
93	6-14-06L	*B. garinii*	Babite	2006	31	80	78	99	81	39	79	87
94	6-10-09L	*B. garinii*	Babite	2006	42	27	29	92	29	36	27	33
95	6-23-06L	*B. garinii*	Babite	2006	43	28	30	90	87	36	28	34
96	6-25-06L	*B. garinii*	Babite	2006	43	28	30	90	87	36	28	34
97	6-21-18L	*B. garinii*	Kemeri	2006	43	28	30	90	87	36	28	34
98	6-38-09L	*B. garinii*	Babite	2006	44	29	31	40	31	87	80	77
99	6-10-112L	*B. garinii*	Babite	2006	46	76	73	43	34	40	31	37
100	6-23-03L	*B. garinii*	Babite	2006	47	32	33	42	91	86	32	36
101	6-23-09L	*B. garinii*	Babite	2006	47	32	33	42	91	86	32	36
102	6-12-09L	*B. garinii*	Babite	2006	47	32	33	42	35	88	32	36
103	6-29-09L	*B. garinii*	Babite	2006	48	34	34	44	27	42	33	39
104	6-92-12L	*B. garinii*	Babite	2006	48	76	29	43	34	42	31	37
105	6-66-12L	*B. garinii*	Babite	2006	99	77	36	91	88	84	75	33
106	6-13-06L	*B. valaisiana*	Babite	2006	49	35	35	45	38	43	35	40
107	6-18-09L	*B. valaisiana*	Babite	2006	50	36	36	45	38	44	35	40
108	6-35-03L	*B. valaisiana*	Babite	2006	50	36	36	45	38	44	35	40
109	6-36-03L	*B. valaisiana*	Babite	2006	50	39	36	45	38	44	35	40
110	6-21-03L	*B. valaisiana*	Babite	2006	50	37	37	45	39	45	36	40
111	6-16-18L	*B. valaisiana*	Kemeri	2006	50	39	74	45	38	44	35	40
112	6-10-412L	*B. valaisiana*	Babite	2006	96	75	36	98	84	44	78	86
113	6-39-09L	*B. valaisiana*	Babite	2006	96	75	36	98	84	44	78	86
114	6-84-12L	*B. valaisiana*	Babite	2006	96	75	36	98	84	44	78	86

115	7-50-21L	*B. afzelii*	Jaunciems	2007	36	24	23	86	22	27	23	28
116	7-14-24L	*B. afzelii*	Jaunciems	2007	36	24	23	31	92	27	23	29
117	7-05-15L	*B. afzelii*	Kemeri	2007	36	24	23	31	85	27	23	29
118	7-59-18L	*B. afzelii*	Kemeri	2007	37	24	24	31	22	92	23	28
119	7-46-18L	*B. afzelii*	Kemeri	2007	39	24	24	31	22	92	23	28
120	7-16-18L	*B. afzelii*	Kemeri	2007	109	24	23	89	22	27	23	28
121	7-2-721L	*B. afzelii*	Jaunciems	2007	51	24	23	28	85	27	23	29
122	7-29-15L	*B. afzelii*	Kemeri	2007	51	24	23	86	85	91	23	29
123	7-19-24L	*B. afzelii*	Jaunciems	2007	109	24	24	85	90	91	24	29
124	7-15-09L	*B. burgdorferi* s.s.	Babite	2007	15	9	12	8	1	17	8	16
125	7-43-06L	*B. burgdorferi* s.s.	Babite	2007	15	9	12	8	1	17	8	16
126	7-25-06L	*B. burgdorferi* s.s.	Babite	2007	15	9	12	8	1	83	8	16
127	7-09-03L	*B. garinii*	Babite	2007	42	27	29	38	29	36	27	33
128	7-42-15L	*B. garinii*	Kemeri	2007	42	27	29	38	29	36	27	33
129	7-36-06L	*B. garinii*	Babite	2007	44	29	31	40	31	37	29	35
130	7-53-09L	*B. garinii*	Babite	2007	44	29	31	40	31	37	29	35
131	7-16-09L	*B. garinii*	Babite	2007	45	30	32	41	32	38	30	36
132	7-08-09L	*B. garinii*	Babite	2007	31	80	78	99	81	39	79	87
133	7-64-18L	*B. garinii*	Kemeri	2007	42	27	29	38	81	36	27	33
134	7-45-03L	*B. garinii*	Babite	2007	43	28	30	39	30	78	28	34
135	7-41-09L	*B. garinii*	Babite	2007	43	28	30	39	88	87	28	34
136	7-04-06L	*B. garinii*	Babite	2007	44	29	31	40	31	37	80	35
137	7-39-12L	*B. garinii*	Babite	2007	47	73	33	42	91	76	32	36
138	7-24-09L	*B. garinii*	Babite	2007	48	34	34	44	27	42	33	39
139	7-44-15L	*B. garinii*	Kemeri	2007	95	74	34	96	83	78	77	85
140	7-58-03L	*B. garinii*	Babite	2007	95	29	34	91	89	78	77	85
141	7-44-03L	*B. garinii*	Babite	2007	95	74	34	96	89	78	77	85
142	7-14-12L	*B. lusitaniae*	Babite	2007	101	21	20	27	86	85	74	81
143	7-20-09L	*B. valaisiana*	Babite	2007	49	35	35	45	38	43	35	40

144	7-30-09L	*B. valaisiana*	Babite	2007	50	36	36	45	38	44	35	40
145	7-40-03L	*B. valaisiana*	Babite	2007	50	36	36	45	38	44	35	40
146	7-25-09L	*B. valaisiana*	Babite	2007	49	35	72	45	38	43	35	40
147	7-2-609L	*B. valaisiana*	Babite	2007	50	35	37	45	38	44	35	40
148	7-38-15L	*B. valaisiana*	Kemeri	2007	50	35	37	45	38	44	35	40
149	7-39-15L	*B. valaisiana*	Kemeri	2007	50	35	37	45	38	44	35	40
150	7-15-12L	*B. valaisiana*	Babite	2007	50	36	37	45	38	44	35	40
151	7-56-18L	*B. valaisiana*	Kemeri	2007	50	36	37	45	38	44	35	40
152	7-22-09L	*B. valaisiana*	Babite	2007	110	39	36	45	38	81	35	40
153	7-24-03L	*B. valaisiana*	Babite	2007	96	37	37	45	39	79	36	86
154	7-15-15L	*B. valaisiana*	Kemeri	2007	96	37	37	45	39	79	36	86
155	7-34-09L	*B. valaisiana*	Babite	2007	96	75	36	98	38	82	78	86
156	10-26-29	*B. garinii*	Babite	2010	260	218	81	91	88	249	82	33
157	10-22-22	*B. garinii*	Babite	2010	43	28	30	39	30	36	28	34
158	10-22-04	*B. garinii*	Babite	2010	42	27	29	92	29	36	27	33
159	10-24-07	*B. valaisiana*	Babite	2010	50	38	36	45	38	44	35	40
160	9-26-41	*B. garinii*	Babite	1999	24		29	91				33
161	9-26-37	*B. garinii*	Babite	1999	43	28	30	90	32			
162	9-26-29	*B. valaisiana*	Babite	1999		35			38	43		40
163	9-26-28	*B. afzelii*	Babite	1999	35	135	22	32	20	23		
164	9-26-26	*B. afzelii*	Babite	1999			25		90	29	24	
165	9-26-02	*B. afzelii*	Babite	1999		24			26			
166	9-25-54	*B. garinii*	Babite	1999	99	77	29			39		33
167	9-25-39	*B. garinii*	Babite	1999	47	177						
168	9-25-36	*B. burgdorferi* s.s.	Babite	1999		1	1	1	168		1	1
169	9-25-33	*B. afzelii*	Babite	1999	36	23	24			27	23	
170	9-22-20	*B. garinii*	Kemeri	1999	112	80	78	31	23			87
171	9-20-28	*B. afzelii*	Jaunciems	1999	37	40	24		98			131

172	9-20-22	*B. afzelii*	Jaunciems	1999	36	24	23		98	27	23	28
173	9-12-33	*B. afzelii*	Babite	1999					163	52		28
174	9-25-35	*B. afzelii*	Babite	1999						27		
175	9-12-22	*B. lusitaniae*	Babite	1999						22		
176	9-25-20	*B. afzelii*	Babite	1999			24				23	
177	0-4-12	*B. afzelii*	Jaunciems	2000	109		24		90		24	29
178	0-4-18	*B. afzelii*	Jaunciems	2000	36			88		27		28
179	0-8-09	*B. afzelii*	Kemeri	2000	109	24	24	31	97	27		30
180	0-8-19	*B. valaisiana*	Kemeri	2000	49			46	38	43		
181	0-8-30	*B. afzelii*	Kemeri	2000	36	23	22	203	26	96	23	
182	0-8-36	*B. garinii*	Kemeri	2000	42		29	43	29	36	77	37
183	0-14-03	*B. afzelii*	Babite	2000		24	24		22		23	28
184	0-14-10	*B. garinii*	Babite	2000		33	34	36	36	38		38
185	0-14-11	*B. valaisiana*	Babite	2000			37	45	39	79	35	86
186	0-14-22	*B. garinii*	Babite	2000	43	28			87	36		34
187	0-14-38	*B. afzelii*	Babite	2000	95			96	89	78		
188	0-4-37	*B. valaisiana*	Jaunciems	2000				45	38			
189	0-5-10	*B. afzelii*	Jaunciems	2000			24			92		
190	0-5-19	*B. afzelii*	Jaunciems	2000	38							
191	0-5-23	*B. afzelii*	Jaunciems	2000	109				23			78
192	0-7-02	*B. afzelii*	Kemeri	2000				86		27		
193	1-5-10	*B. afzelii*	Jaunciems	2001	109	82		88		52		28
194	1-5-28	*B. afzelii*	Jaunciems	2001	37		37	31				
195	1-5-34	*B. afzelii*	Jaunciems	2001	109	24	24	31	92			
196	1-5-28	*B. afzelii*	Jaunciems	2001	37			31				
197	1-5-32	*B. afzelii*	Jaunciems	2001	109			88	23	30		
198	1-5-2	*B. afzelii*	Jaunciems	2001	131							
199	1-8-45	*B. garinii*	Kemeri	2001	44	156	31		38	40		35
200	1-8-27	*B. afzelii*	Kemeri	2001	36		23		38	27	24	

201	1-8-39	*B. valaisiana*	Kemeri	2001	50			45		44	35	
202	1-8-47	*B. valaisiana*	Kemeri	2001			37			79	35	
203	1-8-18	*B. garinii*	Kemeri	2001			31		34	40		
204	1-8-28b	*B. afzelii*	Kemeri	2001	37		81	31	98	91		
205	1-8-46	*B. garinii*	Kemeri	2001	116			39	31	36		
206	1-8-54	*B. afzelii*	Kemeri	2001	109		25					
207	1-8-39b	*B. burgdorferi* s.s.	Kemeri	2001	15		12	8	1	17	8	16
208	1-29-28	*B. garinii*	Jaunciems	2001	40	25	26	36	27	34	25	31
209	1-29-27	*B. afzelii*	Jaunciems	2001	36		23		95	27	23	
210	1-29-43b	*B. garinii*	Jaunciems	2001	47	73	33		91	76	32	
211	1-29-17b	*B. valaisiana*	Jaunciems	2001	96	37				44		
212	1-29-30	*B. garinii*	Jaunciems	2001	112			99		39		
213	1-28-17	*B. afzelii*	Jaunciems	2001	51				85	27		29
214	1-28-25	*B. afzelii*	Jaunciems	2001	109	24	24			91	24	29
215	1-28-12b	*B. garinii*	Jaunciems	2001	47		33		91		32	
216	1-28-05	*B. afzelii*	Jaunciems	2001	37	24		31	96	92		
217	1-4-07	*B. afzelii*	Jaunciems	2001	37		24		98	27		
218	1-13-20	*B. burgdorferi* s.s.	Babite	2001	14							10
219	1-13	*B. afzelii*	Babite	2001	35							221
220	1-13-42	*B. garinii*	Babite	2001	42	27				39		33
221	1-13-16	*B. garinii*	Babite	2001	42							
222	1-14-44	*B. garinii*	Babite	2001	99				108	202		33
223	1-14-45	*B. garinii*	Babite	2001	95		34		89	78		154
224	1-14-48	*B. afzelii*	Babite	2001	51			86		27		29
225	1-14-47	*B. afzelii*	Babite	2001	36					27	23	28
226	1-14-39	*B. burgdorferi* s.s.	Babite	2001		9	12	8				
227	1-14-42	*B. garinii*	Babite	2001	46		29		34	39		33
228	1-14-37	*B. afzelii*	Babite	2001			36	49		91		
229	1-14-43	*B. garinii*	Babite	2001				90	29	39		33

230	1-15-48	*B. garinii*	Babite	2001	43	28		90	87			34
231	1-15-32	*B. valaisiana*	Babite	2001	96			45				
232	1-19-11	*B. afzelii*	Babite	2001		82	23	31	92	27	23	28
233	1-19-18	*B. garinii*	Babite	2001	42	28		39	30	36	28	34
234	1-19-41	*B. burgdorferi* s.s.	Babite	2001	14	1	11	1	1	1	1	
235	1-19-34	*B. afzelii*	Babite	2001	39				85	92		
236	1-19-43	*B. afzelii*	Babite	2001	36	24				27		
237	1-18-20	*B. garinii*	Babite	2001				112	108	54		33
238	1-18-21	*B. garinii*	Babite	2001	116					125		33
239	1-18-30	*B. valaisiana*	Babite	2001				45	38			40
240	1-18-15	*B. garinii*	Babite	2001	42							33
241	1-18-16	*B. afzelii*	Babite	2001		24	23	31			24	
242	1-18-39	*B. afzelii*	Babite	2001	36		24	88			23	
243	1-41-09	*B. afzelii*	Babite	2001	131	24	23	85	22	27		28
244	1-41-11	*B. garinii*	Babite	2001		24		42				136
245	1-41-22	*B. valaisiana*	Babite	2001		39		45		44	39	86
246	1-41-19	*B. valaisiana*	Babite	2001		75				44		86
247	1-40-20	*B. afzelii*	Babite	2001	39	24	24		22	92	23	
248	1-40-31	*B. afzelii*	Babite	2001	36		23	49	26	91	23	
249	1-40-26	*B. garinii*	Babite	2001	112	80				38		50
250	1-40-18	*B. garinii*	Babite	2001	44							
251	1-42-05	*B. afzelii*	Babite	2001	109	24		85	22	91		29
252	1-42-14	*B. afzelii*	Babite	2001		82						28
253	1-42-07	*B. afzelii*	Babite	2001		24		88				
254	1-39-03	*B. afzelii*	Babite	2001	131	24	23	31	22	27		
255	1-39-12	*B. afzelii*	Babite	2001	36		24	86		91		
256	1-38	*B. garinii*	Babite	2001	47			38				
257	1-38-04	*B. garinii*	Babite	2001	42	27				39		33
258	1-38-19	*B. afzelii*	Babite	2001	36	24		31		27		

259	1-38-21	*B. garinii*	Babite	2001	99					54		
260	1-38-28	*B. garinii*	Babite	2001	116	34				42		
261	1-37-11	*B. garinii*	Babite	2001	44	156	31		31	37	35	35
262	1-37-10	*B. valaisiana*	Babite	2001	50	36	36		38	44		40
263	1-37-05	*B. garinii*	Babite	2001	99	77	81		88	84	82	33
264	1-32-08	*B. burgdorferi* s.s.	Kemeri	2001	14		11	8	1	1	1	10
265	1-32-25	*B. afzelii*	Kemeri	2001	51	24	23		85	27		29
266	1-32-31	*B. burgdorferi* s.s.	Kemeri	2001	14	1		1	168	1	1	10
267	1-32-19	*B. valaisiana*	Kemeri	2001		35	37	45				
268	1-31-06	*B. garinii*	Kemeri	2001		27	29	38	29	39	27	39
269	1-31-59	*B. garinii*	Kemeri	2001	43	28	30		87	38		34
270	1-31-49	*B. garinii*	Kemeri	2001	44	156	31	40	22	37		
271	1-31-46	*B. garinii*	Kemeri	2001	44							
272	1-31-03	*B. valaisiana*	Kemeri	2001							35	
273	1-31-07	*B. garinii*	Kemeri	2001		29	31		208			
274	1-30-12	*B. afzelii*	Jaunciems	2001	37	24		31	96	92	23	30
275	1-30-22	*B. garinii*	Jaunciems	2001	42		29				27	
276	1-9-18	*B. afzelii*	Kemeri	2001	131		23			92	23	29
277	1-13	*B. afzelii*	Babite	2001	35							41
278	1-13-42	*B. garinii*	Babite	2001	42	27				39		
279	1-13-20	*B. burgdorferi* s.s.	Babite	2001	14							10
280	1-13-45	*B. garinii*	Babite	2001					29			
281	1-7-46	*B. garinii*	Kemeri	2001	42							
282	1-13-23	*B. garinii*	Babite	2001					29			
283	1-13-38	*B. garinii*	Babite	2001					87			
284	1-4-05	*B. garinii*	Jaunciems	2001					37			
285	1-13-08	*B. valaisiana*	Babite	2001				45				
286	1-14-43	*B. garinii*	Babite	2001					29	39		33
287	1-15-39	*B. afzelii*	Babite	2001			23		22			

288	1-15	*B. afzelii*	Babite	2001				86				28
289	1-15-11	*B. valaisiana*	Babite	2001								86
290	1-17-26	*B. garinii*	Babite	2001	42							
291	1-17-30	*B. garinii*	Babite	2001		156			31			
292	1-18	*B. garinii*	Babite	2001	43					38		34
293	1-18-11	*B. burgdorferi* s.s.	Babite	2001						17		
294	1-31-03	*B. valaisiana*	Kemeri	2001				45	43		35	
295	1-39-10	*B. garinii*	Babite	2001					208			
296	1-14-16	*B. garinii*	Babite	2001	42							
297	3-5-24	*B. garinii*	Kemeri	2003	45	33	49		36	54		33
298	3-17	*B. afzelii*	Kemeri	2003	109	24		87	118	27	23	29
299	3-9-19	*B. afzelii*	Babite	2003	109	24	23		92		23	78
300	3-17-14	*B. garinii*	Kemeri	2003	46		29	43		42	31	37
301	3-9-02	*B. garinii*	Babite	2003	43	23					77	
302	3-8-24	*B. valaisiana*	Kemeri	2003		35	37	45	38			
303	3-9-11	*B. burgdorferi* s.s.	Babite	2003	14	1	11		168	1		
304	3-3-41	*B. afzelii*	Jaunciems	2003		24	25	87	96			
305	3-02-03	*B. afzelii*	Jaunciems	2003				88				
306	3-6-75	*B. afzelii*	Kemeri	2003		24			98			
307	3-6-77	*B. afzelii*	Kemeri	2003				88	92		23	
308	3-9-17b	*B. afzelii*	Babite	2003	109							
309	3-8-17	*B. garinii*	Kemeri	2003							28	
310	3-5-11	*B. garinii*	Kemeri	2003						38		
311	3-12-24	*B. bavariensis*	Babite	2003						35		
312	3-9-13	*B. garinii*	Babite	2003						39		
313	3-5-13	*B. garinii*	Kemeri	2003	48							
314	10-17-17	*B. garinii*	Babite	2010	44	156	31	40	31	37	80	
315	10-15-17	*B. afzelii*	Babite	2010	109	24	24	86		91	23	28

Table S2: Total number of samples used for spatio-temporal distribution analysis

Species	Babite	Jaunciems	Kemeri	Total
B. afzelii	51	47	33	131
B. bavariensis	1	1	0	2
B. garinii	76	4	27	107
B. lusitaniae	2	0	1	3
B. valaisiana	30	5	13	48
B. burgdorferi s.s.	15	0	9	24
Total	175	57	83	315

Table S3-A: Samples used for spatio-temporal distribution analysis from Babite

Babite	1999	2000	2001	2002	2003	2006	2007	2010	Total
B. afzelii	9	6	26	5	2	1		2	51
B. bavariensis					1				1
B. garinii	7	3	33	1	2	14	12	4	76
B. lusitaniae	1						1		2
B. valaisiana	1		10	1		8	9	1	30
B. burgdorferi s.s.	4	1	6		1		3		15
Total	22	10	75	7	6	23	25	7	175

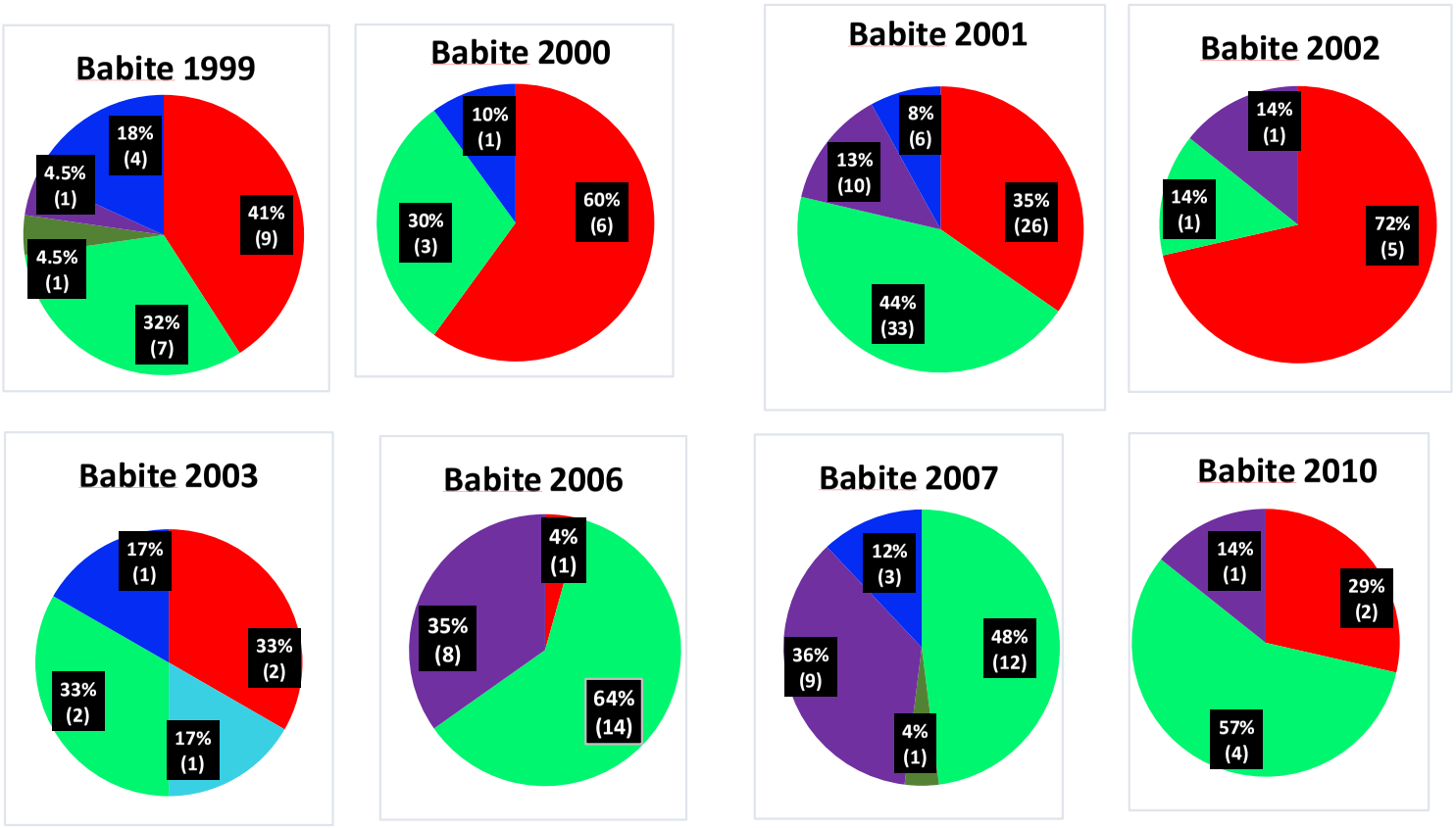
Table S3-B: Samples used for spatio-temporal distribution analysis from Jaunciems

Jaunciems	1999	2000	2001	2002	2003	2006	2007	2010	
B. afzelii	2	6	21	4	4	6	4		47
B. bavariensis	1								1
B. garinii			4						4
B. lusitaniae									0
B. valaisiana		1	4						5
B. burgdorferi s.s.									0
Total	3	7	29	4	4	6	4	0	57

Table S3-C: Samples used for spatio-temporal distribution analysis from Kemeri

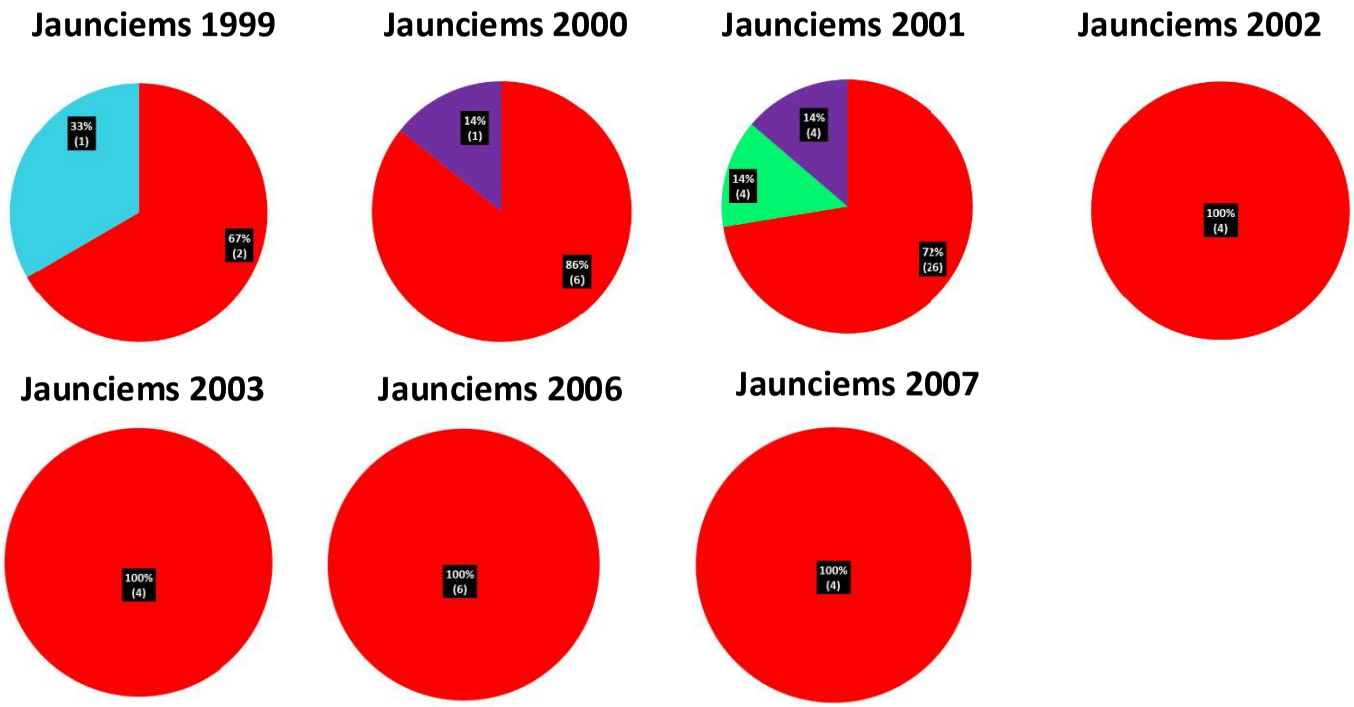
Kemeri	1999	2000	2001	2002	2003	2006	2007	2010	
B. afzelii	3	3	13		5	4	5		33
B. bavariensis									0
B. garinii	4	1	12		6	1	3		27
B. lusitaniae				1					1
B. valaisiana	1	1	5		1	1	4		13
B. burgdorferi s.s.	2	1	4		2				9
Total	10	6	34	1	14	6	12	0	83

Fig. S3: Temporal distribution of *B. burgdorferi* s.l. in Babite



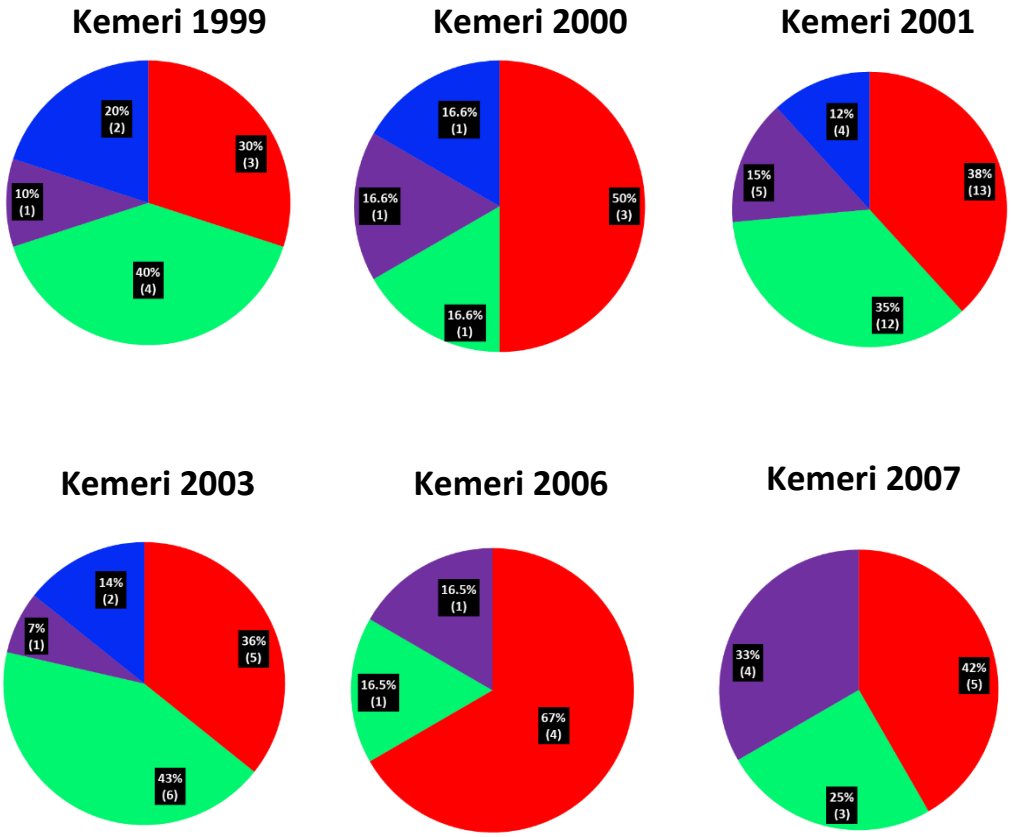
Six different species were identified. The percentage of species is given; the number in parenthesis refers to the actual number. The colors were assigned to the respective species as follows:  *B. afzelii*,  *B. bavariensis*,  *B. burgdorferi* s.s.,  *B. garinii*,  *B. lusitaniae* and  *B. valaisiana.* In 1999, 2002 and 2002 *B. afzelii* dominated, in 2001, 2007 and 2010 *B. garinii* dominated. In 2006 and 2007 *B. valaisiana* was also well represented (> 30 %). *Borrelia burgdorferi* s.s. was not found in the years 2002, 2006 and 2010 whilst *B. valaisiana.*

Fig. S4: Temporal distribution of *B. burgdorferi* s.l. in Jaunciems



Temporal distribution of *B. burgdorferi* s.l. species in Jaunciems. Four species were identified in Jaunciems, whereby *B. afzelii* was the most dominant species in all the years analyzed. The percentage of species is given and number in parenthesis refers to the actual number. The colors were assigned as in fig. S3.

Fig. S5: Temporal distribution of *B. burgdorferi* s.l. in Kemeri



Temporal distribution of *B. burgdorferi* s.l. species in Kemeri. Five species were identified in this region. The percentage of species is given. The number in parenthesis refers to the actual number. The colors were assigned as in fig. S3. *Borrelia afzelii* was the most frequently observed species in 2000, 2001, 2006 and 2007 while *B. garinii* was more abundant than *B. afzelii* in 1999 and in 2003. Other species were also well represented in this region.

Paper 3: Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany

Robert E. Rollins, Zehra Yeyin, Maja Wyczanska, Nikolas Alig, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Noémie S. Becker

Ticks and Tick-borne Diseases (2021), 12 (1): 101589
https://doi.org/10.1016/j.ttbdis.2020.101589

Ticks and Tick-borne Diseases 12 (2021) 101589

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

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

Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany

Robert E. Rollins [a], Zehra Yeyin [a], Maja Wyczanska [a,b], Nikolas Alig [c], Sabrina Hepner [c], Volker Fingerle [c], Gabriele Margos [c], Noémie S. Becker [a,*]

[a] Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Grosshaderner Str. 2, 82152 Planegg-Martinsried, Germany
[b] Department of Pediatrics, Dr. v. Hauner Children's Hospital, LMU Munich, Lindwurmstr. 2a, 80337 Munich, Germany
[c] National Reference Centre for Borrelia, Bavarian Food and Health and Food Safety Authority (LGL), Veterinärstr. 2, 85764 Oberschleissheim, Germany

ARTICLE INFO

Keywords:
Borrelia burgdorferi sensu lato
Ixodid ticks
Drag sampling
Borrelia prevalence

ABSTRACT

Lyme borreliosis (LB) is the most common arthropod-borne disease in Europe and North America and is caused by members of the *Borrelia burgdorferi* sensu lato (*Bbsl*) species complex. These bacteria are transmitted by ixodid tick vectors and therefore human LB risk is influenced by the prevalence and distribution of *Bbsl* genospecies within tick vectors throughout the wild. These distributions can easily change over spatiotemporal scales and, to understand LB risk fully, up to date information on prevalence and distribution of *Bbsl* is required. The last survey of *Bbsl* in southern Germany, including parts of the Munich metropolitan area, was completed in 2006 and new data is needed. Ixodid ticks were collected in seven plots located in and around Munich, Germany, from March to July 2019 and were screened for *Bbsl*. *Borrelia burgdorferi* s. l. positive ticks (52 adults, 158 nymphs) were found in all plots and adults (0–61.5 % *Bbsl* positive/plot) and nymphs (17.4–59.5 % *Bbsl* positive/plot) did not differ significantly in their overall *Bbsl* prevalence. The number of *Bbsl* positive nymphs did vary significantly between plots but the number of positive adults did not. In total, six *Bbsl* genospecies were located with *B. afzelii* and *B. garinii* dominating. Additionally, the relapsing-fever species *B. miyamotoi* was found in two sampling plots. Our results highlight the variability in *Bbsl* prevalence and genospecies distribution over short geographic distances and aid in understanding LB risk in and around the Munich metropolitan area.

1. Introduction

Lyme borreliosis (LB) is the most common arthropod-borne human disease in the northern hemisphere (Stanek et al., 2011), caused by certain genospecies belonging to the *Borrelia burgdorferi* sensu lato (hereafter *Bbsl*) species complex (Stanek et al., 2011). *Borrelia burgdorferi* s. l. bacteria are transmitted between vertebrate reservoir hosts via ixodid tick vectors (Gern, 2008; Kurtenbach et al., 2006). The probability to contract LB is influenced, in part, by the prevalence and distribution of different *Bbsl* species in tick vectors (Randolph, 2004; Takumi et al., 2019). Therefore, determining the prevalence and distribution of *Bbsl* genospecies within a geographic area is a crucial first step to understand human LB risk.

Borrelia burgdorferi s. l. prevalence can vary temporally and is thought to have increased over the past few decades (Rizzoli et al., 2011; Rosà et al., 2018; Sykes and Makiello, 2017). This increase is argued to result from changes in *Bbsl* prevalence due to geographic expansion of ixodid tick vectors (Lindgren et al., 2000; Rosà et al., 2018). Additionally, other studies have proposed that increases in temperature (Wallace et al., 2019) and climate change (Fernández-Ruiz and Estrada-Peña, 2020; Rosà et al., 2018) could be responsible for increased *Bbsl* prevalence Additionally, micro-habitat conditions, such as relative humidity, and other abiotic factors are also very important for tick survival and therefore could impact *Bbsl* prevalence over time in a given region (Randolph, 2008).

In addition to temporal variation, substantial spatial variation in *Bbsl* has also been observed (Estrada-Peña et al., 2018; Strnad et al., 2017). This is influenced by various factors such as tick abundance, host availability (Oorebeek and Kleindorfer, 2008; Takumi et al., 2019), or habitat type and composition (Halos et al., 2010; Ruyts et al., 2016). Various vertebrate species can act as competent reservoir hosts for *Bbsl* genospecies including rodent and bird species (Gern et al., 1998; Kurtenbach et al., 2006, 2002a; Norte et al., 2013). Most *Bbsl* genospecies are adapted to specific hosts (Kurtenbach et al., 2002b, 1998) resulting

* Corresponding author at: Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Großhaderner Str. 2, 82152 Planegg-Martinsried, Germany.
E-mail address: nbecker@bio.lmu.de (N.S. Becker).

https://doi.org/10.1016/j.ttbdis.2020.101589
Received 15 June 2020; Received in revised form 22 September 2020; Accepted 23 September 2020
Available online 10 October 2020
1877-959X/

in a heterogeneous distributions of genospecies in relation to availability of competent hosts and the ecological process that affect host abundance (Etti et al., 2003; Kurtenbach et al., 2006; Ostfeld et al., 2006).

In Germany, 9.4 % of tested human adults showed seropositivity for *Bbsl* (Wilking et al., 2015), which shows that the German population is at risk for *Bbsl* exposure and, therefore, LB. The last major study looking at *Bbsl* prevalence including portions of the Munich metropolitan area was performed in 2006 (Fingerle et al., 2008) and new data is needed. To determine the variability of *Bbsl* around Munich and potential changes in comparison to 2006, ixodid ticks were collected near Munich, Germany, and were screened for *Bbsl* bacteria. This sampling scheme allowed us to determine spatial variation in *Bbsl* prevalence and species distribution over a fine scale geographic area in the aim to better understand LB risk in the Munich metropolitan area.

2. Methods

2.1. Tick sample collection and DNA extraction

Ticks were sampled by drag sampling in seven plots (see Supplemental Table 1 for overview and explanation of plot abbreviations) from March-July 2019. Ticks were collected either in one (FOR, FOP, and GRA), two (NPL, ENG, OBE), or three (STA) sampling events. Both Grafrath (GRA) and Englischer Garten (ENG; MIR in Fingerle et al. (2008)) were also sampled from March to May 2006 by Fingerle et al. (2008). Prior to DNA extraction, ticks were morphologically identified to life-stage and genus. DNA from individual ticks was extracted through alkaline hydrolysis according to a previously published protocol (Okeyo et al., 2019).

2.2. Molecular analysis for Bbsl and tick sample identification

Extracted tick DNA was screened for *Borrelia* DNA using a qPCR targeting the 23S intergenic spacer following standard procedure (Strube et al., 2010). This method has been shown to cross-react with *Borrelia* spp. outside the *Bbsl* species complex, such as *Borrelia turcica* (Hepner et al., 2020) and *B. miyamotoi* (Springer et al., 2020). *Borrelia* negative samples were discarded. Ticks positive for *Borrelia* were subjected to a semi-nested PCR amplifying the housekeeping gene *recG* using a previously described protocol (Margos et al., 2008). To determine which tick species are potentially infected with *Borrelia,* positive tick DNA samples (n = 210) were subjected to a PCR amplifying a fragment of the 16S mitochondrial gene using a previously published method (Noureddine et al., 2011).

For all PCR analyses, multiple internal negative controls were included. Either *B. kurtenbachii* (25015) or *B. mayonii* (DSM 102811. MN14-1420) were used as positive controls for all *Borrelia* related analyses. PCR products were sequenced using Sanger sequencing at the Sequencing Service of Ludwig-Maximilians University and were prepared according to the requirements of the sequencing center (http://www.gi.bio.lmu.de/sequencing/help/protocol). For sequence processing and molecular identification of tick 16S sequences see Supplemental Methods and Supplemental Table 2 for GenBank reference sequences used.

2.3. Statistical analysis

All statistical analysis was done in R (Version 3.6.1) using the base R statistical package (R Core Team, 2019). For most tests, significance was tested with Pearson's chi-square tests using Yates' continuity correction (hereafter χ^2) unless zero counts were present for which a two-sided Fisher's exact test was used with a simulated *p*-value based on 5000 simulations (hereafter Fisher test). Linear models were used to calculated correlation coefficients (R^2).

3. Results

All ticks belonged to the genus *Ixodes.* Of all collected tick samples, 210 tested positive for *Bbsl* DNA of which 200 produced a 16S rRNA amplicon for tick species identification. After quality checks, usable 16S sequences were available for 185 tick samples all of which clustered with *I. ricinus* in the median joining network (MJN) analysis (Supplementary Fig. 1; for details Supplemental Methods).

Borrelia burgdorferi s. l. positive nymphs were found in all plots with prevalence ranging from 17.4 to 59.5% whereas positive adults were found in only six plots with prevalence ranging from 0 to 61.5% (Fig. 1). It is important to mention that some prevalence values could be biased due to low sample sizes. Adults and nymphs did not differ significantly in *Bbsl* prevalence (χ^2: $p = 0.33$). *Borrelia burgdorferi* s. l. prevalence in nymphs was significantly different between plots (χ^2: $p < 0.001$) but prevalence in adults was not (Fisher test: $p = 0.07$). *Borrelia burgdorferi* s. l. prevalence was higher in fragmented plots (adults = 41.0 ± 10.3 %; nymphs = 32.7 ± 13.4 %) than in continuous forest plots (adults = 29.2 ± 11.1 %; nymphs = 27.3 ± 5.6 %) but these differences were not significant (χ^2: adults, $p = 0.54$; nymphs, $p = 0.51$). *Borrelia burgdorferi* s. l. prevalence in nymphs and adults was associated with number of ticks collected (nymphs: $R^2 = 0.67$, $p = 0.015$; adults: $R^2 = 0.65$, $p = 0.017$). We observed a significant increase (χ^2: $p < 0.001$) in *Bbsl* prevalence in nymphs collected during 2019 in GRA when compared to the 2006 prevalence reported by Fingerle et al. (2008) (Supplementary Fig. 2) For absolute numbers regarding this comparison see Supplementary Table 3.

In total, six *Bbsl* species were found: *B. afzelii*, *B. bavariensis*, *B. burgdorferi* s. s., *B. garinii*, *B. spielmanii*, *B. lusitaniae* and *B. valaisiana* (Fig. 2). *Borrelia afzelii* and *B. garinii* were the most common species described (Fig. 2). In addition to *Bbsl* genospecies, three *B. miyamotoi*-positive adult ticks were found in two plots (FOP & STA). Plots varied significantly in their species compositions (Fisher test: $p < 0.001$, Fig. 2). For an overview of *Borrelia*-positive ticks including absolute values see Supplementary Table 4.

4. Discussion

A crucial step in understanding human LB risk, is to know the prevalence and distribution of *Bbsl* genospecies. *Borrelia burgdorferi* s. l. is known to vary both spatially (Estrada-Peña et al., 2018; Strnad et al., 2017) and temporally (Coipan et al., 2013; Okeyo et al., 2020; Rauter and Hartung, 2005) making general extrapolations challenging. The last major *Bbsl* prevalence research including parts of the Munich metropolitan area was published in 2008 (Fingerle et al., 2008). In this study, we aimed to update this research and provide a better understanding of human LB risk in the Munich area.

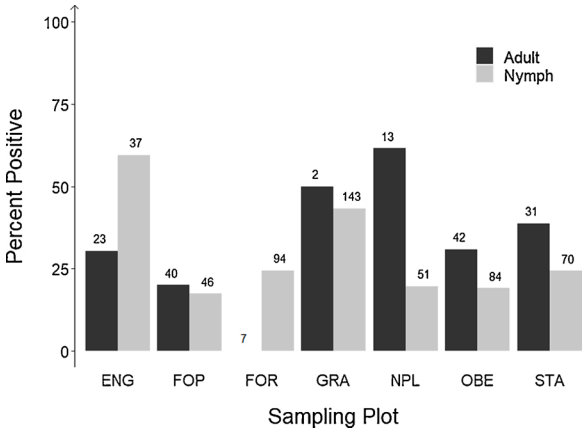


Fig. 1. Prevalence of *Borrelia burgdorferi* sensu lato in *Ixodes ricinus* and *Ixodes* spp. ticks collected in 2019. The number above each bar reflects the total number of ticks screened (*Borrelia* positive and negative) per life stage per plot.

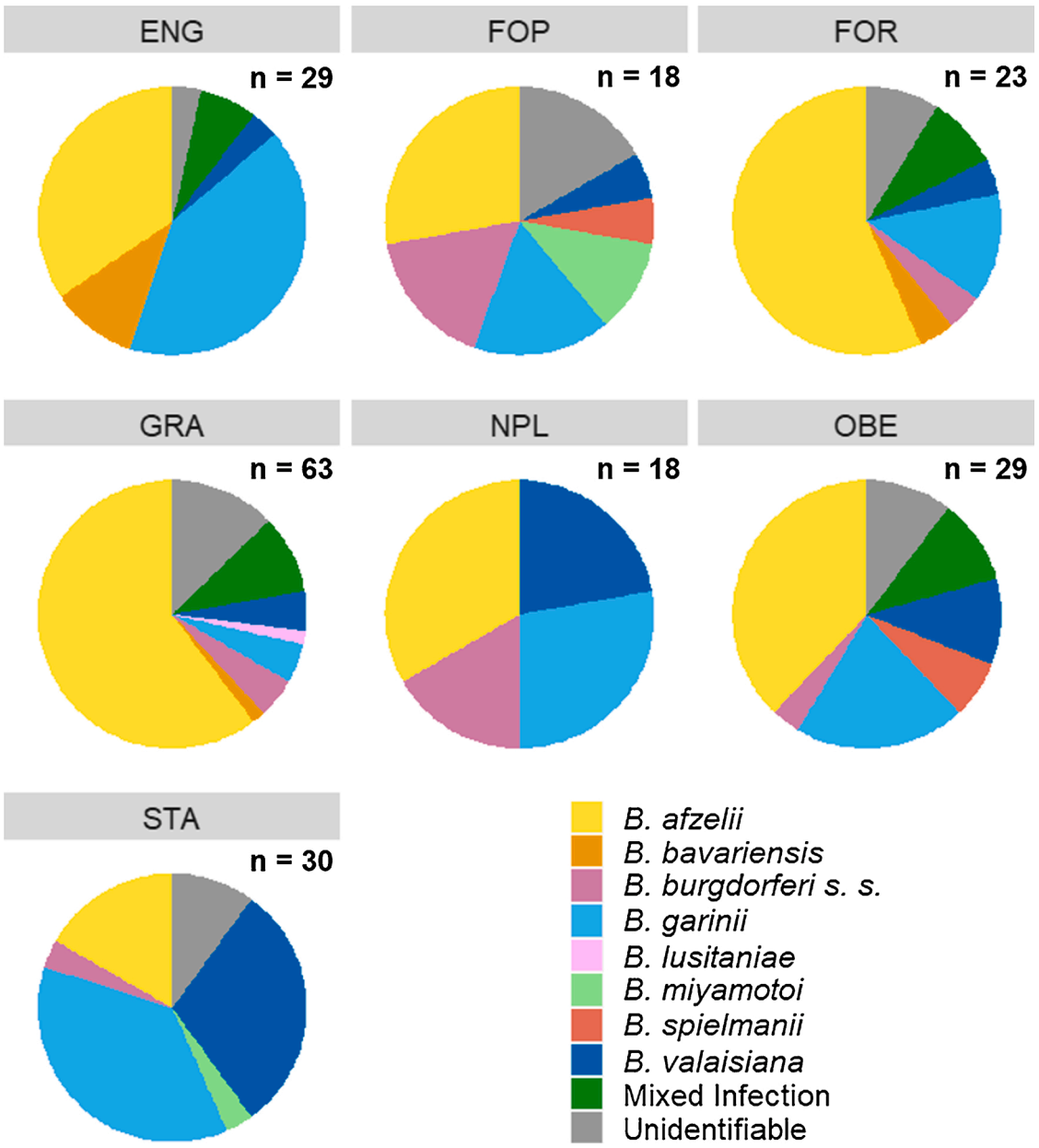


Fig. 2. Distribution of *B. burgdorferi* sensu lato genospecies and *B. miyamotoi* in each study plot in 2019. The number reflects the total number of *Borrelia* positive ticks per plot. Unidentifiable refers to a sample that was positive in the screening PCR but did not amplify *recG* in the semi-nested PCR as described by Okeyo et al. (2019).

All successfully sequenced *Bbsl* positive tick samples clustered with *I. ricinus* in a 16S haplotype network (Supplementary Fig. 1) and where thus molecularly identified as belonging to *I. ricinus*. This is expected as *I. ricinus* is the most common vector of *Bbsl* in Europe (Kurtenbach et al., 2006). As *Bbsl*-negative tick samples were not sequenced for species identification, it is possible that other *Ixodes* species may be present in the seven sampled plots (Estrada-Peña et al., 2018; Strnad et al., 2017).

Borrelia burgdorferi s. l. prevalence in adults and nymphs did not differ significantly in our study. Adults usually tend to have significantly higher *Bbsl* prevalence in comparison to nymphs (Strnad et al., 2017), as adults have had one more blood meal, which increases the probability of *Bbsl* infection (Kurtenbach et al., 2006; Strnad et al., 2017). One argument for higher *Bbsl* prevalence in nymphs, is that many nymphs one year prior to sampling feed on hosts that are non-reservoir hosts of *Bbsl* (Fernández-Ruiz and Estrada-Peña, 2020; Kurtenbach et al., 2006, 2002a, 1998) resulting in fewer infected adults one year later (when ticks were collected). Further sampling years would be needed to determine if this is a general or unique occurrence.

The seven plots in our study did differ significantly in *Bbsl* prevalence as expected from previous literature (Estrada-Peña et al., 2018; Fingerle et al., 2008; Strnad et al., 2017). Fragmented forest plots and continuous forests differed in *Bbsl* prevalence (higher in fragmented plots) and genospecies diversity (higher in continuous forest plots). Although not significant, the observed variation could be due to habitat modifications affecting host population composition and size (Ehrmann et al., 2018; Haddad et al., 2015; Halos et al., 2010). Fragmented plots tend to have increased edge habitats or "ecotones" which have been shown to correlate to increased *Bbsl* prevalence (Ehrmann et al., 2018) but habitat fragmentation is also linked to decreased vertebrate biodiversity (Haddad et al., 2015) which has been argued to decrease *Bbsl* diversity (Ruyts et al., 2016). Together this could explain why fragmented habitats have higher *Bbsl* prevalence but overall lower genospecies diversity. Further

data on host abundance and diversity would be needed to determine if these relationships explain the observed variation in *Bbsl* prevalence and species distributions.

Fingerle et al. (2008) studied the prevalence of *Bbsl* in southern Germany in 2006, including two plots (ENG & GRA) resampled during this study. We did observe variation in *Bbsl* prevalence in both plots when compared to 2006, but only had a large enough sample size of nymphs coming from GRA to test this variation. It is important to mention however, that due to our low sample size, these results would need to be supported with long-term studies within more plots to truly determine the temporal trend in this area. The apparent increase of *Bbsl* in GRA nymphs could be due to use of different screening methods between 2006 and 2019, as this study used a more sensitive screening PCR (Strnad et al., 2017) or the result of climate change as described in other studies (Fernández-Ruiz and Estrada-Peña, 2020; Lindgren et al., 2000; Rosà et al., 2018; Wallace et al., 2019). However, sporadic fluctuations in *Bbsl* are possible as shown by a long-term study in the Netherlands (Coipan et al., 2013). These fluctuations are thought to be linked to host population size including the availability and diversity of *Bbsl* reservoir hosts (Coipan et al., 2013; Kurtenbach et al., 2006; Ostfeld et al., 2006).

At the genospecies level, seven *Bbsl* genospecies were described in 2006: *B. afzelii, B. garinii*, *B. bavariensis* (described as *B. garinii* OspA type 4 as *B. bavariensis* was raised to species level in 2009 (Margos et al., 2009)), *B. burgdorferi* s. s., *B. valaisiana, B. spielmanii*, and *B. lusitaniae* (Fingerle et al., 2008). Five of these genospecies are known to cause LB in humans (Margos et al., 2011; Stanek et al., 2011). All seven genospecies observed in 2006 (Fingerle et al., 2008), were also found in this study. *Borrelia garinii* and *B. afzelii* were the most common species found in both years which is in accordance with published results in Germany (Fingerle et al., 2008; Răileanu et al., 2020; Zubriková et al., 2020) and Europe (Estrada-Peña et al., 2018; Strnad et al., 2017).

In addition to *Bbsl* genospecies, *B. miyamotoi* (n = 3 adult ticks) was also found in two plots during this study (Fig. 2). *Borrelia miyamotoi* belongs to the relapsing-fever spirochetes, not the *Bbsl* species complex, and has been associated with febrile illness in humans (Platonov et al., 2011). The prevalence reported here (1.9 % of all adult ticks) is in agreement with other research coming from Germany (Răileanu et al., 2020). The occurrence of *B. miyamotoi* in two locations used for recreation by the local population is of potential importance to public health.

5. Conclusions

In conclusion, we found substantial geographic variation in *Bbsl* prevalence and species distribution in seven plots located in and around the Munich metropolitan area. Additionally, this work supported that there is a potential risk for LB in these areas as many of the species described are known to be human pathogenic (Platonov et al., 2011; Stanek et al., 2011). This work also displayed that species presence was, for the most part, stable with an apparent increase in *Bbsl* prevalence. A study over a longer time period is necessary to confirm if this trend continues or is potentially a single year event (Coipan et al., 2013; Rosà et al., 2018).

Authors' contributions

R.E.R., S.H., Z.Y., N.A., and M.W. collected data, and R.E.R. performed all data analysis with the guidance of N.S.B., G. M., & V. F. All co-authors helped in developing the study design. R.E.R. wrote the manuscript with input and final approval from all co-authors.

Funding

The project was funded through the German Research Foundation (DFG Grant No. BE 5791/2-1). The National Reference Centre for *Borrelia* was funded by the Robert-Koch-Institut, Berlin.

CRediT authorship contribution statement

Robert E. Rollins: Conceptualization, Investigation, Formal analysis, Writing - original draft. **Zehra Yeyin:** Investigation. **Maja Wyczanska:** Investigation. **Nikolas Alig:** Investigation. **Sabrina Hepner:** Investigation, Writing - review & editing. **Volker Fingerle:** Conceptualization, Supervision, Writing - review & editing. **Gabriele Margos:** Conceptualization, Supervision, Writing - review & editing. **Noémie S. Becker:** Conceptualization, Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgements

We would like to thank all past members of the "Evolutionary Biology" group at the LMU, all lab technicians and colleagues at the "National Reference Centre for *Borrelia*," all members of the "Evolutionary Ecology of Variation" group of the Max Plank Institute for Ornithology and the "Behavioural Ecology" group at LMU for allowing us to collect in their study plots, and all students for help in data collection.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ttbdis.2020.101589.

References

Coipan, E.C., Jahfari, S., Fonville, M., Maassen, C.B., van der Giessen, J., Takken, W., Takumi, K., Sprong, H., 2013. Spatiotemporal dynamics of emerging pathogens in questing *Ixodes ricinus*. Front. Cell. Infect. Microbiol. 4, 1–11.

Ehrmann, S., Ruyts, S.C., Scherer-Lorenzen, M., Bauhus, J., Brunet, J., Cousins, S.A.O., Deconchat, M., Decocq, G., De Frenne, P., De Smedt, P., Diekmann, M., Gallet-Moron, E., Gärtner, S., Hansen, K., Kolb, A., Lenoir, J., Lindgren, J., Naaf, T., Paal, T., Panning, M., Prinz, M., Valdés, A., Verheyen, K., Wulf, M., Liira, J., 2018. Habitat properties are key drivers of *Borrelia burgdorferi* (s.l.) prevalence in *Ixodes ricinus* populations of deciduous forest fragments. Parasit. Vectors 11, 1–15.

Estrada-Peña, A., Cutler, S., Potkonjak, A., Vassier-Tussaut, M., Bortel, W., Zeller, H., Fernández-Ruiz, N., Mihalca, A.D., 2018. An updated meta-analysis of the distribution and prevalence of *Borrelia burgdorferis*.l. in ticks in Europe. Int. J. Health Geogr. 17, 1–16.

Etti, S., Hails, R., Schäfer, S.M., De Michelis, S., Sewell, H.S., Bormane, A., Donaghy, M., Kurtenbach, K., 2003. Habitat-specific diversity of *Borrelia burgdorferi* sensu lato in Europe, exemplified by data from Latvia. Appl. Environ. Microbiol. 69, 3008–3010.

Fernández-Ruiz, N., Estrada-Peña, A., 2020. Could climate trends disrupt the contact rates between*Ixodes ricinus* (Acari, Ixodidae) and the reservoirs of *Borrelia burgdorferi* s. l. ? PLoS One 15, 1–18.

Fingerle, V., Schulte-Spechtel, U.C., Ruzic-Sabljic, E., Leonhard, S., Hofmann, H., Weber, K., Pfister, K., Strle, F., Wilske, B., 2008. Epidemiological aspects and molecular characterization of *Borrelia burgdorferis*.l. from southern Germany with special respect to the new species*Borrelia spielmanii* sp. nov. Int. J. Med. Microbiol. 298, 279–290.

Gern, L., 2008. *Borrelia burgdorferi* sensu lato, the agent of Lyme borreliosis: life in the wilds. Parasite 15, 244–247.

Gern, L., Estrada-Peña, A., Frandsen, F., Gray, J.S., Jaenson, T.G.T., Jongejan, F., Kahl, O., Korenberg, E., Mehl, R., Nuttall, P.A., 1998. European reservoir hosts of Borrelia burgdorferi sensu lato. Zentralblatt für Bakteriol. 287, 196–204.

Haddad, N.M., Brudvig, L.A., Clobert, J., Davies, K.F., Gonzalez, A., Holt, R.D., Lovejoy, T.E., Sexton, J.O., Austin, M.P., Collins, C.D., Cook, W.M., Damschen, E.I., Ewers, R.M., Foster, B.L., Jenkins, C.N., King, A.J., Laurance, W.F., Levey, D.J., Margules, C.R., Melbourne, B.A., Nicholls, A.O., Orrock, J.L., Song, D.X., Townshend, J.R., 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. Sci. Adv. 1, 1–10.

Halos, L., Bord, S., Cotté, V., Gasqui, P., Abrial, D., Barnouin, J., Boulouis, H.J., Vayssier-Taussat, M., Vourc'H, G., 2010. Ecological factors characterizing the prevalence of bacterial tick-borne pathogens in *Ixodes ricinus* ticks in pastures and woodlands. Appl. Environ. Microbiol. 76, 4413–4420.

Hepner, S., Fingerle, V., Duscher, G.G., Felsberger, G., Marosevic, D., Rollins, R.E., Okeyo, M., Sing, A., Margos, G., 2020. Population structure of *Borrelia turcica* from Greece and Turkey. Infect. Genet. Evol. 77.

Kurtenbach, K., Sewell, H.-S., Ogden, N.H., Randolph, S.E., Nuttall, P.A., 1998. Serum complement sensitivity as a key factor in Lyme disease ecology. Infect. Immun. 66, 1248–1251.

Kurtenbach, K., De Michelis, S., Etti, S., Schäfer, S.M., Sewell, H.S., Brade, V., Kraiczy, P., 2002a. Host association of *Borrelia burgdorferi*sensu lato - The key role of host complement. Trends Microbiol. 10, 74–79.
Kurtenbach, K., Schäfer, S.M., Sewell, H., Peacey, M., Hoodless, A., Nuttall, P.A., Randolph, S.E., Scha, S.M., 2002b. Differential survival of Lyme borreliosis spirochetes in ticks that feed on birds. Infect. Immun. 70, 5893–5895.
Kurtenbach, K., Hanincová, K., Tsao, J.I., Margos, G., Fish, D., Ogden, N.H., 2006. Fundamental processes in the evolutionary ecology of Lyme borreliosis. Nat. Rev. Microbiol. 4, 660–669.
Lindgren, E., Tälleklint, L., Polfeldt, T., 2000. Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick *Ixodes ricinus*. Environ. Health Perspect. 108, 119–123.
Margos, G., Gatewood, A.G., Aanensen, D.M., Ra Hanincová, K., Terekhova, D., Vollmer, S.A., Cornet, M., Piesman, J., Donaghy, M., Bormane, A., Hurn, M.A., Feil, E.J., Fish, D., Casjens, S., Wormser, G.P., Schwartz, I., Kurtenbach, K., 2008. MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. Proc. Natl. Acad. Sci. U. S. A. 105, 8730–8735.
Margos, G., Vollmer, S.A., Cornet, M., Garnier, M., Fingerle, V., Wilske, B., Bormane, A., Vitorino, L., Collares-Pereira, M., Drancourt, M., Kurtenbach, K., 2009. A new *Borrelia* species defined by multilocus sequence analysis of housekeeping genes. Appl. Environ. Microbiol. 75, 5410–5416.
Margos, G., Vollmer, S.A., Ogden, N.H., Fish, D., 2011. Population genetics, taxonomy, phylogeny and evolution of *Borrelia bugdorferi* sensu lato. Infect. Genet. Evol. 11, 1545–1563.
Norte, A.C., Ramos, J.A., Gern, L., Núncio, M.S., Lopes de Carvalho, I., 2013. Birds as reservoirs for *Borrelia burgdorferis*.l. in Western Europe: circulation of*B. turdi* and other genospecies in bird-tick cycles in Portugal. Environ. Microbiol. 15, 386–397.
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. Int. J. Parasitol. 41, 183–192.
Okeyo, M., Hartberger, C., Margos, G., Straubinger, R.K., Sing, A., Fingerle, V., 2019. Comparison of methods for economic and efficient tick and *Borrelia* DNA purification. Ticks Tick. Dis. 10, 1041–1045.
Okeyo, M., Hepner, S., Rollins, R.E., Hartberger, C., Straubinger, R.K., Marosevic, D., Bannister, S.A., Bormane, A., Donaghy, M., Sing, A., Fingerle, V., Margos, G., 2020. Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010. Environ. Microbiol. 00, 1999–2010.
Oorebeek, M., Kleindorfer, S., 2008. Climate or host availability: what determines the seasonal abundance of ticks? Parasitol. Res. 103, 871–875.
Ostfeld, R.S., Canham, C.D., Oggenfuss, K., Winchcombe, R.J., Keesing, F., 2006. Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. PLoS Biol. 4, 1058–1068.
Platonov, A.E., Karan, L.S., Kolyasnikova, N.M., Makhneva, N.A., Toporkova, M.G., Maleev, V.V., Fish, D., Krause, P.J., 2011. Humans infected with relapsing fever spirochete *Borrelia miyamotoi*, Russia. Emerg. Infect. Dis. 17, 1816–1823.
R Core Team, 2019. R: a Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Răileanu, C., Tauchmann, O., Vasić, A., Wöhnke, E., Silaghi, C., 2020. Borrelia miyamotoi and Borrelia burgdorferi (sensu lato) identification and survey of tick-borne encephalitis virus in ticks from north-eastern Germany. Parasit. Vectors 13, 1–11.
Randolph, S.E., 2004. Tick ecology: Processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors. Parasitology 129, S37–S65.
Randolph, S.E., 2008. The impact of tick ecology on pathogen transmission dynamics. Ticks: Biology, Disease and Control. Cambridge University Press, pp. 40–72.
Rauter, C., Hartung, T., 2005. Prevalence of *Borrelia burgdorferi* sensu lato in *Ixodes ricinus* ticks in Scandinavia. Appl. Environ. Microbiol. 71, 7203–7216.
Rizzoli, A., Hauffe, H.C., Carpi, G., Vourc'h, G.I., Neteler, M., Rosà, R., 2011. Lyme borreliosis in Europe. Eurosurveillance 16, 1–8.
Rosà, R., Andreo, V., Tagliapietra, V., Baráková, I., Arnoldi, D., Hauffe, H.C., Manica, M., Rosso, F., Blaňarová, L., Bona, M., Derdáková, M., Hamšíková, Z., Kazimírová, M., Kraljik, J., Kocianová, E., Mahríková, L., Minichová, L., Mošanský, L., Slovák, M., Stanko, M., Špitalská, E., Ducheyne, E., Neteler, M., Hubálek, Z., Rudolf, I., Venclikova, K., Silaghi, C., Overzier, E., Farkas, R., Földvári, G., Hornok, S., Takács, N., Rizzoli, A., 2018. Effect of climate and land use on the spatio-temporal variability of tick-borne bacteria in Europe. Int. J. Environ. Res. Public Health 15, 1–15.
Ruyts, S.C., Ampoorter, E., Coipan, E.C., Baeten, L., Heylen, D., Sprong, H., Matthysen, E., Verheyen, K., 2016. Diversifying forest communities may change Lyme disease risk: extra dimension to the dilution effect in Europe. Parasitology 143, 1310–1319.
Springer, A., Raulf, M.K., Fingerle, V., Strube, C., 2020. *Borrelia* prevalence and species distribution in ticks removed from humans in Germany, 2013–2017. Ticks Tick. Dis. 11, 101363.
Stanek, G., Fingerle, V., Hunfeld, K.-P., Jaulhac, B., Kaiser, R., Krause, A., Kristoferitsch, W., O'Connell, S., Ornstein, K., Strle, F., Gray, J., 2011. Lyme borreliosis: Clinical case definitions for diagnosis and management in Europe. Clin. Microbiol. Infect. 17, 69–79.
Strnad, M., Hönig, V., Růžek, D., Grubhoffer, L., Regoa, R.O.M., 2017. Europe-wide meta-analysis of *Borrelia burgdorferi* sensu lato prevalence in questing *Ixodes ricinus* ticks. Appl. Environ. Microbiol. 83, 1–16.
Strube, C., Montenegro, V.M., Epe, C., Eckelt, E., Schnieder, T., 2010. Establishment of a minor groove binder-probe based quantitative real time PCR to detect *Borrelia burgdorferi* sensu lato and differentiation of *Borrelia spielmanii* by ospA-specific conventional PCR. Parasit. Vectors 3, 1–10.
Sykes, R.A., Makiello, P., 2017. An estimate of Lyme borreliosis incidence in Western Europe. J. Public Health (United Kingdom) 39, 74–81.
Takumi, K., Sprong, H., Hofmeester, T.R., 2019. Impact of vertebrate communities on *Ixodes ricinus*-borne disease risk in forest areas. Parasit. Vectors 12, 1–12.
Wallace, D., Ratti, V., Kodali, A., Winter, J.M., Ayres, M.P., Chipman, J.W., Aoki, C.F., Osterberg, E.C., Silvanic, C., Partridge, T.F., Webb, M.J., 2019. Effect of rising temperature on Lyme disease: Ixodes scapularis population dynamics and Borrelia burgdorferi transmission and prevalence. Can. J. Infect. Dis. Med. Microbiol. 2019, 1–15.
Wilking, H., Fingerle, V., Klier, C., Thamm, M., Stark, K., 2015. Antibodies against *Borrelia burgdorferi* sensu lato among Adults, Germany, 2008–2011. Emerg. Infect. Dis. 21, 107–110.
Zubriková, D., Wittmann, M., Hönig, V., Švec, P., Víchová, B., Essbauer, S., Dobler, G., Grubhoffer, L., Pfister, K., 2020. Prevalence of tick-borne encephalitis virus and Borrelia burgdorferi sensu lato in Ixodes ricinus ticks in Lower Bavaria and Upper Palatinate, Germany. Ticks Tick. Borne. Dis. 11, 101375.

Paper 3: Supplementary Material

Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany

Robert E. Rollins, Zehra Yeyin, Maja Wyczanska, Nikolas Alig, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Noémie S. Becker

Ticks and Tick-borne Diseases (2021), 12 (1): 101589
https://doi.org/10.1016/j.ttbdis.2020.101589

Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany

Ticks and Tick-borne Diseases

Robert E. Rollins, Zehra Yeyin, Maja Wyczanska, Nikolas Alig, Sabrina Hepner, Volker Fingerle, Gabriele Margos, & Noémie S. Becker

—Online Supplementary Material—

—Supplemental Methods—

Sequence processing of PCR proudcts

Chromatograms for both *recG* and 16S sequences were manually checked for quality using FinchTV Version 1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com) and *recG* sequences containing ambiguities were marked as mixed infections. Non-mixed sequences were then aligned to the *recG* reference from *B. burdorferi* sensu stricto (s.s.) strain B31 (GenBank: AE000783.1) and trimmed. Trimmed sequences were blasted to the PubMLST database for *Borrelia* (https://pubmlst.org/borrelia) for species determination.

Tick 16S sequences were aligned and trimmed to the *I. ricinus* 16S rRNA GenBank reference (GenBank: L34292.1). The trimmed 16S sequences were then compiled into a dataset of 69 GenBank references (Supplemental Table 2) and analyzed using a median joining network (MJN) to determine tick species. The MJN was calculated in Network 5.0.1.1. (Fluxus Technology Ltd., Stanway, England) using haplotype RTF files generated by DnaSP v.6 which did not consider gaps or missing data and removed invariable sites (108).

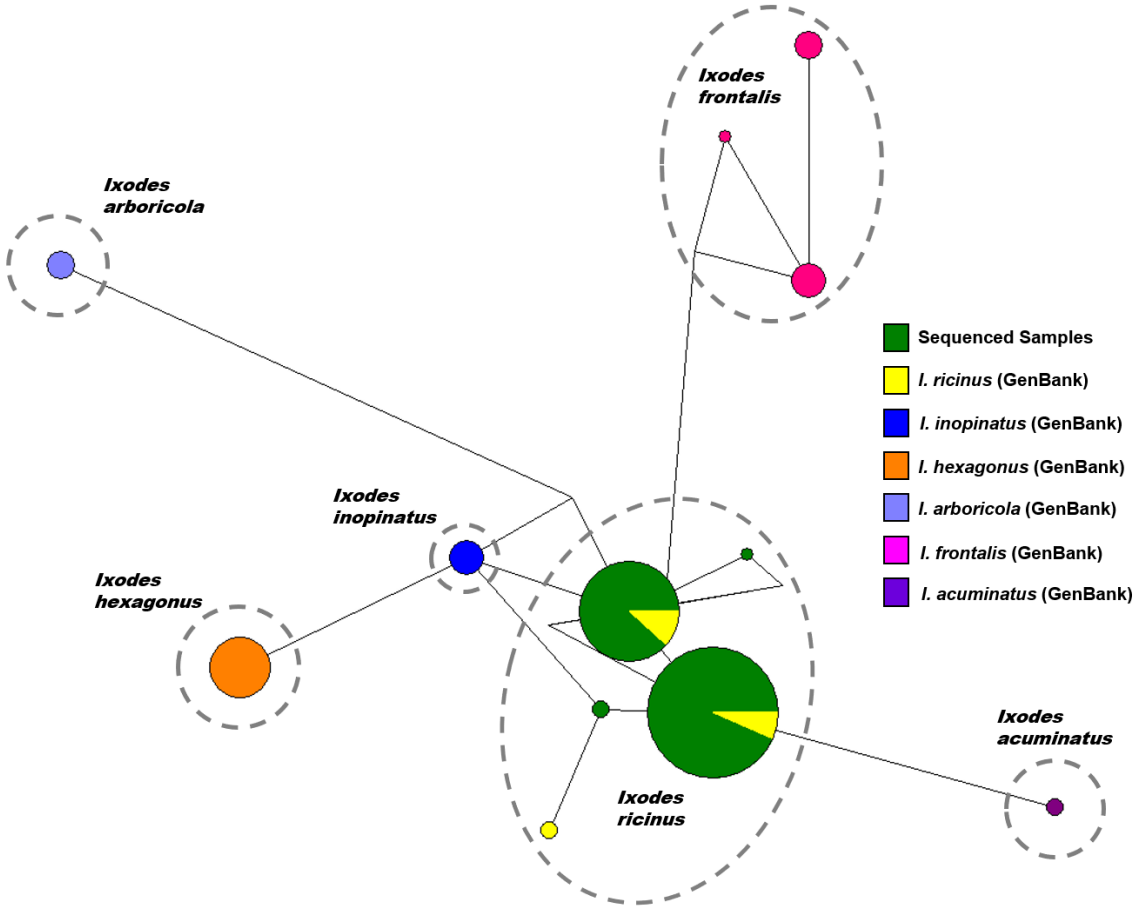
Supplemental Table 1. Sampling plots for ixodid ticks in and around Munich during 2019.

Site	Abbr.	Habitat Type	Altitude (m)	Area (m^2)	GPS Coordinates	Sampling Events	Exact Sampling Dates
Forstenrieder Park 1	FOR	Continuous forest	575	24000	(48.08 N, 11.45 E)	1	29.03.19
Forstenrieder Park 2	FOP	Continuous forest	576	51000	(48.07 N, 11.48 E)	1	03.06.19
Grafrath	GRA	Continuous forest	583	20000	(48.14 N, 11.17 E)	1	18.04.19
Neuperlach	NPL	Field/fragmented forest	542	5000	(48.09 N, 11.66 E)	2	17.04.19; 26.05.19
Englischer Garten	ENG	Fragmented forest	499	10500	(48.18 N, 11.62 E)	2	25.05.19; 13.06.19
Oberschleissheim	OBE	Fragmented forest	492	11000	(48.25 N, 11.56 E)	2	26.05.19; 30.05.19
Starnberg	STA	Continuous forest	685	34000	(48.02 N, 11.34 E)	3	20.06.19; 24.06.19; 03.07.19

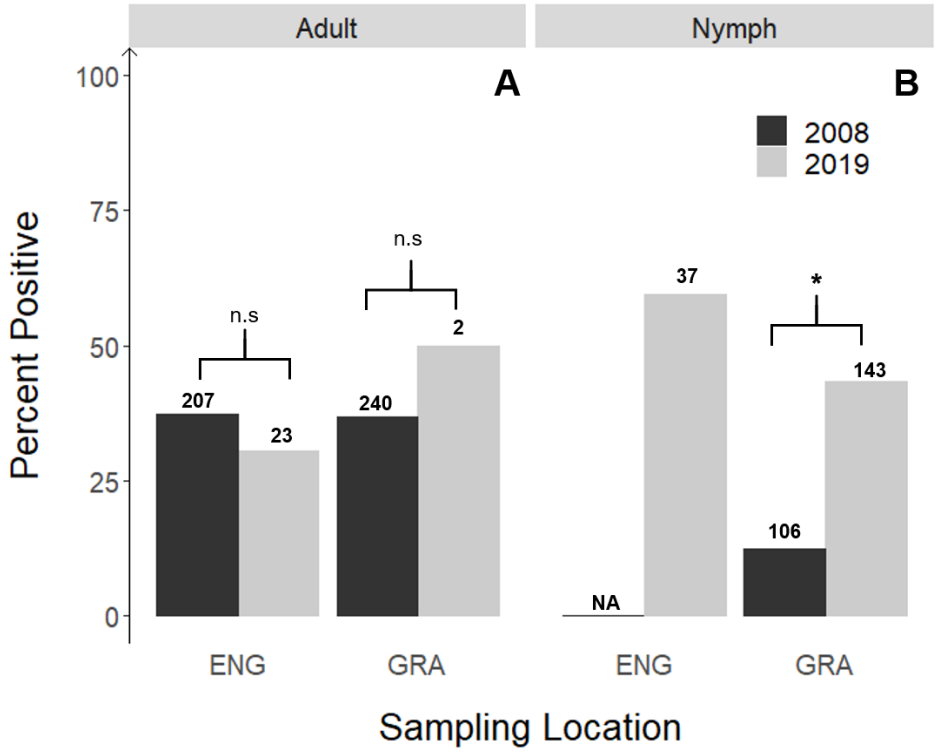
Supplemental Table 2. GenBank accession numbers of 16S references sequences used in the median-joining network analysis on *Borrelia* positive tick specimens collected during this study.

Species Name	*n*	GenBank Accession Numbers
Ixodes arboricola	5	KJ414453.1, KP713676.1, KP713675.1, JF791813.1, JF791812.1
Ixodes acuminatus	2	MH708166.1, MH645515.1
Ixodes frontalis	13	MH645516.1, KJ414455.1, KU170518.1, KP769863.1, MF370645.1, MF370644.1, KJ414454.1, MF688050.1, KP769861.1, KU170519.1, KP769862.1, MF370647.1, MF370646.1
Ixodes hexagonus	22	EU443441.1, KY962076.1, KY962063.1, KY962057.1, LR596328.1, MK731999.1, MK731998.1, MK731997.1, MK731996.1, MK731995.1, KY319189.1, JF928502.1, JF928504.1, KY962058.1, KY962070.1, KY962077.1, JF928501.1, JF928503.1, JF928505.1, U14147.1, AF001400.1, AF001399.1
Ixodes inopinatus	7	KY781393.1, KY781392.1, KY781391.1, KY781390.1, KY781389.1, KY781388.1, KY781387.1
Ixodes ricinus	20	KM211786.1, KY569421.1, EU443401.1, EU443397.1, EU443403.1, EU443402.1, EU443436.1, EU443396.1, EU443400.1, EU443434.1, KY569419.1, KM211787.1, EU443438.1, KM211788.1, KY569420.1, KM211785.1, EU443399.1, EU443432.1, EU443398.1, L34292.1

Supplementary Figure 1. Median-joining haplotype network on the 16S rRNA mitochondrial gene amplified from *Borrelia burgdorferi* s. l. positive tick samples (shown in dark green, n=185). Multiple GenBank reference 16S sequences of various *Ixodes* tick species were included in the analysis: *I. ricinus* (n=20), *I. frontalis* (n=13), *I. hexagonus* (n=22), *I. inopinatus* (n=7), *I. arboricola* (n=5), and *I. acuminatus* (n=2). Accession numbers for reference sequences can be found in Supplemental Table 2.



Supplemental Figure 2. Comparison of 2019 to data from Fingerle et al., (2008). A) *Borrelia burgdorferi* s. l. prevalence in adult ticks collected in Englischer Garten (ENG), Grafrath (GRA); B) *Borrelia burgdorferi* s. l. prevalence in nymphal ticks collected in Englischer Garten (ENG), Grafrath (GRA). No nymphs were reported from Englischer Garten in 2008. * denotes a significant difference in prevalence. It is important to mention that the plot referred to as ENG in our study was denoted as MIR (meadows of the Isar river) in Fingerle et. al., (2008).



Supplementary Table 3. *Borrelia burgdorferi* s. l. prevalence comparison between this study (2019) and Fingerle et. al., (2008) in the two resampled plots: Grafrath (GRA) and Englischer Garten (ENG). It is important to mention that the plot we name ENG is described as MIR (meadows of the Isar river) in Fingerle et. al., (2008).

Site	Year[a]	Tick stage[c]	*n*	Pos[d] (%)
ENG[b]	2008	A	207	74 (35.7)
	2019	A	23	7 (30.4)
	2008	N	0	-
	2019	N	37	22 (59.5)
GRA[b]	2008	A	240	88 (36.7)
	2019	A	2	1 (50.0)
	2008	N	106	13 (12.3)
	2019	N	143	62 (43.4)

[a]2019 data comes from the current study and the data presented for 2008 was published in Fingerle et. al., (2008)
[b]Locations: ENG (Englischer Garten) but this refers to MIR (meadows of the Isar river) in Fingerle et. al., (2008).
[c]Tick life-stage: A (adults both male and female), N (nymphs)
[d]Number of positive ticks of all collected (*n*) followed by the percent positive.

Supplementary Table 4. Ticks collected and prevalence of *Borrelia burgdorferi* s. l. (with genospecies identification) and *B. miyamotoi* in each of the seven collection plots in 2019 (see Table 1 in the main text for explanation of the abbreviations).

Site	Tick stage[b]	*n*	Pos[c] (%)	Genospecies[a]:									
				Bbss (%)	Ba (%)	Bg (%)	Bv (%)	Bb (%)	Bl (%)	Bm (%)	Bs (%)	MI (%)	Und (%)
ENG	A	23	7 (30.4)	0	3 (42.9)	3 (42.9)	1 (14.3)	0	0	0	0	0	0
	N	37	22 (59.5)	0	7 (31.8)	9 (40.9)	0	3 (13.6)	0	0	0	2 (9.1)	1 (4.5)
FOP	A	40	10 (25.0)	1 (10.0)	2 (20.0)	2 (20.0)	1 (10.0)	0	0	2 (20.0)	0	0	2 (20.0)
	N	46	8 (17.4)	2 (0.25)	3 (37.5)	1 (12.5)	0	0	0	0	1 (12.5)	0	1 (12.5)
FOR	A	7	0	0	0	0	0	0	0	0	0	0	0
	N	94	23 (24.5)	1 (4.3)	13 (56.5)	3 (13.0)	1 (4.3)	1 (4.3)	0	0	0	2 (8.7)	2 (8.7)
GRA	A	2	1 (50.0)	0	0	0	0	0	0	0	0	0	1 (100)
	N	143	62 (43.4)	3 (4.8)	38 (61.3)	3 (4.8)	3 (4.8)	1 (1.6)	1 (1.6)	0	0	6 (9.7)	7 (11.3)
NPL	A	13	8 (61.5)	2 (0.25)	2 (0.25)	2 (0.25)	2 (0.25)	0	0	0	0	0	0
	N	51	10 (19.6)	1 (10.0)	4 (40.0)	3 (30.0)	2 (20.0)	0	0	0	0	0	0
OBE	A	42	13 (31.0)	1 (7.7)	4 (30.8)	2 (15.4)	1 (7.7)	0	0	0	0	3 (23.1)	2 (15.4)
	N	84	16 (19.0)	0	7 (43.8)	4 (25.0)	2 (12.5)	0	0	0	2 (12.5)	0	1 (6.2)
STA	A	31	13 (41.9)	1 (7.8)	0	6 (46.2)	4 (30.8)	0	0	1 (7.8)	0	0	1 (7.8)
	N	70	17 (24.3)	0	5 (29.4)	5 (29.4)	5 (29.4)	0	0	0	0	0	2 (11.8)

[a]*Borrelia* genospecies identified: Bbss (*B. burgdorferi* sensu stricto), Ba (*B. afzelii*), Bg (*B. garinii*), Bv (*B. valaisiana*), Bb (*B. bavariensis*), Bl (*B. lusitaniae*), Bm (*B. miyamotoi*), Bs (*B. spielmanii*), MI (Mixed Infections), Und (Unidentifiable).
[b]Tick life-stage: A (adults both male and female), N (nymphs). Larvae were not collected in this study.
[c]Number of positive ticks of all collected (*n*) followed by the percent positive.

Paper 4: High conservation combined with high plasticity: genomics and evolution of *Borrelia bavariensis*

Noémie S. Becker, **Robert E. Rollins**, Kateryna Nosenko, Alexander Paulus, Samantha Martin, Stefan Krebs, Ai Takano, Kozue Sato, Sergey Y. Kovalev, Hiroki Kawabata, Volker Fingerle and Gabriele Margos

BMC Genomics (2020) 21:702
https://doi.org/10.1186/s12864-020-07054-3

Becker et al. BMC Genomics (2020) 21:702
https://doi.org/10.1186/s12864-020-07054-3

BMC Genomics

RESEARCH ARTICLE Open Access

High conservation combined with high plasticity: genomics and evolution of *Borrelia bavariensis*

Noémie S. Becker[1*], Robert E. Rollins[1], Kateryna Nosenko[1], Alexander Paulus[1], Samantha Martin[1,2], Stefan Krebs[3], Ai Takano[4], Kozue Sato[5], Sergey Y. Kovalev[6], Hiroki Kawabata[5], Volker Fingerle[7] and Gabriele Margos[7]

Abstract

Background: *Borrelia bavariensis* is one of the agents of Lyme Borreliosis (or Lyme disease) in Eurasia. The genome of the *Borrelia burgdorferi* sensu lato species complex, that includes *B. bavariensis*, is known to be very complex and fragmented making the assembly of whole genomes with next-generation sequencing data a challenge.

Results: We present a genome reconstruction for 33 *B. bavariensis* isolates from Eurasia based on long-read (Pacific Bioscience, for three isolates) and short-read (Illumina) data. We show that the combination of both sequencing techniques allows proper genome reconstruction of all plasmids in most cases but use of a very close reference is necessary when only short-read sequencing data is available. *B. bavariensis* genomes combine a high degree of genetic conservation with high plasticity: all isolates share the main chromosome and five plasmids, but the repertoire of other plasmids is highly variable. In addition to plasmid losses and gains through horizontal transfer, we also observe several fusions between plasmids. Although European isolates of *B. bavariensis* have little diversity in genome content, there is some geographic structure to this variation. In contrast, each Asian isolate has a unique plasmid repertoire and we observe no geographically based differences between Japanese and Russian isolates. Comparing the genomes of Asian and European populations of *B. bavariensis* suggests that some genes which are markedly different between the two populations may be good candidates for adaptation to the tick vector, (*Ixodes ricinus* in Europe and *I. persulcatus* in Asia).

Conclusions: We present the characterization of genomes of a large sample of *B. bavariensis* isolates and show that their plasmid content is highly variable. This study opens the way for genomic studies seeking to understand host and vector adaptation as well as human pathogenicity in Eurasian Lyme Borreliosis agents.

Keywords: *Borrelia bavariensis*, Lyme Borreliosis, Genome assembly, Plasmids, Genetic plasticity

Background

The *Borrelia burgdorferi* sensu lato (s.l.) species complex contains over 20 genospecies of spirochetal bacteria, among them the agents of human Lyme Borreliosis (LB or Lyme disease). These bacteria are obligate parasites that are transmitted between hosts (mainly rodents and birds) by ticks of the genus *Ixodes* [1–5].

Borrelia bavariensis was raised to species level in 2009 and was thereby separated from its sister species *B. garinii* [6, 7]. Both species are present across Eurasia; their main vectors are *Ixodes persulcatus* in Asia and *I. ricinus* in Europe and both are pathogenic to humans. However, the main hosts of the two species differ, with *B. bavariensis* being found in rodents, while its sister species *B.*

* Correspondence: nbecker@bio.lmu.de
[1]Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Grosshaderner Strasse 2, 82152 Planegg-Martinsried, Germany
Full list of author information is available at the end of the article

garinii is found only in birds [7–9]. Originally, the two species were differentiated genetically by their so-called OspA type (i.e. allele at the gene sequence of the Outer Surface Protein A) [10] but more recent studies have confirmed their species status using multilocus sequence analyses (MLSA) for species delineation and phylogenies based on several genetic sequences [6, 11–13]. *B. bavariensis* is of great interest as it has been isolated from many LB patients in Europe but isolates from questing ticks come almost exclusively from Asia ([6] and Margos, Fingerle, personal communication).

The members of *B. burgdorferi* s.l. are characterized by a very complex and fragmented genome that contains a main linear chromosome of approximately 900 kb and up to 20 different linear or circular plasmids whose repertoire vary between and within species [14–17]. Plasmid types are defined based on the plasmid partition genes they contain, and in particular on the PFam32 gene sequence if present (described below). Each plasmid type can in turn be subdivided into sub-types based on organizational changes [14, 18]. Several plasmids form families of related replicons (cp32 and lp28 families) that share long stretches of their sequences. This makes the reconstruction of *B. burgdorferi* s.l. genomes from Next-Generation Sequencing (NGS) data a challenge [18] and explains why, to date, only 34 fully assembled genomes can be found in NCBI [19] among which more than half (18) belong to the species *B. burgdorferi* sensu stricto (s.s.) that is the main LB pathogen in North America. A fully assembled genome is available for the species *B. bavariensis* for reference strain PBi [20] (Accession number: CP058872) and three strains that are still referenced as *B. garinii* in GenBank (BgVir CP003151.1 [21], SZ CP007564.1 [22] and NMJW1 CP003866.1 [23]), but which are known to belong to the species *B. bavariensis* [11]. However, for the latter, only the main chromosome (strains SZ, NMJW1 and BgVir) and two plasmids (strain BgVir only) are assembled.

The process of reconstruction of *B. burgdorferi* s.l. genomes can be facilitated by the identification of plasmid partition genes on assembled contigs. Five such genes have been described in *B. burgdorferi* s.s. and each replicon is believed to contain no more than one copy of these genes unless it is a fusion of two plasmids [24]. In particular, the sequences of the protein family PFam32 are used to name plasmids in the different species of the complex based on the homology to the sequences in *B. burgdorferi* s.s.. Not all plasmids possess a PFam32 [25, 26] but PFam50 and 57/62 appear also to be unique for each plasmid type and allow for plasmid identification in such cases [26].

Genes encoded on plasmids play an important role in pathogenicity and infection of hosts and vectors [27–29]. Description of the whole plasmid repertoire of different isolates from the same species is thus an important step in searching for genetic factors involved in host and vector adaptation. The species *B. bavariensis* is characterized by differentiation into two populations, one in Asia and one in Europe that utilize different vectors. Previous work has shown that the European population showed very little genetic variability on the main chromosome and on two plasmids and seemed to follow a clonal frame [11]. In contrast, the Asian isolates described so far, showed higher genetic diversity (reviewed in [9]). The origin of the species is still unknown, but this diversity pattern could suggest an Asian origin. In the present study, we combined long read (Pacific Bioscience, hereafter PacBio) and short read (Illumina) data to reconstruct the whole genome sequence of 33 *B. bavariensis* isolates from Europe and Asia (Table 1). We show that the plasmid content varies even in the European population, and that the genome of this species is for one part highly conserved and for the other part highly variable.

Results

Borrelia bavariensis genome reconstruction from next-generation sequencing data

The assembly of *B. burgdorferi s*.l. genomes is known to be difficult due to the fragmentation of the genome and to the presence of highly similar plasmids (like the cp32 plasmid family) [18]. We used a combination of long read (PacBio) and short read (Illumina HiSeq and MiSeq) to overcome this problem.

For three isolates (the *B. bavariensis* type strain PBi from Germany, a second European isolate A104S from the Netherlands and the Japanese isolate NT24: highlighted in gray in Table 1), we used both sequencing techniques. For each isolate an assembly was first reconstructed using PacBio reads and then assembled contigs of Illumina short reads were mapped to the PacBio assemblies (see Methods). For most of the three genomes, the two methods gave very similar results with over 99.99% similarity between the Illumina contigs and the PacBio assemblies. Most differences were point mutations and 1 bp-long indels which are known to occur due to the lower accuracy of the PacBio sequencing method [30]. In such cases, the Illumina version of the sequence was kept.

In one case, the Illumina data allowed us to correct a PacBio assembly. The PacBio assembly for isolate NT24 showed two plasmids of respective sizes of 107,820 bp and 49,218 bp that we originally named plasmids cp32–12 + 5 + 6 and cp32–7 + 7 + 11 due to the presence of the corresponding PFam32 sequences. These two plasmids seemed to be fusions of three cp32 plasmids each. Mapping the Illumina raw reads on these sequences (Suppl. Fig. 1) showed that several regions of those PacBio plasmids were not covered by Illumina reads which

Table 1 Genome content of 33 *Borrelia bavariensis* strains

Strain	Date and Ref	Source	Origin	# plasmids	# PF32	Assembled length	chrom	cp26	lp54	lp36	lp17	lp17 + lp28-4	lp25	lp38	lp56	lp28-2	lp28-3
Arh913	**2012**	***I. persulcatus***	**Russia**	15	18	1255	906.4	27.3	58.6	22.9	21.5		27.2	27.8			19.4
Arh923	**2012**	***I. persulcatus***	**Russia**	15	17	1271	906.2	27.3	59.6	21.9		51.1	27.4	14			16.4
FujiP2	**[87]**	***I. persulcatus***	**Japan**	18	18	1324	906.7	27.3	58.7	22.9	21.6		27.4				19.6
Hiratsuka	**2008 [88]**	**human**	**Japan**	16	17	1319	906.3	27.3	59.7	22		51	27.4		22.9		15.8
J-14	**1995 [88]**	**human**	**Japan**	16	16	1306	904	27.3	58.7	22.9		51.6	27.4				20.1
J-15	**1995 [88]**	**human**	**Japan**	10	13	1190	903.6	27.3	57.7	22.4	17.8		28.7	27.2		26.6	17
J-20T	**1996 [88]**	***I. persulcatus***	**Japan**	15	15	1232	906.7	27.3	57.2	32.1	18.1		29.2				
Konnai17°	**2011**	***I. persulcatus***	**Japan**	15	15	1228	906.3	27.3	58.6	21.8		51.4		10.7			
N346	**[89]**	***I. persulcatus***	**Japan**	15	14	1194	906	27.2	59.8	18.6	20.4		29.1				24.7
NT24	**[90]**	***I. persulcatus***	**Japan**	18	18	1318	906.9	27.2	58.6	23	21.4		27.4				20.5
Prm7019	**2012**	***I. persulcatus***	**Russia**	11	15	1151	906.3	27.3	57.9	18.5	21.9		27	24.4			14.3
Prm7564	**2011**	***I. persulcatus***	**Russia**	17	15	1240	906.4	27.3	57.3	20.4	21.1		25	14			15.7
Prm7569	**2011**	***I. persulcatus***	**Russia**	13	18	1180	906.1	27.3	57.9	19.8	20.6		26.6	25.4			28.5
Prm965	**2013**	***I. persulcatus***	**Russia**	13	16	1171	906.1	27.3	57.9	19.8	20.6		25				
61VB2		**tick**	**Germany**	11	13	1264	905.3	27.3	60.4	21.5	26.3						24.3
A104S	**1996 [11]**	**human**	**Netherlands**	13	14	1315	902.3	27.1	60.4	21.3	26.3						24.1
A91S	**1996 [11]**	**human**	**Netherlands**	12	14	1289	905.3	27.3	60.4	21.4	26.5						24.2
DK6	**1990 [11]**	**human**	**Denmark**	10	13	1251	905.2	27.1	60.2	21.3	26.3						23.9
Lubl25	**1995 [11]**	**human**	**Slovenia**	12	14	1284	905.5	27.3	60.4	21.3	24.5						24.1
PBae I	**1990 [7]**	**human**	**Germany**	10	12	1251	905.4	27.1	60.5	21.4	24.9						32.5
PBae II	**1990 [11]**	**human**	**Germany**	8	13	1139	894.8	26.9	60.4	20.7	25						23.9
PBar	**1988 [6]**	**human**	**Germany**	12	14	1278	904.6	27.1	60.4	21.3	25.1						19.6
PBi	**1984 [91]**	**human**	**Germany**	11	13	1280	905.8	27.1	60.4	21.2	24.6						24.1
PBN	**1999 [7]**	**human**	**Germany**	11	13	1269	902.2	26.7	46.6	21.3	24.6						24.2
PHer I	**1989 [11]**	**human**	**Germany**	12	14	1287	905.2	27.1	60.3	21.5	24.9						24.2
PLad	**2000 [11]**	**human**	**Germany**	12	14	1293	904.9	27.1	60.3	21.3	24.6						24.2
PNeb	**1988 [7]**	**human**	**Germany**	12	14	1289	905.6	27.2	60.6	21.5	25.1						24.3
PNi	**2000 [6]**	**human**	**Germany**	10	12	1240	905.7	27	46.6	21.3	24.6						24.1
PRab	**1994 [7]**	**human**	**Austria**	11	13	1255	904.5	27	59.8	21.1	24.4						24
PRof	**1989 [6]**	**human**	**Germany**	12	14	1288	903.8	27	60.4	21.4	24.6						24.1
PTrob	**1988 [7]**	**human**	**Slovenia**	12	14	1290	905.5	27.2	60.5	21.4	24.4						24.1
PWin	**1987 [6]**	**human**	**Germany**	12	14	1284	901.6	27.1	60.4	21.1	26						24.1
PZwi	**1994 [6]**	**human**	**Germany**	11	13	1273	904.9	27.2	60.4	21.3	24.3						24.2

The length of the main chromosome (chrom) and each assembled plasmid (of at least 5kb length) is shown. The three strains sequenced with both PacBio and Illumina data are shaded in gray. Date is date of isolation and is followed by the original reference (see Bibliography, if blank: this study). #PF32 is the number of Pfam 32 proteins identified by BLAST (see Methods) in the assembled data and in the contigs and * means a plasmid partition gene of the families (PF32, PF49, PF50 or PF57-62) was identified but no plasmid of at least 5kb length could be assembled. Length are in kb. °: isolate Konnai17 was a mixture of *B. afzelii* and *B. bavariensis*, we used clone number 1 which is only *B. bavariensis*

Table 1 Genome content of 33 *Borrelia bavariensis* strains *(Continued)*

Strain	lp28-4	lp28-4 + cp32-1	lp28-6	lp28-7	lp28-8	lp28-9	cp32-1	cp32-3	cp32-3 + lp25	cp32-4	cp32-5	cp32-6	cp32-7	cp32-9	cp/lp32-10	cp32-11	cp32-12
Arh913	30.2			21	10.5		*			*		8.6	9.3	*	32.5	13.2	18.4
Arh923			26.2	27		10.1		18.4		8.3	7.8			*	34.1	15.7	
FujiP2	30.2		29.8	20.7	17.4		21.4			17.7	5.2	*	8.8	20.1	32.7	13.3	18.4
Hiratsuka			29.2	26.8		8	27	16.4					7.9	24.9	29.7	16.5	
J-14			29.1	21	16.9		21.9			20.1		10.1	9.4		32.4	14	19
J-15	29.9									*		*	*		31.9		
J-20T	28.2		28.9			12.7	6	*		5.3	8.4		7.5		46.6	9.4	8.7
Konnai17°			28.2	24.8 / 21			57	12.5		17		9.4	*	17.1	31.3	13.9	16.9
N346					13.6			7.4			5.5	11.4	7.6		14.2	9.3	8.1
NT24	30			21	17.8		22			21.7	7.3	7.3	9.9	30.7	32.5	14	19
Prm7019	29.6				8.6						6		*	*	*	*	9.1
Prm7564	30		8.5	27		10.1		16.4		8.3	7.1			5.5	34.1	5.6	
Prm7569	30.2			*	8.2		12.3			*	*	*	*	5.2	7		5
Prm965	30.2				6.3	27.2	7.3	*		9.8	7.2	*	*	*	20.8		5.9
61VB2		51		21.2	12.9				54.9		27.8		31.1				
A104S		50.9		28.2	8.3				54.7	21.1	29.9		31		29.8		
A91S		50.2		28	10				54.9	21.2	27.4		31.1				
DK6		50.8		21.8					54.8		27.4		31.3				
Lubl25		50.8		28.4	8.2				54.9	21.2	27.8				30		
PBae I		46.8		28.3					54.9	21.2	27.8						
PBae II		49.2		26.5	11.3				*	*	*						
PBar		47.5		27.6	11.3				54.9	21.2	27.5				30		
PBi		50.9		22.8	*	28.4			54.9		28.6		31				
PBN		49.5				33.5			53.6		27.8		30.9				27.9
PHer I		50.6		27.7	12				54.8	21	27.8				30		
PLad		51		22.7	13.6	28.4			54.9		27.5		31.1				
PNeb		50.7		28.4	11.7				54.9	21.2	27.8				30		
PNi		48.8				28.4			55		26.9		31.1				
PRab		50.5		27.8	12				54.9	21.2	27.4						
PRof		50.4		22.6	13.3	26			54.9		27.2		31.1				
PTrob		51.1		28.2	12.9				54.9	21.2	28				30		
PWin		50.4		27.4	11.3				54.4	21.1	28.6				29.8		
PZwi		48.9		28.2					54.3	21	26.7				29.5		

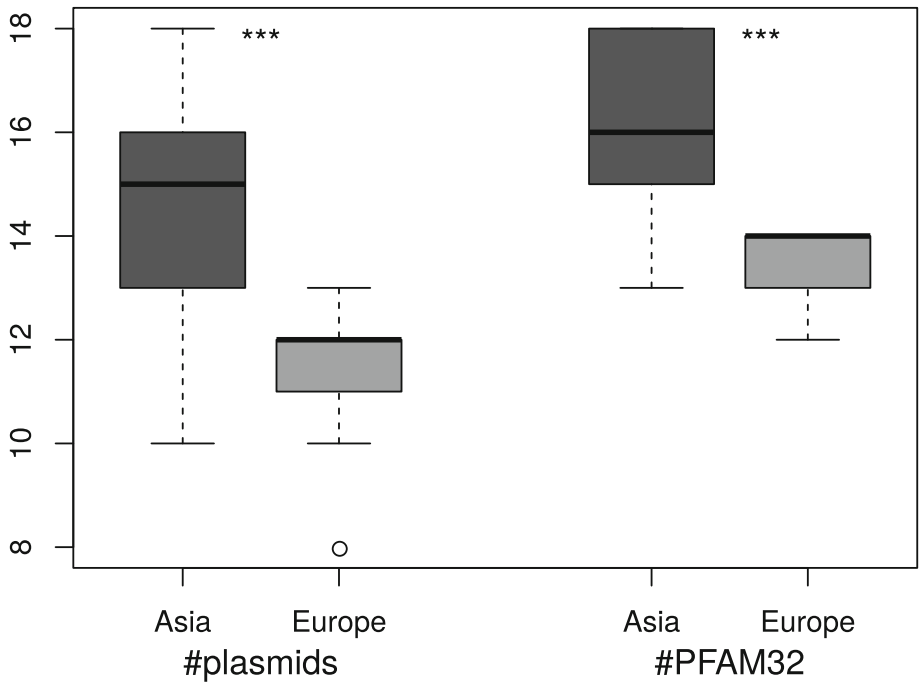


Fig. 1 Asian isolates have more plasmids on average. Boxplots showing the number of plasmids and number of PFam32 proteins identified in the genomes of *B. bavariensis* isolates from Asia (dark grey) and Europe (light grey). ***: Wilcoxon Rank Sum test, *P*-value < 0.001

was not the case for other plasmids. The fusions were thus not supported and were probably an artifact of the PacBio assembly. We used contigs from other isolates as a reference for reconstructing the probable plasmids of the cp32 family in this isolate (see below).

In isolate PBi, unmapped Illumina reads contained sequences similar to lp28–8. This plasmid was not reconstructed in the PacBio assembly. Mapping of PBi Illumina contigs on A104S lp28–8 showed that five contigs mapped to this plasmid with 92–99% similarity to the A104S version. However, the original architecture of the plasmid in PBi was probably different as the five mapping contigs did not cover the full A104S sequence and were themselves not mapped over their whole length. Therefore, the final lp28–8 PBi plasmid sequence could not be reconstructed and, additionally, no PFam32 plasmid partition protein could be found for this plasmid. However, another plasmid partition protein of the family PFam50 for lp28–8 was identified in PBi showing that this plasmid is probably present.

For the remaining 30 isolates which were sequenced with Illumina only, we mapped contigs assembled with SPAdes v. 3.10.1 [31] to the final genomes of the three isolates sequenced with PacBio, as well as to plasmids identified as one full contig in the isolates sequenced with Illumina only, with NUCmer v. 3.1 from package MUMmer [32] (see Methods). Plasmid sequences were kept in the final reconstructed genomes only if they were at least 5000 bp long and were named after the PFam32 protein types identified in their sequence using BLAST v. 2.8.1 [33, 34] or after the reference they were mapped to in case of the absence of a PFam32 sequence (see Methods and Suppl. Table 1). To ensure that the assembly method chosen was good (SPAdes v. 3.10.1 [31]), we also assembled sequence data of 25 isolates with SOAPdenovo v. 1.0 [35] and VelevetOptimizer v. 1.0 [36] (see Methods) and used QUAST v. 4.6 [37] to compare the quality of the three assemblies. As is shown in Supplementary Figure 2, N50 values were significantly higher in SPAdes assemblies compared to assemblies of the two other assemblers (Wilcoxon Rank Sum Tests with each other assembler: Bonferroni-Holm corrected *P*-Value < 0.01) and the number of contigs was significantly smaller (Wilcoxon Rank Sum Tests with each other assembler: Bonferroni-Holm corrected *P*-Value < 10^{-4}). In addition, the total length of the final assembly was largest in SPAdes in 24 out of 25 isolates tested. We conclude that, of the three tested assemblers, SPAdes performed the best.

We also remapped the raw Illumina reads on the final reconstructed genomes to check the quality of our reconstruction (see Methods) and show the relative standard deviation (SD) of coverage as a measure of quality in Supplementary Figure 3. A well assembled genome should have a low coverage variance as reads would map evenly to the contigs. The isolates from Asia showed a

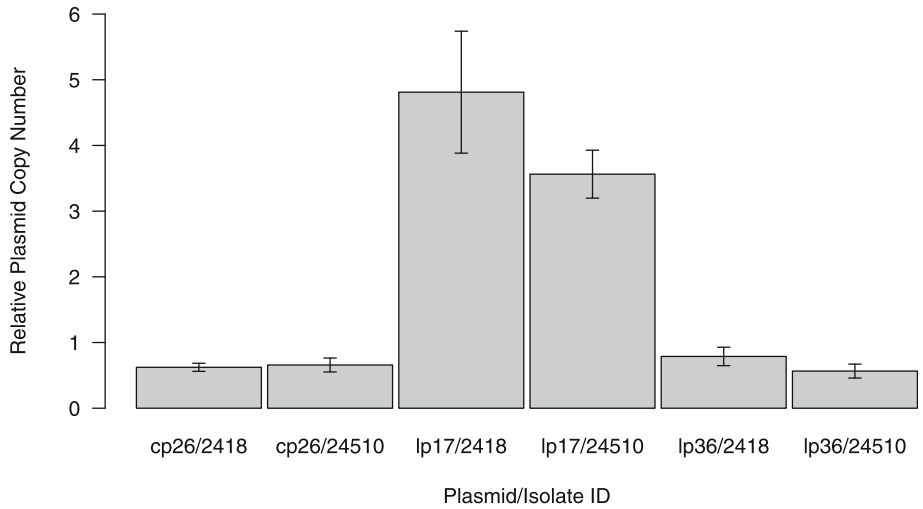


Fig. 2 Relative plasmid copy number based on qPCR results. Relative plasmid copy number was estimated based on qPCR results on the chromosome and plasmids cp26, lp17 and lp36 on PBi isolates 2418 and 24510 ran with three biological and three technical replicates. Error bars represent the standard error of the mean

significantly higher variance in coverage (Wilcoxon Rank Sum Test: *P*-Value $< 10^{-16}$) as compared to the European isolates. This could be due to variation in the quality of the original DNA samples, (DNA samples from the Asian isolates were shipped to Germany), or to the lack of good references for certain plasmids due to the higher diversity observed in the Asian isolates. Indeed, the relative SD was higher for plasmids compared to the main chromosome in Asian isolates even if this difference was not significant (Suppl. Fig. 3b). The quality of the assembly did not depend on the method used for obtaining the final plasmid sequence (either as an own entire contig or with contigs mapped to a reference) (Suppl. Fig. 3a).

Genome composition of 33 *B. bavariensis* isolates

The genomes of the 33 isolates consisted of a main chromosome and a variable number of plasmids (Table 1). Chromosomes were about 900 kb in size (size of reconstructed chromosome varied between 894,779 bp in isolate PBaeII and 906,948 bp in isolate NT24) and made up on average 72.1% of the total assembled genome. Eight to 18 individual plasmid sequences of at least 5000 bp could be assembled per isolate. Additional plasmid sequences were identified in 11 isolates due to the presence of partition genes or as some contigs mapped to plasmids identified in other isolates (Suppl. Table 1). However, these additional plasmids could not be fully assembled or the assembled sequence did not reach the 5000 bp criterion. Several reconstructed plasmid sequences, particularly of the lp28 and cp32 plasmid families, are very short (below 10 kb). It is probable that the sequence reconstructed here for these plasmids does not recover the full plasmid length and that the missing sequences were probably erroneously assembled in other contigs due to similarity. This confirms that short read sequencing alone is not sufficient to reconstruct plasmids from these families. Using long-read sequencing was very helpful in the assembly of plasmids in isolates PBi and A104S. However, even the PacBio assembly pipeline failed to reconstruct properly the cp32 content of isolate NT24. For this isolate we used the same strategy as for the isolates with only Illumina data (see Methods) and mapped Illumina contigs to cp32 plasmids from other isolates. This allowed us to reconstruct plasmids cp32–11 and cp32–12. For plasmids cp32–5, – 6 and – 7 no mapping was possible; we could only use Illumina contigs that were 7.3, 7.3 and 9.9 kb long, respectively, and probably do not represent the full plasmid (Table 1).

The number of plasmids per isolate (Fig. 1) was significantly higher in the Asian population (ranging from 10 to 18 reconstructed plasmids over 5 kb long) as compared to the European population (8 to 13 plasmids). As some plasmid fusions were observed and as some plasmids could not be reconstructed, we also tested for the number of PFam32 gene sequences present in each isolate. This was found to be significantly higher in Asian isolates compared to European isolates (Fig. 1), again implying that fewer plasmids are present in European isolates than in Asian isolates.

We also tested for a deviation in copy number between plasmids with respect to the main chromosome by plotting the coverage of the raw read mapping to each plasmid relative to the chromosome (Suppl. Fig. 4).

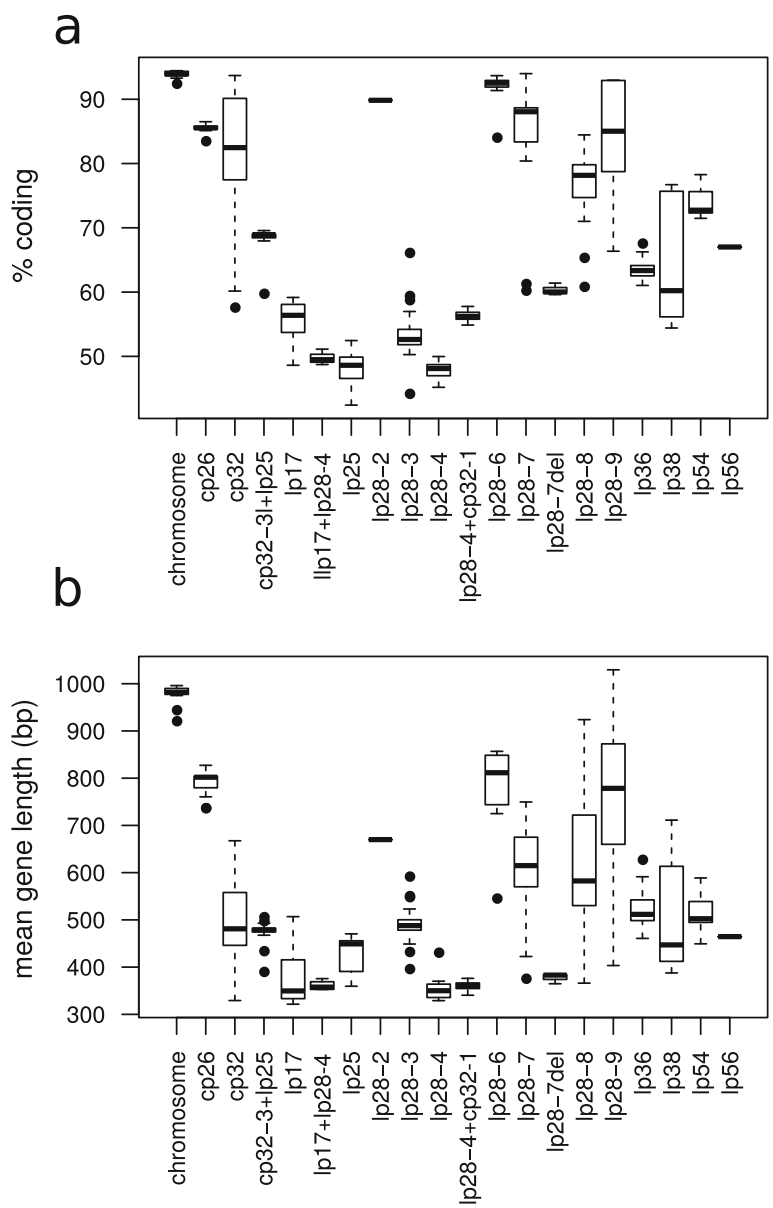


Fig. 3 Gene content of the *B. bavariensis* replicons. Percentage of coding sequence (**a**) and average gene length (**b**) for the chromosome and each plasmid over isolates are shown as boxplots

As the coverage of Asian *B. bavariensis* genomes was more variable, we did this for European isolates only. We found the coverage of plasmids lp17, lp28–7 and lp36 to be significantly higher than for the main chromosome for all European isolates. In particular, based on this coverage measure, there were, on average, about seven copies of lp17 per cell in European isolates.

As several plasmids seemed to have a higher copy number compared to the main chromosome based on the read coverage of the Illumina data, we used a qPCR protocol to directly measure the number of DNA molecules present in a strain relative to the main chromosome. We chose to use plasmids cp26 (which we hypothesized to be present in about the same number as the main chromosome, based on read coverage) and lp17 and lp36 (which seemed to have higher copy numbers). We designed a qPCR protocol following Millan et al. [38] with one PCR per plasmid (see Methods for details) on two low passage isolates of *B. bavariensis* isolate PBi. Each isolate was run using three biological and two technical replicates. As can be seen in Fig. 2, the copy number of plasmid cp26 was estimated to be slightly below one copy per chromosome, that of lp36 was about one copy per chromosome and lp17 plasmid

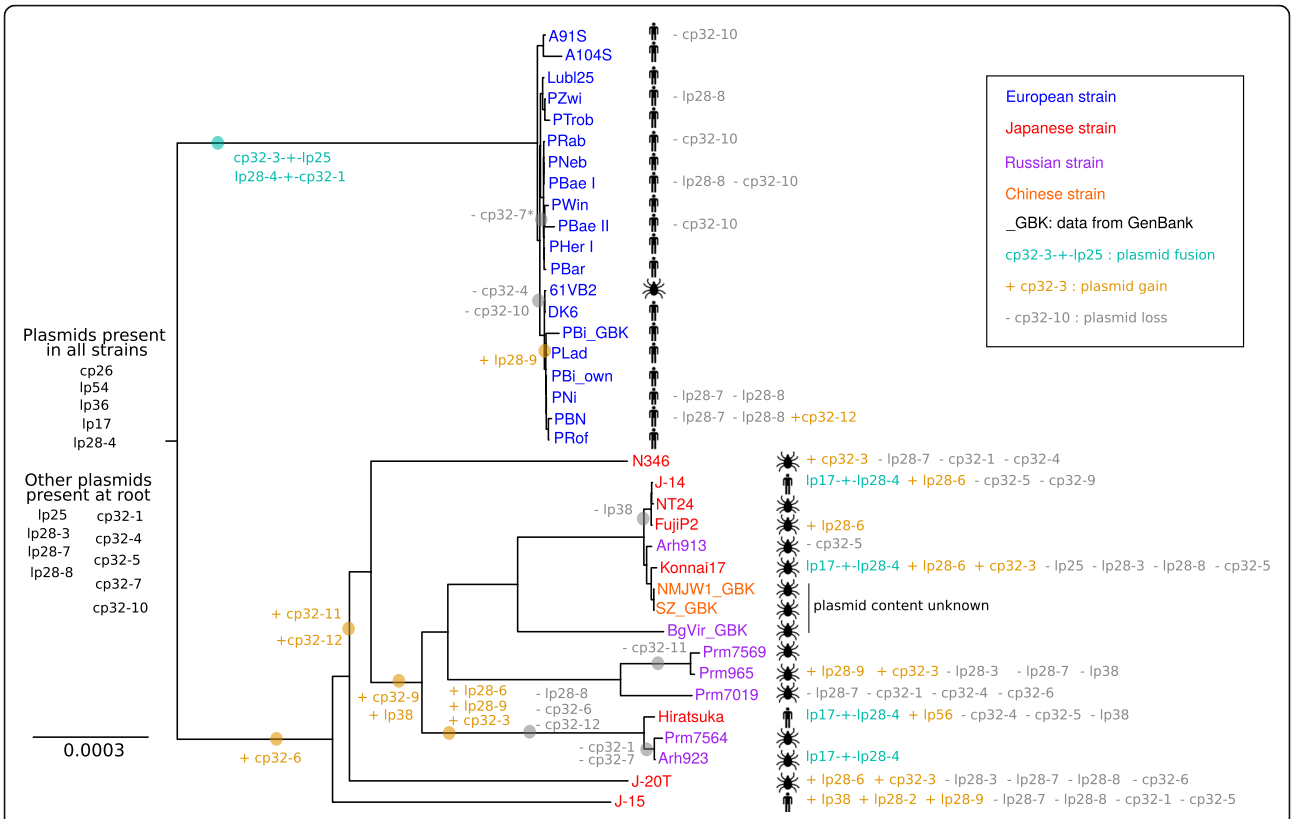


Fig. 4 Phylogeny of *B. bavariensis* reconstructed based on the main chromosome. Phylogeny reconstructed with BEAST v1.8.0 [41] with the following parameters: coalescent model with exponential growth based on doubling time, lognormal-relaxed clock [42], GTR substitution model [43]. A burn-in of 30% of the 100 Million steps chain was removed before selecting the best tree with TreeAnnotator v. 1.10.4 [41]. The scale is in substitutions per site. Node posterior probabilities were above 0.99 for all nodes except in the European clade (very short branches and very low diversity). The gain (orange), loss (gray) and fusion (cyan) events were positioned following maximum parsimony principle. There are indicated on a branch if they concern several isolates and after the isolate name if they concern only one isolate. Isolate origin is indicated by a tick for isolation from a tick (*I. persulcatus* in Asia and species unknown for European isolate 61VB2) and a human for isolation from a human patient. The accession numbers for the sequences coming from public databases can be found in the Methods section. * This plasmid loss event concerns the branch leading to isolates Lubl25, PZwi, PTrob, PRab, PNeb, PBae I, PWin, PBae II, PHer I and PBar

was found to be at a higher relative copy number varying between three and five copies per chromosome. This value is lower than the copy number estimated based on the coverage measure but is probably a more accurate estimate.

Shared versus variable genome components

All *B. bavariensis* isolates sequenced in this study contained, in addition to the main chromosome, plasmids cp26, lp54, lp36, lp17 and lp28–4 (Table 1). In addition, we found in each isolate between 4 and 9 types of cp32 sequences. These were either fused with other plasmids or independent plasmids and their numbers were obtained by counting cp32 PFam32 sequences (as cp32 family plasmids could not be properly assembled in several isolates). Three cases of plasmid fusions were observed in at least two isolates and were thus considered to be true (other cases were not reported as they may have been due to mis-assembly and, in such cases, the plasmids were recorded without the possible fusion). In all European isolates, we observed two cases of fusion of a linear plasmid (lp28–4 or lp25) with a cp32 plasmid (cp32–1 and cp32–3, respectively). These fusions were found to be fixed in European *B. bavariensis* isolates but were absent from Asian isolates. In addition, plasmid lp17 and lp28–4 were found to be fused in four Asian isolates, but not in any of the European isolates. Interestingly, these isolates were found in independent clades in the phylogeny of the species (see below).

Supplementary Figure 5 shows a schematic representation of the fusions involving plasmids lp28–4, lp17 and cp32–1 with a precise description of the different plasmid types as well as plasmid lp28–7 as we found that translocations occurred within the European population between lp28–7 and lp17. To produce Supplementary Figure 5, we first had to determine plasmid types for the four plasmids under study. Following Casjens et al. [14] we counted a new plasmid type each time a deletion or

Table 2 Within and between population genetic diversity for the main chromosome and plasmid orthologous regions

Genomic region	# Asia	# Europe	Length (bp)	# SNP	*π Asia*	*π* Europe	F_{ST}
chromosome	17	19	920,528	42,039	$8.79*10^{-3}$	$1.72*10^{-4}$	0.56
cp26	15	19	29,623	1979	$1.54*10^{-2}$	$1.99*10^{-4}$	0.50
lp17	14	19	13,732	1331	$1.98*10^{-2}$	$4.99*10^{-4}$	0.49
lp25	13	18	27,833	3232	$2.97*10^{-2}$	$7.03*10^{-4}$	0.36
lp28-3	11	19	11,152	1572	$6.80*10^{-2}$	$8.23*10^{-4}$	0.50
lp28-4	14	18	31,849	4144	$2.05*10^{-2}$	$2.62*10^{-3}$	0.52
lp36	14	19	9819	1081	$2.34*10^{-2}$	$3.66*10^{-4}$	0.69
lp54	15	19	67,261	8167	$2.06*10^{-2}$	$3.47*10^{-4}$	0.59

Genetic diversity (π [44]) within populations and genetic distance (F_{ST} [45]) between populations were estimated on orthologous sequences aligned with MAFFT v 7.407 [46, 47]. The number of single nucleotide polymorphisms (SNP) is indicated for both populations mixed and the length is the length of the alignment

insertion of at least 400 bp was observed and for each translocation or inversion of at least 400 bp (see Methods for more details). We were able to identify nine lp28–7 types, 12 lp17 types including two fusions with lp28–4, five other lp28–4 types, six cp32–1 types and six versions of the fused plasmid lp28–4 + cp32–1. There was no case of two Asian isolates sharing the same plasmid types for each one of these four plasmids (i.e. lp17; lp28–4; lp28–7 and cp32.-1) and in the European population we could identify only three groups of two isolates and one group of three isolates that shared the same plasmid types for lp28–7, lp17 and lp28–4 + cp32–1. Even if many short indels were observed on plasmid lp28–4, we could identify an almost 20 kb-long sequence that is shared by all types with or without fusions. The fusion of plasmids lp17 and lp28–4 in four Asian isolates was found to have occurred without any other big rearrangements. However, we identified two different architectures for this fusion. In isolate J-14 (and in isolates FujiP2 and Hiratsuka that were mapped to it) we observed a fusion of the 5′ ends of plasmids lp17 and lp28–4, thus lp17 appeared to be flipped. In isolate Arh923, the two plasmids were fused by their 3′ ends. Of course, this could have been due to mis-assembly. The fusion of lp28–4 and cp32–1, that is fixed in the European population, was shown to be an insertion of cp32–1 into lp28–4. There were two very different types of cp32–1 plasmids in the Asian population, with only about 10 kb homology. The fused plasmid observed in the European isolates seems to have occurred using the cp32–1 type carried by Asian isolate Hiratsuka (or a related cp32–1 type), which does not have more than 2 kb homology with the other Asian type of cp32–1. Apart from these two fusions, we could also observe a reciprocal translocation that occurred between plasmids lp28–7 and lp17 in the European population. Five European isolates including the reference strain PBi carry at the end of plasmid lp17 a 2.5 kb-long sequence that is found at the beginning of plasmid lp28–7 in all other isolates. And reciprocally, plasmid lp28–7 of three of these five isolates (the other two do not have a lp28–7) carry at their beginning a 5 kb-long sequence that is found at the end of lp17 in all other isolates. Both regions contained genes encoding outer membrane proteins.

We used RAST [39, 40] to annotate the reconstructed *B. bavariensis* genomes and, following the method by Mongodin and colleagues [17], kept all detected genes of at least 50 amino-acid length. The main chromosome was found to contain on average 816.4 genes that met this criterion (range 812–842) and on average 94% of the chromosome sequences were coding with very low variation among isolates (standard deviation 0.41 – see Fig. 3a). This was significantly higher than in plasmids (Welsh T test $T = 32.0$, df = 427, P-value < 0.001). Circular plasmids had a significantly higher percentage of coding sequence compared to linear plasmids (average circular: 82.4%, average linear: 66.1%, Welsh T test, $T = 14.6$, df = 366, P-value < 0.001; fusions between circular and linear plasmids were excluded). As shown in Fig. 3b, annotated genes were also significantly longer on the chromosome compared to the plasmids (mean 981 bp, Welsh T test $T = 65.0$, df = 434, P-value < 0.001). Circular plasmids had significantly longer genes compared to linear plasmids (average circular: 562 bp, average linear: 514 bp, Welsh T test, $T = 3.4$, df = 326, P-value < 0.001; fusions between circular and linear plasmids were excluded). We used BLAST v. 2.8.1 [33, 34] at the amino-acid level (algorithm BLASTp) to compare each of the 33 isolates with the 32 others for gene content. A hit between protein sequences in two different isolates was kept if the hit had at least half the length of the original gene and if the identity between the two sequences was as least 90%. Using these criteria, we found that at least 93% of the genes located on each chromosome had a hit on every other chromosome. This confirmed that the chromosome was highly conserved within the species *B. bavariensis*, even between Asian and European isolates. Indeed the best hits between isolates for each chromosomal gene had on average 98.8% sequence identity

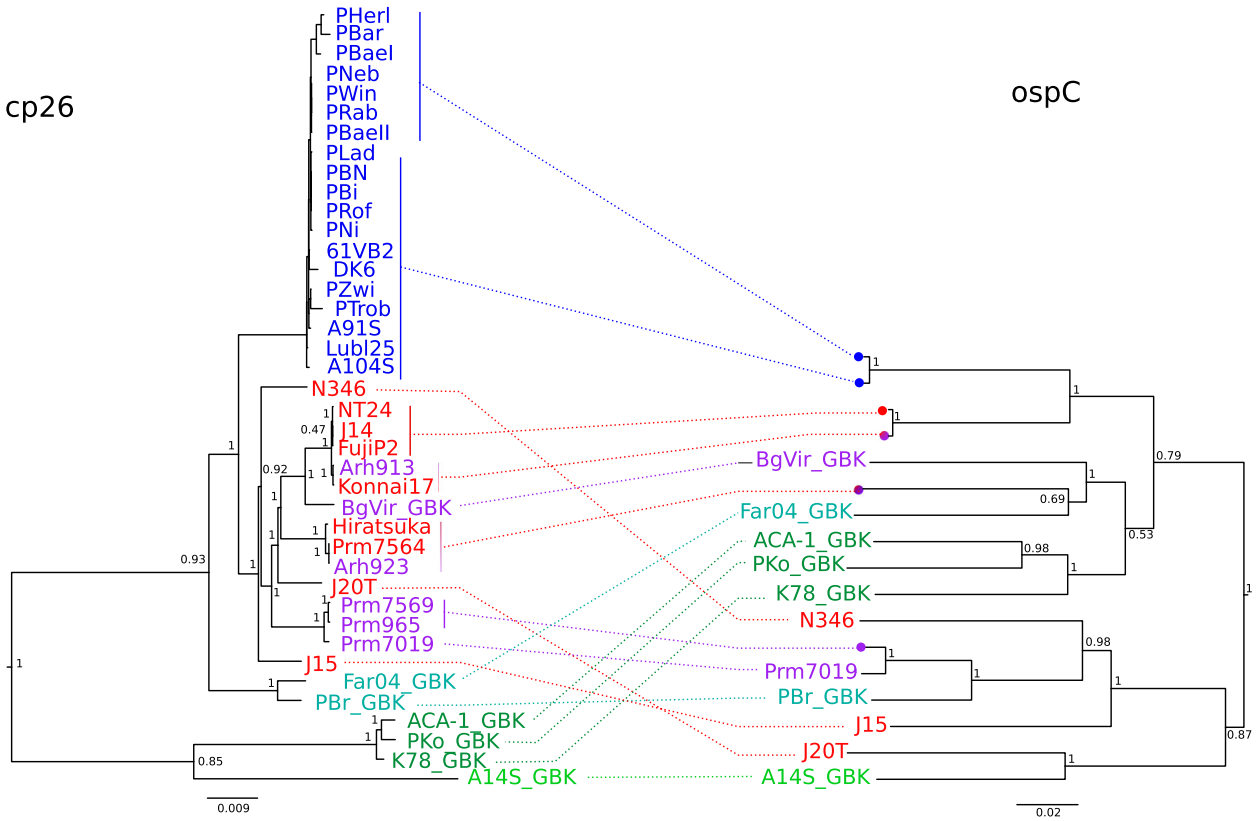


Fig. 5 Comparison of cp26 and *ospC* phylogenies. Sequences for the *ospC* gene and the cp26 plasmid without *ospC* (cutting out 200 bp upstream and downstream the gene) were aligned with MAFFT v7.407 [46, 47] and BEAST v1.8.0 [41] was run for 100 Million states for cp26 and 20 Million states for *ospC* each in triplicate. Best trees were reconstructed after removing a burnin-in of 10% of the chain and all three runs showed very similar results for each tree. Both trees were plotted using FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and manually rotated. Color code for isolates: Light green: *B. spielmanii*, dark green: *B. afzelii*, cyan: *B. garinii*, purple: *B. bavariensis* Russia, red: *B. bavariensis* Japan, marine blue: *B. bavariensis* Europe. Dots on the *ospC* phylogeny represent several isolates having exactly the same sequence. Scale bars are in substitutions per site. Values next to nodes indicate node posterior probability (not shown within the European *B. bavariensis* clade for the sake of clarity)

when the compared genes were from isolates from within the same continent and 97.5% when the compared isolates were from different continents. Plasmid cp26 was also found to be highly conserved with on average 91.1% of its 26 to 28 genes being shared with the cp26 plasmids of each other isolate and the identity of the best hit in each isolate being on average 99.0% for isolates from the same continent and 92.7% for isolates from the other continent.

Out of the 24 different plasmids assembled from the genomic data of the 33 *B. bavariensis* isolates (without taking fusions into account), 19 were not found in all isolates. This estimated variable portion represented on average 19.2% of the total reconstructed genomic content of each isolate and 68.3% of the total assembled plasmid content. These size estimates of the variable genome represent only a lower bound because some plasmids found in all isolates are nevertheless not similar over their whole length and some plasmids were not successfully assembled. The greatest degree of diversity was observed on the two plasmid families lp28 and cp32 which were represented by seven and ten members, respectively, over all isolates with only lp28–4 found to be present in every isolate.

Evolution of the species

We used BEAST v1.8.0 [41] to reconstruct the phylogeny of the main chromosome (see Methods for more details) for all of our 33 isolates as well as four additional isolates for which chromosomal sequences have been published in GenBank (under accession numbers CP000013 for strain PBi from Germany, CP003151 for strain BgVir from Russia and CP003866 and CP007564 for strains NMJW1 and SZ from China). We used *B. garinii* strain 20047 as an outgroup to root the tree (GenBank accession number CP028861). The resulting phylogeny, presented in Fig. 4, shows that the two continental populations are clearly divergent with a deep branching. The European population is characterized by a very short-time divergence and an almost clonal

recent evolution as has already been noted [11]. The Asian population, even if showing greater overall divergence, does not show any geographical structure: isolates from Japan, China and Russia are found in the same terminal clades. Asian isolates also did not cluster by origin of the isolate (questing tick or patient). In Europe, only one isolate from a tick was available and this had no special position in the phylogeny. Both chromosome assemblies for the PBi type strain (ours and that published as CP000013) were both located in the same clade. We compared RAST [39, 40] annotation results for both PBi chromosome sequences and found that there was perfect synteny between the two (Suppl. Fig. 6).

In this phylogeny, we also indicated gains, losses and fusions of plasmids based on the reconstructed genomes using maximum parsimony (Table 1 and Suppl. Table 1). This showed that, in addition to five plasmids present in all isolates, four other linear plasmids and five cp32 plasmids could have been present at the root of the tree in the ancestral *B. bavariensis*. These plasmids would then have been subsequently lost in some derived isolates. Nine gain and ten loss events could be placed on internal branches and thus were shared by at least two isolates. In the European clade, three plasmid loss events on internal branches and ten plasmid loss event on terminal branches were found, whereas only two gain events were identified (plasmid lp28–9 shared by five isolates and plasmid cp32–12 found only in isolate PBN). This shows that the plasmid repertoire of the European population is rather stable with only plasmid losses that could have been due to isolate cultivation in the laboratory rather than to real evolutionary change. In the Asian population, according to our maximum parsimony reconstruction, plasmid gains were as frequent as plasmid losses on internal branches (eight gain events for seven loss events) but there were twice as many losses as gains on terminal branches (13 gains for 29 losses).

Genetic diversity within and between the Asian and European populations was estimated by nucleotide diversity (π [44]) and genetic distance (F_{ST} [45]) for the main chromosome and seven plasmids with orthologous regions in at least five isolates in each population (see Methods, Table 2). Diversity was found to be lower in the European population compared to the Asian population by one to two orders of magnitude depending on the genomic segment and to be lower for the main chromosome compared to plasmids. Genetic distance between Asian and European populations was lowest for lp25 (0.36) and highest on lp36 (0.69).

We also estimated genetic diversity along the main chromosome and for plasmids cp26 and lp54, in which alignments were possible over the whole length (Suppl. Figs. 7, 8 and 9). For all three replicons, we identified peaks of diversity either between populations from the two continents (peak only when considering all isolates) or in one or both regional populations. We found high diversity in several chromosomal genes coding for proteins located in the outer membrane of the bacteria (OppA, ABC transporter, Lmp1, PTS system). This was also true for lp54, particularly in the Asian population, with diversity peaks located in the genes encoding OspA, OspB, DbpA and in the PFam54 gene array. On cp26, the o*spC* gene is well known for having high diversity in several *B. burgdorferi* s.l. species including *B. bavariensis* which is confirmed here for the Asian population [11, 17, 48, 49].

As *ospC* showed a high diversity, and as this locus is known to be a hotspot of recombination in several *B. burgdorferi* s.l. species [11, 48, 50], we reconstructed a phylogeny of this gene and compared it to that of the cp26 plasmid cutting out the *ospC* locus. Several publicly available sequences for *B. bavariensis* (strain BgVir), *B. garinii* (strains Far04 and PBr), *B. afzelii* (strains ACA-1, K78 and PKo) and *B. spielmanii* (strain A14S) (see Methods for details) were additionally included in this analysis. As can be seen in Fig. 5, the cp26 phylogeny followed the known species tree with *B. bavariensis* and *B. garinii* being sister species as are *B. afzelii* and *B. spielmanii*. The phylogeny of plasmid cp26 within *B. bavariensis* was very similar to the phylogeny reconstructed for the main chromosome (Fig. 4), except for minor differences in clustering of Japanese isolates. However, the phylogeny reconstructed for *ospC* was quite different and showed two major clades. One clade contained all European *B. bavariensis* and all *B. afzelii* as well as some Asian *B. bavariensis* and one of the two *B. garinii* strains. The second clade contained *B. spielmanii*, the other *B. garinii* strain and the rest of the Asian *B. bavariensis* haplotypes. Apart from the European *B. bavariensis* clade (where we observed only two different *ospC* haplotypes with only one non-synonymous difference between them) and the *B. afzelii* clade, all other species or populations with several isolates were found not to be monophyletic.

Discussion

Strategies for genome reconstruction of *B. burgdorferi* sensu lato

In this article, we present genome reconstructions for 33 *B. bavariensis* isolates from Eurasia. Following other studies (see for example [18]), we used a combination of long-read (Pacific Bioscience) and short-read (Illumina) sequencing. We show that the PacBio long-read assembly allowed the reconstruction of most plasmids even from the cp32 and lp28 families. It had been reported

before that PacBio assemblies contain inaccuracies [51] and in one out of the three isolates, the PacBio assembly created two, probably spurious, fusions of plasmids belonging to the cp32 family. This occurred in one Japanese isolate that possessed nine cp32 plasmids, the maximum of cp32s observed in our sample set. It shows that proper assembly of sequences carrying so many cp32 plasmids remains challenging even when using long-read data. However, fusions of cp32 plasmids have been observed in other species of the *B. burgdorferi* s.l. complex [52, 53] and it remains an unresolved question whether these were real in isolate NT24. In isolate PBi, Illumina reads were identified that mapped to plasmid lp28–8 and carried the lp28–8 PFam32 sequence but no contigs for this plasmid were found in the PacBio assembly. The Illumina data for this plasmid was too fragmented to reconstruct the plasmid sequence via mapping. Thus, it is possible that this plasmid was not present in each cell of the isolate or was in the process of decaying or being lost while cultivating the isolate for DNA extraction as has been described in many *Borrelia burgdorferi* s.l. isolates [54–56]. Although circular consensus sequencing (CCS) improved the accuracy of PacBio data, it has been established that long-read data is more prone to sequencing errors [30]. It is therefore advisable to complement and correct them using more accurate short-read data. Reassuringly, for each replicon, the similarity between PacBio and Illumina reads was above 99.98%.

For the 30 isolates for which no long-read sequencing data was available, our strategy was to perform de novo assembly of the Illumina reads and then use the three long-read isolates as a reference for mapping if required. For some replicons, the mapping step was not necessary as single contigs were available that covered whole plasmids. This was the case for five out of 30 chromosomes and for numerous plasmids (as an example, all but five cp26 plasmids were each covered by a single contig). It made no noticeable difference for assembly accuracy (Suppl. Fig. 3), whether the data was mapped or assembled directly as one contig. Such contigs that assembled as full plasmids were successfully used as references for other isolates. Despite all this, for 11 isolates, a total of 27 plasmids were missing from, or incomplete in, the final assembly. These replicons were known to be present as plasmid partition gene sequences for them were identified or as contigs mapped to them, but we could not reconstruct a full plasmid. Perhaps not surprising, this happened more frequently in the Asian isolates (in nine isolates a total of 23 plasmids were missing) than in the European isolates (two isolates and four plasmids). Whether this was due to a lower data quality in the Asian isolates and/or challenges to find an appropriate reference (due to the higher diversity in plasmid content observed in this population) is currently unclear. In addition, several reconstructed plasmids were very short and it is probable that part of their sequence was not assembled.

The use of only short-read sequencing thus resulted in a good global description of the plasmid content, but proper full genome reconstruction was only possible in those isolates for which a close reference was available, as was the case for the European isolates. This was also the case in previous studies using Illumina short-read sequencing in *B. burgdorferi* s.s. (see for example [57]).

The *B. bavariensis* genome shows a high degree of conservation

The core genome of the species complex *B. burgdorferi* s.l. is considered to be composed of the main chromosome and plasmids cp26 and lp54 [17]. In addition, all the *B. bavariensis* isolates sequenced here share sequence stretches of three other plasmids: lp17, lp28–4 and lp36. Interestingly, 14 strains of *B. burgdorferi* s.s. have also been shown to share these same five plasmids (cp26, lp17, lp28–4, lp36 and lp54) [14]. For plasmids lp17 and lp28–4, the shared sequence stretches made up about 12 kb and 18 kb, respectively, and for plasmid lp36 a fragment of about 13 kb was found to be shared among all isolates. These sequences can thus be considered as belonging to the core genome of *B. bavariensis* which thus adds up to 1027 kb; being made up of 900 kb of chromosomal sequence plus 127 kb of plasmid content (with 27 kb on cp26 and 57 kb on lp54). The chromosome and cp26 sequences are, in particular, highly conserved as seen when comparing gene content between isolates and as already described [14, 17]. A very high proportion of the genes on these two replicons (93% for the main chromosome and 91.1% for cp26) are found in all isolates.

The main chromosome sequences also allowed us to reconstruct a phylogeny for the species (Fig. 4). We had already published a similar phylogeny using a subset of these isolates [11]. However, the Russian isolates are new to the present paper and allow us to see that the Asian clade shows no detectable geographic clustering. Asian *B. bavariensis* are vectored by *I. persulcatus*, whereas the European vector is *I. ricinus* (see [9] for a review). As these two tick species co-occur and can even hybridize in their overlapping zone in Estonia, Latvia and Western Russia [58], we expected that Russian *B. bavariensis* samples, might be genetically closer to the European isolates than the Japanese isolates, perhaps even showing that the European population might have diverged from a Russian lineage, but this was not the case. The lack of spatial structure in the Asian *B. bavariensis* genomes over such a large geographical scale can be

explained either (i) by the co-occurrence over a long evolutionary period of many strains in the same populations due to specialization to some specific niches (like reservoir hosts) or (ii) by recurrent migration of strains, for example carried by ticks attached to birds. However, this last hypothesis seems less likely as *B. bavariensis* is rodent-adapted and does not survive in bird complement active immune serum [7, 59].

Another conserved pattern was the elevated coverage of the sequence data observed on certain plasmids and particularly on plasmid lp17 with respect to the main chromosome. The coverage of lp17 was higher than that of the chromosome in all isolates (European isolates are shown in Suppl. Fig. 4). This suggests that *B. bavariensis* normally carries a higher copy number of plasmid lp17 than is the case for other plasmids or the main chromosome. In another study, the coverage of a plasmid, lp28–6, in one *B. burgdorferi* s.s. strain was also found to be about ten times higher than then rest of the genome [25] but, to our knowledge, no study reported such a pattern for a plasmid in many isolates of the same species. We experimentally confirmed that the copy number of plasmid lp17 was three to five fold that of the main chromosome for isolate PBi grown under lab conditions (Fig. 2). This finding contradicts the current view of plasmid partitioning in *B. burgdorferi* s.l. according to which each plasmid is expected to contain at maximum one or two copies of each plasmid per cell [25, 60]. The only other study we could find that experimentally tested for copy-number of plasmids in *B. burgdorferi* s.l. was performed on three plasmids of the *B. burgdorferi* s.s. reference strain B31 via relative hybridizations of replicon-specific DNA probes [61]. These three plasmids were found to be present at about one copy per chromosome and this was shown to be stable when the strain was kept in culture. Outer membrane vesicles (OMVs) produced by *B. burgdorferi* s.l. bacteria could provide an explanation for DNA extracted from cultures possessing more copies of certain plasmids than the chromosome. OMVs are membrane-enclosed spheres that many bacteria, including *B. burgdorferi* s.l., fill with different molecules and release into their surroundings [62], often as a response to stress [63] that can be induced by cultivation conditions [64]. OMVs produced by *B. burgdorferi* s.s. have been found to contain both circular and linear DNA [65]. More recently, *B. burgdorferi* s.s. OMVs were also found to contain RNA preferentially transcribed from plasmid sequences but not specifically from lp17 [66]. It is known from other bacterial species that such vesicles can be involved in toxin delivery, cell-cell signal trafficking, protein transfer, and horizontal gene transfer [67]. Plasmids can be transferred via vesicles, and plasmid identity has been shown to strongly influence the efficiency of their loading into vesicles in *E. coli* [68]. Taking all of this into account, together with the fact that lp17 has been shown to be involved in host tissue colonization and evasion of host immunity in *B. burgdorferi* s.s [69, 70]., it is possible that *B. bavariensis* preferentially packages lp17 plasmids into OMVs and that these extra plasmid copies are the reason for the observed increased plasmid to chromosome coverage ratio in *B. bavariensis* isolates that were cultivated to high density, and thus under stressful conditions. This hypothesis, however, remains to be tested.

A further level of genetic conservation can be seen within populations and particularly in the European isolates. The genetic diversity on the chromosome and on plasmids is very low within the European population (Table 2) and even the ospC locus, which is known to be one of the loci with the highest within-population diversity on the *B. burgdorferi* sensu lato genomes [48, 49, 57, 71], shows very little variation in this population (Fig. 5). All the sequenced European isolates also share the presence of three plasmids (lp28–3, cp32–3 + lp25 and cp32–5) in addition to the 5 plasmids present in all *B. bavariensis* isolates. Two plasmid fusions are also shared by all European isolates. However, the European population is not as clonal as previously thought [6, 11, 72] and several plasmids have evidently been lost or gained during its evolution (Table 1 and Fig. 4). In contrast to the Asian population, the European population shows some degree of geographic structure, with the first node separating the two Dutch isolates (A104S and A91S), that are the most western isolates in our sample, from the rest of the population and with the two Slovenian isolates (Lubl25 and PTrob) also being in the same clade together with a German isolate (PZwi).

The Asian population showed more variability, both at the sequence level and in the plasmid repertoire (we could find no pair of Asian isolates having exactly the same plasmid content based on the distribution of the PFam32 sequences). All the Asian isolates are characterized by a higher number of plasmids compared to European isolates and in particular by a higher number of cp32 plasmids (7.3 on average against 4.6 for the European isolates). This large cp32 repertoire might be associated with the ability to infect a wider range of vertebrate hosts; in *B. burgdorferi* s.s. cp32 plasmids carry several genes essential for host infectivity among which are the loci coding for Erp proteins that have been shown to bind complement proteins in humans (see [73] for a review).

The *B. bavariensis* genome also displays a high degree of plasticity

While part of the *B. bavariensis* genome was found to be highly conserved, we also observed a high diversity, in particular in plasmid content. About two thirds of the plasmid content of each isolate was not shared by the

whole species. This has been observed in *B. burgdorferi* s.s. as well [18, 25]. We placed gains and losses of plasmids on our *B. bavariensis* phylogeny based on the main chromosome using maximum parsimony (Fig. 4). According to this reconstruction, 14 out of 24 plasmids would have been present in the common ancestor of the species. It is important here to remind the reader that plasmid loss can occur while *B. burgdorferi* s.l. bacteria are grown in culture, and that this could be the reason for the absence of some plasmids from certain isolates [54–56]. Thus, some of the apparent plasmid losses during *B. bavariensis* phylogeny may not be real. Nevertheless, it is very unlikely that all the apparent losses of plasmids are artifactual, and gains of plasmids cannot be explained in this way. The complexity of the evolution of plasmid content in *B. bavariensis*, as depicted in Fig. 4, is striking and shows that the plasmid fraction of the genome is very plastic as has also been shown for *B. burgdorferi* s.s. [14]. The ability to exchange plasmids, either via OMVs as described above or using other mechanisms, seems to be very pronounced in *B. bavariensis* and in particular in the Asian population.

The genome plasticity of *B. bavariensis* is further demonstrated by the occurrence of three plasmid fusions shared by at least two isolates. Two of these fusions are fixed in the European population and concern the fusion of a member of the cp32 family with a linear plasmid. Such a fusion between a linear plasmid and a cp32 plasmid has been previously observed in plasmid lp56 of *B. burgdorferi* s.s. type strain B31 [74]. We identified one lp56 plasmid in the Japanese isolate Hiratsuka based on the PFam32 protein (86.56% identity to B31 PFam32 sequence for lp56). However, this probably incomplete plasmid was made only of one 23 kb-long contig and showed only very little sequence similarity with its counterpart in strain B31. The third fusion (lp17 + lp28–4) occurred in several Asian isolates and is not monophyletic in the phylogeny depicted in Fig. 4. It was thus probably inherited horizontally and, as it is present in two out of the three Asian isolates coming from patients, one may speculate that it is linked to specific virulence factors. The presence of two different versions of this fused plasmid that differ in the point of fusion (Suppl. Fig. 5) implies that plasmids lp17 and lp28–4 were involved in at least two different fusion or recombination events. Similar fusion or relocation events have been previously observed in other genospecies. Plasmid lp17, for example, has also been suggested to have been involved in multiple relocations and fusions in *B. burgdorferi* s.s. [14].

Candidate genes for host and vector adaptation in *B. bavariensis*

Whereas the plasmid content in the European population was rather well conserved, plasmid lp28–9 was found only in a single European clade made up of five isolates (including the type strain PBi) and was absent from all other European *B. bavariensis* isolates. Plasmid lp28–9 was however present in five Asian isolates (two of which were isolated from patients) and in the two published strains from the sister species *B. garinii*. Annotation of this plasmid in the European isolates allowed us to identify only one gene with a predicted putative function: it is an ortholog of a lp28–2-located gene, BBG11, from *B. burgdorferi* s.s. strain 297 that has been shown to be upregulated in rodent hosts by the RpoS transcription factor [75] and to have higher expression levels in *B. burgdorferi* s.s. infecting steroid-treated non-human primates compared to immuno-competent animals [76]. This gene was found to be present only on the lp28–9 from European isolates and on some, but not all, of the lp28–7 and lp28–6 plasmids of some Asian isolates. Further research is necessary to find the function of this gene and whether it plays a role in pathogenicity in humans.

Other interesting genes highlighted by our study are those located on genetic diversity peaks (Suppl. Figs. 7, 8 and 9) within or between the two *B. bavariensis* populations. Because all Asian *B. bavariensis* isolates are vectored by *I. persulcatus*, whereas European isolates are found only in *I. ricinus*, it has been hypothesized that it is the adaptation to a new vector species that caused the strong bottleneck observed in the European population (see [9] for a review). Genes that show a high differentiation between the two populations are particularly interesting candidates for playing a role in the adaptation to specific tick vector species. Good examples of such genes are those encoding OspA, OspB and OspC located on lp54 (OspA, OspB) and cp26 (OspC) that showed a high diversity in the Asian population but were not variable at the amino-acid level in the European population. These proteins are known to be involved in the interaction between *B. burgdorferi* s.l. bacteria and their vectors and hosts (see [77] for a review). Topological differences that are observed in phylogenies of *ospC* and the rest of cp26 (Fig. 5) implies that differential evolution processes acted on the *ospC* gene and on plasmid cp26 during *B. bavariensis* evolution. A similar discrepancy has also been shown at the level of the *B. burgdorferi* s.l. species complex [78].

OspA and OspB could possibly be associated with evasion of the tick immune system as both genes are known to be expressed during infection of the tick [79]. Genetic variation at the *ospA* locus has already been observed in *B. garinii* [80]. This diversity, however, does not coincide with different vector use. Other species such as *B. afzelii* and *B. burgdorferi* s.s., which can also use different vectors, have a rather conserved *ospA* [80]. It may be that OspA fulfills different functions within the tick apart from its role as receptor binding molecule to TROSPA [81].

Another peak of genetic diversity was observed between 52 and 60 kb of the aligned plasmid lp54. RAST annotation [39, 40] did not show genes with known function in this region but the end of lp54 is known to be the region encoding the PFam54 protein family [74]. Indeed we were able to identify the genes encoding for CspA-related PFam54 proteins BGA66 and BGA71 on the region showing high diversity in both *B. bavariensis* populations (Suppl. Fig. 9). The CspA protein was first identified in *B. burgdorferi* s.s. and is known to be involved in the evasion of the innate immune response in the human host by binding regulators of the complement system [82, 83]. CspA was later shown to belong to a large protein family (PFam54) that is known to be under fast adaptive evolution [84]. Our results are in accordance with this finding. In particular, the evolution of the *B. bavariensis* PFam54 members BGA66 and BGA71 is of high interest as these proteins have been found to be involved in complement inactivation in *B. bavariensis* reference strain PBi but with a different mechanism as compared to the *B. burgdorferi* s. s. CspA protein [85].

Conclusions

Reconstruction of almost complete genomes of 33 *B. bavariensis* isolates from Eurasia showed that this species is characterized by a high degree of genetic conservation combined with plasticity. Asian isolates were found to have a high diversity in plasmid content and showed no geographic structuring. The European population was less diverse, appearing to have undergone a genetic bottleneck, but still showed some heterogeneous plasmid content. Two plasmid fusions were fixed in the latter population with respect to the Asian population. Horizontal transfer of genes or whole plasmids and gain and loss of plasmids likely influenced the evolution of this species. This study opens the way to functional genomic research on genes that have specific evolution pattern in this species and are thus good candidates for vector and host adaptation and for human pathogenicity.

Methods

Isolates used and sequencing

Information on origin of the isolates used for this study can be found in Table 1. All the European isolates from the strain bank of the German National Reference Center for *Borrelia* at the Bavarian Health and Food Safety Authority (Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit). Seventeen isolates were isolated from patients and one isolate was from a questing tick. The Asian isolates were isolated from questing ticks or patients in Russia and Japan.

Borrelia bavariensis were cultured in inhouse-made MKP (European samples) or inhouse-made BSK (Russian and Japanese samples) medium using standard procedures [86] to density of at least 10^8 cells per *mL* in order to obtain enough DNA. DNA was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany) and Japanese isolates were purified using Wizard genomic DNA purification kit (Promega). DNA concentrations and quality (260/280) were determined by using a Qubit® fluorometer 3.0 and (Thermo Fisher Scientific, USA) and NanoDrop® 1000 photometer (Thermo Fisher Scientific, USA).

For all 33 isolates, libraries were prepared according to the Nextera DNA sample preparation guide (Illumina, San Diego CA, USA). The samples were diluted to a DNA concentration of 0.2 ng/μl and "tagmented" by simultaneously fragmenting DNA using transposomes as provided by the manufacturer and adding adapters. After tagmentation, samples having adapters on both ends underwent five PCR cycles to amplify the product and to add index primers. The resulting libraries were then validated using an Agilent 2100 Bioanalyzer (Agilent, Germany). We then sequenced using an Illumina MiSeq platform (Illumina, San Diego CA, USA) that produced paired-end reads of 250 bp. Some low quality samples (A104S, DK6, PBae I, PBae II, PBar, PBN, PLad, PWin and PZwi) were repeated on an Illumina HiSeq platform producing 100 bp long paired-end reads.

For isolates PBi, A104S and NT24, Pacific Bioscience SMRT sequencing (hereafter PacBio) was performed using 10 μg of DNA. A library was prepared using Pacific Biosciences 20 kb library preparation protocol. Size selection of the final library was performed using BluePippin with a 10 kb cut-off. The library was sequenced on a Pacific Biosciences RS II instrument using P6-C4 chemistry with 360 min movie time.

Genome assembly and mapping

PacBio reads were assembled using HGAP v3 (Pacific Biosciences, SMRT Analysis Software v2.3.0). Chromosomes and linear plasmids 3′ and 5′ ends were trimmed for removing the pseudo-telomere regions that are known to be present in *B. burgdorferi* s.l. linear replicons [87]. Illumina contigs (see below) for the same isolates were then mapped to the PacBio assembly with NUCmer v. 3.1 from package MUMmer [32]. As PacBio sequencing technology is prone to sequencing errors like point mutations and short indels [30], we combined the data from PacBio and Illumina using the following rules: for each indel of length 5 bp or less keep the Illumina version, for longer indels keep the PacBio version. For point mutations, keep the Illumina version if all contigs mapping on this position agree, else keep PacBio version.

Illumina reads were assembled using SPAdes v. 3.10.1 [31]. As a comparison, we also assembled 25 isolates with SOAPdenovo v. 1.0 [35] and VelevetOptimizer v.

1.0 [36] and used QUAST v. 4.6 [37] to compare the quality of the three assemblies.

Mapping of SPAdes contigs was performed with NUCmer v. 3.1 from package MUMmer [32] on each one of the three isolates sequenced with PacBio that were used as reference. Contigs that were identified as being a whole chromosome (five cases) or a whole plasmid were used as is. For sequences that needed mapping of several contigs, the closest reference was used (highest identity and longest sequence reconstructed). This reference could be from one of the three PacBio isolates but also a contig identified as a whole plasmid in another Illumina isolate (61VB2 lp17, lp28–8 and cp32–5, A91S lp36, Arh923 lp28–7, FujiP2 lp28–6, Hiratsuka cp32–9 and cp32–11, J-14 lp17, lp28–4, lp28–6, cp32–7, cp32–10, cp32–11 and cp32–12, J-20 T lp25 and cp32–4, Lubl25 lp28–7, PBae II lp28–8, PBar lp54 and lp28–4 + cp32–1, PBN lp28–3, PHer I lp36, PLad lp28–8, PNeb lp36, lp17, lp28–7 and lp28–8, Prm7019 lp28–8 andPrm7569 cp32–1 were used as reference for other isolates). Each mapping file was then curated to suppress contigs overlapping other ones with higher identity (often these were very short contigs that mapped with low identity to a region already covered by a longer contig). We also corrected cases where one contig was supposed to map to several plasmids (often from the lp28 or cp32 families) or contigs which did not map over their whole length. In such cases, we kept the contig only the plasmid with the highest identity to the reference and longest mapping. In some rare cases, we used the same contig twice in the same plasmid as the PacBio reference showed that a sequence was repeated on the plasmid and thus it was not surprising that the Illumina reads from the two repeated regions would be assembled to the same contig; or we used the same contig in two different plasmids if the contig mapped with the same identity in both plasmids, again because the two plasmids had very similar sequences. Final chromosome and plasmid files were created based on the SNPs and indels identified with the program show-snps from package MUMmer [32] using following rules: for SNPs keep the Illumina allele if all contigs mapping at this position agree, else keep the reference allele if at least one contig also has it, else replace the base by "N"; keep insertions and deletions if and only if all contigs mapping at this position agree, else keep the reference version.

Final files, either made of an unmapped contig or of several contigs mapped to a reference were kept only if the final sequence length was at least 5000 bp and if unambiguous identification of the plasmid was possible thanks to mapping or the presence of a plasmid partition gene (see below). Shorter sequences were not considered as a plasmid and discarded from the final genomes.

The quality of the final reconstructed genomes was further studied by re-mapping the raw Illumina reads to the final genomes. This was done using BWA-MEM algorithm v. 0.7.17-r1188 [88] and read duplicates that can arise during library preparation by PCR were removed using Picard v. 2.21.6 (http://broadinstitute.github.io/picard). Read manipulation and extraction of coverage data was done with SAMtools v.1.9 [89]. For isolate NT24, the same procedure was repeated using PacBio plasmids to test for the coverage of the fused plasmids cp32–7 + 7 + 11 and cp32–12 + 5 + 6 (Suppl. Fig. 1). The quality of the assembled genomes was tested by comparing the relative standard deviation of the coverage of the raw reads between chromosomes and plasmids, between populations and between types of procedure to obtain the final sequence (full contig, or several contigs mapped to a reference) using Wilcoxon Rank Sum tests (Suppl. Fig. 3). The relative coverage of plasmids were also compared to the main chromosome over all European samples with Wilcoxon Rank Sum tests with *P*-values corrected for multiple testing with Bonferroni-Holm correction. The coverage of each plasmid relative to the chromosome for all European isolates was represented in Supplementary Figure 4.

Plasmid identification and plasmid partition genes

Final genome elements were named after the PFam32 protein family sequences that they contained. We used BLAST v. 2.8.1 [33, 34] (algorithm blastn) to identify the presence of plasmid partition genes of the PFam32, 49, 50 and 57–62 families. In a first BLAST round we used as queries the PFam32 genes sequences of *B. burgdorferi* s.s. strains B31, BOL26, JD1 and 118a and *B. afzelii* strain PKo to cover the whole plasmid diversity and the PFam49, 50 and 57–62 of *B. burgdorferi* s.s. strain MM1. We performed the search both on the final assembled genome and on the SPAdes Illumina contigs of each isolate as some plasmids could not be assembled. We then reiterated the BLAST search using as queries all the hits found in the first search. We then removed from the final hit lists presented in Supplementary Table 1 all hits that were shorter than half the length of the references (reference lengths were around 750 bp for PFam32, 550 bp for PFam49 and PFam50 and 900 to 1100 bp for PFam57–62) and that had no open reading frame over at least half of the length of the reference.

Quantitative PCR for plasmid copy number estimation

We used a qPCR protocol to estimate the copy number of plasmids cp26, lp17 and lp36 relative to the main chromosome following Millan et al. [38]. This was performed on two isolates of strain PBi (named 2418 and 24510) each grown as three biological replicates in MKP medium with standard conditions [86]. DNA was

extracted using a Maxwell automatic purification instrument once cell density reached approximately 10^7 cells/mL. Digestion with the PstI enzyme (NEB R0140S) was done to ensure equal accessibility of linear and circular plasmids during the PCR reaction. Five hundred nanograms of DNA from each extraction were digested for 1 h and 10 min at 37 °C with 0.5 μL of PstI in a final reaction volume of 25 μL, after which the enzyme was inactivated for 20 min at 80 °C. Quantitative PCR primers were designed to be as similar as possible in their specifications in order for them to be used in a single qPCR run. Primer-BLAST [90] was used to produce primer candidates that did not bind multiple times within the *B. bavariensis* PBi genome (Suppl. Table 2). PCR samples were prepared using 1 μM primer concentrations and 10 ng of DNA using the S7 Fusion Polymerase system according to standard protocol for a final reaction volume of 20 μL (IsoGene Scientific). A two-step PCR program was chosen due to the small sizes of the amplified fragments with a thermocycle of 30 s initialization at 98 °C, followed by 30 cycles of 98 °C denaturation for 5 s and 63 °C annealing for 20 s finishing with an elongation step at 72 °C for 7 min. PCR products were visualized using a 1% agarose gel. All PCR produced the expected product size.

All qPCR runs were run using the SsoAdvanced™ Universal SYBR® Green Supermix (Bio Rad) according to standard protocol on a Bio Rad C1000 Touch™ Thermocycler with the same thermoprofile as the two step PCR described above. For each run, two technical replicates from each biological replicate ($n = 3$) were used for a total of 6 qPCR replicates per isolate. A standard curve was calculated per run for both the plasmid and chromosome primers using standards of known DNA concentration (20, 3.3, 2.5, 2.0, 0.3, and 0.04 ng/μL) made from a DNA pool of all samples. Each standard was run in triplicate for each primer set. A negative control was included for each technical replicate of either unknowns or standards ($n = 10$ per plate). Each run included unknowns and standards for one plasmid (cp26, lp36, lp17) and the main chromosome. Cycle threshold (CT) values were recorded for all samples. Primer efficiencies were then calculated according to standard protocol (Bio Rad) from these standard curves. Plasmid copy numbers were calculated for each technical replicate according to the equation described in [91].

Plasmid fusions

We studied the architecture of plasmids lp17, lp28–4, lp28–7 and cp32–1 in detail as different fusions and translocation involving these plasmids were observed. Following Casjens et al. [14] we defined as a new plasmid subtype, a plasmid sequence that had with respect to the other plasmid subtypes either presence of 400 bp or longer indels or obvious evidence of past interplasmid DNA exchanges (translocations). Casjens et al.'s criteria also involved synteny, but our current annotation contains mostly hypothetical proteins and did not allow us to test for synteny. We used BLAST v. 2.8.1 [33, 34] (algorithm blastn) between each of these four plasmids to identify plasmid types.

Genome annotation

Genome annotation was performed with RAST Annotation Server v. 2.0 [39, 40] with default parameters. As an annotation is available online for the main chromosome of reference strain PBi (GenBank accession number CP000013), we compared this annotation with the one obtained for our genome reconstruction of strain PBi based on combining PacBio and Illumina data with The SEED Viewer v. 2.0 [40] and produced a Blast Dot Plot shown in Supplementary Figure 6.

For each one of the 33 isolates, we compared one by one all genes for which the product is at least 50 amino-acids long, with all genes of the 32 others using blastp algorithm from BLAST v. 2.8.1 [33, 34]. We kept all hits that were at least half as long as the query and shared at least 90% sequence identity with the query and recorded on which genomic segment they were located for each isolate.

Phylogeny reconstruction

Phylogeny reconstruction was performed on the main chromosome as it is known to be very stable in the *B. burgdorferi* s.l. species complex [17]. In addition to the 33 isolates published in this study we also used four *B. bavariensis* strains published in GenBank (under accession numbers CP000013 for strain PBi from Germany, CP003151 for strain BgVir from Russia and CP003866 and CP007564 for strains NMJW1 and SZ from China) and the *B. garinii* strain 20047 as an outgroup to root the tree (GenBank accession number CP028861). Alignment was performed with MAFFT v7.407 [46, 47] and phylogeny reconstruction was performed with BEAST v1.8.0 [41] with the following parameters: coalescent model with exponential growth based on doubling time, lognormal-relaxed clock [42], GTR substitution model [43]. The chain was run for 100 Million steps in three independent runs and convergence was checked with Tracer v. 1.4 [92]. One of the runs did not converge and for the other two runs a burn-in of 30 and 40% respectively was found appropriate. We then used TreeAnnotator v. 1.10.4 [41] to identify the best tree after burn-in. The phylogeny presented in Fig. 4 was plotted with FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/). On the phylogeny we added for each branch the gain or loss of plasmids based on the genome reconstructions presented in Table 1 and Supplementary Table 1 (we considered a plasmid as present when either its sequence

or one of its specific plasmid partition gene was present) and using maximum parsimony principle. When two solutions leaded to the same minimum number of events, we chose the solution with the lowest number of gains.

Phylogenies were also reconstructed on plasmid cp26 cutting out the *ospC* locus (200 bp upstream and downstream the gene) and on gene *ospC* with BEAST v. 1.8.0 [41] using the same priors and the same procedures as above except that the coalescent model did not include exponential growth. We included GenBank strains BgVir (*B. bavariensis* CP003201.1), Far04 and PBr (*B. garinii* CP001319.1 and CP001305.1), PKo, K78 and ACA-1 (*B. afzelii* CP002934.1, CP009060.1, CP001250.1) and A14S (*B. spielmanii* CP001467.1). The sequences were aligned with MAFFT v7.407 [46, 47] and the chains were run for 500 million states for cp26 and 20 million states for *ospC* each in triplicate. Best trees were reconstructed after removing a burn-in of 10% of the chain and all three runs showed very similar results for each tree. Both trees were plotted using FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) and manually rotated to produce Fig. 5 comparing *ospC* and plasmid cp26.

Statistical analyses and genetic diversity

All statistical analyses were performed with R v. 3.5.2 [93] and genetic distance and genetic diversity were estimated using packages pegas v. 0.12 [94] and hierfstat v. 0.04–22 [95] on orthologous plasmid sequences aligned with MAFFT v7.407 [46, 47] and along the alignments of the main chromosome as well as plasmids cp26 and lp54 (only segments that could be aligned over their whole length) using windows of 1000 bp sliding every 100 bp.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12864-020-07054-3.

Additional file 1: Supplementary Figure 1. Coverage of raw reads mapping on PacBio fused plasmids cp32–7 + 7 + 11 (a) and cp32–12 + 5 + 6 (b) of isolate NT24. Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 [88] on PacBio fused plasmids cp32–7 + 7 + 11 (a) and cp32–12 + 5 + 6 (b) of isolate NT24. Regions of low to null coverage (marked in red) show that the fusion is not supported by the short-read data.

Additional file 2: Supplementary Figure 2. Comparison of three assemblers for Illumina assembly of 25 *B. bavariensis* isolates. These violin plots compare N50 (a) and total length of contigs (b) obtained with QUAST v. 4.6 [37] on assemblies performed with SPAdes v. 3.10.1 [31], SOAPdenovo v. 1.0 [35] and VelvetOptimizer v. 1.0 [36].

Additional file 3: Supplementary Figure 3. Replicon assembly quality as a function of population, mapping method (a) and type of replicon (b). Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 [88] to the final reconstructed genomes and the relative standard deviation of the coverage of the raw reads was used as a measure of assembly quality. We compare here replicons from European (left bars) and Asian (right bars) genomes depending on (a) whether the replicon was made as one contig (pink) or as several contigs mapped to a reference (purple) and on (b) whether it was a chromosome (orange) or a plasmid (blue). Error bars show standard error of the mean. ***: Wilcoxon Rank Sum Test for Europe against Asia, *P*-value < 0.001. Other tests comparing mapping methods (a) and type of replicons (b) were not significant.

Additional file 4: Supplementary Figure 4. Coverage ratio of European replicons as a proxy for copy number. Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 [88] to the final reconstructed genomes and the ratio of the coverage of each replicon with respect to the chromosome was computed in each European isolate. Error bars show standard error of the mean. Dark blue numbers indicate the number of plasmids of this type in the European sample. Wilcoxon Rank Sum Tests comparing coverage of each plasmid with that of the chromosomes: *P*-Value after Bonferroni-Holm correction *: < 0.05, ***: < 0.001, else: not significant.

Additional file 5: Supplementary Figure 5. Schematic representation of plasmid subtypes and fusion/relocation events on lp17, lp28–4, lp28–7 and cp32–1. The different plasmid subtypes (numbered arbitrarily) are represented as black bars. We defined as a new plasmid subtype, a plasmid sequence that had, with respect to the other plasmid subtypes, either presence of 400 bp or longer indels or obvious evidence of past interplasmid DNA exchanges (translocations). We used BLAST v. 2.8.1 [33, 34] to identify plasmid types and colour-shaded areas represent BLAST hits on the same strand (blue) and inversions (pink). Different shades of color are just used for clarity and have no meaning. Dashed lines represent plasmid fusions. Scale bars above the plots are plasmid lengths in kb. *: specific cases: Arh913 cp32–1 could no be assembled. Konnai17 had two lp28–7 plasmids, the second one has the same subtype as plasmid lp28–7 in FujiP2.

Additional file 6: Supplementary Figure 6. Dotplot comparing annotation of strain PBi between our isolate and a previously published one. Comparison of gene content realized in RAST Annotation Server v. 2.0 [39, 40] on the main chromosome. PBi accession number in RAST: 290434.1.

Additional file 7: Supplementary Figure 7. Genetic diversity along the main chromosome of *B. bavariensis*. Genetic diversity was estimated using R package pegas v. 0.12 [94] on orthologous sequences aligned with MAFFT v7.407 [46, 47] on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 [39, 40].

Additional file 8: Supplementary Figure 8. Genetic diversity along plasmid cp26 of *B. bavariensis*. Genetic diversity was estimated using R package pegas v. 0.12 [94] on orthologous sequences aligned with MAFFT v7.407 [46, 47] on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 [39, 40].

Additional file 9: Supplementary Figure 9. Genetic diversity along plasmid lp54 of *B. bavariensis*. Genetic diversity was estimated using R package pegas v. 0.12 [94] on orthologous sequences aligned with MAFFT v7.407 [46, 47] on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 [39, 40].

Additional file 10: Supplementary Table 1. Plasmid partition genes identified in 33 B. bavariensis strains. **Supplementary Table 2.** Primers used for qPCR.

Abbreviations

LB: Lyme Borreliosis; NGS: Next-generation sequencing; PacBio: Pacific Bioscience SMRT sequencing; SNP: Single nucleotide polymorphism; GTR: General Time Reversible (substitution model)

Acknowledgments

The Pacific Bioscience SMRT sequencing service was provided by the Norwegian Sequencing Centre (www.sequencing.uio.no), a national technology platform hosted by the University of Oslo and supported by the

"Functional Genomics" and "Infrastructure" programs of the Research Council of Norway and the Southeastern Regional Health Authorities.
The authors thank Hilde Lainer for help in qPCR design.

Authors' contributions
NSB, GM and VF designed the study. RER, KN, SK, AT, KS, SYK, HK, VF and GM collected the samples and performed the experiments and sequencing. NSB, KN, AP and SM analyzed the data. NSB wrote the manuscript with input from RER, KN and GM. All co-authors reviewed the manuscript. The author(s) read and approved the final manuscript.

Funding
Robert-Koch-Institut funded strain isolation, cultivation and Illumina sequencing for 33 isolates at the NRZ Borrelia. PacBio sequencing for tree isolates was financed by the ESCMID Study Group for Lyme Borreliosis. qPCR experiments were funded through the German Research Foundation (DFG Grant No. BE 5791/2–1). Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials
The datasets generated and/or analyzed during the current study are available in the NCBI repository BioProjects PRJNA449844 and PRJNA327303.

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Division of Evolutionary Biology, Faculty of Biology, LMU Munich, Grosshaderner Strasse 2, 82152 Planegg-Martinsried, Germany. [2]University of Helsinki, Biomedicum Helsinki, PO Box 63, Haartmaninkatu 8, FIN-00014 Helsinki, Finland. [3]Gene Center, Laboratory for Functional Genome Analysis, LMU Munich, Feodor-Lynen-Strasse 25, 81377 Munich, Germany. [4]Department of Veterinary Epidemiology, Joint Faculty of Veterinary Medicine, Yamaguchi University, Yamaguchi 753–8515, Japan. [5]Department of Bacteriology-I, National Institute of Infectious Diseases, Tokyo 162–8640, Japan. [6]Laboratory of Molecular Genetics, Institute of Natural Sciences and Mathematics, Ural Federal University, Lenin Avenue 51, Yekaterinburg 620000, Russia. [7]National Reference Centre for Borrelia at the Bavarian Health and Food Safety Authority, Veterinärstr 2, 85764 Oberschleissheim, Germany.

Received: 15 May 2020 Accepted: 6 September 2020
Published online: 08 October 2020

References
1. Tilly K, Rosa PA, Stewart PE. Biology of infection with *Borrelia burgdorferi*. Infect Dis Clin N Am. 2008;22:217–34.
2. Gern L. *Borrelia burgdorferi* sensu lato, the agent of Lyme borreliosis: life in the wilds. Parasite. 2008;15:244–7.
3. Price AL, Tandon A, Patterson N, Barnes KC, Rafaels N, Ruczinski I, Beaty TH, Mathias R, Reich D, Myers S. Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. PLoS Genet. 2009;5:e1000519.
4. Eisen L. Vector competence studies with hard ticks and *Borrelia burgdorferi* sensu lato spirochetes: a review. Ticks Tick Borne Dis. 2019;11:101359.
5. Kurtenbach K, Hoen AG, Bent SJ, Vollmer SA, Ogden NH, Margos G. Population biology of Lyme Borreliosis spirochetes. In: Robinson DA, Falush D, Feil EJ, editors. Bacterial population genetics in infectious disease. 1st ed. Hoboken: Wiley; 2010.
6. Margos G, Wilske B, Sing A, Hizo-Teufel C, Cao W-C, Chu C, Scholz H, Straubinger RK, Fingerle V. *Borrelia bavariensis* sp. nov. is widely distributed in Europe and Asia. Int J Syst Evol Microbiol. 2013;63(Pt 11):4284–8.
7. Margos G, Vollmer SA, Cornet M, Garnier M, Fingerle V, Wilske B, Bormane A, Vitorino L, Collares-Pereira M, Drancourt M, Kurtenbach K. A new *Borrelia* species defined by multilocus sequence analysis of housekeeping genes. Appl Environ Microbiol. 2009;75:5410–6.
8. Margos G, Vollmer SA, Ogden NH, Fish D. Population genetics, taxonomy, phylogeny and evolution of *Borrelia burgdorferi* sensu lato. Infect Genet Evol. 2011;11:1545–63.
9. Margos G, Fingerle V, Reynolds S. *Borrelia bavariensis*: vector switch, niche invasion, and geographical spread of a tick-borne bacterial parasite. Front Ecol Evol. 2019;7(October):1–20.
10. Masuzawa T, Wilske B, Komikado T, Suzuki H, Kawabata H, Sato N, Muramatsu K, Isogai E, Isogai H, Johnson RC, Yanagihara Y. Comparison of OspA serotypes for *Borrelia burgdorferi* sensu lato from Japan, Europe and North America. Microbiol Immunol. 1996;40:539–45.
11. Gatzmann F, Metzler D, Krebs S, Blum H, Sing A, Takano A, Kawabata H, Fingerle V, Margos G, Becker NS. NGS population genetics analyses reveal divergent evolution of a Lyme Borreliosis agent in Europe and Asia. Ticks Tick Borne Dis. 2015;6:344–51.
12. Becker NS, Margos G, Blum H, Krebs S, Graf A, Lane RS, Castillo-Ramírez S, Sing A, Fingerle V. Recurrent evolution of host and vector association in bacteria of the *Borrelia burgdorferi* sensu lato species complex. BMC Genomics. 2016;17(1):734.
13. Postic D, Garnier M, Baranton G. Multilocus sequence analysis of atypical *Borrelia burgdorferi* sensu lato isolates - description of *Borrelia californiensis* sp. nov., and genomospecies 1 and 2. Int J Med Microbiol. 2007;297:263–71.
14. Casjens SR, Gilcrease EB, Vujadinovic M, Mongodin EF, Luft BJ, Schutzer SE, Fraser CM, Qiu WG. Plasmid diversity and phylogenetic consistency in the Lyme disease agent *Borrelia burgdorferi*. BMC Genomics. 2017;18(1):165.
15. Brisson D, Drecktrah D, Eggers CH, Samuels DS. Genetics of *Borrelia burgdorferi*. Annu Rev Genet. 2012;46:515–36.
16. Fraser CM, Casjens S, Huang WM, Sutton GG, Clayton R, Lathigra R, White O, Ketchum KA, Dodson R, Hickey EK, Gwinn M, Dougherty B, Tomb JF, Fleischmann RD, Richardson D, Peterson J, Kerlavage AR, Quackenbush J, Salzberg S, Hanson M, van Vugt R, Palmer N, Adams MD, Gocayne J, Weidman J, Utterback T, Watthey L, McDonald L, Artiach P, Bowman C, et al. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature. 1997;390:580–6.
17. Mongodin EFEF, Casjens SRSR, Bruno JFJF, Xu Y, Drabek EF, Riley DRDR, Cantarel BL, Pagan PE, Hernandez YYA, Vargas LCLC, Dunn JJJ, Schutzer SSE, Fraser CCM, Qiu W-GGW, Luft BBJ, Piesman J, Clark K, Dolan M, Happ C, Burkot T, Steere A, Coburn J, Glickstein L, Radolf J, Salazar J, Dattwyler R, Dennis D, Nekomoto T, Victor J, Paul W, et al. Inter- and intra-specific pan-genomes of *Borrelia burgdorferi* sensu lato: genome stability and adaptive radiation. BMC Genomics. 2013;14:693.
18. Margos G, Hepner S, Mang C, Marosevic D, Reynolds SE, Krebs S, Sing A, Derdakova M, Reiter MA, Fingerle V. Lost in plasmids: next generation sequencing and the complex genome of the tick-borne pathogen *Borrelia burgdorferi*. BMC Genomics. 2017;18(1):422.
19. Genome. Bethesda: National Library of Medicine (US), National Center for Biotechnology Information; 2004. Available from: https://www.ncbi.nlm.nih.gov/gene/. Accessed 31 Mar 2020.
20. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh SM, Oskam C, Fingerle V. The genus Borrelia reloaded. PLoS One. 2018;13:1–14.
21. Brenner EV, Kurilshikov AM, Stronin OV, Fomenko NV. Whole-genome sequencing of *Borrelia garinii* BgVir, isolated from Taiga ticks (*Ixodes persulcatus*). J Bacteriol. 2012;194:5713.
22. Wu Q, Liu Z, Li Y, Guan G, Niu Q, Chen Z, Luo J, Yin H. Genome sequence of *Borrelia garinii* strain SZ, isolated in China. Genome Announc. 2014;2(4):e00010–4.
23. Jiang B, Yao H, Tong Y, Yang X, Huang Y, Jiang J, Cao W. Genome sequence of *Borrelia garinii* strain NMJW1, isolated from China. J Bacteriol. 2012;194:6660–1.
24. Chaconas G, Norris SJ. Peaceful coexistence amongst *Borrelia* plasmids: getting by with a little help from their friends? Plasmid. 2013;70:161–7.
25. Casjens SR, Mongodin EF, Qiu W-G, Luft BJ, Schutzer SE, Gilcrease EB, Huang WM, Vujadinovic M, Aron JK, Vargas LC, Freeman S, Radune D, Weidman JF, Dimitrov GI, Khouri HM, Sosa JE, Halpin RA, Dunn JJ, Fraser CM. Genome stability of Lyme disease spirochetes: comparative genomics of *Borrelia burgdorferi* plasmids. PLoS One. 2012;7:e33280.
26. Casjens SR, Di L, Akther S, Mongodin EF, Luft BJ, Schutzer SE, Fraser CM, Qiu W-G. Primordial origin and diversification of plasmids in Lyme disease agent bacteria. BMC Genomics. 2018;19:218.
27. Brisson D, Zhou W, Jutras BL, Casjens S, Stevenson B. Distribution of cp32 prophages among Lyme disease-causing spirochetes and natural diversity of their lipoprotein-encoding erp loci. Appl Environ Microbiol. 2013;79:4115–28.

28. Lin T, Gao L, Zhang C, Odeh E, Jacobs MB, Coutte L, Chaconas G, Philipp MT, Norris SJ. Analysis of an ordered, comprehensive STM mutant library in infectious *Borrelia burgdorferi*: insights into the genes required for mouse infectivity. PLoS One. 2012;7:e47532.
29. Norris SJ. The vls antigenic variation systems of Lyme disease *Borrelia*: eluding host immunity through both random, segmental gene conversion and framework heterogeneity. Microbiol Spectr. 2014;2. https://doi.org/10.1128/microbiolspec.MDNA3-0038-2014.
30. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012;13:341.
31. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.
32. Delcher AL. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res. 2002;30:2478–83.
33. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
34. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. BMC Bioinformatics. 2009;10:421.
35. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res. 2010;20:265–72.
36. Zerbino DR, Birney E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res. 2008;18:821–9.
37. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.
38. Millan AS, Santos-Lopez A, Ortega-Huedo R, Bernabe-Balas C, Kennedy SP, Gonzalez-Zorn B. Small-plasmid-mediated antibiotic resistance is enhanced by increases in plasmid copy number and bacterial fitness. Antimicrob Agents Chemother. 2015;59:3335–41.
39. Aziz RK, Bartels D, Best A, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.
40. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 2014;42(Database issue):D206–14.
41. Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 2007;7:214.
42. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. Relaxed phylogenetics and dating with confidence. PLoS Biol. 2006;4:e88.
43. Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. Am Math Soc Lect Math Life Sci. 1986;17:57–86.
44. Nei M. Molecular evolutionary genetics. Book. New York: Columbia University Press; 1987.
45. Weir BS, Cockerham CC. Estimating F-statistics for the analysis of population structure. Evolution (N Y). 1984;38:1358–70.
46. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30:772–80.
47. Katoh K. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30:3059–66.
48. Barbour AG, Travinsky B. Evolution and distribution of the ospC gene, a transferable serotype determinant of *Borrelia burgdorferi*. MBio. 2010;1. https://mbio.asm.org/content/1/4/e00153-10.
49. Wang IN, Dykhuizen DE, Qiu W, Dunn JJ, Bosler EM, Luft BJ. Genetic diversity of ospC in a local population of *Borrelia burgdorferi* sensu stricto. Genetics. 1999;151:15–30.
50. Brisson D, Dykhuizen DE. ospC diversity in *Borrelia burgdorferi*: different hosts are different niches. Genetics. 2004;168:713–22.
51. Kuleshov KV, Margos G, Fingerle V, Koetsveld J, Goptar IA, Markelov ML, Kolyasnikova NM, Sarksyan DS, Kirdyashkina NP, Shipulin GA, Hovius JW, Platonov AE. Whole genome sequencing of *Borrelia miyamotoi* isolate Izh-4: reference for a complex bacterial genome. BMC Genomics. 2020;21:16.
52. Schutzer SE, Fraser-Liggett CM, Casjens SR, Qiu WG, Dunn JJ, Mongodin EF, Luft BJ. Whole-genome sequences of thirteen isolates of *Borrelia burgdorferi*. J Bacteriol. 2011;193:1018–20.
53. Schutzer SE, Fraser-Liggett CM, Qiu WG, Kraiczy P, Mongodin EF, Dunn JJ, Luft BJ, Casjens SR. Whole-genome sequences of *Borrelia bissettii*, *Borrelia valaisiana*, and *Borrelia spielmanii*. J Bacteriol. 2012;194:545–6.
54. Barbour AG. Plasmid analysis of *Borrelia burgdorferi*, the Lyme disease agent. J Clin Microbiol. 1988;26:475–8.
55. Schwan TG, Burgdorfer W, Garon CF. Changes in infectivity and plasmid profile of the Lyme disease spirochete, *Borrelia burgdorferi*, as a result of in vitro cultivation. Infect Immun. 1988;56:1831–6.
56. Grimm D, Elias AF, Tilly K, Rosa PA. Plasmid stability during in vitro propagation of *Borrelia burgdorferi* assessed at a clonal level. Infect Immun. 2003;71:3138–45.
57. Tyler S, Tyson S, Dibernardo A, Drebot M, Feil EJ, Graham M, Knox NC, Lindsay LR, Margos G, Mechai S, Van Domselaar G, Thorpe HA, Ogden NH. Whole genome sequencing and phylogenetic analysis of strains of the agent of Lyme disease *Borrelia burgdorferi* from Canadian emergence zones. Sci Rep. 2018;8:1–12.
58. Kovalev SY, Golovljova IV, Mukhacheva TA. Natural hybridization between *Ixodes ricinus* and *Ixodes persulcatus* ticks evidenced by molecular genetics methods. Ticks Tick Borne Dis. 2016;7:113–8.
59. Kurtenbach K, Sewell HS, Ogden NH, Randolph SE, Nuttall PA. Serum complement sensitivity as a key factor in Lyme disease ecology. Infect Immun. 1998;66:1248–51.
60. Casjens S, Huang WM. Linear chromosomal physical and genetic map of *Borrelia burgdorferi*, the Lyme disease agent. Mol Microbiol. 1993;8:967–80.
61. Hinnebusch J, Barbour AG. Linear- and circular-plasmid copy numbers in *Borrelia burgdorferi*. J Bacteriol. 1992;174:5251–7.
62. Jan AT. Outer Membrane Vesicles (OMVs) of gram-negative bacteria: a perspective update. Front Microbiol. 2017;8:1053.
63. Kulp A, Kuehn MJ. Biological functions and biogenesis of secreted bacterial outer membrane vesicles. Annu Rev Microbiol. 2010;64:163–84.
64. Malabirade A, Habier J, Heintz-Buschart A, May P, Godet J, Halder R, Etheridge A, Galas D, Wilmes P, Fritz JV. The RNA complement of outer membrane vesicles from *Salmonella enterica* Serovar Typhimurium under distinct culture conditions. Front Microbiol. 2018;9:2015.
65. Garon CF, Dorward DW, Corwin MD. Structural features of *Borrelia burgdorferi* - the lyme disease spirochete: silver staining for nucleic acids. Scanning Microsc. 1989;3(Suppl. 3):109–15.
66. Malge A, Ghai V, Reddy PJ, Baxter D, Kim TK, Moritz RL, Wang K. mRNA transcript distribution bias between *Borrelia burgdorferi* bacteria and their outer membrane vesicles. FEMS Microbiol Lett. 2018;365(13):fny135.
67. Mashburn-Warren LM, Whiteley M. Special delivery: vesicle trafficking in prokaryotes. Mol Microbiol. 2006;61(4):839–46.
68. Tran F, Boedicker JQ. Genetic cargo and bacterial species set the rate of vesicle-mediated horizontal gene transfer. Sci Rep. 2017;7(1):1–0.
69. Casselli T, Tourand Y, Bankhead T. Altered murine tissue colonization by *Borrelia burgdorferi* following targeted deletion of linear plasmid 17-carried genes; 2012.
70. Casselli T, Crowley MA, Highland MA, Tourand Y, Bankhead T. A small intergenic region of lp17 is required for evasion of adaptive immunity and induction of pathology by the Lyme disease spirochete. Cell Microbiol. 2019;21:e13029.
71. Livey I, Gibbs CP, Schuster R, Dorner F. Evidence for lateral transfer and recombination in OspC variation in Lyme disease *Borrelia*. Mol Microbiol. 1995;18:257–69.
72. Marconi RT, Hohenberger S, Jauris-Heipke S, Schulte-Spechtel U, LaVoie CP, Rossler D, Wilske B. Genetic analysis of *Borrelia garinii* OspA serotype 4 strains associated with neuroborreliosis: evidence for extensive genetic homogeneity. J Clin Microbiol. 1999;37:3965–70.
73. Kraiczy P, Skerka C, Zipfel PF, Brade V. Complement regulator-acquiring surface proteins of *Borrelia burgdorferi*: a new protein family involved in complement resistance. Wien Klin Wochenschr. 2002;114:568–73.
74. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, Rosa P, Lathigra R, Sutton G, Peterson J, Dodson RJ, Haft D, Hickey E, Gwinn M, White O, Fraser CM. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. Mol Microbiol. 2000;35:490–516.

75. Caimano MJ, Iyer R, Eggers CH, Gonzalez C, Morton EA, Gilbert MA, Schwartz I, Radolf JD. Analysis of the RpoS regulon in *Borrelia burgdorferi* in response to mammalian host signals provides insight into RpoS function during the enzootic cycle. Mol Microbiol. 2007;65:1193–217.
76. Narasimhan S, Camaino MJ, Liang FT, Santiago F, Laskowski M, Philipp MT, Pachner AR, Radolf JD, Fikrig E. Borrelia burgdorferi transcriptome in the central nervous system of non-human primates. Proc Natl Acad Sci U S A. 2003;100:15953–8.
77. Templeton TJ. *Borrelia* outer membrane surface proteins and transmission through the tick. J Exp Med. 2004;199(5):603–6.
78. Jacquot M, Gonnet M, Ferquel E, Abrial D, Claude A, Gasqui P, Choumet V, Charras-Garrido M, Garnier M, Faure B, Sertour N, Dorr N, De Goer J, Vourc'h G, Bailly X. Comparative population genomics of the *Borrelia burgdorferi* species complex reveals high degree of genetic isolation among species and underscores benefits and constraints to studying intra-specific epidemiological processes. PLoS One. 2014;9:e94384.
79. Tilly K, Bestor A, Rosa PA. Functional equivalence of OspA and OspB, but not OspC, in tick colonization by *Borrelia burgdorferi*. Infect Immun. 2016;84:1565–73.
80. Will G, Jauris-Heipke S, Schwab E, Busch U, Rossler D, Soutschek E, Wilske B, Preac-Mursic V. Sequence analysis of ospA genes shows homogeneity within *Borrelia burgdorferi* sensu stricto and *Borrelia afzelii* strains but reveals major subgroups within the *Borrelia garinii* species. Med Microbiol Immunol. 1995;184:73–80.
81. Pal U, Li X, Wang T, Montgomery RR, Ramamoorthi N, Desilva AM, Bao F, Yang X, Pypaert M, Pradhan D, Kantor FS, Telford S, Anderson JF, Fikrig E. TROSPA, an *Ixodes scapularis* receptor for *Borrelia burgdorferi*. Cell. 2004;119:457–68.
82. Cordes FS, Roversi P, Kraiczy P, Simon MM, Brade V, Jahraus O, Wallis R, Skerka C, Zipfel PF, Wallich R, Lea SM. A novel fold for the factor H-binding protein BbCRASP-1 of *Borrelia burgdorferi*. Nat Struct Mol Biol. 2005;12:276–7.
83. Cordes FS, Kraiczy P, Roversi P, Simon MM, Brade V, Jahraus O, Wallis R, Goodstadt L, Ponting CP, Skerka C, Zipfel PF, Wallich R, Lea SM. Structure-function mapping of BbCRASP-1, the key complement factor H and FHL-1 binding protein of *Borrelia burgdorferi*. Int J Med Microbiol. 2006;296(Suppl):177–84.
84. Wywial E, Haven J, Casjens SR, Hernandez YA, Singh S, Mongodin EF, Fraser-Liggett CM, Luft BJ, Schutzer SE, Qiu WG. Fast, adaptive evolution at a bacterial host-resistance locus: the PFam54 gene array in *Borrelia burgdorferi*. Gene. 2009;445:26–37.
85. Hammerschmidt C, Klevenhaus Y, Koenigs A, Hallström T, Fingerle V, Skerka C, Pos KM, Zipfel PF, Wallich R, Kraiczy P. BGA66 and BGA71 facilitate complement resistance of *Borrelia bavariensis* by inhibiting assembly of the membrane attack complex. Mol Microbiol. 2016;99:407–24.
86. Preac-Mursic V, Wilske B, Schierz G. European *Borrelia burgdorferi* isolated from humans and ticks culture conditions and antibiotic susceptibility. Zentralbl Bakteriol Mikrobiol Hyg A. 1986;263:112–8.
87. Chaconas G, Kobryn K. Structure, function, and evolution of linear replicons in *Borrelia*. Annu Rev Microbiol. 2010;64:185–202.
88. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009;25:1754–60.
89. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. The sequence alignment/map format and SAMtools. Bioinformatics. 2009;25:2078–9.
90. Ye J, Coulouris G, Zaretskaya I, Cutcutache I, Rozen S, Madden TL. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinformatics. 2012;13:134.
91. San Millan A, Heilbron K, MacLean RC. Positive epistasis between co-infecting plasmids promotes plasmid survival in bacterial populations. ISME J. 2014;8:601–12.
92. Rambaut A, Suchard MA, Xie D, Drummond AJ: Tracer v1.6. 2014. Available from http://beast.bio.ed.ac.uk/Tracer.
93. R Development Core Team: R: a language and environment for statistical computing. 2013.
94. Paradis E. pegas: an R package for population genetics with an integrated-modular approach. Bioinformatics. 2010;26:419–20.
95. Goudet J. hierfstat, a package for r to compute and test hierarchical F-statistics. Mol Ecol Notes. 2005;5:184–6.

Publisher's Note

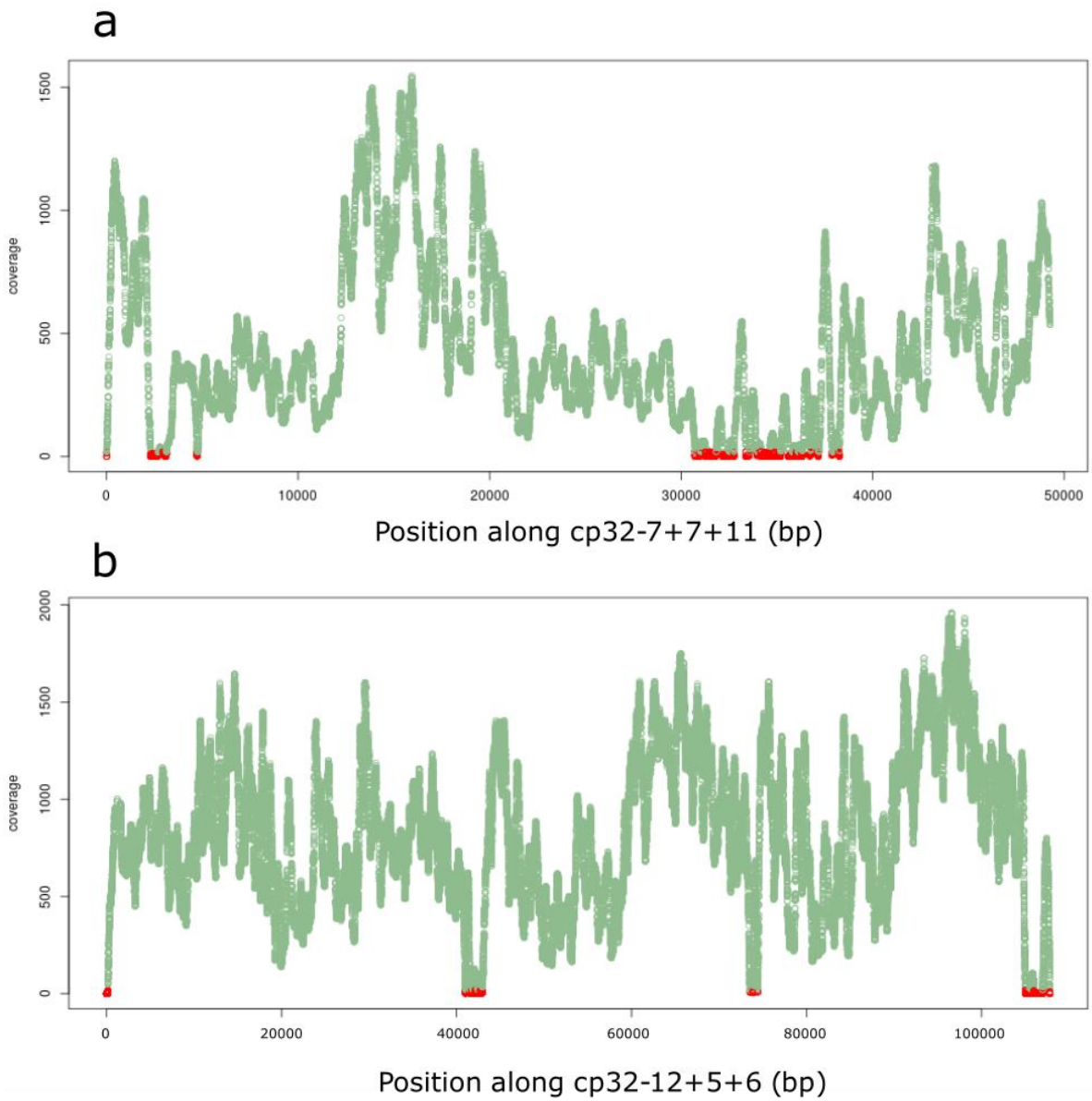
Paper 4: Supplementary Material

High conservation combined with high plasticity: genomics and evolution of *Borrelia bavariensis*

Noémie S. Becker, **Robert E. Rollins**, Kateryna Nosenko, Alexander Paulus, Samantha Martin, Stefan Krebs, Ai Takano, Kozue Sato, Sergey Y. Kovalev, Hiroki Kawabata, Volker Fingerle and Gabriele Margos

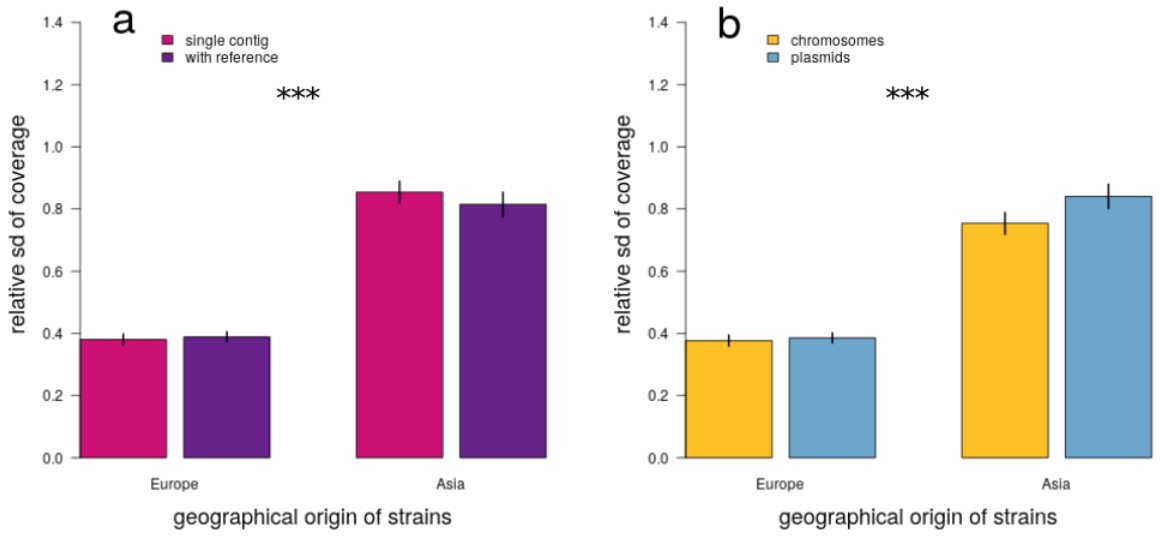
BMC Genomics (2020) 21:702
https://doi.org/10.1186/s12864-020-07054-3

Supplementary Figure 1. Coverage of raw reads mapping on PacBio fused plasmids cp32–7 + 7 + 11 (a) and cp32–12 + 5 + 6 (b) of isolate NT24. Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 (109) on PacBio fused plasmids cp32–7 + 7 + 11 (a) and cp32–12 + 5 + 6 (b) of isolate NT24. Regions of low to null coverage (marked in red) show that the fusion is not supported by the short-read data.

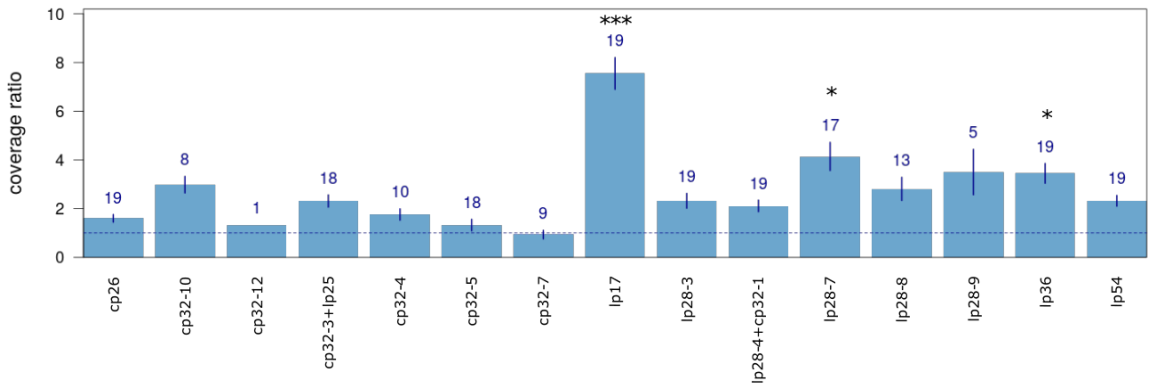


Supplementary Figure 2. Comparison of three assemblers for Illumina assembly of 25 B. bavariensis isolates. These violin plots compare N50 (a) and total length of contigs (b) obtained with QUAST v. 4.6 (110) on assemblies performed with SPAdes v. 3.10.1 (111), SOAPdenovo v. 1.0 (112) and VelvetOptimizer v. 1.0 (113).

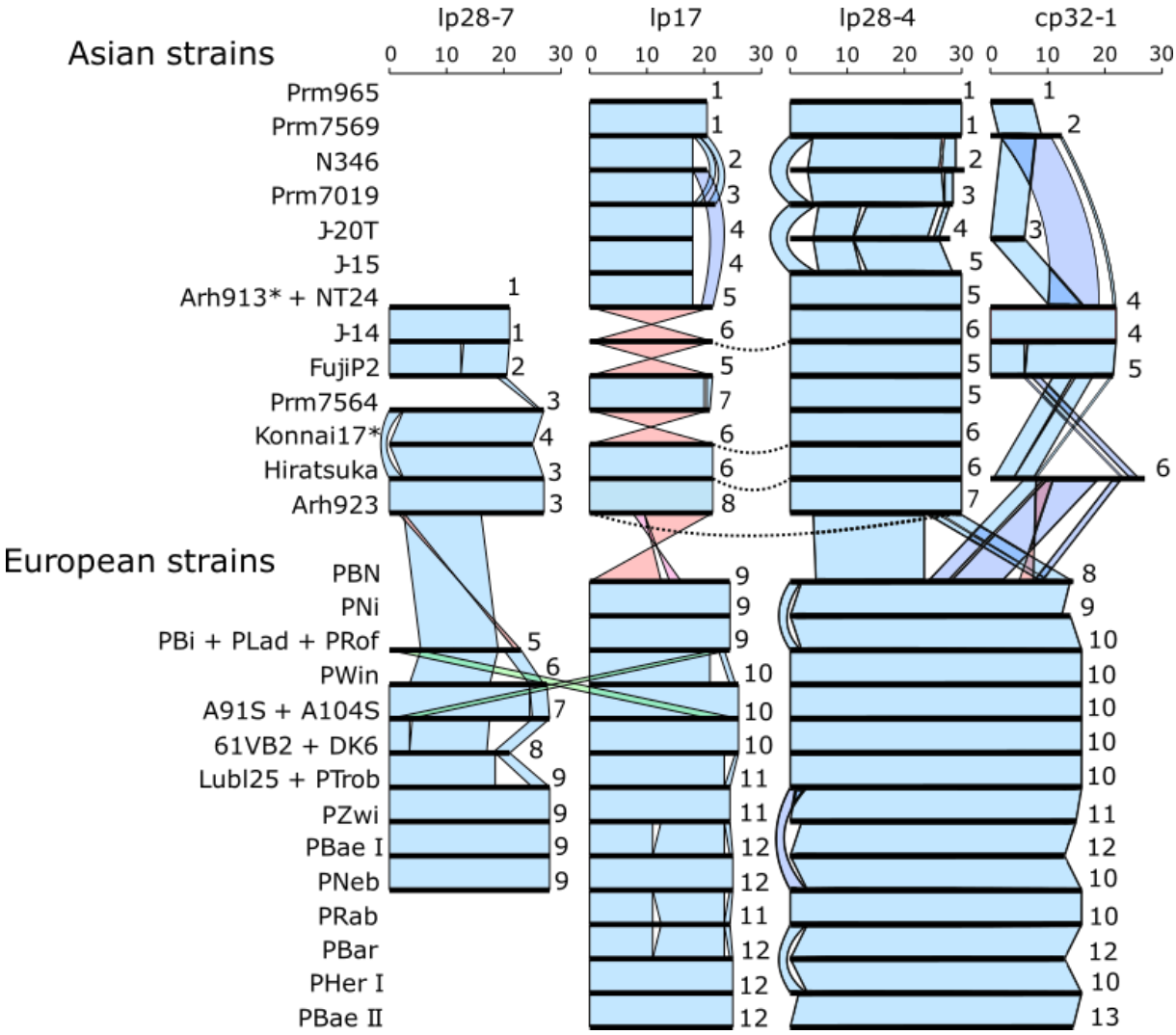
Supplementary Figure 3. Replicon assembly quality as a function of population, mapping method (a) and type of replicon (b). Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 (109) to the final reconstructed genomes and the relative standard deviation of the coverage of the raw reads was used as a measure of assembly quality. We compare here replicons from European (left bars) and Asian (right bars) genomes depending on (a) whether the replicon was made as one contig (pink) or as several contigs mapped to a reference (purple) and on (b) whether it was a chromosome (orange) or a plasmid (blue). Error bars show standard error of the mean. ***: Wilcoxon Rank Sum Test for Europe against Asia, P-value < 0.001. Other tests comparing mapping methods (a) and type of replicons (b) were not significant.



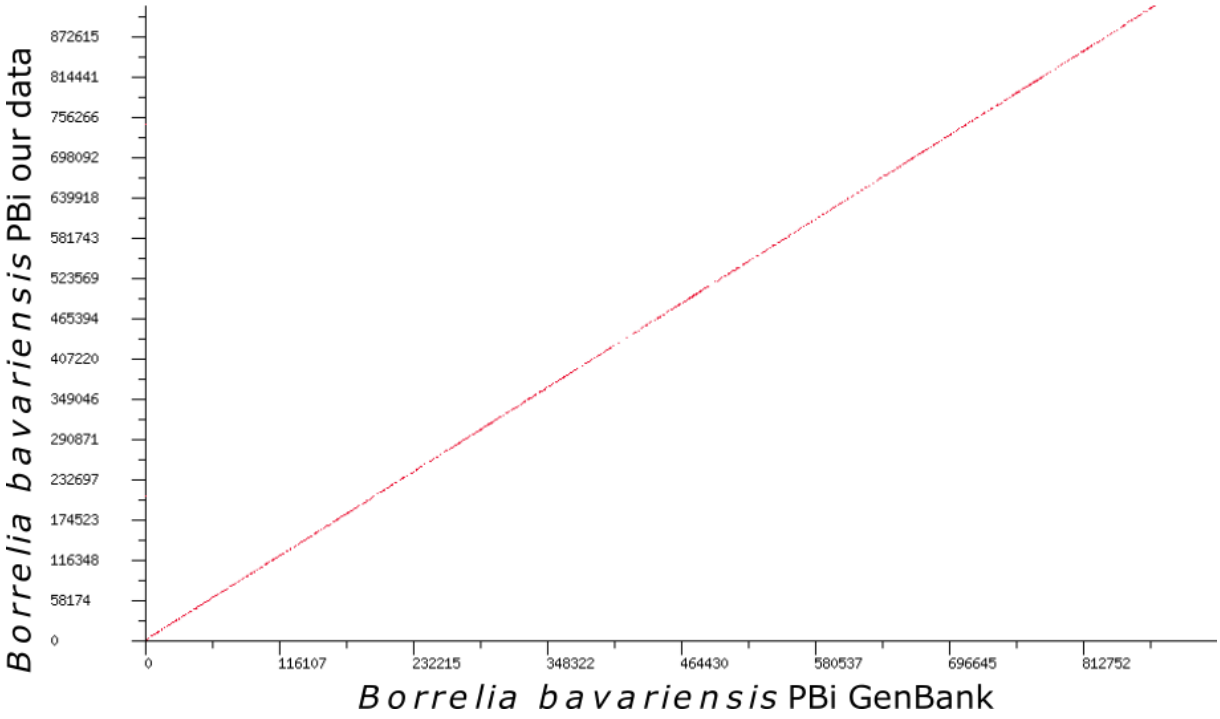
Supplementary Figure 4. Coverage ratio of European replicons as a proxy for copy number. Illumina raw reads were mapped with BWA-MEM algorithm v. 0.7.17-r1188 (109) to the final reconstructed genomes and the ratio of the coverage of each replicon with respect to the chromosome was computed in each European isolate. Error bars show standard error of the mean. Dark blue numbers indicate the number of plasmids of this type in the European sample. Wilcoxon Rank Sum Tests comparing coverage of each plasmid with that of the chromosomes: P-Value after Bonferroni-Holm correction *: < 0.05, ***: < 0.001, else: not significant.



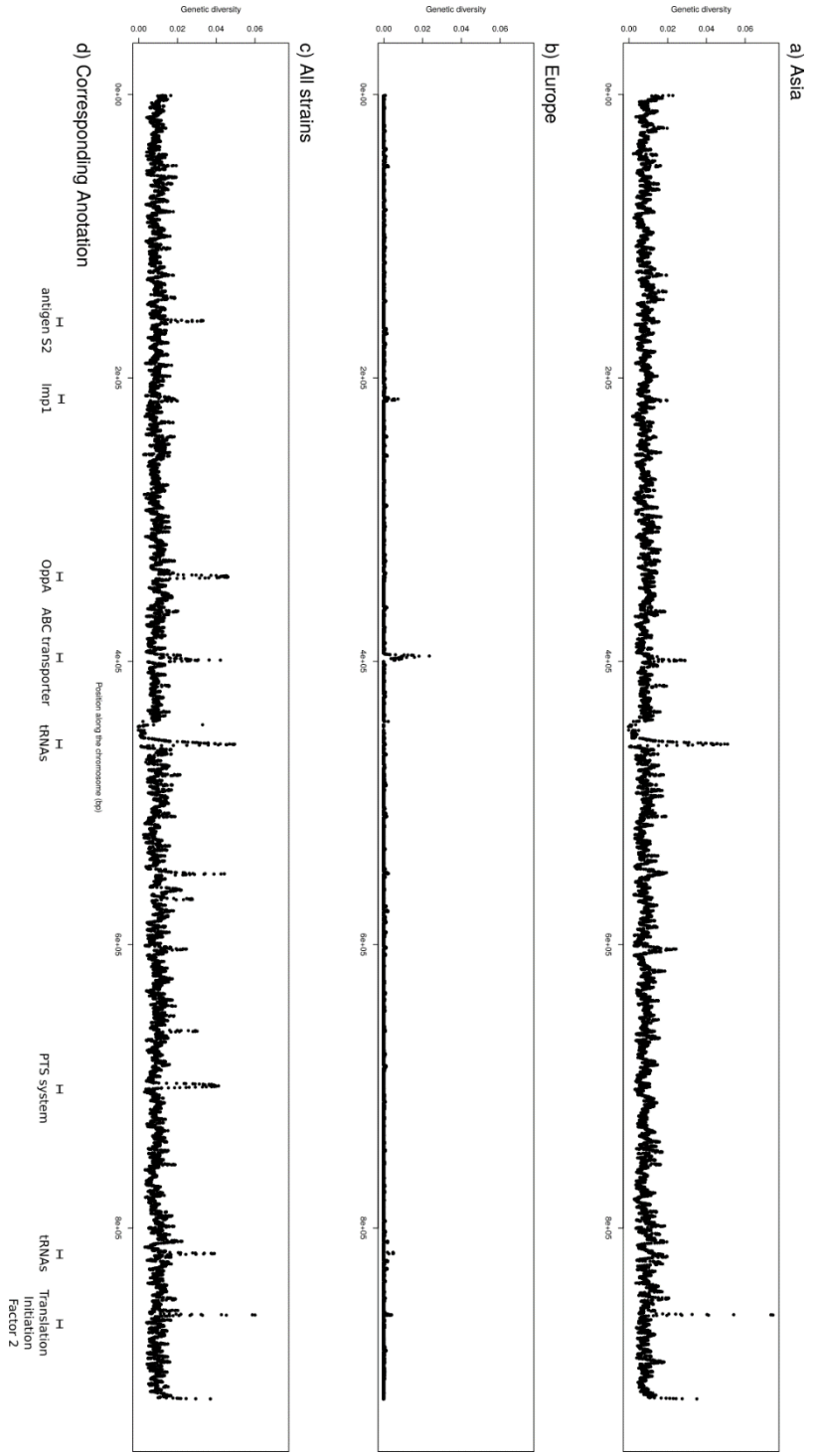
Supplementary Figure 5. Schematic representation of plasmid subtypes and fusion/relocation events on lp17, lp28–4, lp28–7 and cp32–1. The different plasmid subtypes (numbered arbitrarily) are represented as black bars. We defined as a new plasmid subtype, a plasmid sequence that had, with respect to the other plasmid subtypes, either presence of 400 bp or longer indels or obvious evidence of past interplasmid DNA exchanges (translocations). We used BLAST v. 2.8.1 (114, 115)to identify plasmid types and colour-shaded areas represent BLAST hits on the same strand (blue) and inversions (pink). Different shades of color are just used for clarity and have no meaning. Dashed lines represent plasmid fusions. Scale bars above the plots are plasmid lengths in kb. *: specific cases: Arh913 cp32–1 could not be assembled. Konnai17 had two lp28–7 plasmids, the second one has the same subtype as plasmid lp28–7 in FujiP2.



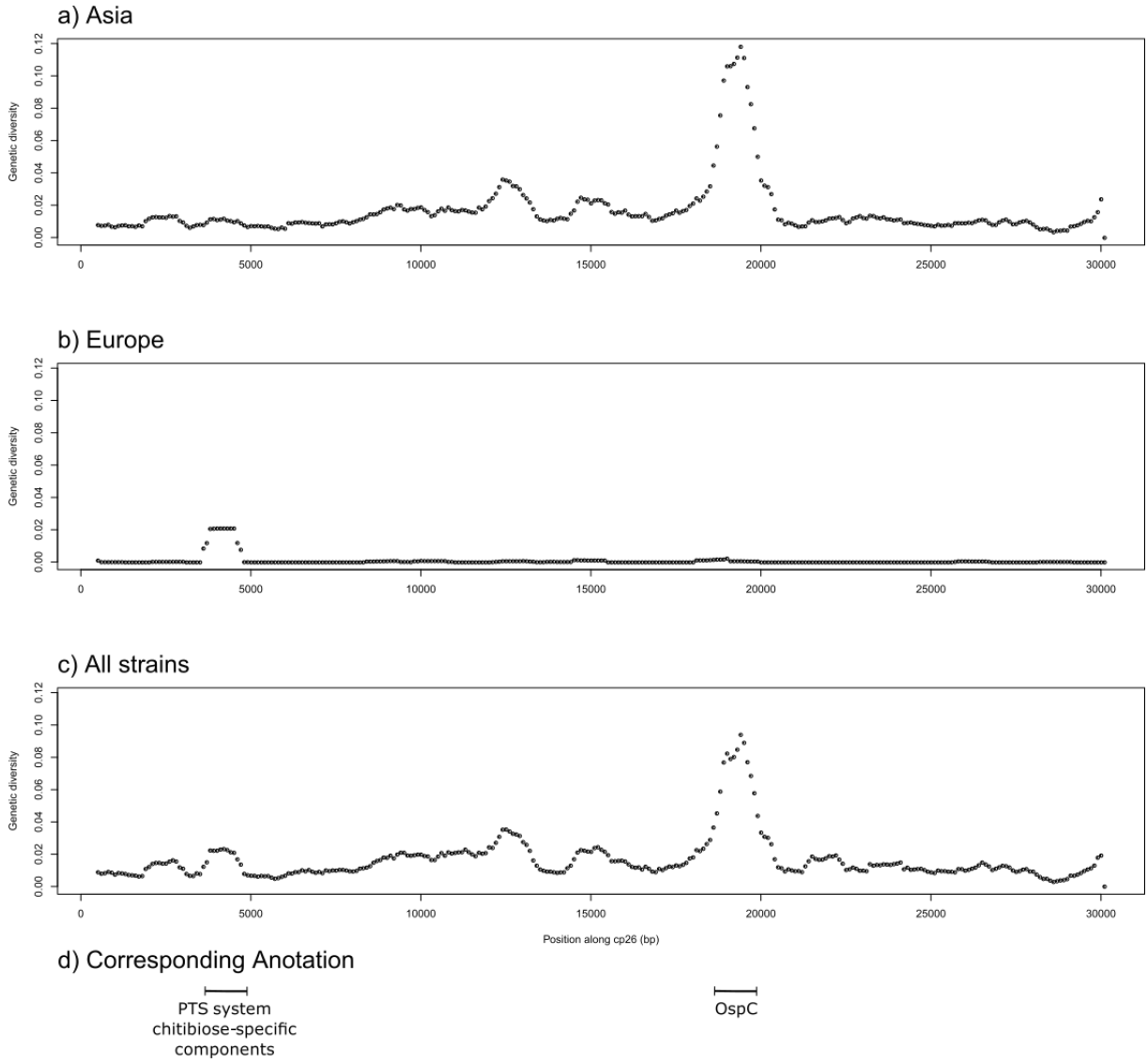
Supplementary Figure 6. Dotplot comparing annotation of strain PBi between our isolate and a previously published one. Comparison of gene content realized in RAST Annotation Server v. 2.0 (116, 117) a on the main chromosome. PBi accession number in RAST: 290434.1.



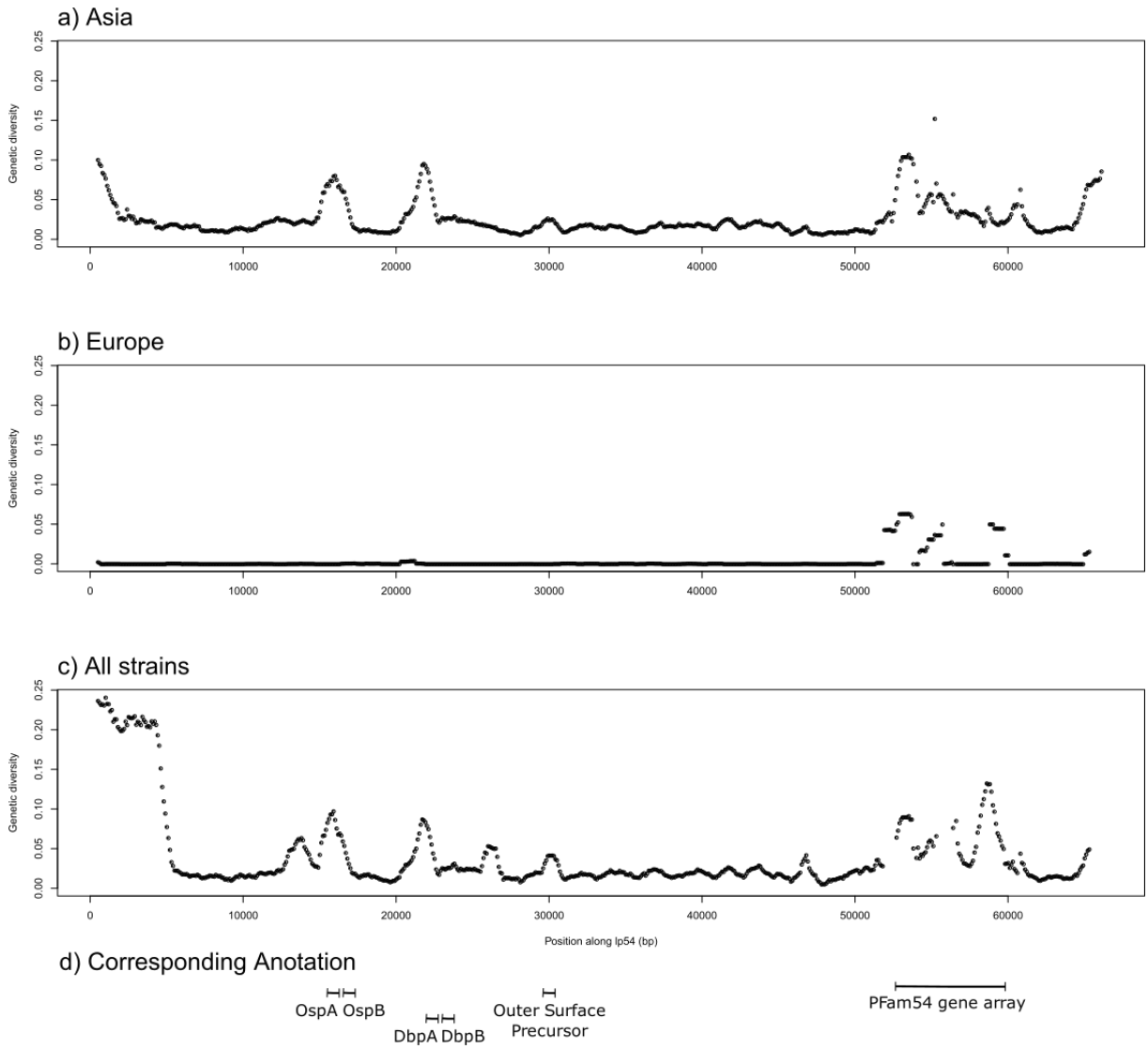
Supplementary Figure 7. Genetic diversity along the main chromosome of B. bavariensis. Genetic diversity was estimated using R package pegas v. 0.12 (118) on orthologous sequences aligned with MAFFT v7.407 (105, 119) on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 (116, 117).



Supplementary Figure 8. Genetic diversity along plasmid cp26 of B. bavariensis. Genetic diversity was estimated using R package pegas v. 0.12 (118) on orthologous sequences aligned with MAFFT v7.407 (105, 119) on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 (116, 117).



Supplementary Figure 9. Genetic diversity along plasmid lp54 of B. bavariensis. Genetic diversity was estimated using R package pegas v. 0.12 (118) on orthologous sequences aligned with MAFFT v7.407 (105, 119) on 1000 bp windows sliding every 100 bp in Asian isolates only (a), European isolates only (b) and all isolates (c). Genes located on diversity peaks (d) come from RAST Annotation Server v. 2.0 (116, 117).



Strain	cp26	lp54	lp36	lp17	lp17+lp28-4	lp25	lp38	lp56	lp28-2	lp28-3	lp28-4	lp28-4+cp32-1	lp28-6	lp28-7	lp28-8	lp28-9	cp32-1	cp32-3	cp32-3+lp25	cp32-4	cp32-5	cp32-6	cp32-7	cp32-9	cp/lp32-10	cp32-11	cp32-12
Arh913	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57	32/49/50			32/49/57	32/49/50/57			57	32/49/50		32/49			32/49/57		32/57	32/49/50/57	32	32/49/50/57	32/49/57	32/49/50
Arh923	32/49/57	32/49/50/57	32/49/50/57		32/49/50/57 *	32/49/50/57	32/49/50			32/49/50/57			57	32		32		32/49/50/57		32/49/50/57	32/49/57			32/49/50/57	32/49/50/57	32/49/57	
FujiP2	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57				32/49/57	32/49/50/57		57	57	32/49/50		32/49/50/57			32/49/57	32/49/57	32	32/49/50/57	32/49/50/57	32/49/50/57	32/49/57	32/49/50/57
Hiratsuka	32/49/57	32/49/50/57	32/49/50/57		32/49/50/57 *	32/49/50/57		32		32/49/50/57			57	32		32	32/49/50/57	32/49/50/57					32/49/50/57	32/49/50/57	32/49/50/57	32/49/57	
J-14	32/49/57	32/49/50/57	32/49/50/57		32/49/50/57 *	32/49/50/57				32/49/57			57	57	32/49/50		32/49/50/57			32/57		32/50/57	32/49/50/57		32/49/50/57	32/49/57	32/49/50/57
J-15	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57	32/49/50		32/57	32/49/50/57	32/49/50/57									32/49		32	32/49		32/49/50/57		
J-20T	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57					32/49/50/57		57			32	32/49/50/57	32/49/57		32/49/50/57	32/49/50/57		32/49/50/57		32/49/57	32/49/57	32/49/50/57
Konnai17	32/49/57	32/49/50/57	32/49/50/57		32/49/50/57 *		32/49						57	32 + 57 *			32/49/50/57			32/49/57		32/50/57	32/49	32/49/50/57	32/49/50/57	32/49/57	32/49/50/57
N346	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57				32/49/57					32/49/50			32/49/57			32/49/57	32/50/57	32/49/50/57		32/49/50/57	32/49/57	49/57
NT24	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57				32/49/57	32/49/50/57			57	32/49/50		32/49/50/57			32/49/57	32/49/57	32/50/57	32/49/50/57	32/49/50/57	32/49/50/57	32/49/57	32/49/50/57
Prm7019	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57	32/49/50			32/49/57	32/49/50/57				32/49/50						32/49/57		32/49/50	32/49/50/57	32/49/50/57	32	32/49/50/57
Prm7564	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57	32/49/50			32/49/50/57	32/49/50/57		57	32				32/49/50/57		32/49/50/57	32/49/57			32/49/50/57	32/49/50/57	32/49/57	
Prm7569	32/49/57	32/49/50/57	32/49/50/57	32		32/49/50/57	32/49/50			32/49/57	32/49/50/57			32	32/49/50		32/49/50/57			32/49/50/57	32/49/57	32	32/49/50	32/49/50/57	32/49/50/57		32/49/57
Prm965	32/49/57	32/49/50/57	32/49/50/57	32		32/49/57					32/49/50/57				32/49/50	32	32/49/50/57	32/49/50		32/49/50/57	32/49/57	32	32	32/49	57		32/49/57
61VB2	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57		32/49/57		32/49/50/57				
A104S	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57 *	32/49	32/49		32/49/50/57		32/49/50/57		
A91S	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57		32/49/50/57				
DK6	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32 50				32/49/57		32/49/57		32/49/50/57				
Lubl25	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57				32/49/50/57		
PBae I	32/49/57	32/49/50/57	32/49/50/57	32						32/49/50/57		32/49/50/57		32					32/49/57	32/49/50/57	32/49/57						
PBae II	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/50/57 *	32/49/50/57	32/49						
PBar	32/49/57	32/49/50/57	32/49/50/57	32						49/57 *		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57				32/49/50/57		
PBi	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	50	32			32/49/57		32/49/57		32/49/50/57				
PBN	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57				32			32/49/57		32/49/57		32/49/50/57				32/50/57
PHer I	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57				32/49/50/57		
PLad	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50	32			32/49/57		32/49/57		32/49/50/57				
PNeb	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57				32/49/50/57		
PNi	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57				32			32/49/57		32/49/57		32/49/50/57				
PRab	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57						
PRof	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50	32			32/49/57		32/49/57		32/49/50/57				
PTrob	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/57	32/49/50/57	32/49/57				32/49/50/57		
PWin	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32	32/49/50				32/49/50/57 *	32/49/50/57	32/49/57				32/49/50/57		
PZwi	32/49/57	32/49/50/57	32/49/50/57	32						32/49/57		32/49/50/57		32					32/49/57 *	32/49/57	32/49				32/49/50/57		

Supplementary Table 1. Plasmid partition genes identified in 33 *B. bavariensis* strains

Plasmid partition genes were searched for using BLAST v. 2.8.1 on assembled plasmids and on Illumina contigs. The queries were sequences from *B. burgdorferi* sensu stricto strains B31, BOL26, JD1, 118a and MM1 and *B. afzelii* PKo for PFam protein families 32, 49, 50 and 57. All hits on the *B. bavariensis* genomes were then subsequently used as query for a second BLAST run. Final hits that were shorter than half the original gene sequence were removed. Gray cells represent a reconstructed plasmid of at least 5 kb length (shown in Table 1).

* specific cases:

lp17l+p28-4 fusion: lp17 part has only one PFam32 hit whereas the lp28-4 part has all four PFam;

lp28-3 PBar: the PFam32 hit was present but not full as stretched over two contigs;

lp28-7 Konnai17: Konnai17 has two lp28-7 plasmids, one carrying PFam32 and one PFam57;

cp32-3lp25: one European strain had no PFam49 for cp32-3 (A104S) and three others had two lp25 hits for PFam50 (PBae II, PWin, PZwi).

Target	Sense	Sequence (5’-3’)	Size (bp)	Tm (°C)	Product size (bp)
chromosome	F	gaaagcaaggcaacaaggg	19	57	133
	R	tgccctttgagcttacagaag	21	59	
cp26	F	tcacaccagaaagtgcaagc	20	58	130
	R	cctcccattacgctcatttg	20	58	
lp17	F	gaagggctaacggcactcaa	20	60	132
	R	agatgtgaaggaggagcatca	21	59	
lp36	F	cgggtgcattagagcaggat	20	60	127
	R	accaatagcaccccacgttt	20	58	

Supplementary Table 2. Primers used for qPCR

Paper 5: Out of Asia? Recurrent vector switches leading to the expansion of Eurasian Lyme borreliosis bacteria

Robert E. Rollins, Kozue Sato, Minoru Nakao, Mohammed T. Tawfeeq, Fernanda Herrera-Mesías, Ricardo J. Pereira, Sergey Kovalev, Gabriele Margos, Volker Fingerle, Hiroki Kawabata, and Noémie S. Becker

Unpublished Manuscript

Out of Asia? Vector switches leading to the expansion of Eurasian Lyme disease bacteria

Robert E. Rollins, Kozue Sato, Minoru Nakao, Mohammed T. Tawfeeq, Fernanda Herrera-Mesías, Ricardo J. Pereira, Sergey Kovalev, Gabriele Margos, Volker Fingerle, Hiroki Kawabata, and Noémie S. Becker

Abstract

Vector-borne pathogens exist in obligate transmission cycles between vector and reservoir host species. Host shifts can lead to geographic expansion and the emergence of new diseases. Three etiological agents of human Lyme borreliosis (*Borrelia afzelii*, *Borrelia bavariensis*, and *Borrelia garinii*) predominantly utilize two distinct tick species as vectors in Asia (*Ixodes persulcatus*) and Europe (*Ixodes ricinus)* but how and in which order they colonized each continent remains unknown. Here, by reconstructing the evolutionary history of 142 Eurasian isolates, we show that all three *Borrelia* genospecies evolved from an Asian origin, suggesting that successful expansion into Europe resulted through invading a novel vector. The pattern of gene flow between continents is different between genospecies and most likely conditioned by reservoir host association and their dispersal. Our results highlight that Eurasian Lyme borreliosis agents are all capable of geographic expansion through vector shifts, but potentially differ in their capacity as emergent pathogens.

Introduction

Lyme borreliosis (LB), also termed Lyme disease, is the most common vector-borne disease in the Northern hemisphere (30, 120), caused by certain genospecies of *Borrelia* bacteria (30, 39, 43). These spirochete bacteria are maintained naturally in obligatory transmission cycles between tick vectors and specific vertebrate reservoir hosts (43). In North America human LB is predominantly caused by *Borrelia burgdorferi* sensu stricto (*Bb*ss) while three additional genospecies act as causative agents across Eurasia (*Borrelia afzelii*, *Borrelia bavariensis*, *Borrelia garinii*) (30, 39, 43). Genomic analyses have already shown a complex ancestral spread of *Bb*ss across North America (121) which is also observed in specific Eurasian genospecies (84, 91). The Eurasian genospecies offer a unique opportunity to understand the geographic expansion of *Borrelia* spirochetes using comparative genomics. However, no study has

integrated genomic data from the different genospecies. In particular, no study at the population-level of Asian *B. afzelii* has been published to date.

Borrelia genospecies cannot transmit successfully through all tick species (39, 122) and can only infect specific vertebrate classes (i.e. rodents, passerines, sea-birds, etc.) while being easily cleared by the immune systems of others (39, 43, 48). Eurasian *Borrelia* genospecies currently exist in separate transmission cycles vectored predominately by two generalist tick species in Asia (*Ixodes persulcatus*) and Europe (*Ixodes ricinus*) (43) (Figure 2). This suggests that each genospecies successfully invaded a novel tick vector resulting in the expansion into a new continental transmission cycle. However, how and in which order this expansion occurred is still unknown (Figure 2). For *B. bavariensis,* an Asian origin was already hypothesized as the Asian population displays a higher genetic diversity compared to the almost clonal European population (39, 84, 93). European *B. bavariensis* is thought to have undergone a selective bottleneck while colonizing the European tick vector, *I. ricinus*, resulting in the observed clonal structure (39, 84, 93). Whether or not the other genospecies also underwent this bottleneck has never been studied so far. Both *B. afzelii* and *B. bavariensis* utilize rodents as reservoir hosts (43, 45, 94) (Figure 2). In comparison, *B. garinii* is adapted to avian host species (45, 98), which includes interconnected terrestrial and marine transmission cycles (Figure 2). This association in *B. garinii* is thought to allow for migration between the European and Asian populations which is not accessible to rodent adapted genospecies. (Figure 2).

Each of these genospecies has successfully established into multiple transmission cycles and offers an opportunity to study how *Borrelia* expanded across Eurasia through comparative genomics. Although, no study to date has integrated genomic data from all three Eurasian-distributed genospecies. Here we report the reconstructed evolutionary history of 142 *B. afzelii*, *B. garinii*, and *B. bavariensis* Eurasian isolates based on full genome sequences including the first Japanese *B. afzelii* genomes sequenced. Our results highlight that these genospecies share an Asian origin with support for migration from an ancestral Asian population vectored by *I. persulcatus* into a novel European vector, *I. ricinus*. Post-colonization gene flow appears to be associated with the dispersal range of the respective reservoir host species. Our results provide new information on the ability of three *Borrelia* genospecies to colonize new environments and how this could relate to the further expansion of human LB.

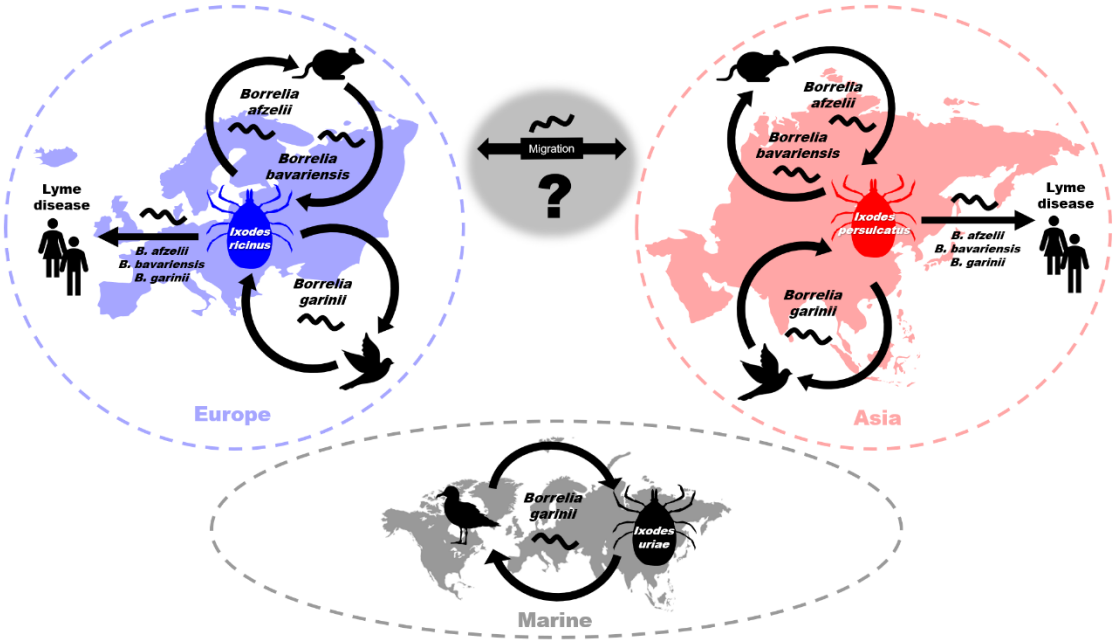


Figure 2. Schematic overview of the transmission cycles of *B. afzelii*, *B. bavariensis*, and *B. garinii* across Eurasia. These three *Borrelia* genospecies are maintained predominately by the tick vector *I. ricinus* in Europe and *I. persulcatus* in Asia in a transmission cycle utilizing either rodents (*B. afzelii* and *B. bavariensis*) or birds (*B. garinii*) as reservoir hosts (39, 43, 123). *Borrelia garinii* specifically utilizes interconnected terrestrial and marine based transmission cycles (45, 98, 124). In marine systems, this species is maintained by seabird reservoir host species and the vector *I. uriae* (45). In both Europe and Asia, all three genospecies can be transmitted to humans through *I. ricinus* or *I. persulcatus* and can manifest as Lyme disease (30, 43).

Materials & Methods

Isolates used and sequencing

For all information on isolates, including origin and source material refer to Table S1. This study utilized DNA of 136 *Borrelia* isolates coming from three human pathogenic species: *B. afzelii* (n=33), *B. garinii* (n=57), and *B. bavariensis* (n =46). Of these, 52 are novel *Borrelia* isolated from ticks collected either in Japan (n=43) or Germany (n=9) (see Text S1). Additionally, 55 European isolates (*B. afzelii,* n=11; *B. garinii*, n=25; and *B. bavariensis*, n=19) were provided by the German National Reference Center for *Borrelia* at the Bavarian Food and Health Safety Authority. All isolates, expect one tick isolate, were isolated from humans. DNA for additional Japanese human and tick isolates (n = 12) was provided by the National Institute of Infectious Disease in Tokyo, Japan. Finally, previously sequenced Russian *B. bavariensis* (n=7) (84) and DNA from 12 additional Russian *B. garinii* tick isolates were included in the study (see Text S1).

Borrelia isolates were cultured either in inhouse-made MKP (125) (all European isolates) or inhouse-made BSK-H (126) (all Russian and Japanese isolates) medium according to standard procedures (125, 126) until the cultures reached a density of at least 10^8 cells per mL at which point whole genomic DNA was extracted. Genomic DNA from all European isolates was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany) and from all Japanese and Russian isolates using the Wizard® Genomic DNA Purification Kit (Promega, Germany). DNA quality (260/280) and concentration were measured using a NanoDrop® 1000 photometer (Thermo Fisher Scientfic, USA) and a Qubit® 3.0 fluorometer (Thermo Fisher Scientific, USA), respectively.

For all samples, libraries were produced according to the Nextera XT sample preparation guide (Illumina, San Diego, CA, USA). Library quality was checked using an Agilent TapeStation 2200 (Agilent, Germany) before being sequenced using an Ilumina MiSeq platform according to standard protocol (Illumina, San Diego, CA, USA) that produced paired end reads of 250bp.

Chromosome assembly and phylogeny reconstruction

Illumina reads were first trimmed for Illumina MiSeq adapter sequences using Trimmomatic v. 0.38 (127) before being assembled using SPAdes v. 3.13.0 (111), which has been shown to be the best option for *de novo* assemblies of *Borrelia* genomes (84). Pacific Bioscience sequences were obtained for three *B. bavariensis* isolates (PBi, A104S, and NT24) (84) and three *B. garinii* isolates (PHeI, PBr, and NT31; see Suppl. Met.). Additionally, three *B. afzelii* chromosomes were downloaded from GenBank for use as references and inclusion in all analyses: PKo (CP009058.1), K78 (CP002933.1), and ACA-1 (NZ_ABCU00000000.2). SPAdes contigs were then mapped to reference chromosomes using NUCmer v. 3.23 from the package MUMmer (128, 129). Final chromosomes were produced according to the mapping protocol outlined in Becker et al. (2020) (see Suppl. Met.). Three additional *B. bavariensis* chromosomes were downloaded from GenBank and used in further analyses: SZ (CP007564.1), BgVir (CP003151.1), and NWJW1 (CP003866.1).

Final assembled chromosomes were aligned using MAFFT v. 7.407 (105, 119). Recombination is known to be low on the *Borrelia* chromosome (93) but as recombinant regions could bias the phylogenetic signal, we searched for areas of the chromosome violating the four-gamete condition (130) (as described in Gatzmann et al. (2015); see Suppl. Met.). Regions with strong violation of the four-gamete condition

were assumed to be recombinant and were removed from the final alignments. Phylogeny reconstruction was done in MrBayes v. 3.2.6 (131, 132) with ploidy set to haploid and a GTR (133) substitution model with gamma distributed rate variation. Three independent runs were launched and ran for 5 million generations at which point convergence of parameters was checked with Tracer v. 1.7.1 (134). Consensus trees were built using the *sumt* command from MrBayes using a respective burn-in of 25%. Convergence to a single topology in all three independent runs was checked manually in FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) which was also used to plot the tree shown in Figure 3. Trees were midpoint rooted on the longest branch, which corresponded to the well-established delineation between *B. afzelii* and the monophyletic group containing *B. bavariensis* and *B. garinii* (101, 123).

Identification of plasmid content through plasmid partitioning genes

Plasmid content was approximated by the number of plasmid partitioning genes present in each assembly, which have been shown to be unique to specific plasmid types and exist as single copies in *Borrelia* (82, 85, 135). Identification of plasmid partitioning genes was performed as outlined in Becker et al. (2020) (see Suppl. Met.). Briefly, we used BLAST v.2.8.1 (114, 115) (algorithm: *blastn*) to search for the presence of plasmid partitioning genes of the PFam32, 49, 50, and 57.62 families in the assembled SPAdes contigs. Hits were removed if they did not cover more than half the length of the references and had lower than 80% percent identity. After curation, we defined a plasmid being present if at least one of the partitioning genes was present in the assembled contigs.

Statistical and population genetic analyses

All statistical analysis was performed in R v. 3.6.1 (136). Genetic diversity (π) (137) and Tajima's *D* test statistic (138) were estimated in the package *pegas* (118). Analysis of molecular variance (AMOVA) (139) was performed using the package *poppr* (140) whereas F_{ST} (137) and D_{XY} (141) were estimated with the package *PopGenome* (142).

Standard two-side, unpaired t-tests were run on plasmid number between genospecies comparing the two geographic populations using the function *t.test* from the base R package (136). Classical multidimensional scaling (MDS) was run using the *cmdscale* function using the base R package on a distance matrix calculated from the

binary presence/absence plasmid data per isolate. Further effects on plasmid content were tested using a generalized linear mixed effects model assuming a Poisson error distribution using the *glmer* function from the package *lme4* (143). Fixed effects were included for sample origin (Asia vs. Europe) and source (human vs. tick isolate) and genospecies was fitted as a random effect. Mean estimates and their 95% credible intervals were estimated based on 5000 simulations using the *sim* function from the package *arm* (144). Residual error was calculated according to Nakagawa & Schielzeth (2010).

Results

For phylogenetic and population genetics analyses, we focused on the linear chromosome as it is a core genomic compartment present in all *Borrelia* and is generally used to reconstruct the evolutionary history between genospecies (84, 85, 87). *Borrelia* genomes are highly fragmented and can contain over 20 unique linear and circular plasmids (82, 85) which can be highly plastic even within a single genospecies (84, 86, 87). *Borrelia* plasmids contain genes related to host and vector adaptation and absence of certain plasmid types have been linked to reduced infectivity (66, 86, 146–150). Therefore, their presence could influence the evolution of these bacteria and should be considered. As each plasmid carries specific, partitioning genes (categorized to PFam32, 49, 50, 57/62) and generally exists in single copies per cell (82, 85, 135), we were able to approximate plasmid content in each isolate by searching for these partitioning genes using BLAST v.2.8.1 (114, 115) (algorithm: *blastn*)

In total 142 full chromosome sequences were used for further population genetics analysis, of which 136 were assembled *de novo* from Illumina MiSeq data. Plasmid content could only be estimated for the 136 samples for which raw MiSeq data was available. For full isolate information see Table S1.

All genospecies display a probable Asian origin

For both *B. afzelii* and *B. garinii* the oldest node separates a clade containing only Asian isolates from a clade containing isolates from both continents (Figure 3A,C) suggesting that the Asian population is ancestral for both genospecies. *Borrelia bavariensis* displays a deep branching between the two continents, and European isolates are characterized by a low divergence and almost clonal expansion as previously described (84, 93) (Figure 3B). Our original analysis did not include Russian

B. afzelii isolates. A single Russian *B. afzelii* isolate exists with a full chromosome in GenBank, Tom3107 (Accession Number: NZ_CP009212.1). We re-ran the phylogeny utilizing all *B. afzelii* chromosome sequences including Tom3107 and PBi as an outgroup to root the tree (see Suppl. Met.). Tom3107 was basal to the monophyletic European *B. afzelii* clade (Figure S1) suggesting a stepwise colonization from far-east Asia through Russia into Europe, which was not observed in the other two genospecies.

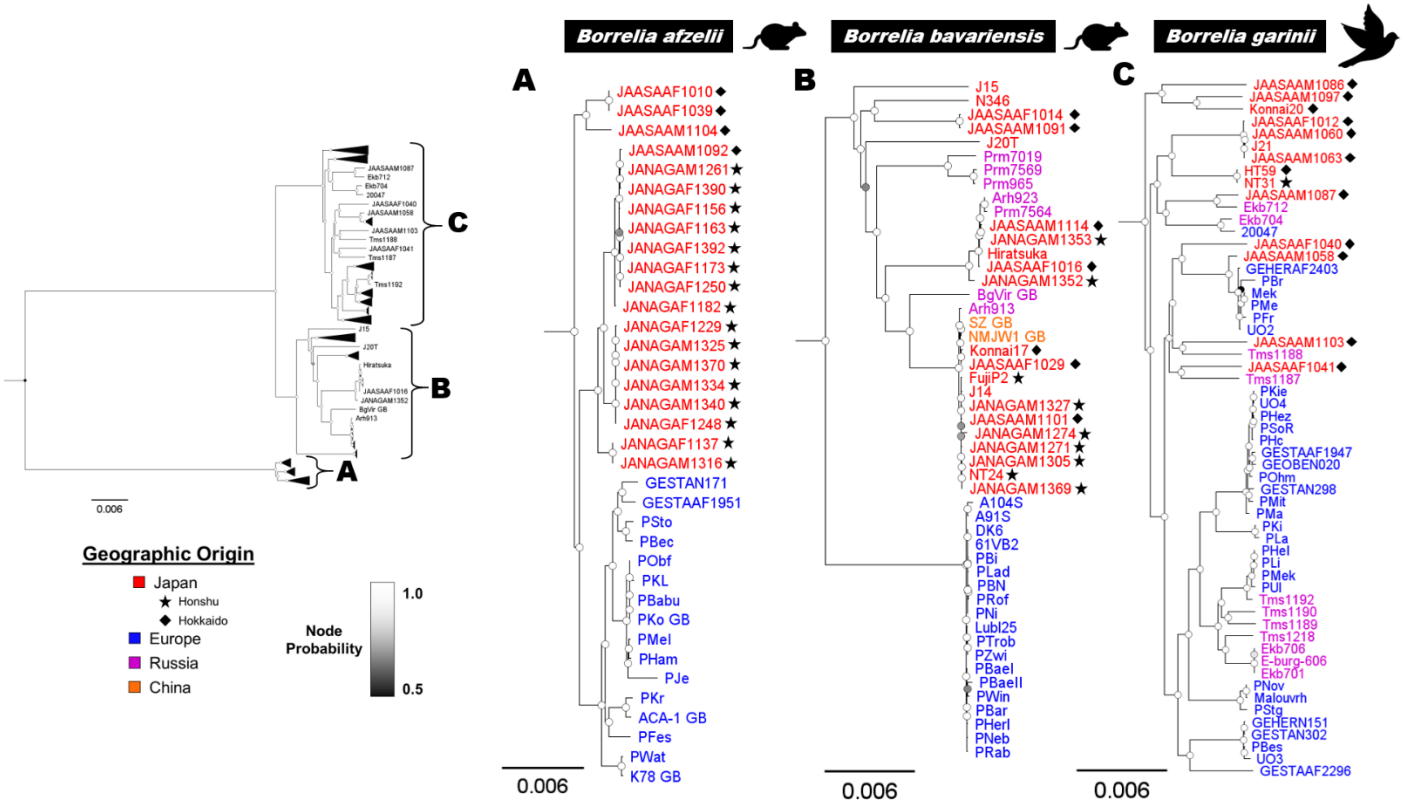


Figure 3. Phylogeny of *B. afzelii*, *B. bavariensis*, and *B. garinii* based on the main chromosome corrected for recombining regions (see Suppl. Met.). The phylogeny was reconstructed with MrBayes v. 3.2.6 (131, 132) with ploidy set to haploid and a GTR (133) substitution model with gamma distributed rate variation. Three independent runs were launched and ran for 5 million generations each at which point convergence of parameters was checked with Tracer v. 1.7.1 (134). Consensus trees were built using the *sumt* command from MrBayes using a respective burn-in of 25%. The collapsed tree displays the full phylogeny (where monophyletic groups are collapsed if all isolates come from the same geographic origin) and then the expanded tree is shown independently for *B. afzelii* (A), *B. bavariensis* (B), and *B. garinii* (C). Colors correspond to geographic origin of the isolates: Europe (blue), Japan (red), purple (Russia), orange (China). For Japanese tick isolates, the island of origin is shown either as a diamond (Hokkaido) or star (Honshu) when known. The scale bar is in substitutions per site.

Higher genetic diversity (π (137)) was found in Asian *B. bavariensis* and *B. garinii* in comparison to their European counterparts (Table 1). Genetic diversity was similar between Asian and European *B. afzelii* isolates (Table 1). In all cases, the *Borrelia* populations showed negative Tajima`s *D* (138) values (Table 1) as expected

for bacteria due to the influence of population expansion (93, 138). The European samples always showed more negative values (Table 2), suggesting a more recent expansion into Europe. *Borrelia bavariensis* displayed the largest difference in Tajima's *D* and also had the largest absolute divergence value (D_{xy} (141)) in comparison to the other two genospecies hinting that *B. bavariensis* branching is potentially the oldest and that *B. afzelii* is the youngest with the lowest absolute divergence and difference in Tajima's *D* (Table 1).

Table 1. Population genetics statistics for full population samples of *B. afzelii*, *B. bavariensis*, and *B. garinii*. The Asian populations for *B. garinii* and *B. bavariensis* contain all Russian samples. These calculations include non-randomly sampled isolates (both tick and human), but values calculated for randomly sampled isolates showed similar statistics (see Text S2 & Table S2)

Genospecies	Population	n	π	Tajima's *D*	F_{ST}	D_{XY}
Borrelia afzelii	Asian	20	0.00193	-3.932	0.570	0.00379
	European	16	0.00217	-4.193		
Borrelia bavariensis	Asian	30	0.00784	-2.616	0.744	0.0141
	European	19	0.000170	-4.138		
Borrelia garinii	Asian	25	0.00900	-2.302	0.130	0.00694
	European	32	0.00619	-3.353		

Our dataset, as with many others, includes non-randomly sampled isolates which could lead to biased estimates of population level statistics (137). As our data set includes randomly sampled isolates as well (see Text S1) we were able to test for potential sampling biases. Interestingly, we did not observe strong bias in any of these statistics (π, F_{ST}, D_{XY}, Tajima's *D*) when calculated on datasets containing random and non-random samples (Text S2).

Each genospecies display unique structuring

Borrelia bavariensis displayed the strongest geographic structuring between the European and Asian samples (F_{ST} (137) = 0.744; $AMOVA_{continent}$ (139) = 69.7% of molecular variance (σ)) followed by *B. afzelii* (F_{ST} = 0.570; $AMOVA_{continent}$ = 40.2% of σ) (Table 1 & 2). Regions (defined as country or sampling locality if known) within continents further explained variation in *B. afzelii* samples ($AMOVA_{Region}$ = 23.6% of σ; Table 2) and structuring was observed between randomly sampled *B. afzelii* isolates

from the islands of Hokkaido (ASA) and Honshu (NAG) (F_{ST} = 0.379; Table S2). Honshu and Hokkaido *B. afzelii* isolates do form two reciprocally monophyletic clades, with the notable exception of one Hokkaido isolate belonging to the Honshu clade (Figure 3A) suggesting some level of migration. Of interest however, this trend was not observed for *B. bavariensis* ($AMOVA_{Region}$ = 0.99% of σ; Table 2) and, indeed, randomly sampled *B. bavariensis* isolates from the islands of Hokkaido and Honshu did not show geographic structuring (F_{ST} = 0.057; Table S2) even though both *B. bavariensis* and *B. afzelii* are rodent adapted (39, 43, 123). Furthermore, Asian *B. bavariensis* displayed a low divergence clade containing samples from Japan (including isolates from distinct islands), China, and Russia (Figure 3B) suggestive of relatively high migration between Asian regions. Less geographic structuring by continent was observed in *B. garinii* (F_{ST} = 0.13; $AMOVA_{Continent}$ = 8.7% of σ; Table 1 & 2) as expected as *B. garinii* displayed little geographic structuring throughout the phylogeny with mixing of samples from different geographic origins (Figure 3C).

Table 2. Hierarchical AMOVA(139) of *B. afzelii*, *B. bavariensis*, and *B. afzelii* populations coming from Europe and Asia describing the percentage of genetic variation (σ) attributed to each hierarchical level. Regions within continent (Europe, Asia) are defined as country or sampling locality if known. The Asian populations for *B. garinii* and *B. bavariensis* contain all Russian samples.

Genospecies	Level	σ (%)
Borrelia afzelii	Between continents	40.223
	Regions within continent	23.612
	Within samples	36.165
Borrelia bavariensis	Between continents	69.654
	Regions within continent	0.988
	Within samples	29.358
Borrelia garinii	Between continents	8.749
	Regions within continent	9.272
	Within samples	81.979

Plasmid content is generally homogenous between genospecies

Borrelia afzelii and *B. bavariensis* both differed significantly in plasmid numbers between Europe and Asia, with European *B. afzelii* (two-sided unpaired t-test, p = 0.03) and Asian *B. bavariensis* ($p < 0.001$) having significantly more plasmids in comparison to the other geographic population (Figure 4A). *Borrelia garinii* isolates

did not differ in overall plasmid number between Asia and Europe (p = 0.08) but had significantly fewer plasmids in comparison to both *B. afzelii* populations (Asian, p = 0.003; European, p < 0.001) and to Asian *B. bavariensis* (p < 0.001) (Figure 4A). When we look at the absolute plasmid number for a population, defined as the number of unique plasmid types present in at least one isolate from the population, only European *B. bavariensis* showed a lower absolute plasmid number (black circle; Figure 4A) in comparison to the other species such as *B. garinii*. *Borrelia garinii* also had on average lower plasmid numbers per isolate (comparable to European *B. bavariensis*), but the absolute number of plasmid types present in the population (i.e. diversity of plasmid types) is comparable to *B. afzelii* and Asian *B. bavariensis* (Figure 4A).

Based on the plasmid presence/absence matrix for all samples (n = 136), we further ran a multi-dimensional scaling (MDS) analysis to test if plasmid content corresponds to factors such as continent (i.e. vector) or genospecies. Plasmid content appears more homogenous between Asian isolates (Figure 4B) versus European isolates, which display clusters based on genospecies (Figure 4B). This could result from European isolates representing a subset of available plasmid combinations which are all present in the Asian populations. Even so, no plasmid types were more frequently associated with factors such as genospecies or geography (Figure S2). It is important however to note, as *Borrelia* can lose plasmids due to long-term culturing (83). Many human isolates (Table S1) have been potentially kept in culture longer suggesting that the current plasmid results could be biased even though sequencing focused on low passage isolates (<10 passages). Tick isolates did indeed have on average higher plasmid content (Table S3; mean: 1.19; 95% CI: 0.16, 2.22) suggesting that this bias may be present for our human isolates.

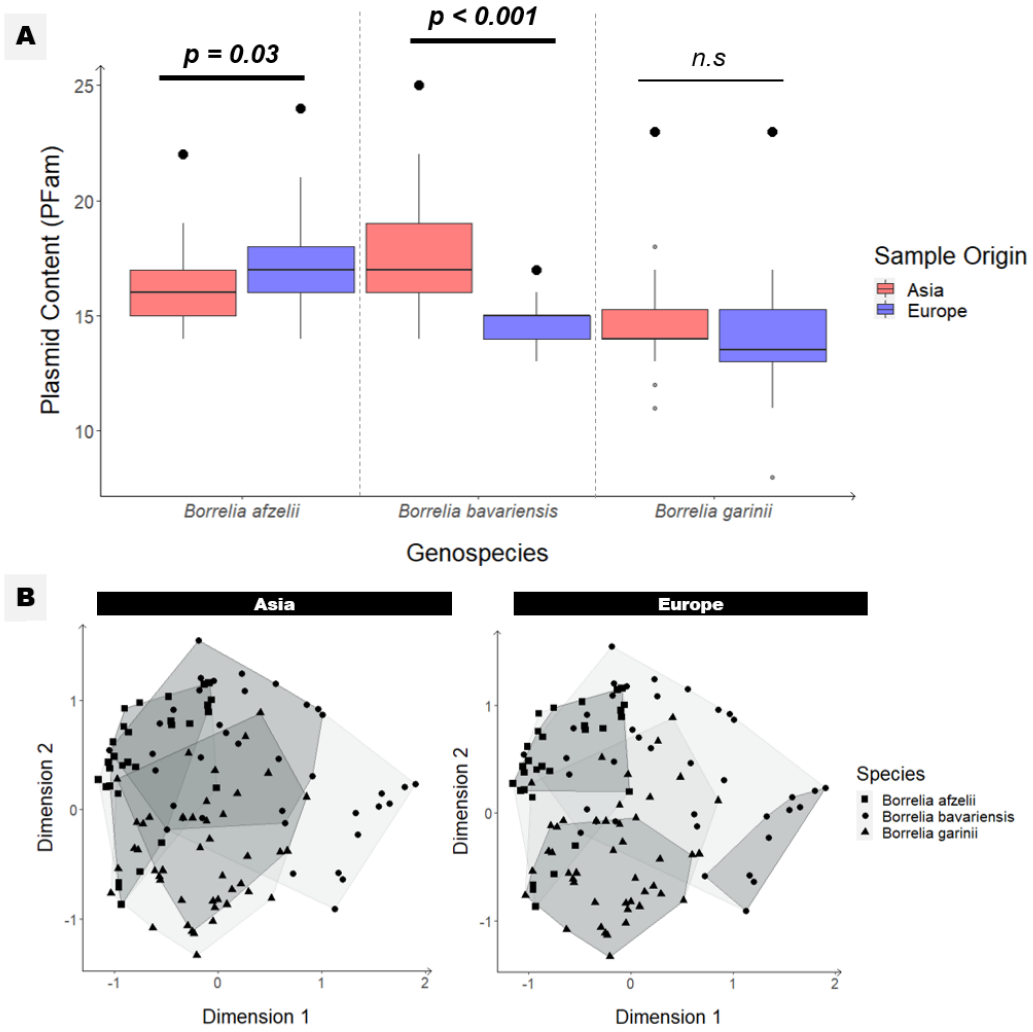


Figure 4. Analysis of plasmid content for sequenced strains estimated by the unique number of plasmid partitioning genes (PFam32, 49, 50, and 57.62) present in the assembled contigs. A plasmid was considered present if at least one of the partitioning genes was present. A) Boxplot of all plasmids present in isolates from Asia or Europe. The black circles represent the absolute number of unique plasmid types found in the geographic population defined as the plasmid type being observed in at least one isolate. P-values refer to an unpaired, two-sided t-test run on plasmid number between the European and Asian populations of each species individually. B) MDS analysis on plasmid presence/absence matrix for all samples. This figure shows the same MDS twice with emphasis on Asia (left) and Europe (right) by outlining isolates from Asia or Europe in a dark grey. Shapes correspond to genospecies: *B. afzelii* (square), *B. bavariensis* (circle), *B. garinii* (triangle).

Discussion

The expansion of vector-borne pathogens is inherently linked to their ability to infect and transmit through reservoir host and vector populations. This fact can be observed as the current major etiological agents of human Lyme borreliosis (LB) in Eurasia (*B. afzelii*, *B. bavariensis*, and *B. garinii*) are vectored mainly by two different tick species: *I. persulcatus* in Asia, and *I. ricinus* in Europe. This means that each of these genospecies has, at least once during its evolution, successfully invaded a novel

tick species and consequently a local population of vertebrate hosts. Yet how this invasion occurred and in which order is currently not known due to a lack of data. Here we report a reconstructed phylogeny of 142 Eurasian isolates belonging to the genospecies *B. afzelii*, *B. bavariensis*, and *B. garinii*. All three genospecies appear to share an Asian origin, suggesting a repeated expansion into Europe in relation to successfully invading a novel tick vector, *I. ricinus*. However, all three genospecies display unique sub-structuring which could be linked to ecological variability in their specific reservoir hosts. The results further show that the observed bottleneck in European *B. bavariensis* isolates argued to be in connection to invading *I. ricinus* (39, 84, 93), is not shared by the other two genospecies. This all suggests that Eurasian Lyme borreliosis agents were all capable of geographic expansion through vector shifts but differ in their capacity as emergent pathogens in relation to potential, future expansions into novel transmission cycles.

Borrelia bavariensis was already argued to have an Asian origin due to the deep branching observed between European and Asian isolates and that the majority of diversity exists in the Asian population (39, 84, 93). This finding is supported by the expanded analysis reported here. One point that warrants consideration is that all European *B. bavariensis* isolates come from humans and that the observed pattern in diversity could potentially be an artifact of sampling only a low diversity sub-set of European *B. bavariensis* better adapted to human infection. As our study includes European human isolates for two additional genospecies, we were able to disprove that this pattern is due to sampling bias through showing that human *B. garinii* or *B. afzelii* isolates do not display a similar reduction in genetic diversity (Text S2). In addition to this, a search of the *Borrelia* MLST (multiple locus sequencing typing) database (151) shows eight *B. bavariensis* samples coming from *I. ricinus* DNA which do not differ from patient isolates on the eight MLST loci, which can roughly proxy the full chromosome diversity (Figure S3). These data support that the observed pattern in *B. bavariensis* is genuine. Compared to *B. bavariensis*, no research has focused on the geographic origin of *B. afzelii* or *B. garinii*. Previous work raised the hypothesis of an Asian origin for *B. afzelii* but based on very few samples (152), whereas for *B. garinii* only partial structuring between continents was previously reported (91). Here though, we show that both *B. afzelii* and *B. garinii* are characterized by a basal node which splits a fully Asian clade from a clade of mixed geographic origin, suggesting for the first time that all three of these pathogenic genospecies originate in Asia and that

through successful colonization of *I. ricinus* were able to expand into Europe. MDS clustering based on plasmid profiles further supported this by suggesting that the plasmid profiles present in Europe are a subset of available profiles present in Asia (Figure 4B). This could further show that the European population stems from the Asian population. *Borrelia afzelii* was the only genospecies which showed a step-wise colonization pattern from far-east Asia through Russia and into Europe (Figure S1) which has been observed in other tick-borne pathogens (153) suggesting differences in migratory patterns between species. We further hypothesized about which genospecies colonized Europe first through calculating absolute divergence (D_{XY}) and Tajima's *D* (138). *Borrelia bavariensis* shows the highest D_{XY} suggesting that this colonization is the oldest of the three with *B. afzelii* then the youngest with the lowest value of D_{XY}. Additionally, as expected from bacterial populations (93, 138), Tajima's *D* values are consistent with population expansion (negative Tajima's *D*) but the European expansion for each species is younger (more negative values). The magnitude of difference in Tajima's *D* mirrors that of D_{XY} with *B. bavariensis* showing the lowest difference in Tajima's *D* (less recent) and *B. afzelii* showing the highest difference (most recent).

It is apparent from our analysis that, after the colonization of Europe, each genospecies experienced variable levels of gene flow which we argue can be related back to their host associations. The fact that *B. garinii* showed little to no geographic structuring is in accordance with previous results (45, 91, 98). *Borrelia garinii* utilizes birds as reservoir hosts and exists in overlapping terrestrial and marine transmission cycles, where it is vectored by different tick species (terrestrial: *I. ricinus* and *I. persulcatus*; marine: *I. uriae*) (45, 99, 154, 155) (Figure 2). The lack of geographic structure observed in *B. garinii* is thought to be a result of this, as birds could aid in the migration of this genospecies (45, 91, 98). This would explain why we cannot differentiate between distinct geographic locations. This pattern for *B. garinii* was already observed on a European level (91), but we are now able to show that it occurs over the whole distribution range of the genospecies. *Borrelia afzelii* and *B. bavariensis* displayed structured populations in our analysis. Within-continent structuring for European *B. afzelii* was previously attributed to utilizing rodents as reservoir hosts (90, 156), which we now propose to also occur in Asian *B. afzelii* populations. Even though our analysis does show some level of migration is possible along the geographic scale of this project as one Hokkaido isolate does cluster within

the Honshu clade (Figure 3A). As *B. bavariensis* also associates with rodents (94, 157), we would expect to also observe geographic structuring. As previously reported, there does not appear to be gene flow between the European and Asian populations suggesting genetic isolation (high F_{ST} and D_{XY}; Table 1), but within Asia *B. bavariensis* is not structured as expected for a rodent adapted genospecies (94, 157). Our analysis builds upon previous work which observed migration between Asian regions (i.e. Japan, China, Russia) (84), but by further adding randomly sampled isolates from distinct Japanese islands: Honshu and Hokkaido. Unlike *B. afzelii*, where we observe lower migration between the islands (F_{ST}=0.379; Table S2), *B. bavariensis* isolates do not seem to have the same barrier to migration (F_{ST}=0.057; Table S2). This brings forward the question, what mechanism(s) could result in this unexpected migration of Asian *B. bavariensis* isolates? One suggestion could be that Asian *B. bavariensis* utilize secondary hosts besides rodents which increase effective dispersal rate. Recently, *B. bavariensis* DNA was found far afield of its Eurasian range in seabird associated ticks (*I. uriae*) in Canada (158). As there are similarities in the structuring of Asian *B. bavariensis* to *B. garinii* from our results (low F_{ST}, high π, AMOVA with low σ due to geography; Table 1 & 2), it could be that in rare cases *B. bavariensis* may successfully transmit through avian hosts although rodent adapted. This fact had been previously observed where rodent-associated genospecies (i.e. *B. afzelii*) appeared to transmit through avian hosts in Europe (159). Although the extent of transmission appears to be different between *B. bavariensis* and *B. afzelii* based on our analyses. Until 2009 (94), *B. bavariensis* was considered a sub-type of *B. garinii* which utilized rodents as reservoir hosts (152, 160). This association with rodents was experimentally shown for two isolates (PBi, European; NT29, Asian) where they were exposed to rodent or avian immune sera and were susceptible to avian sera only (47, 48). In this case, as in many studies, immune serum resistance is taken as a proxy of reservoir host associations (47, 48). This result was used to support that *B. bavariensis* is not able to transmit through avian hosts. As the Asian population is quite diverse (84, 93) it is possible that a single isolate will not be representative for the entire population. Previous work did indeed suggest that similar genotypes of *B. bavariensis* (described as rodent adapted *B. garinii*) whichwere isolated from infected mice in Japan shared unique sequence components to a bird isolated strain from the Korean Peninsula, suggesting that *B. bavariensis* could spread from mainland Asia to Japan through migratory birds (161), as we are proposing. Additionally, a study based on restriction fragment length

polymorphism (RFLP) analysis described a novel RFLP type (type IVa) (162) which is now known to belong to *B. bavariensis* (Dr. Minoru Nakao & Dr. Hiroki Kawabata, personal communication). The isolates belonging to this RFLP type were isolated from rodents, humans, but also birds (162). Whole genome sequencing of these isolates would allow us to confirm if these bird isolates truly belong to *B. bavariensis*.

The results presented here suggest some answers to how LB spirochetes (*B. afzelii*, *B. bavariensis*, and *B. garinii*) expanded across Eurasia, through showing that all currently known pathogenic Eurasian *Borrelia* genospecies expanded into Europe from an ancestral Asian population through successful colonizing a novel tick vector (*I. ricinus*). Recently, *B. garinii* was found in *I. uriae* ticks in seabird colonies along the Atlantic coast of North America (45, 98, 163). As *B. garinii* was shown to be rarely transmitted through the North American tick vector (*I. scapularis*) in lab based studies (122) and here we show that *B. garinii* expanded into Europe through colonization of *I. ricinus*, potentially another expansion into the North American transmission cycle is possible if other requirements, such as reservoir host availability, are met. Outside of this, we further observed that post-colonization gene flow appears to relate to host association and were able to make further testable hypotheses regarding the ecology of the populations. Our analysis provides novel information to the spread of LB-causing spirochetes across Eurasia with applications to how adaptation to novel vector species can facilitate geographic expansion and thus potentially aid in the spread of emergent human pathogens.

Paper 5: Supplementary Material

Out of Asia? Recurrent vector switches leading to the expansion of Eurasian Lyme disease bacteria

Robert E. Rollins, Kozue Sato, Minoru Nakao, Mohammed T. Tawfeeq, Fernanda Herrera-Mesías, Ricardo J. Pereira, Sergey Kovalev, Gabriele Margos, Volker Fingerle, Hiroki Kawabata, and Noémie S. Becker

Unpublished Manuscript

Out of Asia? Vector switches leading to the expansion of Eurasian Lyme disease bacteria

Robert E. Rollins, Kozue Sato, Minoru Nakao, Mohammed T. Tawfeeq, Fernanda Herrera-Mesías, Ricardo J. Pereira, Sergey Kovalev, Gabriele Margos, Volker Fingerle, Hiroki Kawabata, and Noémie S. Becker

— Supplementary Methods —

Generating high quality Borrelia garinii reference genomes

As previously described for *B. bavariensis* (PBi, A104S, and NT24)(32, 84), Pacific Bioscience SMRT sequencing (hereafter PacBio) was performed using 10 µg of DNA from the *B. garinii* isolates PHei, PBr, and NT31. DNA was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany). Libraries were prepared using Pacific Biosciences 20 kb library preparation protocol. Size selection of the final library was performed using Blue-Pippin with a 10 kb cut-off. The library was sequenced on a Pacific Biosciences RS II instrument using P6-C4 chemistry with 360 min movie time. Additionally, these samples were sequenced with Illumina MiSeq short-read sequencing as described by Methods in the main text.

PacBio reads were assembled using HGAP v3 (Pacific Biosciences, SMRT Analysis Software v2.3.0). Chromosome ends were trimmed to remove the pseudo-telomere regions present in *B. burgdorferi* s. l. linear replicons(164). Assembled Illumina contigs (see Methods) for the same isolates were then mapped to the PacBio contigs NUCmer v 3.23 from the package MUMmer(128, 129). As PacBio sequencing technology is prone to sequencing errors like point mutations and short indels(83, 165), we combined the data from PacBio and Illumina using the following rules: for each indel of length 5 bp or less keep the Illumina version, for longer indels keep the PacBio version. For point mutations, keep the Illumina version if all contigs mapping on this position agree, else keep PacBio version. This dual approach (long and short-read sequencing) has been shown to be the most effective to produce high quality reference genomes for *B. burgdorferi* s. l. samples(83). The final chromosome sequences were then utilized along with the *B. bavariensis* PacBio(32, 84) and *B. afzelii* chromosomes from GenBank (PKo, K78 and ACA-1) as references in the

mapping and assembly of final chromosomes for isolates with only Illumina MiSeq data.

Mapping analysis to produce final chromosome sequences

Assembled SPAdes contigs were mapped using NUCmer v 3.23 from the package MUMmer(128, 129) to each reference chromosome and the closest reference (highest percent ID with the full chromosome mapped) was chosen. Each mapping file was then curated to remove contigs overlapping contigs with higher identity. Single nucleotide polymorphisms (SNPs) and indels for the curated mapping files were called using the program *show-snps* from the package MUMmer(128, 129). Final chromosome files were then created according to the following rules outlined in Becker et al. (2020): for SNPs keep the Illumina allele if all contigs mapping at this position agree, else keep the reference allele if at least one contig also has it, else replace the base by "N"; keep insertions and deletions if and only if all contigs mapping at this position agree, else keep the reference version. These final files were then used in all further analysis.

Correcting chromosome alignments for recombining regions

The four-gamete condition introduced by Hudson (1985) detects recombining sites under an infinite-sites model, under the assumption when considering two polymorphic sites in a population that mutation alone can only produce three allele combinations. If the fourth combination is observed, the four-gamete condition is violated and either recombination or back mutation must have occurred. As our data does not fit an infinite sites model, these signals can be ambiguous and should only be considered as signals of recombination if they cannot be explained through a few back mutations or double hits.

We applied the four-gamete condition to our full chromosome alignment as described in Gatzmann et al. (2015). The ordered list of chromosome segregating sites was divided into blocks containing the same number of sites (n = 12). Each pair of SNPs in each block was then assessed if the four-gamete condition was violated or not and given a score of 1 (violated) or 0 (not violated). The within block average and standard deviation was then calculated and averaged across all blocks and used as a measure of background violation due to double hits or back mutations. To single out SNP blocks which were most likely under recombination, we then calculated all pairwise

comparisons between blocks and recorded the violation score. This score was then averaged over all comparisons for a specific block. Blocks were considered recombinant if:

$$x_i \geq \mu_{within} + 2sd_{within}$$

And non-recombinant if:

$$x_i < \mu_{within} + 2sd_{within}$$

Where x_i is the average violation of block i over all comparisons, μ_{within} is the average within-block violation score, and sd_{within} is the standard deviation of within-block violation score. Recombinant blocks were removed from the final alignments prior to phylogenetic reconstruction.

Phylogeny of Borrelia afzelii including Russian isolate Tom3107

Our original data set did not include any *Borrelia afzelii* isolates coming from Russia. As this could influence our conclusions regarding the evolution of the genospecies, we re-ran the *B. afzelii* phylogeny including all *B. afzelii* chromosome sequences and adding a publicly available Russian *B. afzelii* chromosome sequence from GenBank (Tom3107; Accession number: NZ_CP009212.1). The chromosome sequence from *B. bavariensis* (isolate: PBi) was also included as an outgroup to root the tree. All chromosomes were aligned using MAFFT v. 7.407(105, 119). Recombinant regions were identified through searching for areas which violated the four-gamete condition(130) (as described in Gatzmann et al. (2015); see Suppl. Met.) and were removed from the final alignment. Phylogeny reconstruction was done in MrBayes v. 3.2.6(131, 132) with ploidy set to haploid and a GTR(133) substitution model with gamma distributed rate variation. Three independent runs were launched and ran for 10 million generations at which point convergence of parameters was checked with Tracer v. 1.7.1(134). Consensus trees were built using the *sumt* command from MrBayes using a respective burn-in of 25%. Convergence to a single topology in all three independent runs was checked manually in FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) which was also used to plot the tree shown in Figure S3.

Identification of plasmid partitioning genes for proxy of plasmid content

We used BLAST v. 2.8.1 (algorithm *blastn*) (114, 115) to identify the presence of plasmid partitioning genes of the PFam32, 49, 50 and 57/62 families. In a first BLAST

round we used as queries the PFam32 genes sequences of *B. burgdorferi* s. s. strains B31, BOL26, JD1 and 118a and *B. afzelii* strain PKo to cover the whole plasmid diversity and the PFam49, 50 and 57/62 of *B. burgdorferi* s. s. strain MM1. We performed the search on the assembled SPAdes Illumina contigs of each isolate. We then reiterated the BLAST search using as queries all the hits found in the first search. We then removed all hits with percent identity lower than 80% and that were shorter than half the length of the references (reference lengths were around 750 bp for PFam32, 550 bp for PFam49 and PFam50 and 900 to 1100 bp for PFam57–62) and that had no open reading frame over at least half of the length of the reference. Plasmids were considered present if at least one of the partitioning genes was present. In this we did not account for potential plasmid fusions(83, 84), so each plasmid was individually counted as present or absent, so the actual number could be lower as seen in other work(84).

— Supplementary Texts —

Text S1: Isolation of novel *Borrelia* tick isolates and choice for inclusion in genomic study

With the advent of sequencing technologies (both long and short read) and the prices of obtaining whole genomes decreasing(166, 167), there has been a surge of genomic studies looking into bacteria, including *Borrelia burgdorferi* sensu lato (*Bb*sl)(83–85, 87, 168).

With these new genomes, researchers have been able to apply population genetics methods to better understand the overall evolution of *Bb*sl(168), but many studies include non-randomly sampled isolates which can violate basic assumptions of many population genetic statistics and lead to inherent biases(137). Therefore, a goal of this study was to produce randomly sampled *Bb*sl isolates coming from three genospecies: *B. garinii*, *B. bavariensis*, and *B. afzelii*. For this, isolates were gathered from the same geographic location and if possible isolated from the same pool of collected ticks to reduce biasing factors.

Tick sampling and novel Borrelia *isolation*

Ticks were collected by flagging, where a 1×1 m flag was dragged along the understory foliage for 10m before being turned over and assessed for attached ticks. In Germany, all adults and nymphs were collected whereas in Japan only adult ticks were

collected. All collected ticks were stored in vials containing grass to maintain humidity. Japanese ticks were collected from two sampling sites, one on the northern Japanese island of Hokkaido (near the city of Asahikawa) and the other on the southern Japanese island of Honshu (near the city Matsumoto). In both cases, all ticks later used were collected in a single sampling event during May 2019 and represent a random sampling of distinct Japanese tick populations. German ticks were collected in seven plots (described in Rollins et al. 2021) during the Spring and Summer of 2018-2019. For all information on sampling sites, see Supplemental Table 4. Prior to processing, ticks were morphologically identified to species level using published taxonomic keys(18, 170). Only *Ixodes ricinus* (Europe) or *Ixodes persulcatus* (Japan) ticks were used in further analyses.

Japanese tick samples were washed in batches of 5-10 ticks in H_2O_2 (4%), 80% ethanol, and distilled water for 5 minutes each. After washing, legs and mouth parts were removed with a sterile scalpel. Using a fresh, sterile scalpel, ticks were bifurcated along the longitudinal axis. One half was placed in BSK-H media(126) containing antibiotics (rifampicin, 32 µg/mL; phosphomycin, 12.8 µg/mL; amphotericin B, 10 µg/mL) the other half was used for direct DNA extraction using the DNeasy Blood & Tissue kit (Qiagen, Germany) according to standard procedure for a separate project. German tick samples were processed similarly to the Japanese samples but instead were washed in distilled water, absolute ethanol, and sterile 1×PBS. The whole tick was homogenized after removal of mouth parts and legs in 50µL of PBS which was then transferred to fresh MKP media(125) containing antibiotics (phosphomycin 28.6 µg/mL; sulfamethoxazole 14.3 µg/mL; amikacin 4.3 µg/mL). All cultures were incubated at 37°C with 5% CO_2.

All cultures were checked for *Borrelia* bacteria using dark field microscopy and *Borrelia* positive samples were cultured according to standard procedure (125, 126) until the cultures reached a density of at least 10^8 cells per mL at which point whole genomic DNA was extracted. Genomic DNA from all European isolates was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany) and DNA from all Japanese isolates was extracted using the Wizard® Genomic DNA Purification Kit (Promega, Germany).

Aliquots of all Japanese and Germany isolates were stored at -80°C at either the National Institute for Infectious Diseases in Tokyo (all Japanese isolates) or the

German National Reference Center for *Borrelia* at the Bavarian Health and Food Safety Authority (all German isolates).

Genospecies determination of novel isolates

To determine which *Borrelia* genospecies had been isolated, genomic DNA was subjected to a semi-nested PCR amplifying the housekeeping gene *recG* using a previously described protocol (151). Either *B. kurtenbachii* (25015) or *B. mayonii* (DSM 102811.MN14-1420) was used as a positive control. PCR products were sequenced using Sanger sequencing at the Sequencing Service of Ludwig-Maximilians University and were prepared according to the requirements of the sequencing center (http://www.gi.bio.lmu.de/sequencing/help/protocol). Chromatograms for *recG* sequences were manually checked for quality using FinchTV v. 1.4.0 (Geospiza, Inc.; Seattle, WA, USA; http://www.geospiza.com) and *recG* sequences containing ambiguities were marked as mixed infections. Non-mixed sequences were then aligned to the *recG* reference from *B. burdorferi* sensu stricto (s.s.) strain B31 (GenBank: AE000783.1) and trimmed. Trimmed sequences were blasted to the PubMLST database for *Borrelia* (https://pubmlst.org/borrelia) for genospecies determination.

Additional Russian isolates

Additional ticks were collected in Russia by flagging in the regions of Tomsk (2013) and Yekaterinburg (2011 & 2019). Ticks were identified to species level according to published taxonomic keys(18, 170) and washed for 5 min H_2O_2 (4%), 5 min 80% ethanol, and 5 min distilled water. After washing, the tick was homogenized and placed in BSK-H media(126). Cultures were checked for *Borrelia* using dark field microscopy.

In total 11 *B. garinii* isolates were successfully isolated from Tomsk in 2013 (n=6) and Yekaterinburg in 2011 (n=4) and 2019 (n=1). Genomic DNA was extracted using the Wizard® Genomic DNA Purification Kit (Promega, Germany). All samples were then included in sequencing for the project.

Isolation success in Japan and Germany

In total, 393 Japanese *I. persulcatus* (all adults) and 1252 *I. ricinus* (280 adults, 972 nymphs) ticks were collected and processed during this study. The sampling sites in Japan had the highest positive rates with 38.33% (ASA) and 50.92% (NAG) of all

processed ticks (Supplemental Table 5). In Germany, positive rates were normally much lower ranging from 0-10.71% positive (Supplemental Table 5). In fact, isolates were only obtained from three of the seven sampled plots: OBE, STA, and HER (for location definition see Supplemental Table 4).

Not all positive samples produced isolates however as some were lost due to contamination or the culture crashing. In the end, 95 *B. afzelii*, 49 *B. bavariensis*, 11 *B. garinii*, and four mixed tick isolates were obtained from the two locations in Japan (Supplemental Table 6). In Germany, 7 *B. burgdorferi* sensu stricto, 9 *B. garinii*, 2 *B. afzelii*, and one mixed tick isolate were obtained.

Criteria for inclusion in genomic study

Our study aimed to study the evolution of Eurasian distributed *Borrelia* species, so we limited our sequencing to the species *B. afzelii*, *B. bavariensis*, and *B. garinii*. The major criteria for inclusion in the genomic study were DNA concentration over 0.2ng/µL (minimum amount needed for producing MiSeq library with a Nextera XT preparation kit), a high purity (260/280 > 1.80), and the *recG* sequencing did not display any ambiguous base pairs (signs of mixed isolate). After filtering for these criteria, 20 isolates were chosen from the available *B. afzelii* and *B. bavariensis* isolates for sequencing. For *B. afzelii* this corresponded to the four available Hokkaido (ASA) samples and 16 randomly chosen from the available Honshu (NAG) samples. For *B. bavariensis*, 10 isolates each were randomly chosen from ASA and NAG. Only 11 *B. garinii* isolates were available (all from Hokkaido) and all were included in sequencing.

As for German samples, all available *B. garinii* (n=7) and *B. afzelii* (n=2) were included in the study. Seven of these samples arise from the same geographic plot (STA) isolated in the years 2018 and 2019. As the current isolate collection of *B. bavariensis* (n=19) are all isolated from humans and this could potentially bias the results of our analyses, tick collection in 2018-2019 focused on plots where ticks positive for *B. bavariensis* DNA were previously reported(169, 171). Even so, we were not successful in isolating *B. bavariensis* from tick samples and further work is needed to determine what method should be used to isolate this species in Europe.

Filtering low quality assemblies and mixed strains

All German samples sequenced produced usable assemblies and were therefore included in all further analyses. In the end, 20/20 Japanese *B. afzelii*, 10/11 Japanese

B. garinii, and 13/20 Japanese *B. bavariensis* isolates produced usable assemblies. These 52 assemblies were then used for the main study. All additional Russia *B. garinii* isolates produced usable assemblies.

The single *B. garinii* had low DNA concentration and purity, so it is not unexpected that it did not produce usable sequence data. The seven *B. bavariensis* samples that did not produce usable assemblies are a bit more complex. Three samples produced low quality assemblies, resulting in an inability to reconstruct the full chromosome. The other four samples produced high quality assemblies but appeared to be a mixture of two *Borrelia* species (in all cases a mix of *B. bavariensis* and *B. afzelii*). As both of these species utilize rodents as reservoir hosts (39, 43, 101), it is not surprising to find them together in a mixed infection. All four of these samples, however, did not show signs of mixed infection on the *recG* sequences. As *Borrelia* genomes pose challenges to assembly when the DNA samples are pure (83, 84) let alone when they are mixed with a highly similar species, these four samples were removed from the analysis to avoid any potential biases due to mis-assembly.

Text S2: Potential biases to population genetics statistics due to non-random sampling

Non-random sampling in population genetic studies can lead to biased estimates of population level statistics(137). Random sampling design has been considered in some studies on *Borrelia* genomics(168), with also many studies biased towards human isolates instead of tick isolates. Theoretically in the case of human isolates, there could be a selective bottleneck of which isolates can infect humans biasing the samples obtained although some results suggest this bias does not occur(172).

The design of our study allows us to test for such biases as we report both random and non-random samples. For *B. garinii*, this extends to being able to compare a random German (European) and Japanese (Asian) sample. Here we see that nucleotide diversity (π)(137) does not differ between randomly (π_{Europe} = 0.006333; π_{Asia} = 0.009499; Supplementary Table 2) and non-randomly (π_{Europe} = 0.006185; π_{Asia} = 0.009003; Table 2) sampled *B. garinii* isolates. A similar trend for F_{ST}(137) of *B. garinii* populations was observed for random (F_{ST} = 0.1003; Supplemental Table 2) and non-random (F_{ST} = 0.1318; Table 1) samples. *Borrelia afzelii* and *B. bavariensis* samples however show us that there can be local level variation in population level

statistics (Supplemental Table 2) highlighting the importance of samples coming from known regions to understand micro-scale evolutionary forces acting on these bacteria.

A further finding is the low diversity of European *B. bavariensis* samples, of which all are human isolates. Therefore, it could be the case that low diversity is a hallmark of European human samples, potentially suggesting a selective bottleneck to infect humans. As this study additionally includes European human isolates of *B. garinii* (n = 21) and *B. afzelii* (n = 11), this allowed us to see if this reduced diversity is unique to *B. bavariensis*. Both *B. garinii* (π_{Human} = 0.005956) and *B. afzelii* (π_{Human} = 0.002107) human isolates showed slightly lower nucleotide diversity in comparison to all isolates (Table 1), but they did not differ by an order of magnitude as observed with *B. bavariensis* (Supplemental Table 2; Table 1). Furthermore, in the *Borrelia* MLST(151) database (https://pubmlst.org/borrelia/) there are eight samples identified in ticks (*I. ricinus*) that have the same MLST sequence type as human isolates. From our analysis we can show that MLST sequences capture the nucleotide diversity calculated from the full chromosome sequence (Figure S2). These results support that the reduced diversity of European *B. bavariensis* indeed is a unique phenomenon and not a result of human samples biasing the result.

Taken altogether, the results here show that there could be slight biases to basic population genetic statistics but nothing drastic that could modify conclusions. Even so, we highlight the fact that within-region, random sampling allows for researchers to potentially understand local level evolution of *Borrelia* bacteria and showing that there are small scale influences dependent on geographic origin.

— Supplementary Tables and Figures —

Table S1. Meta-data for all isolates sequenced and utilized in the study. Samples marked with GB were downloaded from GenBank and included in the analysis, under the accession numbers: CP007564.1 (SZ), CP003866.1 (NMJW1), CP003151.1 (BgVir), CP002933.1 (K78), CP009058.1 (PKo), NZ_ABCU00000000.2 (ACA-1).

Isolate	Species	Year	Continent	Region	Source	Information
FujiP2	*B. bavariensis*	-	Asia	Japan (Shizuoka)	Tick	*I. persulcatus*
Hiratsuka	*B. bavariensis*	2008	Asia	Japan (Tick bite in Niigata)	Human	erythrema migrans
J14	*B. bavariensis*	1995	Asia	Japan (Hokkaido)	Human	erythrema migrans
J15	*B. bavariensis*	1995	Asia	Japan (Hokkaido)	Human	erythrema migrans
J20T	*B. bavariensis*	1996	Asia	Japan (Hokkaido)	Human	erythrema migrans
Konnai17	*B. bavariensis*	2011	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
N346	*B. bavariensis*	-	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
NT24	*B. bavariensis*	-	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JAASAAF1014	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAF1016	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAF1029	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1091	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1101	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1114	*B. bavariensis*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JANAGAM1271	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1274	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1305	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1327	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1352	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1353	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1369	*B. bavariensis*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
SZ (GB)	*B. bavariensis*	-	Asia	China	Tick	*Dermacentor* spp.
NMJW1 (GB)	*B. bavariensis*	-	Asia	China	Tick	*I. persulcatus*
BgVir (GB)	*B. bavariensis*	-	Asia	Russia	Tick	*I. persulcatus*
Arh913	*B. bavariensis*	2012	Asia	Russia (Kotlas)	Tick	*I. persulcatus*
Arh923	*B. bavariensis*	2012	Asia	Russia (Kotlas)	Tick	*I. persulcatus*
Prm7019	*B. bavariensis*	2012	Asia	Russia (Kungur)	Tick	*I. persulcatus*
Prm7564	*B. bavariensis*	2011	Asia	Russia (Cherdyn)	Tick	*I. persulcatus*
Prm7569	*B. bavariensis*	2011	Asia	Russia (Cherdyn)	Tick	*I. persulcatus*
Prm965	*B. bavariensis*	2013	Asia	Russia (Kudymkar)	Tick	*I. persulcatus*
A104S	*B. bavariensis*	1996	Europe	Netherlands	Human	-
A91S	*B. bavariensis*	1996	Europe	Netherlands	Human	-
DK6	*B. bavariensis*	1990	Europe	Denmark	Human	-
Lubl25	*B. bavariensis*	1995	Europe	Slovenia	Human	-

61VB2	*B. bavariensis*	-	Europe	Germany	Human	Skin condition
PBael	*B. bavariensis*	1990	Europe	Germany	Human	skin condition
PBaeII	*B. bavariensis*	1990	Europe	Germany	Human	neuroborreliosis
PBar	*B. bavariensis*	1988	Europe	Germany	Human	skin condition
PBi	*B. bavariensis*	<1993	Europe	Germany	Human	neuroborreliosis
PBN	*B. bavariensis*	1999	Europe	Germany	Human	neuroborreliosis
PHerl	*B. bavariensis*	1989	Europe	Germany	Human	skin condition
PLad	*B. bavariensis*	2000	Europe	Germany	Human	skin condition
PNeb	*B. bavariensis*	1988	Europe	Germany	Human	skin condition
PNi	*B. bavariensis*	2000	Europe	Germany	Human	lymphoma
PRab	*B. bavariensis*	1994	Europe	Austria	Human	Lyme arthritis
PRof	*B. bavariensis*	1989	Europe	Germany	Human	skin condition
PTrob	*B. bavariensis*	1988	Europe	Slovenia	Human	skin condition
PWin	*B. bavariensis*	1987	Europe	Germany	Human	skin condition
PZwi	*B. bavariensis*	1994	Europe	Germany	Human	skin condition
JAASAAF1012	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAF1040	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAF1041	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1058	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1060	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1063	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1086	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1087	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1097	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1103	*B. garinii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
HT59	*B. garinii*	-	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
J21	*B. garinii*	1996	Asia	Japan (Hokkaido)	Human	erthyma migrans
NT31	*B. garinii*	-	Asia	Japan (Nagano)	Tick	*I. persulcatus*
Konnai20	*B. garinii*	-	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
E-Burg-606	*B. garinii*	2019	Asia	Russia (Yekaterinburg)	Tick	*I. persulcatus*
Ekb701-11	*B. garinii*	2011	Asia	Russia (Yekaterinburg)	Tick	*I. persulcatus*
Ekb704-11	*B. garinii*	2011	Asia	Russia (Yekaterinburg)	Tick	*I. persulcatus*
Ekb706-11	*B. garinii*	2011	Asia	Russia (Yekaterinburg)	Tick	*I. persulcatus*
Ekb712-11	*B. garinii*	2011	Asia	Russia (Yekaterinburg)	Tick	*I. persulcatus*
Tms1187-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. pavlovskyi*
Tms1188-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. pavlovskyi*
Tms1189-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. pavlovskyi*
Tms1190-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. pavlovskyi*
Tms1192-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. persulcatus*
Tms1218-13	*B. garinii*	2013	Asia	Russia (Tomsk)	Tick	*I. pavlovskyi*
PNov	*B. garinii*	1990	Europe	Yugoslavia	Human	skin condition
PMa	*B. garinii*	1989	Europe	Yugoslavia	Human	skin condition
Malouvrh	*B. garinii*	-	Europe	Slovenia	Human	skin condition

20047	*B. garinii*	-	Europe	France	Tick	*I. ricinus*
GEOBEN020	*B. garinii*	2018	Europe	Germany (Oberschleißheim)	Tick	*I. ricinus*
GEHERN151	*B. garinii*	2018	Europe	Germany (Herrsching)	Tick	*I. ricinus*
GESTAN298	*B. garinii*	2018	Europe	Germany (Starnberg)	Tick	*I. ricinus*
GESTAN302	*B. garinii*	2018	Europe	Germany (Starnberg)	Tick	*I. ricinus*
GESTAAF1947	*B. garinii*	2019	Europe	Germany (Starnberg)	Tick	*I. ricinus*
GESTAAF2296	*B. garinii*	2019	Europe	Germany (Starnberg)	Tick	*I. ricinus*
GEHERAF2403	*B. garinii*	2019	Europe	Germany (Herrsching)	Tick	*I. ricinus*
PBr	*B. garinii*	1985	Europe	Germany	Human	neuroborreliosis
PHel	*B. garinii*	1987	Europe	Germany	Human	neuroborreliosis
Mek	*B. garinii*	1992	Europe	Germany	Human	neuroborreliosis
PBes	*B. garinii*	1989	Europe	Germany	Human	neuroborreliosis
PFr	*B. garinii*	1995	Europe	Germany	Human	neuroborreliosis
PHc	*B. garinii*	1996	Europe	Germany	Human	neuroborreliosis
PHez	*B. garinii*	1994	Europe	Germany	Human	neuroborreliosis
PKi	*B. garinii*	1992	Europe	Germany	Human	neuroborreliosis
PLa	*B. garinii*	1988	Europe	Germany	Human	neuroborreliosis
PLi	*B. garinii*	1988	Europe	Germany	Human	neuroborreliosis
PMe	*B. garinii*	1988	Europe	Germany	Human	skin condition
PMek	*B. garinii*	1992	Europe	Germany	Human	neuroborreliosis
PMit	*B. garinii*	1997	Europe	Germany	Human	neuroborreliosis
POhm	*B. garinii*	1991	Europe	Germany	Human	neuroborreliosis
PSoR	*B. garinii*	1989	Europe	Germany	Human	neuroborreliosis
PStg	*B. garinii*	1996	Europe	Germany	Human	Lyme arthritis
PUl	*B. garinii*	1999	Europe	Germany	Human	neuroborreliosis
U02	*B. garinii*	-	-	-	-	-
U03	*B. garinii*	-	-	-	-	-
U04	*B. garinii*	-	-	-	-	-
PKie	*B. garinii*	1993	Europe	Germany	Human	neuroborreliosis
JAASAAF1010	*B. afzelii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAF1039	*B. afzelii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1092	*B. afzelii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JAASAAM1104	*B. afzelii*	2019	Asia	Japan (Hokkaido)	Tick	*I. persulcatus*
JANAGAF1137	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1156	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1163	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1173	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1182	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1229	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1248	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1250	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1261	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1316	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*

JANAGAM1325	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1334	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1340	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAM1370	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1390	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
JANAGAF1392	*B. afzelii*	2019	Asia	Japan (Nagano)	Tick	*I. persulcatus*
GESTAN171	*B. afzelii*	2018	Europe	Germany (Starnberg)	Tick	*I. ricinus*
GESTAF1951	*B. afzelii*	2019	Europe	Germany (Starnberg)	Tick	*I. ricinus*
PWat	*B. afzelii*	1994	Europe	Germany	Human	neuroborreliosis
PMel	*B. afzelii*	1990	Europe	Germany	Human	neuroborreliosis
PBabu	*B. afzelii*	2001	Europe	Germany	Human	osteomyelitis
PHam	*B. afzelii*	1987	Europe	Germany	Human	skin condition
PJe	*B. afzelii*	1992	Europe	Germany	Human	neuroborreliosis
PBec	*B. afzelii*	1988	Europe	Germany	Human	skin condition
PSto	*B. afzelii*	1986	Europe	Germany	Human	skin condition
PFes	*B. afzelii*	1988	Europe	Germany	Human	skin condition
PKL	*B. afzelii*	1993	Europe	Germany	Human	neuroborreliosis
PKr	*B. afzelii*	1992	Europe	Germany	Human	neuroborreliosis
PObf	*B. afzelii*	1998	Europe	Germany	Human	skin condition
K78 (GB)	*B. afzelii*	-	Europe	Austria	Human	erthyma migrans
PKo (GB)	*B. afzelii*	1984	Europe	Germany	Human	skin condition
ACA-1 (GB)	*B. afzelii*	-	Europe	Sweden	Human	skin condition
Tom3107 (GB)	*B. afzelii*	-	Europe	Sweden	Human	skin condition

Table S2. Population genetics statistics calculated on randomly sampled *Borrelia* populations in both Japan and Germany.

Genospecies	Population	n	π	Tajima's *D*	F_{ST}	D_{XY}
Borrelia afzelii	Hokkaido (JA)	4	0.003084	-9.708	0.379	0.00303
	Nagano (JA)	16	0.001117	-3.959		
Borrelia bavariensis	Hokkaido (JA)	6	0.008685	0.060	0.057	0.00680
	Nagano (JA)	7	0.004679	-4.302		
Borrelia garinii	Hokkaido (JA)	10	0.009499	-1.580	0.188	0.00904
	Munich (GE)	7	0.006333	-4.089		

Table S3. Effect size estimates (β) and 95% credible intervals (CIs) for plasmid content of all isolates *de novo* assembled for this study (n=136).

	PFam Number
Fixed Effects	**β (95% CI)**
Intercept	14.88 (13.10, 16.68)
Sample Origin[a]	-0.08 (-1.10, 0.92)
Source[b]	1.19 (0.16, 2.22)
Random Effects	**σ^2 (95% CI)**
Genospecies	1.62 (0.80, 2.76)
Residual	1.86 (1.86, 1.86)

[a]Differences in origin of samples (Reference = Asia)

[b]Differences in sample source (Reference = Human)

Table S4. Sampling locations for novel *Borrelia* isolates.

Country	Location	Abbr.	Year (s)	GPS
Japan	Nagano	NAG	2019	36.2094 N, 138.1131 E
	Hokkaido	ASA	2019	43.7550 N, 142.3930 E
Germany	Oberschleißheim	OBE	2018-2019	48.2424 N, 11.5824 E
	Herrsching	HER	2018-2019	47.9802 N, 11.1589 E
	Starnberg	STA	2018-2019	48.0195 N, 11.3391 E
	Grafrath	GRA	2019	48.1350 N, 11.1653 E
	Perchting	PER	2019	47.9933 N, 11.2797 E
	Forstenrieder Park	FOR	2019	48.0828 N, 48.0744 E
	Englischer Garten	ENG	2019	48.1758 N, 11.6244 E

Table S5. Total number of ticks collected from each sampling plot and the number positive for *Borrelia* under dark field microscopy (DFM) in Japan and Germany.

Country	Region	Ticks Collected*	*Borrelia* positive (DFM)	% *Borrelia* positive (DFM)
Japan	ASA	120	46	38.33
	NAG	273	139	50.92
Germany	OBE	56	6	10.71
	HER	192	10	5.21
	STA	134	8	5.97
	GRA	275	0	0.00
	PER	114	0	0.00
	FOR	273	0	0.00
	ENG	63	0	0.00

*Refers to both adult (♂,♀) and nymphal ticks

Table S6. Total number of isolated *Borrelia* per genospecies from each sampling site in Germany and Japan.

Country	Region	Genospecies	Final Isolates
Japan	NAG	*Borrelia afzelii*	87
		Borrelia bavariensis	36
		Mixed infections	2
	ASA	*Borrelia afzelii*	8
		Borrelia bavariensis	13
		Borrelia garinii	12
		Mixed infections	2
Germany	OBE	*Borrelia burgdorferi* sensu stricto	3
		Borrelia garinii	2
	HER	*Borrelia burgdorferi* sensu stricto	4
		Borrelia garinii	2
		Mixed infections	1
	STA	*Borrelia afzelii*	2
		Borrelia garinii	5

Figure S1. Phylogeny of *Borrelia afzelii* based on the main chromosome corrected for recombining regions (see Suppl. Met.) a Russian isolate (Tom3107) downloaded from GenBank (Accession Number: NZ_CP009212.1). The phylogeny was reconstructed with MrBayes v. 3.2.6(131, 132) with ploidy set to haploid and a GTR(133) substitution model with gamma distributed rate variation. Three independent runs were launched and ran for 10 million generations each at which point convergence of parameters was checked with Tracer v. 1.7.1(134). Consensus trees were built using the *sumt* command from MrBayes using a respective burn-in of 25%. Colors correspond to geographic origin of the isolates: Europe (blue), Japan (red), or purple (Russia). For Japanese tick isolates, the island of origin is shown either as a diamond (Hokkaido) or star (Honshu) when known. The scale bar is in substitutions per site.

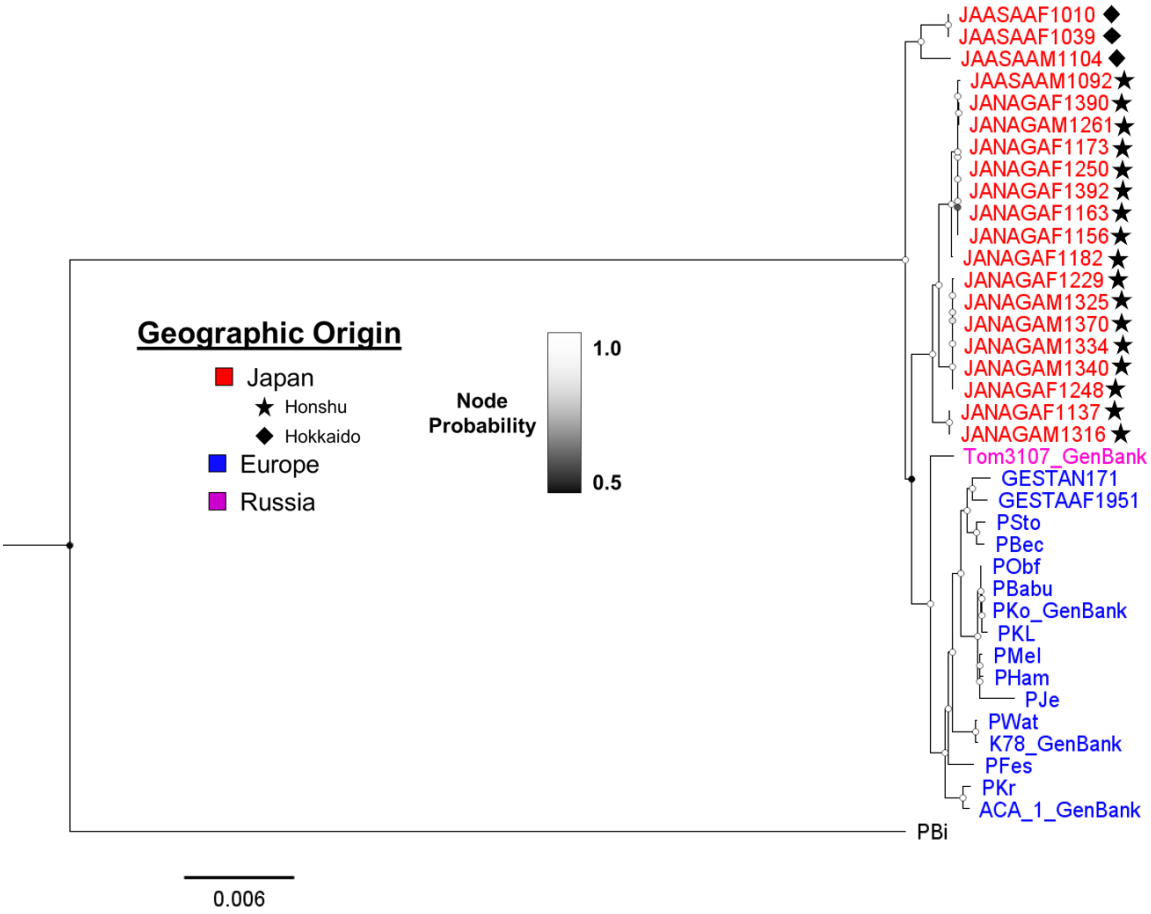


Figure S2. Heat map of plasmid frequency based on partitioning genes (PFam32, 49, 50, 57/62) per tick vector species (*I. persulcatus* and *I. ricinus*) and *Borrelia* genospecies split between Asian and European isolates. Frequency is calculated as the total number of isolates per group which contain the plasmid divided by all isolates in the group.

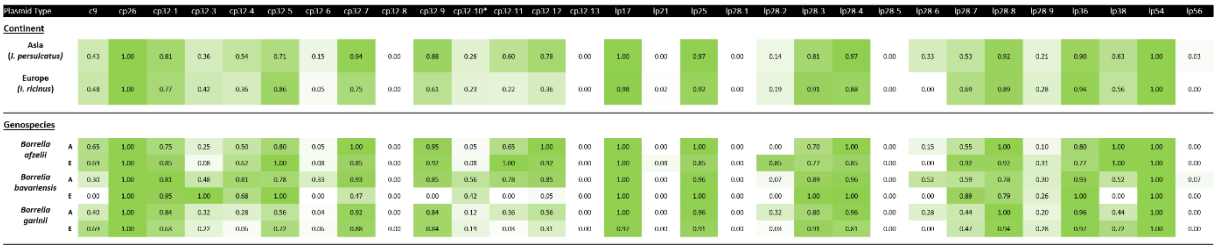
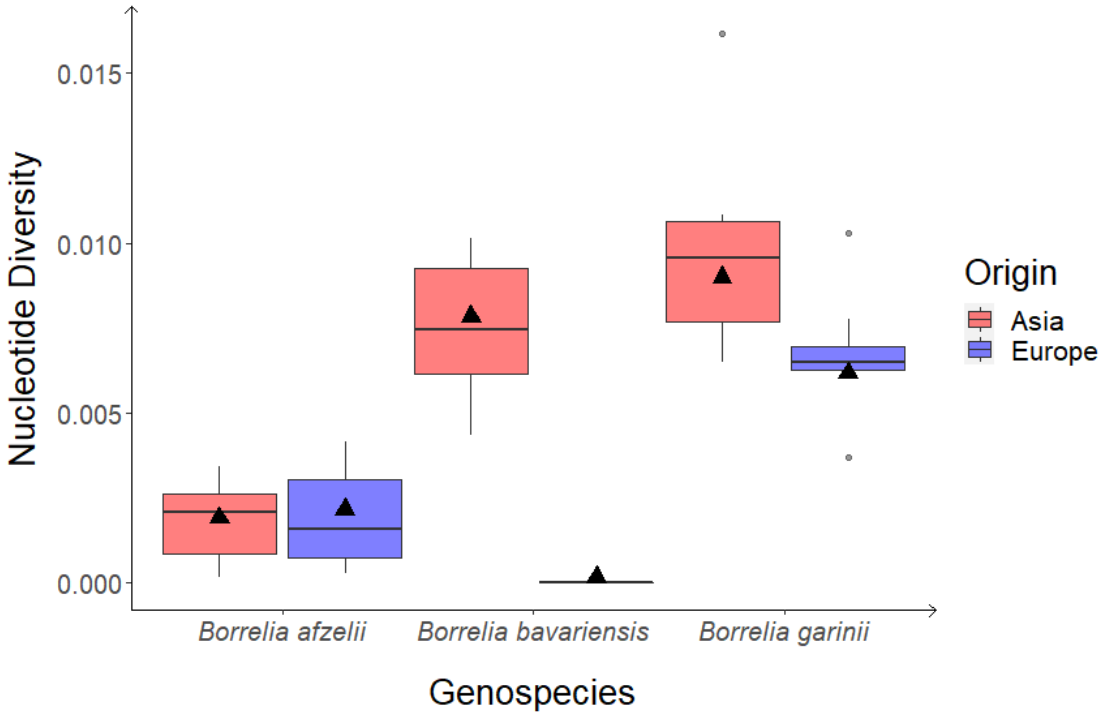


Figure S3. Boxplots shown for nucleotide diversity (π, (137)) calculated per MLST gene alignment described by Margos et al. (2008) for each geographic population of the three *Borrelia* genospecies researched here (*B. afzelii*, *B. bavariensis*, *B. garinii*). Black triangles show the nucleotide diversity (π, (137)) calculated for full chromosomes alignments for each geographic population of the three *Borrelia* genospecies researched here (*B. afzelii*, *B. bavariensis*, *B. garinii*).



Paper 6: Disentangling the role of *Borrelia bavariensis* PFam54 proteins *in vitro* and *in vivo* using two strains naturally lacking the PFam54 gene array

Robert E. Rollins, Janna Wülbern, Florian Röttgerding, Tristan Nowak, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Yi-Pin Lin, Peter Kraiczy, Noémie S. Becker

Unpublished Manuscript

Disentangling the role of *Borrelia bavariensis* PFam54 complement-interacting proteins in vitro and in vivo using two strains naturally lacking the PFam54 gene array

Robert E. Rollins, Janna Wülbern, Florian Röttgerding, Tristan Nowak, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Yi-Pin Lin, Peter Kraiczy, Noémie S. Becker

Abstract

Lyme borreliosis is the most common vector-borne disease in the Northern hemisphere, caused by spirochetes belonging to the *Borrelia burgdorferi* sensu lato (*Bb*sl) species complex which are transmitted by ixodid ticks. *Bb*sl species produce a family of proteins on the linear plasmid 54 (PFam54), some of which confer the functions of cell adhesion and inactivation of complement, the first line of host defense. However, the impact of PFam54 in promoting *Bb*sl pathogenesis remains unclear because of the hurdles to simultaneously knock out all PFam54 proteins in a spirochete. Here, we found two *Borrelia bavariensis* (*Bbav*) strains isolated from patients, PBN and PNi, naturally lacking PFam54 but maintaining the rest of the genome with greater than 95% identity to the type-strain PBi. We found that PBN and PNi less efficiently survive in human serum than PBi. Such defects were restored by pre-incubating serum with two *Bbav* PFam54 recombinant proteins, BGA66 and BGA71, confirming the role of these proteins in providing complement evasion of *Bbav*. Further, we found that all three isolates remain detectable in heart tissue 21 days post subcutaneous mouse infection, supporting the non-essential role of *Bbav* PFam54 in promoting spirochete persistence in hosts. However, we found differences in tissue tropism between the three strains. This study identified and utilized isolates deficient in PFam54 to associate the defects with the absence of these proteins, building the foundation to further study the role of each PFam54 protein in contributing to *Bb*sl pathogenesis.

Importance

To establish infections, Lyme borreliae utilize various means to overcome the host's immune system. Proteins encoded by the PFam54 gene array play a role for spirochete survival *in vitro* and *in vivo*. Moreover, this gene array has been described in all currently available Lyme borreliae genomes. By investigating the first two *Borrelia bavariensis* isolates naturally lacking the entire PFam54 gene array, we showed that

both patient isolates display an increased susceptibility to human serum, which can be rescued in the presence of two PFam54 recombinant proteins. However, both isolates remain infectious in mice after intradermal inoculation, suggesting a non-essential role of PFam54 during long-term infection, but do show a decreased colonization of joint tissue. Furthermore, these isolates display high genomic similarity to type-strain PBi (>95%) and thus could be used in future studies investigating the role of each PFam54 protein in Lyme borreliosis pathogenesis.

Introduction

Lyme borreliosis (LB, also termed Lyme disease in North America) is the most common vector-borne disease in the northern hemisphere (29, 30) with estimated annual cases between 65,000 and 200,000 in Europe (173, 174) and between 30,000 and 476,000 estimated cases in the USA (175–177). This disease is caused by spirochetes belonging to the *Borrelia burgdorferi* sensu lato (*Bb*sl) species complex (30, 43, 46, 178) and are maintained in an enzootic transmission cycle between ixodid ticks, normally of the genus *Ixodes*, and various vertebrate reservoir host species (13, 43, 46, 179, 180). Most LB cases in North America are caused by the genospecies *B. burgdorferi* sensu stricto (s.s.) while additional genospecies, such as *B. afzelii*, *B. garinii*, and *B. bavariensis* are causative agents of LB across Eurasia (29, 30, 43, 46). *Borrelia bavariensis* is of particular interest in Europe as it has a high propensity to infect humans and is associated with severe LB manifestations (i.e. neuroborreliosis) (30, 181), while additionally being rarely recovered from *Ixodes* ticks collected in the field (52, 169, 171). *Borrelia bavariensis* was proposed as a genospecies in 2009 and validated in 2013 (94, 157), prior to which, it was considered to be a subtype (OspA serotype 4 or NT29-like) of its sister species *B. garinii* (157). *Borrelia bavariensis* appears to primarily be a rodent-adapted genospecies (39, 47, 94, 100), which is distributed across Eurasia (157) where it utilizes either the tick species *I. ricinus* (Europe) or *I. persulcatus* (Asia) as a vector (39). This genospecies exists in two distinct populations with a high diversity, ancestral population in Asia (39, 84, 93, 101) and a genetically homogenous, almost clonal population in Europe (39, 84, 93, 101).

To establish an infection, *Bb*sl must evade complement, an important pillar of innate immunity, either indirectly through the acquisition of complement regulators or directly through interactions with complement proteins (61, 62, 65–67). The complement system consists of three distinct pathways (classical, lectin, and

alternative) all leading to the cleavage of C3 to form activated C3b (14). This initiates the activation of other complement components ending in the assembly of the membrane attack complex (MAC) through recruitment of late stage complement proteins (C6, C7, C8, and C9); ultimately leading to bacterial cell lysis (14). Additionally, cleavage of C3 and C5 by the C3 and C5 convertases, respectively, leads to the release of anaphylatoxins C3a and C5a, which can recruit additional immune cells and, therefore, are integral in mounting further host immune responses to infection (14). Host cells control complement damage by utilizing membrane-bound or fluid-phase regulatory proteins (14). All complement regulators can terminate the complement cascade at specific activation levels to protect self-cells from complement-mediated damage (14).

Lyme borreliae produce diverse outer surface proteins that bind distinct host complement components resulting in complement inactivation (12, 61, 62, 68, 182, 183). In fact, several *Borrelia* proteins named CRASPs (**c**omplement **r**egulator **a**cquiring **s**urface **p**roteins) are capable of binding complement regulators belonging to the factor H protein family and thereby allow spirochetes to overcome the host's innate immune system (12, 61, 184). One well studied factor-H binding protein, CspA, belongs to the large paralogous protein family, PFam54, with members capable of binding complement regulatory proteins or even of direct interactions with complement components (65, 69, 71, 185, 186). Members of the PFam54 are encoded by genes predominantly arranged in a multi-gene array located at the terminal end of the linear plasmid (lp) 54 in all *Bb*sl genomes studied so far (85, 96, 187). The PFam54 gene array can be separated into five major lineages where lineages I, II, III, and V share one to one orthology among genospecies (187). Lineage IV, however, contains a variable number of paralogs and many genospecies displaying unique PFam54 paralogs not found in other genospecies (187). The *B. bavariensis* type strain PBi contains PFam54 paralogs belonging to all of these lineages, although most do not have a described function (187). Seven of these PBi PFam54 paralogs belong to lineage IV including *bga66* and *bga71* (187). These genes encode the proteins that bind to late stage complement proteins, C7, C8, and C9, and thereby inhibiting the assembly of a functional MAC, which confers resistance to human complement (69). In other genospecies, paralogs belonging to lineage IV have also been found to facilitate human and non-human factor H binding (12, 71, 72, 188). Specific non-lineage IV paralogs have been studied in *B. burgdorferi* s.s. where some even have proposed functions. For

example, both *bba64* (PBi paralog, *bga63*) and *bba66* (PBi paralog, *bga65*) are required for tick-to-host transmission (189, 190). Even so, there are still many open questions regarding what role the proteins encoded by the PFam54 gene array play during the enzootic cycle of *Bb*sl spirochetes, especially in genospecies besides *B. burgorferi* s.s..

We recently characterized the genomes of 33 Eurasian *B. bavariensis* isolates of which two European isolates, PBN and PNi, were found to contain a shorter lp54 in comparison to the type strain PBi (84). In this study, we show that both PBN and PNi are naturally lacking the entire PFam54 gene array. Such findings raise the possibility of using these strains to study the roles of these genes. In this study, we compared the complement evasion activity and infectivity conferred by a *B. bavariensis* isolate that has intact PFam54 genes (PBi) and the strains deficient of those genes (PBN, PNi). We provide new insights into the potential contribution of PFam54 proteins to facilitate spirochete survival *in vitro* and *in vivo*.

Table 3. Characteristics of *B. bavariensis* isolates utilized in this study.

Isolate	Genospecies	Year of culturing	Country	Biological origin	Disease Manifestation	lp54 length (kb)
PBi	*Borrelia bavariensis*	<1993	Germany	Human	neuroborreliosis	60.4
PBN	*Borrelia bavariensis*	1999	Germany	Human	neuroborreliosis	46.6
PNi	*Borrelia bavariensis*	2000	Germany	Human	lymphoma	46.6

Methods

Ethics statement

Collection of blood samples and consent documents were approved by the ethics committee at the University Hospital of Frankfurt (control number 160/10 and 222/14), Goethe University of Frankfurt am Main. All healthy blood donors provided written informed consent in accordance with the Declaration of Helsinki. The mouse experiments were performed in strict accordance with all provisions of the Animal Welfare Act, the Guide for the Care and Use of Laboratory Animals, and the PHS Policy on Humane Care and Use of Laboratory Animals. The mouse infection protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of the

Wadsworth Center, New York State Department of Health (Protocol docket number 19-451). All efforts were made to minimize animal suffering.

Bacteria cultivation and DNA extraction

Borrelia bavariensis isolates PBN, PNi, and PBi were provided by the German National Reference Centre for *Borrelia* from the Bavarian Food and Health Safety Authority. Bacterial isolates were maintained according to standard procedures (125, 126) either in inhouse-made MKP (125), in inhouse-made BSK-H medium (126) or a commercially available BSK-H medium (Bio&Sell, Germany) supplemented with 6% rabbit serum (Sigma-Aldrich, Germany). Once cultures reached a density of 1×10^8 cells per mL, whole genomic DNA was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany). DNA quality (260/280) and concentration were measured using a NanoDrop® 1000 photometer (Thermo Fisher Scientific, USA) and a Qubit® 3.0 fluorometer (Thermo Fisher Scientific, USA), respectively.

Next generation sequencing and genome assembly

All isolates were sequenced using both long- and short-read sequencing methods. Short read libraries were produced using the Nextera XT DNA Library Preparation Kit or the Illumina DNA Prep Kit (Illumina, San Diego, CA, USA) according to the manufacturer protocol. Library quality was evaluated using a Fragment Analyzer (Agilent, Germany) or an Agilent TapeStation 2200 (Agilent, Germany). Sequencing was performed using a MiSeq Reagent V2 kit on an Illumina MiSeq platform according to standard protocol (Illumina, San Diego, CA, USA) that produced paired end reads of 250bp. Illumina reads were first trimmed for Illumina MiSeq adapter sequences using Trimmomatic v. 0.38 (127) before being assembled using SPAdes v. 3.13.0 (111), which has been shown to be the best option for *de novo* assemblies of *Borrelia* genomes (83, 84). SPAdes contigs were then mapped to references using NUCmer v.3.23 from the package MUMmer (128, 129) as described in Becker et al. (2020).

Long read sequence data of PBN and PNi were generated using Pacific Biosciences single-molecule, real-time (SMRT) technology by the Norwegian Sequencing Center (www.sequencing.uio.no). Libraries were prepared from genomic DNA sheared to 12kb using Pacific Biosciences' protocol for SMRTbell™ Libraries and PacBio® Barcoded Adapters for Multiplex SMRT® Sequencing. Libraries were size

selected using 0.45 Amoure PB beads. The library was sequenced on a Pacific Biosciences Sequel instrument using Sequel Polymerase v3.0, SMRT cells v3 LR, and sequencing chemistry v3.0 (movie time: 15h). Reads were demultiplexed using the Demultiplex Barcodes pipeline on SMRT Link v8.0.0.80529 (SMRT Tools v8.0.0.80502). A minimum barcode score of 40 was used. Reads were assembled using the Microbial Assembly application on SMRT Link (v8.0.0.80529, SMRT Link Analysis Services and GUI v 8.0.0.80501). Long read sequence data of PBi is available at NCBI (Accession Number: CP028873.1) (32). In short, the library was generated using Pacific Biosciences 20 kb library preparation protocol. Size selection of the final library was performed using Blue-Pippin with a 10 kb cut-off. The library was sequenced on a Pacific Biosciences RS II instrument using P6-C4 chemistry with 360 min movie time. PacBio reads were assembled using HGAP v3 (Pacific Biosciences, SMRT Analysis Software v2.3.0). Overlapping contig regions due to the circularization during the PacBio Library preparation were removed. To polish indels and sequencing errors that may be present in the PacBio contigs, the Illumina short reads were mapped to the PacBio contigs and a consensus sequence was extracted using the CLC Genomic Workbench 11.

Characterizing PFam54 gene array in PBN, PNi, and PBi

Sequences for all PFam54 paralogs described for PBi in Wywial etal (2009) (*bga63-bga73*) were downloaded from GenBank (Accession Numbers: PBi, CP000015.1) and used as queries. We used BLAST v.2.8.1 (114, 115) (algorithm: *blastn*) to search for the PFam54 paralogs described above in all assembled contigs. Blast hits shorter than 500bp and with a percentage identity lower than 80% compared to PBi were not considered paralogous to their reference. Further BLAST hits were removed if they were overlapping with regions previously assigned to a result of higher quality. Presence or absence of PFam54 orthologs were further checked through paralog specific PCRs with primers designed for the paralogs present in PBi. All PCR products were cleaned using a Zymo DNA Clean and Concentrator-5 kit (Zymo Research, USA) and sequenced at the Sequencing Service of Ludwig-Maximilians University according to standard protocol for Sanger sequencing (http://www.gi.bio.lmu.de/sequencing/help/protocol). For further detail on PCR analyses see Supplementary Materials and Table S7. Gene orthology was confirmed through phylogenetic reconstruction performed in MrBayes (131, 132) based on all

GenBank references and extracted PFam54 paralog sequences from PBi, PBN, and PNi (see Supplementary Materials).

Human serum, proteins, and antibodies

Non-immune human serum (NHS) obtained from healthy volunteers was initially tested for anti-*Borrelia* IgM and IgG antibodies as previously described (191) and for complement activity employing the Wieslab® functional complement assays (SVAR, Malmö, Sweden). Only serum samples considered to be negative for anti-*Borrelia* antibodies and displaying complement activity were used to form the serum pool. Polyclonal anti-C3 antisera were obtained from Merck Biosciences (Bad Soden, Germany) and the neoepitope-specific monoclonal anti-C5b-9 antibody was purchased from Quidel (San Diego, CA, USA). Generation and purification of His-tagged BGA66 und BGA71 have been previously described (191). Alexa Fluor 488-conjugated anti-goat IgG and Alexa Fluor 488-conjugated anti-mouse IgG were purchased from ThermoFisher (Langenselbold, Germany).

Serum susceptibility assays

To test serum susceptibility of PBN and PNi to NHS, 1×10^7 highly viable *Borrelia* cells suspended in 50 µl BSK-H medium (Bio&Sell, Germany) were incubated with 50 µl NHS at 33 °C as described previously (192). *Borrelia bavariensis* type strain PBi was included as a serum-resistant control and *B. garinii* strain G1 as a serum-sensitive control. The number of motile cells were counted at different time points (0, 1, 2, 4, 6 h) using dark-field microscopy. Nine microscopy fields were counted for each time point per strain. Each assay was conducted at least three times.

BGA66 or BGA71 are known to facilitate serum resistance of *B. bavariensis* PBi (69), and to further explore if they could rescue serum sensitive *Borrelia* isolates from complement-mediated killing, a modified serum bactericidal assay was conducted. For this, 50 µl NHS was pre-incubated with either 10 µM purified BGA66, 10 µM purified BGA71, 10 µM bovine serum albumin (BSA), or a combination of BGA66 and BGA71 (5 µM each) for 15 minutes at 37 °C. Thereafter, 1×10^7 *Borrelia* cells suspended in 50 µl BSK-H medium were added to 50 µl of pre-incubated NHS and further incubated at 33°C. The number of motile cells were counted at different time points (0, 1, 2, 4, 6 h) using dark field microscopy. As further controls, reaction mixtures containing heat-

inactivated NHS (hiNHS), NHS alone, and wash buffer (50 mM Tris/HCl, pH 8.0) were also included. Each assay was conducted in triplicate.

Genetic characterization of serum survivors

To further characterize PBN and PNi cells which survived treatment with 50% NHS, spirochetes (labeled as PBN-ST or PNi-ST for post-serum treatment) were re-cultured in BSK-H medium (Bio&Sell, Germany) until they reached a density of 1 □ 10^8 cells per mL (125, 126) at which point whole genomic DNA was extracted using a Maxwell® 16 LED DNA kit (Promega, Germany). *Borrelia* DNA was then sequenced on a MiSeq platform using the protocol described above for short-read library preparation and assembly. We used BLAST v.2.8.1 (114, 115) (algorithm: *blastn*) to again search for the PFam54 paralogs described above and paralog specific PCRs to support the presence or absence of PFam54 paralogs in PBN-ST and PNi-ST (see Supplementary Materials). All assembled contigs of PBN-ST or PNi-ST were then mapped back to the original PBN or PNi assembly (84) using NUCmer v. 3.23 from the package MUMmer (128, 129). Single nucleotide polymorphisms (SNPs) and indels between the two assemblies were then called using the program *show-snps* from the package MUMmer (128, 129). The location of SNPs were compared to a previously published annotation for PBN and PNi (84) to determine if they were located in proposed open reading frames. Gene annotations without proposed function (i.e. hypothetical proteins) were subjected to a protein structure similarity search using the online-based HHpred server (193, 194).

Immunofluorescence microscopy

To determine deposition of activated complement components on the *Borrelia* surface, an immunofluorescence assay was conducted as previously described (195, 196). For this, 6 × 10^6 cells were suspended in 150 µl GVB^{++} and either 50 µl NHS or 50 µl hiNHS was added. Following incubation for 1 h at 37 °C, 10 µl of the suspension was transferred to a glass slide, air-dried overnight, and then fixed by using a glyoxal solution (Merck, Germany). Slides were then incubated for 1 h at 37 °C with either an anti-C3 (1:1000) or a neoepitope-specific anti-C5b-9 antibody (1:70) to detect C3 and the assembled MAC, respectively. After washing with PBS, Alexa Fluor 488 conjugated antibodies (1:1000) were applied and the slides were incubated for 1 h at 37 °C in the dark. After washing, the slides were overlayed with a DAPI solution (1:500) and

incubated for 10 min at 4 °C. Finally, the glass slides were enclosed with a coverslip using fluorescence mounting medium (Agilent Technologies Denmark ApS, Glostrup, Denmark). The evaluation was carried out with an Axio Imager M2 fluorescence microscope (Zeiss, Germany) equipped with a Spot RT3 camera (Visitron Systems, Germany).

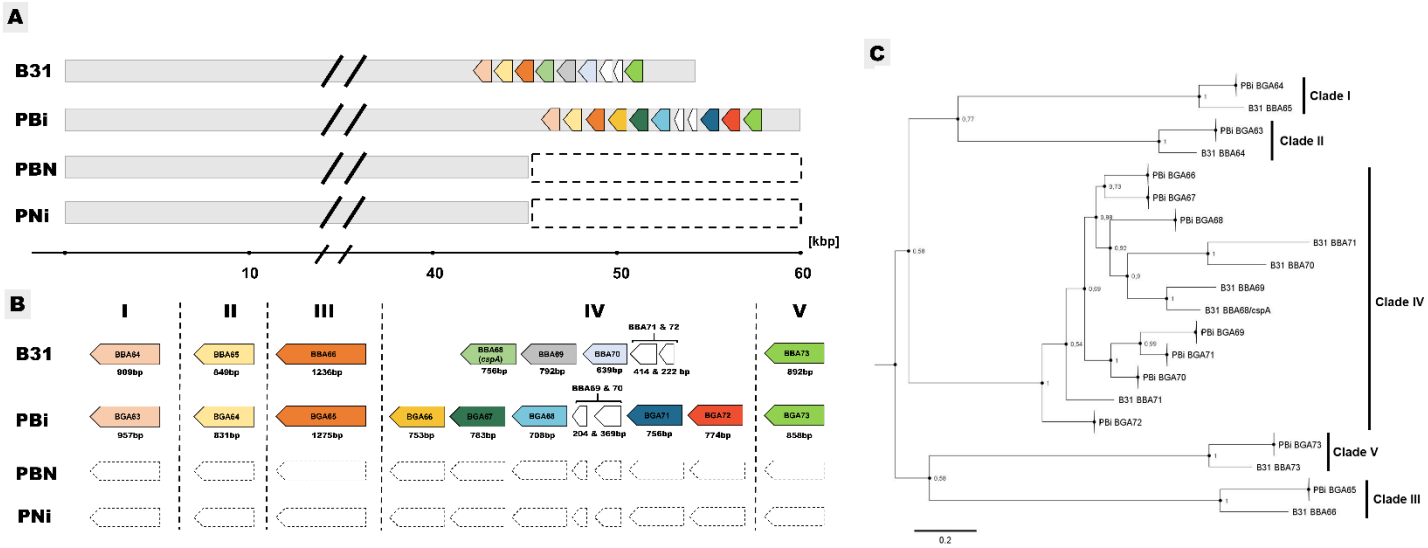


Figure 5. A) Schematic overview of the aligned lp54 sequences for the *B. bavariensis* isolates PBi, PBN, PNi, and *B. burgdorferi* s.s. strain B31. Both, PBi and B31, have longer lp54 sequences and contain the PFam54 gene array at the 3' end while PBN and PNi have a shorter lp54 and lack ~14kb from the 3' end of the aligned sequences, where the PFam54 gene array is located. Individual genes are shown as arrows and colors denote orthologous genes between isolates. Arrow direction denotes on which DNA strand the gene is found. B) Synteny map of the PFam54 gene array. Genes are colored according to orthology and broken down into the five clades originally described by (187). These colors correspond to panel A. PBN and PNi are missing all orthologs belonging to all five clades (shown as dashed outlines in reference to the PBi paralogs). C) Phylogeny to check orthology of PBi and B31 PFam54 paralogs run in MrBayes (see Supplementary Material for further details). References for PBi and B31 were downloaded from GenBank for all PFam54 paralogs present in Wywial et al., (2009) (B31, Accession Number: AE000790.2; PBi, Accession Number: CP000015.1). PBi clades include the orthologs found in the PacBio assembly from Margos et al., (2018). All PBi sequences were the same and are displayed as a single node.

Intradermal infection of Mus musculus mice and the quantification of bacterial burdens in tissues

Four-week-old female BALB/c mice were intradermally inoculated with 1×10^5 of *B. bavariensis* strains PBi, PBN, or PNi as described previously (71). Control mice were inoculated in parallel with BSK-H medium unmodified with serum. As a plasmid profiling procedure for these isolates is not currently available, the isolates were cultured at less than ten passages to avoid decreased infectivity due to potential plasmid loss events during *in vitro* cultivation. At 21 days post-infection, mice were sacrificed and harvested for organs and tissues including the inoculation site of the skin, the ankle joints, and heart. These were collected and processed for quantitatively

assessment of bacterial burdens during infection. DNA was extracted from tissues using the EZ-10 Genomic DNA kit (Biobasic, Canada), and the quantity and quality of DNA was assessed using a Nanodrop 1000 UV/Vis spectrophotometer (ThermoFisher Scientific, USA). The 280:260 ratio of these samples was between 1.75 and 1.85, indicating no contamination of RNA or proteins in those DNA samples. Quantitative PCR was performed as described previously (71). In brief, spirochete genomic equivalents were calculated using an ABI 7500 Real-Time PCR System (ThermoFisher Scientific, USA) in conjunction with PowerUp SYBR Green Master Mix (ThermoFisher Scientific, USA), based on amplification of the Lyme borreliae 16S rRNA gene using primers 16SrRNAfp and 16SrRNArp (Table S7) as described previously (74, 197). Cycling parameters for SYBR green-based reactions were 50°C for 2min, 95°C for 10min, and 45 cycles of 95°C for 15s, 52°C for 30s, and 60°C for 1min. The number of 16S rRNA copies was calculated by establishing a threshold cycle (Cq) standard curve of a known number of 16S rRNA gene extracted from each *B. bavariensis* strain PBi, PBN, or PNi, then comparing the Cq values of the experimental samples for each respective strain.

Statistical analysis.

Differences in serum sensitivity including trials utilizing pre-incubated NHS were tested using pairwise t-tests with a Bonferroni multiple testing correction in R v. 3.6.1 (136), while for the mouse experiments, differences were tested using the the Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli. For all tests, a P-value < 0.05 (*) was considered significant.

Results

PBN and PNi naturally lack the entire PFam54 gene array

In both long (PacBio) and short-read (MiSeq) assemblies PBN and PNi had a shorter lp54 of approximately 46.6 kbp in comparison to type strain PBi with a lp54 of approximately 60.4 kbp (Table 3). Blast searches conducted for PBN and PNi returned no hits for the PBi PFam54 paralogs either on the reconstructed lp54 or when blasted to all assembled contigs, whereas all paralogs (n=11) were found in our PBi assembly (Figure 5). These findings were further confirmed by using a paralog-specific PCR approach which showed no amplicon for all PFam54 paralogs in PBN and PNi except for the primer pair targeting *bga68* (Figure S4). When using genomic DNA of PBN and

PNi as a template and primers specific to amplification of *bga68*, we observed an unexpected PCR product of approximately 1800bp (much larger than the expected product of 1369bp). Following sequencing of these amplicons, sequences corresponded to a region on the *B. bavariensis* chromosome but not to any of the PFam54 paralogs. When DNA from PBi was used as a positive control, bands of expected size were observed for all PFam54 paralogs (Figure S4) and were confirmed through Sanger sequencing. All primer pairs were designed for PBi as the PFam54 paralogs are present in this isolate. As European *B. bavariensis* are characterized by a low-diversity and are almost clonal (39, 84, 93), primers designed for PBi should amplify the genes if present in PBN and PNi. Therefore, the lack of observed products is not a result of potential sequence polymorphisms in primer binding sites.

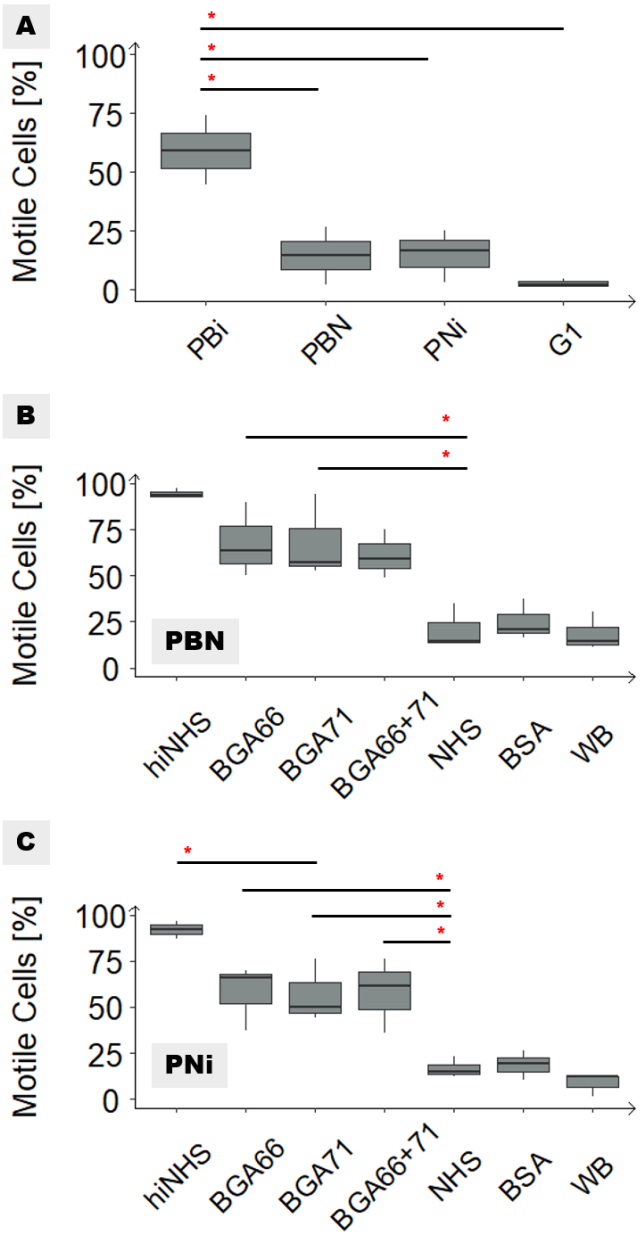


Figure 6. Serum susceptibility of *B. bavariensis* strains PBN and PNi lacking the PFam54 gene array after 6 h incubation at 33°C. *Borrelia bavariensis* strain PBi and *B. garinii* strain G1 were included as serum-resistant and susceptible controls, respectively. A) Percentage of motile PBN, PNi, PBi and G1 after 6 hours of incubation in 50% NHS. B) Percentage of motile PBN cells after 6 hours incubation with NHS which was pre-incubated with 10µM purified BGA66, BGA71 or a combination of BGA66 and BGA71. C) Percentage of motile PNi cells after 6 hours incubation with NHS which was pre-incubated with 10µM purified BGA66, BGA71 or a combination of BGA66 and BGA71. For both cases PBN (B) and PNi (C) controls were included where cells were incubated with NHS, heat-inactivated NHS (hiNHS), wash buffer (WB), or bovine serum albumin (BSA). All tests were done in triplicate. Significance was tested using pair-wise t-tests with a Bonferonni multiple testing correction and significant differences are marked with red *. For all time points (0, 1, 2, 4, 6 hours) see Figure S5.

PBN and PNi display an increased serum susceptibility to human serum

Having demonstrated the absence of the entire PFam54 gene array in PBN and PNi, including *bga66* and *bga71*, serum bactericidal assays were conducted *in vitro* to explore the strains capability to overcome complement-mediated killing. PBi and *B. garinii* strain G1 were included either as a serum resistant or serum sensitive control, respectively. Both PBN and PNi had a significantly higher susceptibility to NHS in comparison to PBi, with approximately 20% of PBN or PNi cells surviving after 6 h as compared to around 60% of PBi cells surviving (Figure 6A). Additionally, PBN and PNi did not significantly differ from G1, where under the same conditions around 90% of cells were killed after 6 hours (Figure 6A). When NHS was pre-incubated with purified recombinant proteins of PBi PFam54 paralogs BGA66 and BGA71 (69), both PBN (Figure 6B) and PNi (Figure 6C) were rescued to serum resistance levels comparable to PBi after 6 hours of incubation (Figure 6B-C). Only in the case of PNi incubated with BGA71-treated NHS were there significantly fewer motile cells in comparison to PBi (Figure 6C). Moreover, each recombinant protein was able to confer serum resistance while a combination of both did not increase the resistance level compared to the individual molecule (Figure 6B-C). Refer to Figure S5 for all serum assay time points. To further support PBN and PNi's increased susceptibility to complement, deposition of activated complement components of C3 (i.e. C3b) and the MAC was investigated by employing immunofluorescence microscopy. These complement activation assays revealed that, in contrast to the serum-resistant PBi, the majority of PBN and PNi cells showed deposition of C3 including activated C3b on the spirochetal surface (Figure 7A to C). Moreover, the staining pattern for PBN and PNi was similar to serum-sensitive *B. garinii* G1 (Figure 7D). The same staining pattern could be observed in relation to the pore-forming MAC, when cells were stained for late stage complement proteins (C5b to C9) (Figure S6). #

Both PBN and PNi were isolated from patients and are hypothesized to be clonal populations representing a single isolate. To determine though if the serum susceptibility pattern of PBN and PNi could result from these isolates representing mixed cultures of two unique isolates (one serum-sensitive and one serum-resistant), the remaining spirochetes which survived NHS treatment after 6 h were transferred to fresh BSK-H medium and grown until they reached the exponential phase. After DNA isolation, PBN-ST and PNi-ST were sequenced using the Illumina MiSeq platform (see Methods). The genomes of PBN-ST and PNi-ST were identical to the original PBN and

PNi genomes, including all plasmids and the shortened lp54 (84). Only very few variants were observed in repeat regions or regions which are known to be challenging to assemble using Illumina data (83, 84). The only notable exception was a single non-synonymous mutation in a hypothetical protein found on lp28-9 of PBN-ST only. A HHPred structure based search (193, 194) suggested similarities with Dynamin-like proteins from *Bacillus* species. No BLAST hits for PFam54 paralogs were found in PBN-ST and PNi-ST and they did not produce amplicons the PFam54 paralog specific PCRs (Figure 5; Figure S5).

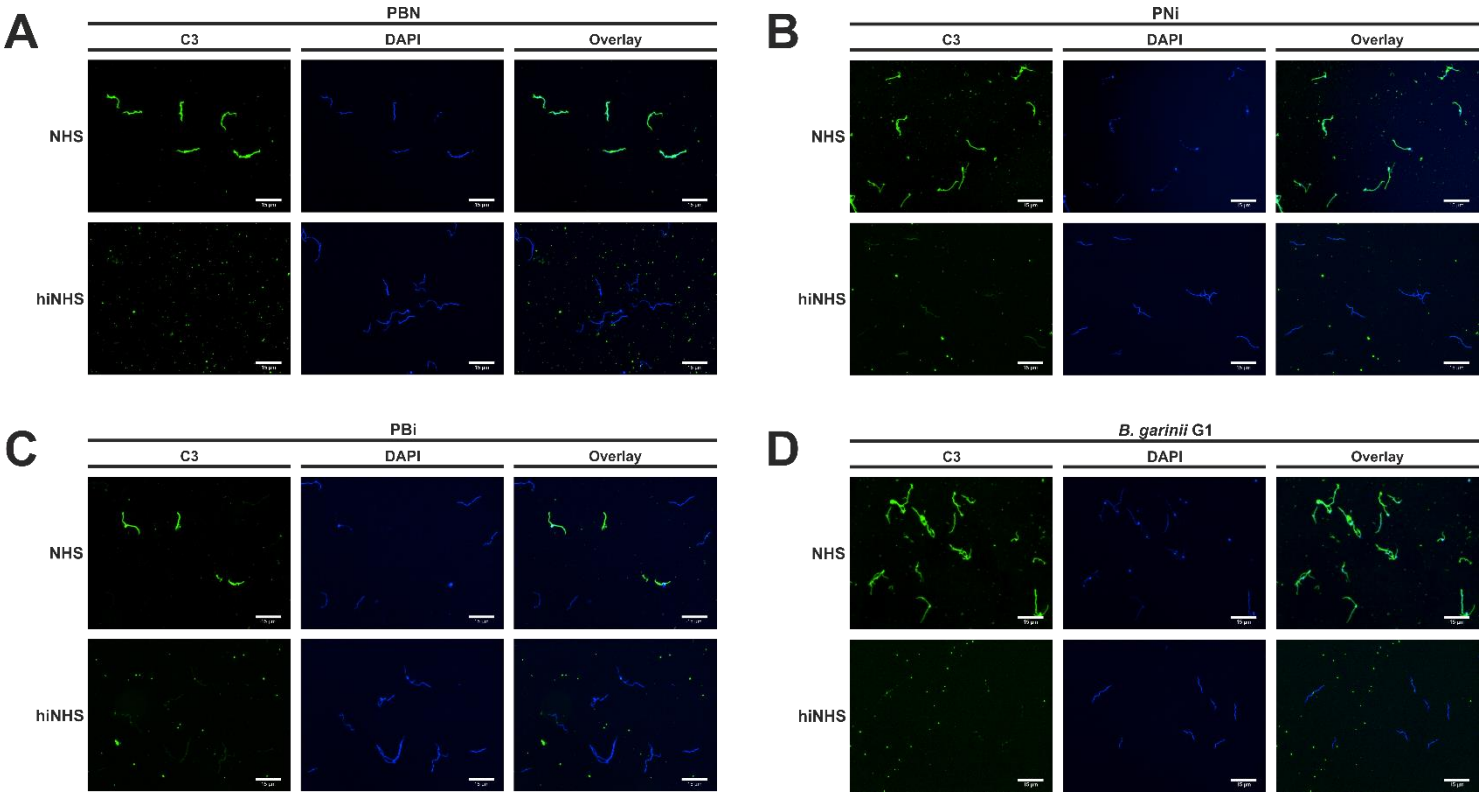


Figure 7. Deposition of complement component C3 after challenging viable spirochetes from PBN (A), PNi (B), PBi (C), and *B. garinii* G1 (D) with NHS (top row) and heat-inactivated NHS (hiNHS, bottom row). After fixation, deposition of C3 (green) were visualized with a polyclonal anti-C3 antibody (1:1000) and spirochetal DNA (blue) was stained by using DAPI. Shown are representative data from two independent experiments. All scale bars are equal to 15µm. The spirochetes were observed at a magnification of 1,000x. The data were recorded with an Axio Imager M2 fluorescence microscope (Zeiss) equipped with a Spot RT3 camera (Visitron Systems).

PBN and PNi remain infectious to Balb/c mice, but show a trend for less efficient joint colonization after intradermal inoculation

We further determined the infectivity of PBi, PBN, and PNi *in vivo* by intradermally inoculating each of these isolates into Balb/c mice, and a mock infected group with culture medium as a control. At 21 days post inoculation, we found that the

spirochete burdens at the inoculation site of skin from the mice infected with PBi, PBN, or PNi were indistinguishable from those in the mock-infected group (Figure 8A). Conversely, the spirochete burdens in heart tissue derived from each of these three isolates did not statistically differ from each other but was significantly higher than those from mock-infected mice (Figure 8B). Further, the spirochete burdens at joints from PBi-infected mice but not PBN- or PNi-infected mice were significantly greater than those burdens from mock-infected mice, although the three isolates did not significantly differ among each other (Figure 8C). These results demonstrate the ability of PBi, PBN, and PNi to yield productive infections in mice, suggesting that PFam54 is not essential for spirochete persistence in mice after intradermal infection. However, the differences in the spirochete burdens at joints from PBi-, PBN-, and PNi-infected mice in comparison to those in mock-infected mice implies a strain-specific joint colonization at late stages of infection.

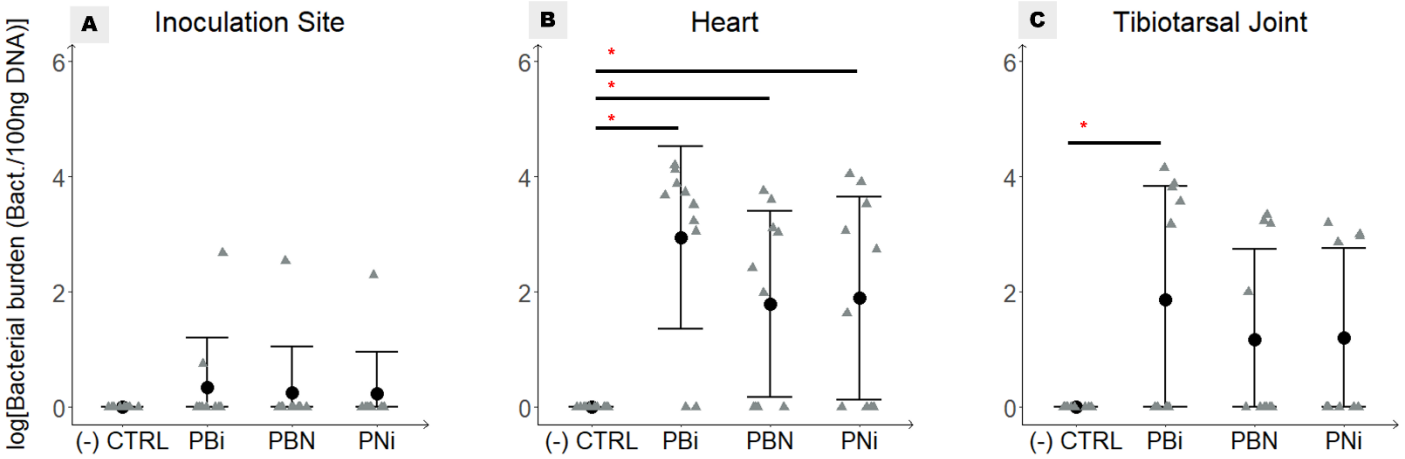


Figure 8. Bacterial burden of mice at the inoculation site (A), heart (B), and tibiotarsal joint (C) 21 days past needle inoculation with 1×10^5 of PBN, PNi, PBi or BSK-II medium not supplemented with serum as a negative control. After 21 days, mice were sacrificed, and bacterial burden was determined using a qPCR targeting the 16S rRNA gene of *Borrelia* which was normalized to the total amount of DNA. Differences between bacterial burdens were tested for significance using the Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli. Significant differences are marked with red *.

Discussion

Lyme borreliae utilize a number of sophisticated strategies to successfully infect and colonize a host, including evading diverse host immune responses to survive in the hosts bloodstream, tissues, and organs (12, 61, 62, 184). Proteins encoded by the PFam54 gene array play an important role in some of these processes (12, 71, 185, 186, 198, 199), and these genes have been found in all sequenced *B. burgdorferi* s. l.

genomes to date (85, 86, 96). Only a few of these proteins have been functionally characterized and despite their high sequence similarity they were found to have different functions (69, 199, 200). Simultaneously knocking out PFam54 proteins from a spirochete to study the functions additional PFam54 proteins is not feasible because each spirochete carries more than five Pfam54 proteins in any given Lyme borreliae species or strain (187). In a recent study using Illumina MiSeq assemblies of *B. bavariensis* genomes (84), we found that isolates PBN and PNi contain a shorter lp54 plasmid of 46.6 kb compared to lp54 of type strain PBi (60.4 kb) (32, 84). For confirmatory purposes, long-read sequencing (PacBio) was conducted as a combination of long and short sequencing technologies enables proper plasmid reconstruction (83). Comparative analyses of those genomes clearly indicated that the entire PFam54 gene array was missing in PBN and PNi. Such a novel identification of two isolates deficient in the entire PFam54 provides a tool to study the functions of PFam54 *in vitro* and *in vivo* by comparing the phenotypes of these strains with the strains that carry the entire PFam54 (eg. PBi).

We hypothesized that the absence of these genes in PBN and PNi could lead to an increased susceptibility to human complement, which was supported by our analyses. Compared to PBi, we found that PBN and PNi displayed decreased survivability in human serum and increased deposition of the major complement component, C3. This result is consistent with the fact that PFam54 from PBi encodes BGA66 and BGA71, two proteins that bind to C7, C8, and C9 to inactivate terminal complement pathway in promoting the serum survivability in a gain-of-function spirochete background in *B. garinii* (69). In fact, we found that the presence of BGA66 and BGA71 rescued the serum sensitivity and complement deposition of PBN and PNi to levels indistinguishable from those from PBi. These results strongly support the role of PFam54 in promoting complement evasion (12, 69, 71, 199, 201).

Both PBN and PNi, were isolated from patients with confirmed LB (Table 3) indicating that both isolates were able to establish human infection. Unfortunately, the time point at which PBN and PNi lost this region of the lp54 could not be determined. The loss could have occurred before or during the infection process or during *in vitro* cultivation as described previously for other plasmids (83, 202, 203). Interestingly, spirochetes that survived serum treatment after 6 h of incubation (PBN-ST and PNi-ST) did not differ along their entire genome, including the shortened lp54, indicating that spirochetes who are killed and those who survived have the same genetic makeup.

Serum survival of a subpopulation, as also observed in the serum sensitive control isolate *B. garinii* G1, could be a result of a stochastic process involving transcriptional changes due to external signals which equip spirochetes with complement-interfering factors. *Borrelia* transcriptomes, including members of the PFam54 family, do change in response to stressors stemming from both reservoir hosts (temperature, pH) and tick vectors (low nutrients, oxidative stress) (46, 77, 204). Utilizing a transcriptional approach could therefore highlight which genes are transcribed by these two isolates while under serum stress to understand which mechanisms PBN and PNi, or even a subpopulation thereof, utilize to overcome complement-mediated killing. *Bb*sl genomes are quite redundant, containing a number of genes with similar functions (82, 85, 205). Therefore, the presence and potential variation in other known anti-complement lipoproteins may contribute to such a strain-to-strain variation in complement evasion (12, 68, 71, 182). Additionally, considering the majority of *Bb*sl plasmid encoded proteins are either hypothetical or uncharacterized (96), it is likely that there are further gene products which confer serum resistance through yet uncharacterized immune evasion mechanisms (12, 61, 62, 66) and are contributing to the observed phenotype of PBN and PNi.

Though PBN and PNi are more susceptible to human complement-mediated killing in vitro than PBi, when we compared the late stages colonization (21 days post intradermal infection in mice) of these strains with that of PBi, we did not notice differences among the isolates in heart and the initial infection site of the skin. Our results still suggest that Pfam54 is not essential for persistent survival of spirochetes, and that the persistence of PBi, PBN, and PNi is independent on the ability of these isolates to evade complement. In fact, such a complement-independent persistence is consistent with several documented studies showing that the role of complement in discriminating strain-to-strain differences of infectivity is more apparent at early infection onsets than at the later stages of infection (206, 207). Hart et al., (2018) reported that CspA and its orthologous proteins from several isolates of *B. burgdorferi* s.s. or *B. afzelii* confer spirochete survival in fed ticks, and thus promote tick-to-host transmission (71). Other *B. burgdorferi* s.s. PFam54 paralogs (i.e., *bba64*, *bba73*) are additionally upregulated during tick feeding suggesting that PFam54 genes, and encoded products, may be important to tick-to-host transmission (77). As PBi contains paralogs of these upregulated genes which PBN and PNi are lacking, this could hint at an additional possibility that PBN and PNi may differ in their transmissibility during

tick infection or ability to successfully colonize the tick vector, requiring additional investigation.

We found that PBN, PNi, and PBi colonized heart at indistinguishable levels, with burdens greater than the burdens in mock-infected mice, but the spirochete burdens of all three tested isolates at the initial infection site in the skin did not differ from those derived from that tissue of mock-infected mice. Such a heart-specific colonization by *B. bavariensis* is consistent with previous findings for PBi (208). This also suggests complete clearance of these *B. bavariensis* isolates at the initial infection site, unlike isolates from other Lyme borreliae genospecies (*B. burgdorferi* s.s., *B. afzelii*, and *B. garinii*) (74, 208–210). We did find, however, that spirochete burdens of PBi but neither PBN nor PNi in the tibiotarsal joints were significantly higher compared to control mice. Such a strain-to-strain difference in tissue tropism is similar to findings obtained with strains of other Lyme borreliae genospecies (49, 211–213). A possible contributing factor for such a strain-specific tissue tropism could be the presence or absence of spirochete adhesins or the polymorphisms in said adhesins. For example, some PFam54 proteins, including *bga71* from *B. bavariensis*, were shown to confer spirochete attachment to mammalian cells (185, 186, 200). Additionally, some spirochete polymorphic adhesins, e.g. outer surface protein C (OspC) and decorin binding protein A (DbpA), promote differential levels of tissue colonization among *Borrelia* isolates and genospecies (74, 75, 184). Although PBN, PNi, and PBi do share the same *ospC* (84) and *dbpA* sequences (unpublished data) suggesting these are most likely not causing the observed phenotype. Taken together, our results suggest that the PFam54 proteins have other functions in addition to complement resistance and could be adhesins that contribute to specific tissue tropism.

The PFam54 gene array is found across the *Bb*sl species complex and therefore has been maintained throughout the evolutionary history of these spirochetes (88, 187). The genospecies besides *B. burgdorferi* s.s., including *B. bavariensis*, have been challenging to be genetically modified (214). Thus, isolates PBN and PNi, with almost identical genetic make-ups to type strain PBi (>95% identity over all genomic compartments) (84) and naturally lacking the PFam54 array, offer a unique opportunity to study the role of these proteins in infection and tick-to-host transmission in humans and reservoir hosts. Taken together, the natural loss of the PFam54 gene array in *B. bavariensis* is associated with an increase of serum susceptibility and complement deposition and tissue-dependent variable persistence

after intradermal inoculation in mice. Such findings pave the road to further study the function of each PFam54 protein in *B. bavariensis* and other spirochete species and their contribution to the Lyme borreliae infectivity.

Paper 6: Supplementary Material

Disentangling the role of *Borrelia bavariensis* PFam54 proteins in vitro and in vivo using two strains naturally lacking the PFam54 gene array

Robert E. Rollins, Janna Wülbern, Florian Röttgerding, Tristan Nowak, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Yi-Pin Lin, Peter Kraiczy, Noémie S. Becker

Unpublished Manuscript

Disentangling the role of *Borrelia bavariensis* PFam54 proteins *in vitro* and *in vivo* using two strains naturally lacking the PFam54 gene array

Unpublished Manuscript

Robert E. Rollins, Janna Wülbern, Florian Röttgerding, Tristan Nowak, Sabrina Hepner, Volker Fingerle, Gabriele Margos, Yi-Pin Lin, Peter Kraiczy, Noémie S. Becker

— Supplementary Methods —

Primer design and PCR analysis to confirm absence of PFam54 paralogs in PBN and PNi

Primers for PBi PFam54 paralogs were either taken from (69) including genes *bga66*, *bga68*, *bga71*, *bga72*, and *bga73* or designed in this study (*bga63*, *bga64*, *bga65*, *bga67*). For primer design, GenBank PBi PFam54 genes (Accession Number: CP000015.1) were used as references and primers were designed in Primer3 (215–217) with default setting and maximum oligonucleotide length set to 35bp. For each gene, a maximum of five primer pairs were kept and blasted using BLAST v.2.8.1 (114, 115) (algorithm: *blastn,* task: short-blast) back to the full assembled MiSeq contigs of either PBN, PNi, or PBi. The plasmid pair with the highest specificity was ordered. PBi was chosen as a reference for the PFam54 genes as they are all present in this isolate (187). As European *B. bavariensis* is characterized by a low-diversity, almost clonal distribution (39, 84, 93), primers designed for PBi should amplify the gene if present in PBN and PNi.

For all genes, PCR samples were prepared using 1 µM primer concentrations and 10 ng of DNA using the S7 Fusion Polymerase system according to standard protocol for a final reaction volume of 20 µL (IsoGene Scientific, Netherlands). Each test included DNA from PBN, PNi, PBN-ST, PNi-ST (where ST stands for post-serum treatment). PBi DNA was included as a positive control and for each gene a non-template (water) negative control was included. For all genes besides *bga68* and *bga73*, PCR conditions were initialization at 95°C for 2 minutes, followed by 40 cycles of each 95°C for 15 sec, 60°C for 45 sec, and 72°C for 60 sec, followed by an elongation

step at 72°C for 10 minutes. For *bga68* and *bga73*, PCR conditions were initialization at 95°C for 2 minutes, followed by 40 cycles of each 95°C for 15 sec, 63°C for 45 sec, and 72°C for 60 sec, followed by an elongation step at 72°C for 10 minutes. Additionally, a PCR for the housekeeping gene for *recG* was included as a positive control for all DNA samples using the primers and protocol described in (151).

All PCR products were visualized on a 1% agaraose gel in 1×TAE buffer. Products for primers designed in this project (*bga63*, *bga64*, *bga65*, *bga67*) were cleaned using the Zymo DNA Clean & Concentrator™-5 kit (Zymo Research, USA) according to standard protocol for PCR amplicons. Cleaned PCR products were sequenced using Sanger sequencing at the Sequencing Service of Ludwig-Maximilians University and prepared in accordance with their requirements (http://www.gi.bio.lmu.de/sequencing/help/protocol). PCR products were sequenced with both forward and reverse primers. Consensus sequences were then assembled in UGENE (218) using CAP3 (219). Sequences were then aligned back to their respective genes to determine adherence to the expected sequence.

Phylogenetic analysis of located PFam54 paralogs

Sequences for our PBi assembly were complied with references for PFam54 paralogs as described in (187) from *B. bavariensis* (PBi) and *B. burgdorferi* s.s. from GenBank and were aligned as translated amino acids using Muscle v. 3.8.425 (220, 221) in AliView v. 1.26 (222). Phylogeny reconstruction was done in MrBayes v. 3.2.6 (131, 132) with ploidy set to haploid and a GTR (133) substitution model with inverse gamma distributed rate variation. Three independent runs were launched and ran for 10 million generations at which point convergence of parameters was checked with Tracer v. 1.7.1 (134). Consensus trees were built using the *sumt* command from MrBayes using a respective burn-in of 25%. Convergence to a single topology in all three independent runs was checked manually in FigTree v. 1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/) which was also used to plot the tree shown in Figure 5.

Table S6. Overview of all primers used in this study.

Gene	Reference	Primer	Sequence (5' to 3')	Amplicon (bp)	Citation
BGA63	PBi	*BGA63_PBi_F1*	AACTGGGCTAATTTTGCTTTC	1161	This study
		BGA63_PBi_R1	TTACTGAATTGGGGCAAGAA		
BGA64	PBi	*BGA64_PBi_F3*	TAACATTGGGGATAATAACATTT	686	This study
		BGA64_PBi_R3	ATCGTATTGCAGCTCTAAGG		
BGA65	PBi	*BGA65_PBi_F1*	GCTCAACAGATGATCAAGCAAAGA	1481	This study
		BGA65_PBi_R1	AAGCTGTGATTTTGATTCTCCTG		
BGA66	PBi	*BGA66_F_PK*	CGTTGCACTTGATATTTTTTAAGAAGAGAAGC	1395	(Hammerschmidt et al., 2016)
		BGA66_R_PK	GGCTATGCACTTTAAAGGTATTAATGATTTAATTTTCAAGATG		
BGA67	PBi	*BGA67_PBi_F1*	CATCTCATGTGCCGTTAATAAAAATT	1001	This study
		BGA67_PBi_R1	AGTTGTTCTGCACTGGTTTTAATGT		
BGA68	PBi	*BGA68_F_PK*	CATTAAAAACCTATGTGTACGAACTAGCAGCATATGGAG	1369	(Hammerschmidt et al., 2016)
		BGA68_R_PK	GTTAATGTCGTTGTAAGAATATTAAGCTTAATT		
BGA71	PBi	*BGA71_F_PK*	GGCAGATATATGAAGTTTGTAAAGAACTTGGTACTTTTAAGCCTATTG	1406	(Hammerschmidt et al., 2016)
		BGA71_R_PK	GTTTAGTTTTTGGCAAAATCCTTTCCTTATAACTAATATTTTAACC		
BGA72	PBi	*BGA72_F_PK*	CTAAAAGCCTGTTGTGATCACATACACTACTAATTGG	1298	(Hammerschmidt et al., 2016)
		BGA72_R_PK	CTTTAATTATATTTAGTTTAGTTTTTGTCAAAATACTCTCC		
BGA73	PBi	*BGA73_F_PK*	GTTCTACAACATTGGATTGAGATTGTAGAAACGCTTAC	1779	(Hammerschmidt et al., 2016)
		BGA73_R_PK	GTATCAATTTGATCAACAAGTGAGCATGAGATGCAC		
recG	-	*recF890*	CGCTTGTTGCCTTGCTTTC	741	(Margos et al., 2008)
		recR1694	GAAAGTCCAAAACGCTCAG		
16S rRNA	-	16SrRNAfp	GCTTCGCTTGTAGATGAGTCTGC	-	(Lin et al., 2014; Marcinkiewicz et al., 2018)
		16SrRNArp	TTGCAGTGTGACCGTTCACC		

Figure S4. Agarose gel of PCR products supporting the presence or absence of genes in PBN, PNi, and PBi. In all panels, numbers correspond to PBN (1), PBN-ST (2), PNi (3), PNi-ST (4), PBi as positive control (5), no template negative control (6). For all gels, a 1-kb Plus DNA ladder (ThermoFisher, USA) was used and band sizes are shown at the left in (A). Absence of PFam54 genes were tested for *bga63* (B), *bga64* (C), *bga65* (D), *bga66* (E), *bga67* (F), *bga68* (G), *bga71* (H), *bga72* (I), and *bga73* (J). The MLST housekeeping gene, *recG* (J), was further included as a positive control.

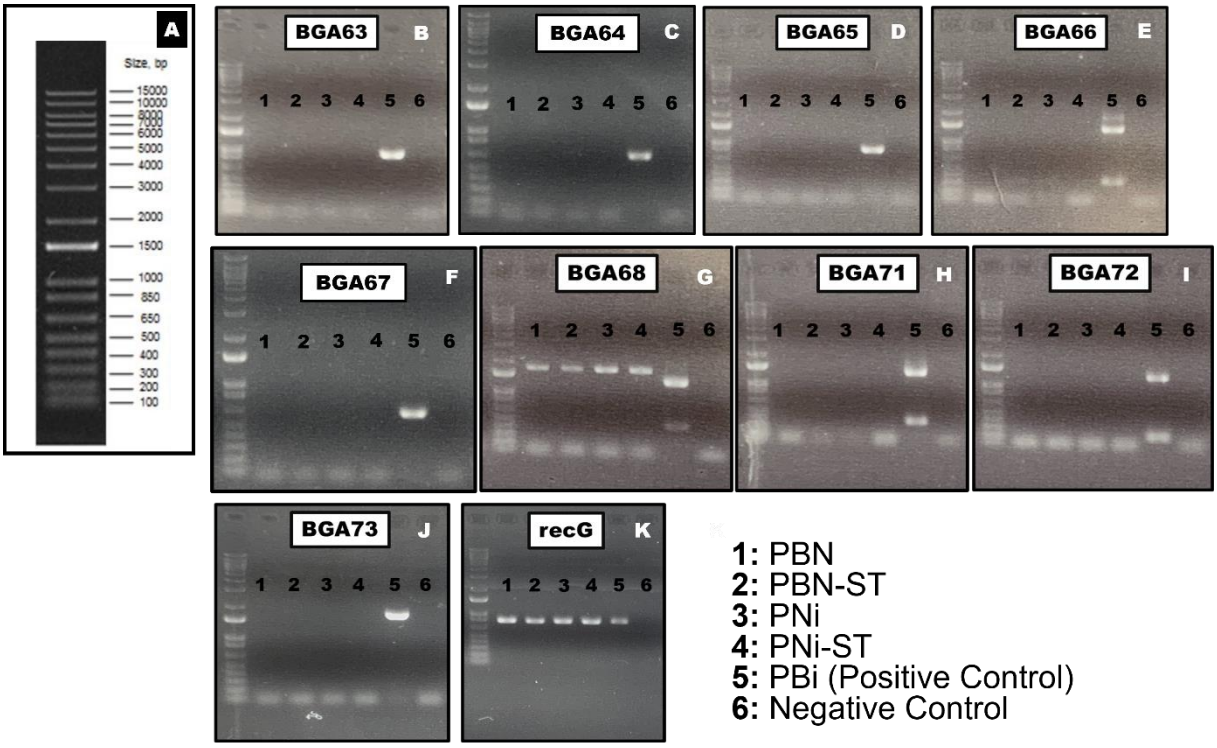


Figure S5. Serum susceptibility assays of PBN and PNi lacking the PFam54 gene array. A) Survival of *B. bavariensis* PBN, PNi, and PBi as well as *B. garinii* G1 in 50% NHS was monitored by dark field microscopy. The number of motile cells after incubation of *Borrelia* cells were determined at 0, 1, 2, 4, and 6 h. At least three independent experiments were conducted, each with similar result. Data from all experiments are shown. B-C) BGA66 and BGA71 rescue PBN and PNi from complement-mediated killing. NHS was pre-incubated with 10µM purified BGA66, BGA71, or a combination of BGA66 and BGA71. PBN and PNi were then incubated in 50% pre-incubated NHS and the number of motile PBN (B) and PNi (C) were recorded. Cells incubated with NHS, heat-inactivated NHS (hiNHS), wash buffer (WP), and bovine serum albumin (BSA) were included as controls. All tests were done in triplicate and error bars reflect standard deviation.

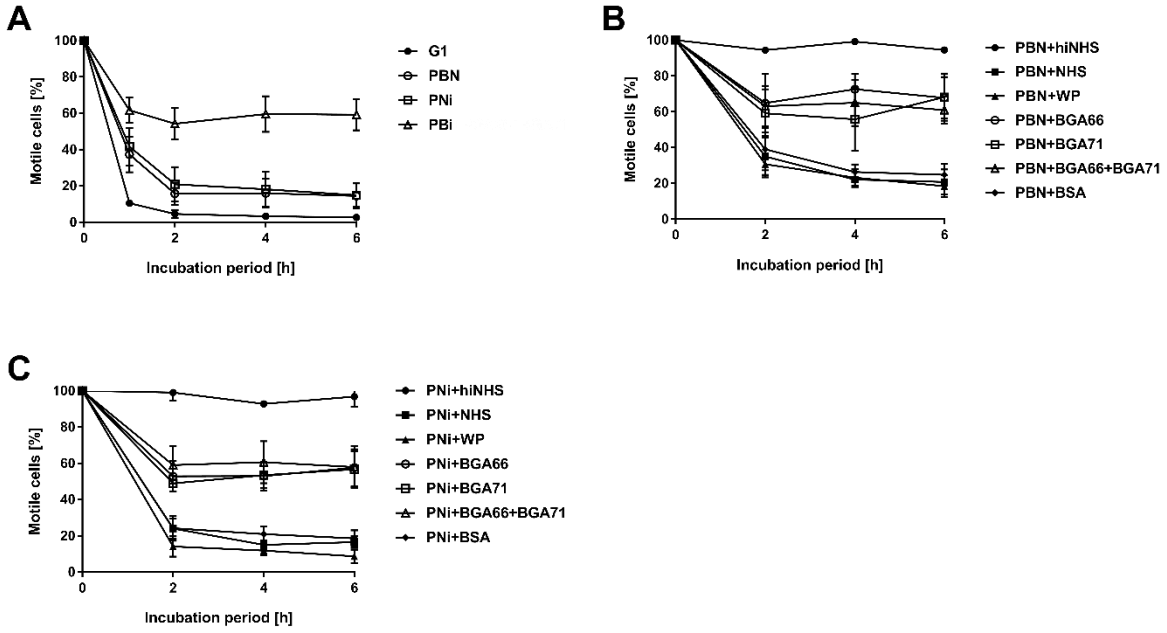
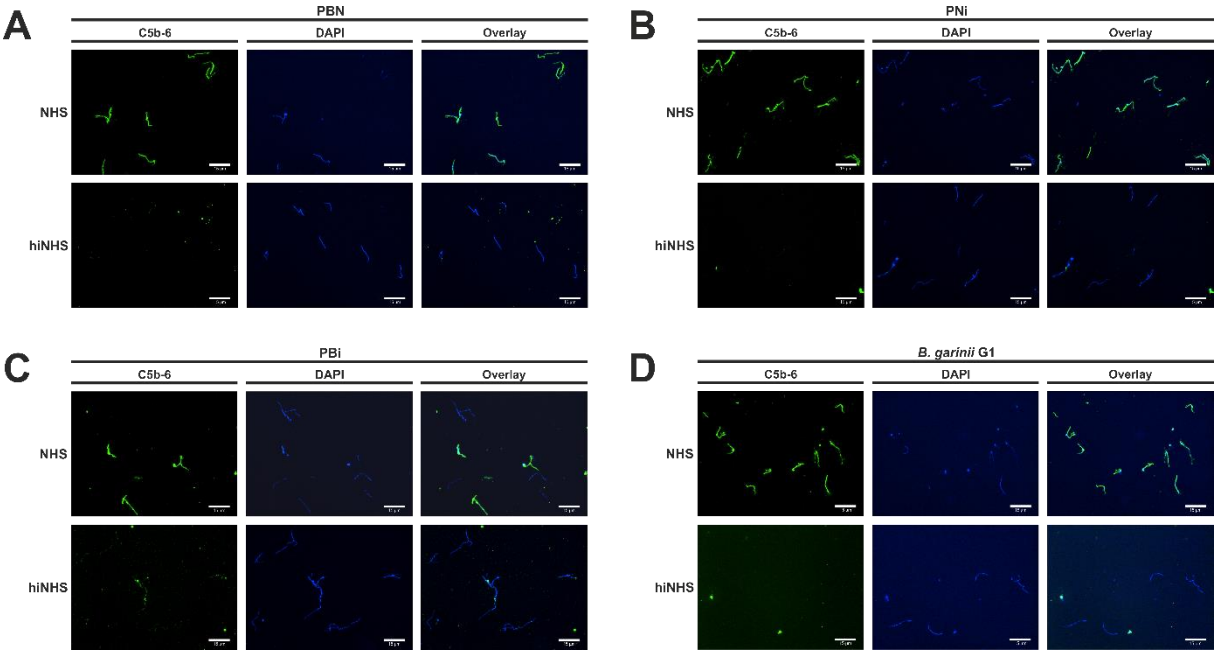


Figure S6. Immunofluorescence images of cells of *B. garinii* G1 (A), PBN (B), PNi (C), and PBi (D) incubated with normal human serum (NHS, top row) and heat-inactivated normal human serum (hiNHS, bottom row). The photos for each isolate (from left to right) are as follows: staining for the complement protein C5b-C9 (green), DAPI staining for nucleic acids (blue), and a merge of the C3 and DAPI images. All scale bars are equal to 15µm.



General Discussion

Obligate vector-borne pathogens, such as spirochetes belonging to the *Borrelia burgdorferi* sensu lato (*Bb*sl) species complex, can establish into novel transmission cycles through adaptation to novel hosts and/or vector species. Through comparative genomics it is possible to estimate the evolutionary forces leading to this adaptation and how known ecological interactions influence these processes. The Eurasian *Bb*sl system allows for a fascinating opportunity to study both host and vector adaptation. Three human-pathogenic *Bb*sl genospecies (*Borrelia afzelii*, *Borrelia bavariensis*, and *Borrelia garinii*) all share a Eurasian distribution and currently exist in separate transmission cycles vectored predominately by two generalist tick species in Asia (*Ixodes persulcatus*) and Europe (*Ixodes ricinus*) (43), creating a natural system in which to study vector adaptation in three independent *Bb*sl genospecies. Additionally, these three genospecies differ in their host associations with both *B. bavariensis* and *B. afzelii* utilizing rodents as reservoir hosts while *B. garinii* is associated with birds (39, 43, 45, 98–100). *Borrelia garinii* and *B. bavariensis* are closely related sister taxa (39, 101), which suggests that at the split between these two genospecies there was also a switch in host association (39, 101).

A current impediment to these studies though is a lack of biological samples and genomic data for specific populations of spirochetes (i.e. Asian) including also a bias in current samples towards human isolates from LB patients, which may not represent the full diversity of *Bb*sl genospecies. Most of what is currently known regarding the molecular mechanisms of *Bb*sl transmission, is based on studies including very few type strains of various genospecies (11, 13, 44, 92). This does not take natural variation into account, which has been shown in a few recent studies to play an important role in transmission efficiency (71, 75). Through taking these factors into account and producing novel isolates, we can then study both host and vector adaptation from an evolutionary ecology perspective (11, 16).

Proposed host associations appear to drive spatial and temporal variation in *Borrelia burgdorferi* sensu lato community structure

To complete their life-cycle, *Bb*sl spirochetes are dependent on the presence and interaction of both a competent vector and reservoir host (43, 46). Furthermore, each

*Bb*sl genospecies has a limited number of vectors and hosts which the spirochetes can successfully infect (42, 43). Therefore, ecological factors which influence both host and vector species will have indirect effects on *Bb*sl spirochetes by modifying prevalence and genospecies distribution (43, 46, 223). In Papers 2 and 3, we do observe a high level of spatial variability in *Bb*sl presence and species diversity with differences in which *Bb*sl genospecies are present suggesting something is modifying *Bb*sl community structure (52, 169). Spatial structure within genospecies is known to exist (55, 89) and is generally associated with microspatial variation in host-community structure (56, 90, 91). Community structure generally refers to the assemblage of species in an ecological community (1). This includes then characteristics of species richness, composition, and abundance (1). In our studies we can discuss host community structure referring to the assemblage and relative abundances of hosts which can serve as tick and/or *Bb*sl hosts in our study area. Additionally, we can discuss *Bb*sl community structure which refers to genospecies diversity and relative abundance of those genospecies in the vectors and hosts. It is important to note though that in Papers 2 and 3 we only quantified the *Bb*sl community structure in questing ticks and not in hosts. Genospecies appeared to remain stable over years, with presence and frequency of a genospecies in year *t* being positively correlated to these factors in the year *t-1* (52). Even so, Paper 2 showed that *Bb*sl community structure can both change or remain stable over the time frame of our study (52). This variability in *Bb*sl communities arises most likely due to direct ecological impacts to host populations (44, 50, 224). Factors such as habitat fragmentation can influence host community structure and behavior resulting in modifications to tick-host-*Bb*sl interaction rates (225–228). It has been observed in malaria systems, another vector-borne pathogen, that measures to reduce human-mosquito contact have resulted in a change in mosquito feeding habits towards non-human animals (229). This shift resulted in major modifications to malaria transmission cycles as these non-human animals were not amplifying hosts for *Plasmodium* parasites (229). Human disturbance without habitat modification can also modify host habitat use resulting in differing susceptibility to tick, and consequently, pathogen infection (230). Ecological changes to climate or vegetation could also negatively influence tick survival, potentially excluding *Bb*sl spirochetes from an environment (23, 25, 231–233). We were able to show in Paper 2 that changes in *Bb*sl community structure was not due to a decrease in tick population size (52). This again suggests that spatial and temporal variability in

*Bb*sl is a result of host community structure. Even so, our studies did not include quantifications of host communities, so we are not able to confirm if the interactions described above are occurring in our study plots. Future studies would need to include information on host diversity and relative abundance to definitively say if host community structure is driving the observed variability in *Bb*sl.

We observed that host community structure appears to drive variability in *Bb*sl prevalence and genospecies distributions. This is hypothesized to also relate to migration of spirochetes (56, 90, 91). Some *Bb*sl genospecies are bird adapted and show little to no geographic structure, suggesting they migrate easily between different geographic locations (56, 90, 91). We were able to confirm this fact as in Paper 5 population structure between continents was reduced in bird-associated genospecies and Paper 1 confirmed that birds contribute to the movement of tick-borne pathogens (234). Of interest however, bird-associated *Bb*sl sequence types were not shared as commonly between distinct locations in Paper 2 suggesting very little migration between our study plots (52). The fact that we do not observe a higher proportion of shared sequence types for bird-associated *Bb*sl genospecies between study plots in Paper 2, could support that micro-scale variation in *Bb*sl migration is possible. Focusing on the bird-associated *Bb*sl genospecies found in Paper 2, these can utilize various passerine birds as reservoir hosts in their terrestrial based cycles (99, 159). Passerine bird species are known to differ in their site fidelity, or how likely they are to stay and/or to defend a given territory (235, 236). Depending on which bird species are therefore present in the environment and acting as local reservoirs of these *Bb*sl genospecies, there could be variability in bird movement between sites based on bird behavior. We were also able to see further in Paper 5, that sequenced isolates of *B. garinii*, a bird adapted genospecies, gathered from plots on a much smaller scale than in Paper 2, do not form monophyletic clades based on plot suggesting that in this area movement over short distances is possilbe. For this reason, as stated previously, future work will need to take host ecology and behavior into account to address these open questions.

Recently, it has been discussed that *Bb*sl prevalence, and through this LB incidence, could be increasing due to climate change and the associated influences on tick vector geographic expansion (2, 3, 7, 237) although this effect is still debated (17, 52, 54). In our studies, Paper 3 supported a potential increase in overall *Bb*sl prevalence in comparison to previous research in the same geographic area (169, 171), but the

opposite (decrease in *Bb*sl prevalence) was observed longitudinally in Paper 2 (52). The results in Paper 2 are in agreement with other longitudinal studies which also displayed a decrease in *Bb*sl prevalence over their 15 year study (22, 54). The trend observed in Paper 3 could easily be a single year event caused by effects of host population fluctuations, as a response to event such as mast years (238, 239). Masting in Europe, normally refers to when beech (*Fagus sylvatica*) trees produce a high number of seeds in a given year which can influence vertebrate populations (238) and through this both tick and *Bb*sl populations as well (22, 54). Considering that 2017 was a masting year (240), and that effects of masting to *Bb*sl only appear two years after masting (54), increased prevalence would be expected in 2019 when all ticks in Paper 3 were collected (169). This shows that longitudinal studies, such as Paper 2, are required to make conclusions regarding temporal change to pathogen communities as major fluctuations in prevalence are possible in individual years based on extraneous ecological factors.

So far, we have seen that host community and various other ecological factors lead to both stability and variability in *Bb*sl community structure. In theory, this could also maintain diversity within *Bb*sl genospecies although what drives diversity in *Bb*sl communities is still a major unanswered question. Two major hypotheses exist with varying support to describe how *Bb*sl diversity is maintained in nature: 1) negative frequency dependent selection, or 2) multiple niche polymorphism (11, 44, 92). Selective forces imposed by host immunity could lead to negative frequency-dependent selection, where common genotypes are selected against resulting in fluctuations of genotypes over time (241). This would then result in a high diversity of genotypes being maintained with hypothesized fluctuations in dominating genotypes (11, 241, 242). As our results do clearly show an influence of host community on *Bb*sl prevalence and genospecies assemblage, this hypothesis could hold true for our environments. Although, based on the MLST data collected in Paper 2, we did not find support for yearly fluctuations in sequence types which would suggest that common genotypes are not selected against in favor of rarer genotypes (52). This would argue against negative frequency-dependent selection in our system and for other mechanisms which maintain overall diversity, such as multiple niche polymorphism (44, 92, 241). In the case of multiple niche polymorphism, diversity is maintained through a heterogenous host and vector environment imposing a form of balancing selection where multiple genotypes are maintained as no single genotype has the highest overall fitness (11, 92).

Work done on *B. burgdorferi* s. s. in North America used whole genome data and simulation-based modeling to determine which evolutionary process maintained spirochete diversity in the wild (242). A similar approach could be taken utilizing the full *B. bavariensis* genomes introduced in Paper 4 (84) or for all three LB genospecies in Paper 5 once their full genomes are assembled. This would allow us to test if there is support for negative frequency-dependent selection or multiple niche polymorphism maintaining *Bb*sl diversity in Eurasia. Taken together though, it appears that in our study system *Bb*sl prevalence and diversity is predominantly a product of host community structure even if the exact selective mechanism behind this diversity requires further work before it can be fully understood.

Geographic expansion is facilitated through vector shifts

As stated previously, *Bb*sl spirochetes and vector-borne pathogens are dependent on the presence of both a competent vector and reservoir host to complete their life cycle (43, 46). Vector populations are able to expand into new geographic ranges due to changes in vector host preference but also modifications to environmental factors, such as temperature, argued to be a result of climate change (2, 9, 243–245). This can also influence the geographic expansion of vector-borne pathogens (2, 9, 243–245). Vector-borne pathogens, like *Bb*sl spirochetes, could also adapt to new vectors potentially creating an opportunity to establish into new vector transmission cycles. Indeed, the *Bb*sl phylogeny shows multiple instances of vector shifts (101), but how these vector shifts relate to *Bb*sl geographic expansion has not been thoroughly studied (9). Paper 4 re-asserted that *B. bavariensis* consists of two distinct populations in Asia and Europe (39, 84, 93). By including Russian isolates we were able to show that the Asian population showed little to no geographic structuring (extended further in Paper 5 to include randomly sampled Japanese isolates) (39, 84, 93). The European population has undergone a selective bottleneck which is attributed to the colonization of the novel vector (*I. ricinus*) from an ancestral Asian population most likely vectored by *I. persulcatus* (39, 84, 93). In Paper 5, we were able to extend this Asian origin to two additional *Bb*sl genospecies: *B. afzelii* and *B. garinii*. Together, this showed that colonization of Europe by these three genospecies occurred through adaptation to a novel tick vector. Geographic expansion in relation to the colonization of a novel vector has been observed in other vector-borne pathogens such as West Nile

virus where expansion across North America resulted from novel adaptation to new mosquito vectors (9, 243).

A major difference between the genospecies which we studied is that both *B. afzelii* and *B. garinii* did not show reductions in diversity in relation to colonization of the European tick vector. This means that, although these three pathogenic species all share a Eurasian distribution, they each have experienced different evolutionary pressures. Pathogens can face various barriers to transmission including biological (e.g., inability to colonize host/vector) or geographic barriers (e.g. spatial separation of host/vector and pathogens) (246). In the case of a geographic barrier, it is hypothesized that transmission could happen if the pathogen and the host co-occurred in a geographic region but due to spatial separation, they are not able to (246). This can be observed through the establishment of exotic diseases into a new environment after, for example, introduction by humans (9, 246). This occurs through a process known as ecological fitting via resource tracking (247). In other words, the pathogen is able to establish in the new host population due to similarities to previous hosts so that adaptation to the ancestral host facilitates colonization (247). For both *B. afzelii* and *B. garinii*, our data seems to support that previous barriers to vector shifts where geographic as we do not observe a drastic change in overall diversity (11, 13, 16). This would need to be tested further though, especially including more genomic compartments besides just the chromosome. Even so, the results for *B. afzelii* and *B. garinii* are a stark contrast in comparison to *B. bavariensis* where it is apparent that the barrier to colonization of *I. ricinus* was most likely a biological one (39, 93). With our current data we can only hypothesize what could have happened to facilitate the colonization of *I. ricinus* by *B. bavariensis*. Two hypotheses that exist currently could be 1) *B. bavariensis* colonization of *I. ricinus* was facilitated by another *Bb*sl genospecies (39), or 2) rapid mutation or horizontal gene transfer affecting genes related to vector colonization (i.e. OspA) occurred in the ancestor to the European *B. bavariensis* clade. Co-infections between *Bb*sl genospecies are common (seen in Paper 2, 3, and 5) and there is a chance that few *B. bavariensis* spirochetes were able to colonize *I. ricinus* due to beneficial molecules secreted or produced by an *I. ricinus* adapted genospecies (39, 52, 169). In the case of other vector-borne diseases, namely West Nile virus, it has also been shown that very few mutations along genes related to vector colonization can lead to vector shifts and, consequently, geographic expansion (9, 243). This could also be the case for *B. bavariensis*, although we see a high level of

diversity between the Asian and European populations suggesting stronger divergence than just a few mutations (84). Recombination events between *Bb*sl spirochetes are possible although they appear to strongly influence specific loci while recombination on the whole genome remains low (84, 205, 242, 248, 249). Even so, some work has shown that intra-specific recombination may occur 50 times more frequently than inter-specific recombination (168). The "ecotype model" of bacterial evolution states that in the case of low recombination within bacterial populations, recombination at a selectively advantageous loci could lead to a genome wide selective sweep (250–252). This would occur through the bacteria with the recombinant allele or locus quickly dominating in an environment due to higher fitness prior to other individuals gaining the allele or locus through recombination or mutation (250–252). A by-product of this clonal expansion would be the purging of genetic diversity within that population (250–252). If the ancestor to the European *B. bavariensis* population had experienced a recombination event at a vector-related loci (e.g., OspA, BBE31, OspC) it could have led to a genome wide selective sweep resulting in the low diversity, almost clonal population we observed in our studies. Future work would need to extract known vector-related loci but also identify novel loci which may have undergone selection or recombination events to determine if this could be the explanation of what occurred during *B. bavariensis* evolution.

Borrelia afzelii, *B. bavariensis*, and *B. garinii* are all genospecies of interest due to their ability to cause LB in humans (29, 30), although a discussion of the Eurasian system, especially in relation to evolution, would not be complete without discussing the additional *Bb*sl genospecies found across this geographic range. One additional genospecies, *Borrelia turdi*, has been found both in Europe and Asia while all other species are isolated to single continents (39, 41, 253–255). In total three additional *Bb*sl genospecies have been found in Europe (*B. spielmanii*, *B. valaisiana*, *B. lusitaniae*) and four only distributed in Asia (*B. japonica, B. yangtzensis*, *B. sinica*, and *B. tanuki*) (39, 41, 253–255). In the same way that a vector switch could lead to geographic expansion, vector specialization could make it so that genospecies are isolated to a certain region. All Asian genospecies are vectored by *Ixodes* species which are found only in Asia (42, 160), although information on the distribution of these species is sparse. Previous theoretical work and also within fungal pathogens showed that host (or in the case of vector-borne pathogens, vector shifts) can be predicted by phylogenetic distance between host species, with the probability of a shift decreasing

when hosts are distantly related (244, 256). This could easily be related to vector shifts if these Asian tick species are distantly related to European tick species reducing the probability of a vector shift. Unfortunately, in recent phylogenetic reconstruction work, Asian species besides *I. persulcatus* were not included so it is not possible to definitively say if this relationship is true (257). Along this line though, *I. ricinus* and *I. persulcatus* were found to be closely related suggesting a vector shift, as those observed in Paper 4 & 5, would potentially be possible due to low phylogenetic distance (244, 256, 257). For the three *Bb*sl genospecies only found across Europe, all are vectored by *I. ricinus*, which, according to the argument above, could mean that a vector shift into Asia would be possible. For these genospecies, there is most likely a barrier to transmission which could be related to tick-host interactions. As an example, *B. lusitaniae* is associated with lizards and is distributed along the Mediterranean basin (258, 259) but has also been found in parts of central Europe (169, 171). This genospecies most likely evolved in Europe, even though its closet relative is *B. japonica* which is only found in Asia (39, 101). One argument for *B. lusitaniae* distribution could be that association with lizards reduces its effective migration rate therefore isolating this genospecies to its current range, even though this genospecies is vectored by a generalist tick. Indeed, this genospecies appears to show structuring based on tick populations suggesting very little to no gene flow even on a European scale (259). As for *B. spielmanii* and *B. valaisiana*, these genospecies also most likely evolved in Europe after an ancestral *Bb*sl spirochete colonized the continent. These are rodent-adapted (*B. spielmanii*) or bird-adapted genospecies (*B. valaisiana*) (39, 99, 123). We can only hypothesis if these genospecies are isolated to Europe based on potential tick-host interactions as in the case of *B. lusitaniae* (42). This poses an interesting concept if these genospecies could eventually migrate back into the Asian transmission cycle through adaption to *I. persulcatus* or if that would not be possible. Future work could test if these two genospecies are able to be transmitted by *I. persulcatus* under lab conditions and monitor various geographic locations across Europe, Russia, and Asia to see if these genospecies maintain a stable or expanding range. Taken together, our results show that geographic expansion, at least on a large spatial scale, is determined by vector adaptation with vector associations further playing an integral role in determining where genospecies are found. Host association or tick-host specialization also are important factors but seem to relate to micro-scale variation in *Bb*sl community structure as discussed in the following section.

Host associations explain population structure and potentially influence speciation events

Even though it does appear that vector associations and adaptation to novel vector species governs larger scale (i.e. continent-level) *Bb*sl distributions, we were able to use our results to argue that gene-flow between geographic locations (Paper 5) and *Bb*sl community structure (Papers 2-3) are most likely a product of host-associations (52, 84, 169, 234). The concept of host community structure influencing pathogen communities is not new as obligate pathogens (i.e., those without a free-living stage) cannot survive in an environment without their host (43, 46, 50). This is also true for *Bb*sl spirochetes as discussed in the section "Proposed host associations appear to drive spatial and temporal variation in *Borrelia burgdorferi* sensu lato community structure." In Papers 4 & 5, we extend the ecological factors to include genomic information which allows us to study *Bb*sl evolution from an evolutionary ecology approach (11, 16). Focusing for now on the information gained from just the chromosome, we were able to compare the effects of proposed host association to gene flow and geographic structuring. In our study we utilized two rodent-associated genospecies (*B. bavariensis* and *B. afzelii*) and one bird-associated genospecies (*B. garinii*) (39, 43, 100, 123). Two interesting trends arose from this analysis 1) *B. afzelii* and *B. garinii* follow our expectations based on their accepted host-associations but *B. bavariensis* does not, 2) chromosomal diversity seems to correlate with host-adaptation types.

In Paper 4, we noticed that Asian *B. bavariensis* samples including Japanese and Russian isolates did not display the characteristic geographic structuring associated with rodent-adapted *Bb*sl genospecies (56, 84, 90, 91). Paper 5 extended this to include randomly sampled isolates of *B. bavariensis* and *B. afzelii* coming from the same tick populations either on the island of Hokkaido or Honshu (Japan). Utilizing this information, we were able to directly compare two proposed rodent-associated genospecies and observe that they do indeed structure differently. The type-strain of *B. bavariensis*, PBi, has been shown to transmit successfully through rodents (100) and further was shown along with Asian NT29-like *B. bavariensis* to be susceptible to bird immune serum, which is used as a proxy to determine host associations (47, 48, 70). Although, to the best of our knowledge, no transmission experiments with *B. bavariensis* and bird hosts have been performed to date. *Borrelia*

afzelii has been shown to be rodent-associated and additionally to not survive in bird hosts (47, 48, 70), although transmission of *B. afzelii* between ticks can occur on birds when many ticks feed on the same blood pool (so called co-feeding transmission) (159). To explain this difference between the two genospecies, we argue that Asian *B. bavariensis* may rarely utilize birds as secondary hosts. In Paper 5, we noticed that there were notable similarities between *B. garinii*, a bird-associated genospecies, and the Asian *B. bavariensis* population. *Borrelia bavariensis* DNA was also found recently in sea-bird ticks (*I. uriae*) along the Atlantic coast of North America suggesting migration out of Eurasia potentially by utilizing sea-birds (158). Even though an Asian *B. bavariensis* isolate was shown to be susceptible to avian immune serum (47), we argue that the high diversity observed in Asian *B. bavariensis* cannot be described by a single isolate and that this diversity could include novel host infection mechanisms allowing *B. bavariensis* to transmit through birds. This hypothesis will need to be tested further both in lab-based transmission experiments and in field-based studies to see if Asian bird species are found to be infected with *B. bavariensis* or not. This fact alone displays the value of studying these bacteria from an evolutionary ecology perspective (11), as through comparison of ecological and genomic data we were able to make testable ecological hypotheses regarding *Bb*sl spirochete transmission. In future sampling and sequencing of less studied *Bb*sl spirochetes comparative genomics could allow for clarification regarding ecology based on population structure, filling gaps in our existing knowledge on *Bb*sl transmission cycles in non-human pathogenic genospecies.

If we accept the above argument that Asian *B. bavariensis* may utilize birds as secondary hosts in addition to rodents, we can observe a general trend that chromosome diversity, and potentially plasmid diversity, relates to host-association types. *Borrelia garinii* and Asian *B. bavariensis* display the highest genetic diversity while *B. afzelii* displays lower diversity as does European *B. bavariensis*. This could mean that bird-association is related to higher diversity while rodent-association is associated with low diversity. Host-association does influence gene flow between populations and therefore can modify how susceptible populations are to stochastic divergence processes, such as genetic drift (11, 16, 56, 90, 91). Genetic drift leads to a reduction in standing variation in the absence of gene flow which could result in lower overall diversity due to local geographic structure from host-associations (11, 241, 260, 261). Rodent-association or association with small vertebrates (e.g., lizards,

hedgehogs, etc.) is found quite commonly across *Bb*sl genospecies while bird-association is found in fewer genospecies (42, 101). As we suggest that rodent-association leads to higher susceptibility to drift, this association could promote speciation events through passive divergence due to isolation and consequent stochastic divergence (11, 241, 260, 261). Host-association is already thought to influence *Bb*sl speciation (39, 42, 44, 92, 101), as in the case of sister taxa *B. garinii* and *B. bavariensis*, which appear to have diverged due to a host shift between birds and rodents (although the direction is still not fully known) (39, 93, 94). It is apparent from previous phylogenetic reconstruction that included 15 *Bb*sl genospecies that host switches are quite common in the evolutionary history of *Bb*sl genospecies and the *B. bavariensis/B. garnii* split is not an isolated event. We would need to test some of this further, including determining in other *Bb*sl genospecies if the trend between host-association and diversity is universal or not. This would require sampling of genospecies, many of which are currently monotypic or described by very few isolates. Regardless though, our results appear to show that vector-association is the main driver of geographic expansion (argued for in the section "Geographic expansion is facilitated through vector shifts") while host-association is potentially a main driver of speciation through influences on *Bb*sl community structure and consequent to the process of genetic divergence.

Unstudied diversity in Eurasian Lyme borreliosis genospecies does influence host-pathogen interactions

Our studies showed again the considerable within-species diversity known to exist in *Bb*sl genospecies (83–86, 101, 168). Even though this within-species diversity exists, most functional studies to date related to *Bb*sl transmission dynamics rely on type-strains, with little focus on the impact of within-species diversity to transmission dynamics. The few functional studies to include within-strain variation, have been limited to *B. burgdorferi* s.s., the main LD causative agent in North America, and even there have focused on only a few loci (e.g., OspC, dbpA) (74, 75, 262, 263). Functional validation of transmission dynamics can be time-consuming and expensive which explains the focus on type-strains to make general conclusions regarding *Bb*sl transmission phenotypes (11, 42). Comparative genomics can offer a solution to this issue by providing guidance to functional studies in determining sample subsets which

capture global diversity (11, 16, 92). This is what we have started to do with our data set described in Papers 4 and 5.

Just based on the chromosome data presented in Paper 5, we can see that *B. afzelii*, *B. bavariensis*, and *B. garinii* all show within-species diversity. This extends even to the European population of *B. bavariensis* which showed variation in plasmid content (Paper 4) (84) even though this population was thought to be clonal (93, 157). Currently, we only have the full genomes of the 33 *B. bavariensis* isolates analyzed in Paper 4. Utilizing this information, we were able to start analyzing the impacts of natural variation to *Bb*sl transmission phenotypes, presented in Paper 6. Paper 4 highlighted a few hotspots of nucleotide diversity in the *B. bavariensis* genomes, including the region located between 52 and 60kb on the aligned lp54 plasmids (84). This section is proposed to carry the PFam54 gene array, which contains genes related to host adaptation (i.e., CRASPs, complement evasion, host cell adhesion etc.) (69, 71, 88, 187, 199). In Paper 4, we described that two European *B. bavariensis* isolates (PBN and PNi) were found to contain shorter lp54 sequences in comparison to type strain PBi (84). In Paper 6, we went on to show that the shorter lp54 sequences corresponded to these isolates naturally lacking the entire PFam54 gene array. This gene-array can be separated into five major lineages (I-V) (187). Lineage IV contains a variable number of paralogs and many genospecies displaying unique paralogs belonging to this lineage not found in other genospecies (187). The *B. bavariensis* type strain PBi contains PFam54 paralogs belonging to all of these lineages, although most do not have a described function (187). Seven of these PBi PFam54 paralogs belong to lineage IV including *bga66* and *bga71* (187), which produce proteins which bind late stage complement proteins (C7, C8, C9) and thereby inhibit the assembly of a functional membrane attack complex (MAC) (Figure 1); conferring resistance to human complement (69). The protein encoded by *bga71* has also been implicated in binding human brain microvascular endothelial cells and therefore has been implicated in host tissue tropism (186). We hypothesized that the loss of these genes in PBN and PNi would affect their sensitivity to immune serum and potentially mouse (i.e., host) infection. We were able to show an increased susceptibility to human complement but that PBN and PNi could establish viable infections after intradermal needle inoculation in lab mice. However, we did observe tissue specific colonization with lower efficiency in joint tissue, which we argued in Paper 6 to be due to PFam54 gene products potentially acting as adhesins.

Paper 6 proves the fact that within-species diversity is important to understand the full breadth of *Bb*sl transmission dynamics and that variation even within an almost clonal population of *Bb*sl spirochetes can modify infectivity and resistance phenotypes (84). This brings forward the point that type-strains, even in highly similar populations, cannot universally represent all isolates. Even so, our results highlight the fact that comparative genomic studies, especially those with full genome assemblies, can guide functional studies in isolate choice. Our future work will need to involve utilizing the full *B. bavariensis* genomes presented in Paper 4 and the genomes, once completed, for the additional genospecies presented in Paper 5 to search for novel candidates of adaptation. This includes genes of interest for vector-adaptation (*B. bavariensis*, *B. afzelii*) and host-adaptation (*B. bavariensis* vs. *B. garinii*). One example of future studies could be further analyzing the PFam54 gene family, but in isolates that contain the genes. It has already been shown that isolates of the same genospecies can differ not only allelically but also in presence/absence of specific PFam54 paralogs (187). Allelic variation between genospecies also has been shown to have major impacts to host immune evasion (71), and comparative analyses of our isolates could highlight mutations which are of interest for both host and vector adaptation or human LB manifestation. We showed in Paper 5 that additional LB-causing genospecies share a high level of diversity along the chromosome and do also differ in their plasmid content. Although, plasmid presence or absence was not associated with factors such as continent (i.e., vector) or genospecies (i.e., host association). This suggests that functionally relevant variation occurs at the gene level either in the form of the presence/absence of genes (as shown for PBN and PNi in Paper 6) or allelic variation among isolates. Through characterizing the full genome assemblies of the isolates presented in Paper 5, we will be able to better understand what genes relate to host adaptation, including variability in candidate genes (e.g., PFam54 gene array, OspA, etc.) as described previously (12, 13, 66, 74, 187). What is potentially of greater interest is how comparative genomics could identify novel genes, such as *bbg11* in Paper 4 (84), which may influence the *Bb*sl life-cycle in yet uncharacterized ways (84). Paper 4 also opened opportunities to study additional factors, such as plasmid copy number which could influence gene-product dosages and potentially modify host-spirochete or vector-spirochete interactions. Plasmid copy number over one copy per cell has not been observed before (85, 135), suggesting this could be a fully novel aspect of how *Bb*sl spirochetes adapt to their environment. All of

this can be studied further utilizing the data collected and presented within this dissertation and will help guide functional studies in choosing representative isolates to better describe the full range of *Bb*sl transmission phenotypes.

Final remarks and outlook

This dissertation aimed to study the evolution of Eurasian *Bb*sl spirochetes focusing on three human-pathogenic genospecies utilizing an evolutionary ecology approach. We were able to support many current hypotheses regarding the influences of ecological factors such as host community structure on the prevalence and distribution of *Bb*sl genospecies in a geographic area (Papers 1-3). We then integrated genomic and molecular based methodologies to study the drivers of *Bb*sl evolution (Papers 4-6). This has allowed us to show that adaptation to tick vectors governs the geographic distribution of *Bb*sl genospecies (Paper 4-5), although this will need to be tested in well sampled population level datasets for additional *Bb*sl genospecies, especially those currently described based on monotypic isolates (32, 101, 264). We utilize our results then to argue that, after expansion into a novel transmission cycle, host communities and selective pressures associated with successful colonization of these hosts, govern the local *Bb*sl community structure and genetic diversity (Paper 2-3). Host adaptation further influences the level of gene flow between geographic locations based most likely on host mobility (Paper 1, 4-5). This could modify the susceptibility of a population to genetic drift (241, 260, 261), potentially affecting divergence and speciation. Finally, we could start the process of showing that type-strains cannot always be taken to represent a full population (Paper 6) and that comparative genomic studies can guide which isolates are chosen for functional studies to best represent global variation. All of this advances our knowledge of how *Bb*sl genospecies have evolved across Eurasia and fills gaps regarding unstudied or neglected *Bb*sl populations (e.g., Asian *B. afzelii* isolates).

Future work which will focus on completing the plasmid assemblies for the isolates introduced in Paper 5 and studying the existing variation in the complete *B. bavariensis* genomes (Paper 4). This will bring many new insights into how natural variation can drive differences in host or vector associations. The results presented in this dissertation show that these factors can influence the process of geographic expansion or even speciation. This knowledge and future studies will lead to a better understanding of how *Bb*sl spirochetes may adapt to changing environments, allowing

for better predictions of pathogen emergence and better tracking human LB risk. The *Bb*sl system presents an amazing opportunity to study the evolutionary constraints on a vector-borne pathogen with a highly complex ecological life cycle. This will lead to better overall understanding of the evolutionary and ecological processes that maintain diversity, lead to diversification, and underlie speciation within vector-borne pathogens.

References

1. Morin PJ. 2011. Community Ecology2nd Editio. Wiley-Blackwell, West Sussex, UK.
2. Ogden NH, Lindsay LR. 2016. Effects of Climate and Climate Change on Vectors and Vector-Borne Diseases: Ticks Are Different. Trends Parasitol 32:646–656.
3. Lindgren E, Tälleklint L, Polfeldt T. 2000. Impact of climatic change on the northern latitude limit and population density of the disease-transmitting European tick *Ixodes ricinus*. Environ Health Perspect 108:119–123.
4. Sonenshine DE. 2018. Range Expansion of Tick Disease Vectors in North America: Implications for Spread of Tick-Borne Disease. Int J Environ Res Public Heal .
5. Ford SE, Chintala MM. 2006. Northward expansion of a marine parasite: Testing the role of temperature adaptation. J Exp Mar Bio Ecol 339:226–235.
6. van den Hurk AF, Ritchie SA, Mackenzie JS. 2008. Ecology and Geographical Expansion of Japanese Encephalitis Virus. Annu Rev Entomol 54:17–35.
7. Ogden NH, Mechai S, Margos G. 2013. Changing geographic ranges of ticks and tick-borne pathogens: Drivers, mechanisms and consequences for pathogen diversity. Front Cell Infect Microbiol 4:1–11.
8. Rochlin I, Toledo A. 2020. Emerging tick-borne pathogens of public health importance: A mini-review. J Med Microbiol 69:781–791.
9. Kilpatrick AM, Randolph SE. 2012. Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. Lancet 380:1946–1955.
10. Petersen LR, Beard CB, Visser SN. 2019. Combatting the Increasing Threat of Vector-Borne Disease in the United States with a National Vector-Borne Disease Prevention and Control System. Am J Trop Med Hyg 100:242–245.
11. O'Keeffe KR, Oppler ZJ, Brisson D. 2020. Evolutionary ecology of Lyme *Borrelia*. Infect Genet Evol 85:104570.
12. Kraiczy P. 2016. Hide and seek: How Lyme disease spirochetes overcome complement attack. Front Immunol 7.
13. Kurokawa C, Lynn GE, Pedra JHF, Pal U, Narasimhan S, Fikrig E. 2020. Interactions between *Borrelia burgdorferi* and ticks. Nat Rev Microbiol 18:587–600.
14. Atkinson JP, Du Clos TW, Mold C, Kulkarni H, Hourcade D, Wu X. 2019. 21 - The Human Complement System: Basic Concepts and Clinical Relevance, p. 299-317.e1. *In* Rich, RR, Fleisher, TA, Shearer, WT, Schroeder, HW, Frew, AJ, Weyand, CM (eds.), Clinical Immunology (Fifth Edition)Fifth Edit. Elsevier, London.
15. Murphy K, Weaver C. 2017. Janeway's Immunology9th Editio. Garland Science, Taylor & Francis Group, LLC, New York and London.
16. Oppler ZJ, O'Keeffe KR, McCoy KD, Brisson D. 2021. Evolutionary genetics of Borrelia. Curr

Issues Mol Biol 42:97–112.

17. Fernández-Ruiz N, Estrada-Peña A. 2020. Could climate trends disrupt the contact rates between *Ixodes ricinus* (Acari , Ixodidae) and the reservoirs of *Borrelia burgdorferi* s . l .? PLoS One 15:1–18.

18. Hillyard PD. 1996. Ticks of North-West Europe. Feild Studies Council, Schrewsbury, London.

19. Estrada-Peña A, Bouattour A, J-L C, Walker A. 2004. Ticks of Domestic Animals in the Mediterranean Region. A Guide to Identification of Species.

20. Latif AA, Putterill JF, de Klerk DG, Pienaar R, Mans BJ. 2012. *Nuttalliella namaqua* (Ixodoidea: Nuttalliellidae): First description of the male, immature stages and re-description of the female. PLoS One 7.

21. Guglielmone AA, Apanaskevich DA, Estrada-Peña A, Robbins RG, Petney TN, Horak IG. 2014. The hard ticks of the world: (Acari: Ixodida: Ixodidae). Springer, Berlin.

22. Bregnard C, Rais O, Voordouw MJ. 2020. Climate and tree seed production predict the abundance of the European Lyme disease vector over a 15-year period. Parasites and Vectors 13:1–12.

23. Tack K, Madder M, Baeten L, Frenne P DE, Verheyen K. 2012. The abundance of *Ixodes ricinus* ticks depends on tree species composition and shrub cover. Parasitology 139:1273–1281.

24. Tack W, Madder M, Baeten L, Vanhellemont M, Verheyen K. 2013. Shrub clearing adversely affects the abundance of *Ixodes ricinus* ticks. Exp Appl Acarol 60:411–420.

25. Akimov IA, Nebogatkin I V. 2016. Ixodid Ticks (Acari, Ixodidae) in Urban Landscapes. A review. Vestn Zool 50:155–162.

26. Uilenberg G. 2006. *Babesia* - A historical overview. Vet Parasitol 138:3–10.

27. Stanek G. 2009. Büchse der Pandora: Krankheitserreger in *Ixodes ricinus*-Zecken in Mitteleuropa. Wien Klin Wochenschr 121:673–683.

28. Rollend L, Fish D, Childs JE. 2013. Transovarial transmission of *Borrelia* spirochetes by *Ixodes scapularis*: A summary of the literature and recent observations. Ticks Tick Borne Dis 4:46–51.

29. Steere AC, Strle F, Wormser GP, Hu LT, Branda JA, Hovius JWR, Li X, Mead PS. 2016. Lyme borreliosis. Nat Rev Dis Prim 2.

30. Stanek G, Fingerle V, Hunfeld K-P, Jaulhac B, Kaiser R, Krause A, Kristoferitsch W, O'Connell S, Ornstein K, Strle F, Gray J. 2011. Lyme borreliosis: Clinical case definitions for diagnosis and management in Europe. Clin Microbiol Infect 17:69–79.

31. Barbour AG, Gupta RS. 2021. The Family Borreliaceae (Spirochaetales), a Diverse Group in Two Genera of Tick-Borne Spirochetes of Mammals, Birds, and Reptiles . J Med Entomol 1–12.

32. Margos G, Gofton A, Wibberg D, Dangel A, Marosevic D, Loh SM, Oskam C, Fingerle V. 2018. The genus *Borrelia* reloaded. PLoS One 13:1–14.

33. Adeolu M, Gupta RS. 2014. A phylogenomic and molecular marker based proposal for the division of the genus *Borrelia* into two genera: the emended genus *Borrelia* containing only the members of the relapsing fever *Borrelia*, and the genus *Borreliella* gen. n. Antonie Van Leeuwenhoek 105:1049–1072.

34. Estrada-Peña A, Cabezas-Cruz A. 2019. Phyloproteomic and functional analyses do not support a split in the genus *Borrelia* (Phylum Spirochaetes). BMC Evol Biol 19:54.

35. Margos G, Fingerle V, Oskam C, Stevenson B, Gofton A. 2020. Comment on: Gupta, 2019, Distinction between *Borrelia*and *Borreliella* is more robustly supported by molecular and phenotypic characteristics than all other neighbouring prokaryotic genera: Response to Margos' et al. "The genus *Borrelia*. Ticks Tick Borne Dis. Netherlands.

36. Barbour AG, Adeolu M, Gupta RS. 2017. Division of the genus *Borrelia* into two genera (corresponding to Lyme disease and relapsing fever groups) reflects their genetic and

phenotypic distinctiveness and will lead to a better understanding of these two groups of microbes (Margos et al. (. Int J Syst Evol Microbiol 67:2058–2067.

37. Gupta RS. 2019. Distinction between *Borrelia* and *Borreliella* is more robustly supported by molecular and phenotypic characteristics than all other neighbouring prokaryotic genera: Response to Margos' et al. "The genus *Borrelia* reloaded" (PLoS ONE 13(. PLoS One 14:e0221397.

38. Margos G, Marosevic D, Cutler S, Derdakova M, Diuk-Wasser M, Emler S, Fish D, Gray J, Hunfeldt K-P, Jaulhac B, Kahl O, Kovalev S, Kraiczy P, Lane RS, Lienhard R, Lindgren PE, Ogden N, Ornstein K, Rupprecht T, Schwartz I, Sing A, Straubinger RK, Strle F, Voordouw M, Rizzoli A, Stevenson B, Fingerle V. 2017. There is inadequate evidence to support the division of the genus *Borrelia*. Int J Syst Evol Microbiol 67:1081–1084.

39. Margos G, Fingerle V, Reynolds S. 2019. *Borrelia bavariensis*: Vector Switch, Niche Invasion, and Geographical Spread of a Tick-Borne Bacterial Parasite. Front Ecol Evol 7:1–20.

40. Wang G, Alje P, Dankert J. 1999. Molecular Typing of *Borrelia burgdorferi* Sensu Lato: Taxonomic, Epidemiological, and Clinical Implications 12:633–653.

41. Margos G, Becker NS, Fingerle V, Sing A, Ramos JA, Carvalho IL de, Norte AC. 2019. Core genome phylogenetic analysis of the avian associated *Borrelia turdi* indicates a close relationship to *Borrelia garinii*. Mol Phylogenet Evol 131:93–98.

42. Wolcott KA, Margos G, Fingerle V, Becker NS. 2021. Host association of *Borrelia burgdorferi* sensu lato: a review. Ticks Tick Borne Dis 12:101766.

43. Kurtenbach K, Hanincová K, Tsao JI, Margos G, Fish D, Ogden NH. 2006. Fundamental processes in the evolutionary ecology of Lyme borreliosis. Nat Rev Microbiol 4:660–669.

44. Tsao JI. 2009. Reviewing molecular adaptations of Lyme borreliosis spirochetes in the context of reproductive fitness in natural transmission cycles. Vet Res 40:36.

45. Comstedt P, Jakobsson T, Bergström S. 2011. Global ecology and epidemiology of *Borrelia garinii* spirochetes. Infect Ecol Epidemiol 1:9545.

46. Radolf JD, Caimano MJ, Stevenson B, Hu LT. 2012. Of ticks, mice and men: Understanding the dual-host lifestyle of Lyme disease spirochaetes. Nat Rev Microbiol 10:87–99.

47. Kurtenbach K, De Michelis S, Etti S, Schäfer SM, Sewell HS, Brade V, Kraiczy P. 2002. Host association of *Borrelia burgdorferi* sensu lato - The key role of host complement. Trends Microbiol 10:74–79.

48. Kurtenbach K, Sewell H-S, Ogden NH, Randolph SE, Nuttall PA. 1998. Serum complement sensitivity as a key factor in Lyme disease ecology. Infect Immun 66:1248–1251.

49. Baum E, Hue F, Barbour AG. 2012. Experimental infections of the reservoir species *Peromyscus leucopus* with diverse strains of *Borrelia burgdorferi*, a Lyme disease agent. MBio 3:e00434-12.

50. Randolph SE. 2004. Tick ecology: Processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors. Parasitology 129:S37–S65.

51. Fabri ND, Sprong H, Hofmeester TR, Heesterbeek H, Donnars BF, Widemo F, Ecke F, Cromsigt JPGM. 2021. Wild ungulate species differ in their contribution to the transmission of *Ixodes ricinus*-borne pathogens. Parasit Vectors 14:360.

52. Okeyo M, Hepner S, Rollins RE, Hartberger C, Straubinger RK, Marosevic D, Bannister SA, Bormane A, Donaghy M, Sing A, Fingerle V, Margos G. 2020. Longitudinal study of prevalence and spatio-temporal distribution of *Borrelia burgdorferi* sensu lato in ticks from three defined habitats in Latvia, 1999–2010. Environ Microbiol 00:1999–2010.

53. Coipan EC, Jahfari S, Fonville M, Maassen CB, van der Giessen J, Takken W, Takumi K, Sprong H. 2013. Spatiotemporal dynamics of emerging pathogens in questing *Ixodes ricinus*. Front Cell Infect Microbiol 4:1–11.

54. Bregnard C, Rais O, Voordouw MJ. 2021. Masting by beech trees predicts the risk of Lyme disease. Parasites and Vectors 14:1–22.

55. Strnad M, Hönig V, Růžek D, Grubhoffer L, Regoa ROM. 2017. Europe-Wide Meta-Analysis of *Borrelia burgdorferi* Sensu Lato Prevalence in Questing *Ixodes ricinus* Ticks. Appl Environ Microbiol 83:1–16.

56. Vollmer SA, Feil EJ, Chu CY, Raper SL, Cao WC, Kurtenbach K, Margos G. 2013. Spatial spread and demographic expansion of Lyme borreliosis spirochaetes in Eurasia. Infect Genet Evol 14:147–155.

57. LoGiudice K, Ostfeld RS, Schmidt KA, Keesing F. 2003. The ecology of infectious disease: effects of host diversity and community composition on Lyme disease risk. Proc Natl Acad Sci U S A 100:567–71.

58. Ostfeld R, Keesing F. 2000. Biodiversity and Disease Risk: The Case of Lyme Disease. Conserv Biol 14:722–728.

59. Randolph SE, Dobson ADM. 2012. Pangloss revisited: A critique of the dilution effect and the biodiversity-buffers-disease paradigm. Parasitology 139:847–863.

60. Hutchinson GE. 1957. Population studies - animal ecology and demography - concluding remarks. Cold Spring Harb Symp Quant Biol 22:415–427.

61. Coburn J, Garcia B, Hu LT, Jewett MW, Kraiczy P, Norris SJ, Skare J. 2021. Lyme disease pathogenesis. Curr Issues Mol Biol 42:473–518.

62. Skare JT, Garcia BL. 2020. Complement Evasion by Lyme Disease Spirochetes. Trends Microbiol 28:889–899.

63. Lone AG, Bankhead T. 2020. The *Borrelia burgdorferi* VlsE Lipoprotein Prevents Antibody Binding to an Arthritis-Related Surface Antigen. Cell Rep 30:3663-3670.e5.

64. Pal U, Kitsou C, Drecktrah D, Büyüktanir Yaş Ö, Fikrig E. 2021. Interactions between ticks and Lyme disease spirochetes. Curr Issues Mol Biol 42:113–144.

65. Kraiczy P, Stevenson B. 2013. Complement regulator-acquiring surface proteins of *Borrelia burgdorferi*: Structure, function and regulation of gene experession. Ticks Tick Borne Dis 4:26–34.

66. Lin YP, Diuk-Wasser MA, Stevenson B, Kraiczy P. 2020. Complement Evasion Contributes to Lyme Borreliae–Host Associations. Trends Parasitol 1–12.

67. Dulipati V, Meri S, Panelius J. 2020. Complement evasion strategies of *Borrelia burgdorferi* sensu lato. FEBS Lett 594:2645–2656.

68. Garcia BL, Zhi H, Wager B, Höök M, Skare JT. 2016. *Borrelia burgdorferi* BBK32 Inhibits the Classical Pathway by Blocking Activation of the C1 Complement Complex. PLOS Pathog 12:e1005404.

69. Hammerschmidt C, Klevenhaus Y, Koenigs A, Hallström T, Fingerle V, Skerka C, Pos KM, Zipfel PF, Wallich R, Kraiczy P. 2016. BGA66 and BGA71 facilitate complement resistance of *Borrelia bavariensis* by inhibiting assembly of the membrane attack complex. Mol Microbiol 99:407–424.

70. Kurtenbach K, Schäfer SM, Sewell H, Peacey M, Hoodless A, Nuttall PA, Randolph SE, Scha SM. 2002. Differential survival of Lyme borreliosis spirochetes in ticks that feed on birds. Infect Immun 70:5893–5895.

71. Hart T, Thien N, Nguyen T, Nowak NA, Zhang F, Linhardt J, Diuk-wasser M, Ram S, Kraiczy P, Lin Y. 2018. Polymorphic factor H-binding activity of CspA protects Lyme borreliae from the host complement in feeding ticks to facilitate tick- to-host transmission. PLoS Pathog 14:e1007105.

72. Bhide MR, Escudero R, Camafeita E, Gil H, Jado I, Anda P. 2009. Complement factor H binding by different Lyme disease and relapsing fever *Borrelia* in animals and human. BMC

Res Notes 2:134.

73. Brissette CA, Gaultney RA. 2014. That's my story, and I'm sticking to it-an update on *B. burgdorferi* adhesins. Front Cell Infect Microbiol 4:1–13.

74. Lin YP, Benoit V, Yang X, Martínez-Herranz R, Pal U, Leong JM. 2014. Strain-Specific Variation of the Decorin-Binding Adhesin DbpA Influences the Tissue Tropism of the Lyme Disease Spirochete. PLoS Pathog 10:e1004238.

75. Pietikäinen A, Åstrand M, Cuellar J, Glader O, Elovaara H, Rouhiainen M, Salo J, Furihata T, Salminen TA, Hytönen J. 2021. Conserved lysine residues in decorin binding proteins of *Borrelia garinii* are critical in adhesion to human brain microvascular endothelial cells. Mol Microbiol 0–3.

76. Pal U, Li X, Wang T, Montgomery RR, Ramamoorthi N, DeSilva AM, Bao F, Yang X, Pypaert M, Pradhan D, Kantor FS, Telford S, Anderson JF, Fikrig E. 2004. TROSPA, an *Ixodes scapularis* receptor for *Borrelia burgdorferi*. Cell 119:457–468.

77. Iyer R, Caimano MJ, Luthra A, Axline D, Corona A, Iacobas DA, Radolf JD, Schwartz I. 2015. Stage-specific global alterations in the transcriptomes of Lyme disease spirochetes during tick feeding and following mammalian host adaptation. Mol Microbiol 95:509–538.

78. Zhang L, Zhang Y, Adusumilli S, Liu L, Narasimhan S, Dai J, Zhao YO, Fikrig E. 2011. Molecular Interactions that Enable Movement of the Lyme Disease Agent from the Tick Gut into the Hemolymph. PLOS Pathog 7:e1002079.

79. Bierwagen P, Sliwiak J, Jaskolski M, Urbanowicz A. 2021. Strong interactions between Salp15 homologues from the tick *I. ricinus* and distinct types of the outer surface OspC protein from *Borrelia*. Ticks Tick Borne Dis 12.

80. Ramamoorthi N, Narasimhan S, Pal U, Bao F, Yang XF, Fish D, Anguita J, Norgard M V, Kantor FS, Anderson JF, Koski RA, Fikrig E. 2005. The Lyme disease agent exploits a tick protein to infect the mammalian host. Nature 436:573–577.

81. Pospisilova T, Urbanova V, Hes O, Kopacek P, Hajdusek O, Sima R. 2019. Tracking of *Borrelia afzelii* transmission from infected *Ixodes ricinus* nymphs to mice. Infect Immun 87:1–13.

82. Fraser C, Casjens S, Huang W. 1997. Genomic sequence of a Lyme disease spirochaete, *Borrelia burgdorferi*. Nature 390:580–586.

83. Margos G, Hepner S, Mang C, Marosevic D, Reynolds SE, Krebs S, Sing A, Derdakova M, Reiter MA, Fingerle V. 2017. Lost in plasmids: next generation sequencing and the complex genome of the tick-borne pathogen *Borrelia burgdorferi*. BMC Genomics 18:422.

84. Becker NS, Rollins RE, Nosenko K, Paulus A, Martin S, Krebs S, Takano A, Sato K, Kovalev SY, Kawabata H, Fingerle V, Margos G. 2020. High conservation combined with high plasticity: Genomics and evolution of *Borrelia bavariensis*. BMC Genomics 21:702.

85. Casjens SR, Mongodin EF, Qiu WG, Luft BJ, Schutzer SE, Gilcrease EB, Huang WM, Vujadinovic M, Aron JK, Vargas LC, Freeman S, Radune D, Weidman JF, Dimitrov GI, Khouri HM, Sosa JE, Halpin RA, Dunn JJ, Fraser CM. 2012. Genome stability of Lyme disease spirochetes: Comparative genomics of *Borrelia burgdorferi* plasmids. PLoS One 7.

86. Casjens SR, Di L, Akther S, Mongodin EF, Luft BJ, Schutzer SE, Fraser CM, Qiu WG. 2018. Primordial origin and diversification of plasmids in Lyme disease agent bacteria. BMC Genomics 19:1–24.

87. Mongodin EF, Casjens SR, Bruno JF, Xu Y, Drabek EF, Riley DR, Cantarel BL, Pagan PE, Hernandez YA, Vargas LC, Dunn JJ, Schutzer SE, Fraser CM, Qiu WG, Luft BJ. 2013. Inter- and intra-specific pan-genomes of *Borrelia burgdorferi* sensu lato: Genome stability and adaptive radiation. BMC Genomics 14.

88. Hart TM, Dupuis AP, Tufts DM, Blom AM, Starkey S, Rego ROM, Ram S, Kraiczy P, Kramer LD, Diuk-Wasser MA, Kolokotronis S-O, Lin Y-P. 2021. Host tropism determination by convergent evolution of immunological evasion in the Lyme disease system. bioRxiv

2021.02.09.430532.

89. Estrada-Peña A, Cutler S, Potkonjak A, Vassier-Tussaut M, Bortel W, Zeller H, Fernández-Ruiz N, Mihalca AD. 2018. An updated meta-analysis of the distribution and prevalence of *Borrelia burgdorferi* s.l. in ticks in Europe. Int J Health Geogr 17:1–16.

90. Vollmer SA, Bormane A, Dinnis RE, Seelig F, Dobson ADM, Aanensen DM, James MC, Donaghy M, Randolph SE, Feil EJ, Kurtenbach K, Margos G. 2011. Host migration impacts on the phylogeography of Lyme Borreliosis spirochaete species in Europe. Environ Microbiol 13:184–192.

91. Norte AC, Margos G, Becker NS, Albino Ramos J, Núncio MS, Fingerle V, Araújo PM, Adamík P, Alivizatos H, Barba E, Barrientos R, Cauchard L, Csörgő T, Diakou A, Dingemanse NJ, Doligez B, Dubiec A, Eeva T, Flaisz B, Grim T, Hau M, Heylen D, Hornok S, Kazantzidis S, Kováts D, Krause F, Literak I, Mänd R, Mentesana L, Morinay J, Mutanen M, Neto JM, Nováková M, Sanz JJ, Pascoal da Silva L, Sprong H, Tirri IS, Török J, Trilar T, Tyller Z, Visser ME, Lopes de Carvalho I. 2020. Host dispersal shapes the population structure of a tick-borne bacterial pathogen. Mol Ecol 29:485–501.

92. Seifert SN, Khatchikian CE, Zhou W, Brisson D. 2015. Evolution and population genomics of the Lyme borreliosis pathogen, *Borrelia burgdorferi*. Trends Genet 31:201–207.

93. Gatzmann F, Metzler D, Krebs S, Blum H, Sing A, Takano A, Kawabata H, Fingerle V, Margos G, Becker NS. 2015. NGS population genetics analyses reveal divergent evolution of a Lyme Borreliosis agent in Europe and Asia. Ticks Tick Borne Dis 6:344–351.

94. Margos G, Vollmer SA, Cornet M, Garnier M, Fingerle V, Wilske B, Bormane A, Vitorino L, Collares-Pereira M, Drancourt M, Kurtenbach K. 2009. A new *Borrelia* species defined by multilocus sequence analysis of housekeeping genes. Appl Environ Microbiol 75:5410–5416.

95. Burgdorfer W, Barbour AG, Hayes SF, Benach JL, Grunwaldt E, Davis JP. 1982. Lyme disease-a tick-borne spirochetosis? Science 216:1317–1319.

96. Schwartz I, Margos G, Casjens SR, Qiu WG, Eggers CH. 2021. Multipartite genome of Lyme disease *Borrelia*: Structure, variation and prophages. Curr Issues Mol Biol 42:409–454.

97. Strnad M, Rego ROM. 2020. The need to unravel the twisted nature of the *Borrelia burgdorferi* sensu lato complex across Europe. Microbiol (United Kingdom) 166:428–435.

98. Comstedt P, Asokliene L, Eliasson I, Olsen B, Wallensten A, Bunikis J, Bergström S. 2009. Complex population structure of Lyme borreliosis group spirochete *Borrelia garinii* in subarctic Eurasia. PLoS One 4.

99. Norte AC, Ramos JA, Gern L, Núncio MS, Lopes de Carvalho I. 2013. Birds as reservoirs for *Borrelia burgdorferi* s.l. in Western Europe: Circulation of *B. turdi* and other genospecies in bird-tick cycles in Portugal. Environ Microbiol 15:386–397.

100. Hu CM, Wilske B, Fingerle V, Lobet Y. 2001. Transmission of*Borrelia garinii* OspA serotype 4 to BALB_c mice by *Ixodes ricinus* ticks collected in the field. J Clin Microbiol 39:1169–1171.

101. Becker NS, Margos G, Blum H, Krebs S, Graf A, Lane RS, Castillo-Ramírez S, Sing A, Fingerle V. 2016. Recurrent evolution of host and vector association in bacteria of the *Borrelia burgdorferi* sensu lato species complex. BMC Genomics 17:734–746.

102. Maggini I, Trez M, Cardinale M, Fusani L. 2020. Stopover dynamics of 12 passerine migrant species in a small Mediterranean island during spring migration. J Ornithol 161:793–802.

103. Bairlein F. 1995. European-African songbird migration network: Manual of Field MethodsRevised Ed. Jade Bright, Wilhelmshaven.

104. Fracasso G, Matthysen E, Dhondt AA, Heylen D. 2019. Experimental study of micro-habitat selection by ixodid ticks feeding on avian hosts. Int J Parasitol 49:1005–1014.

105. Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.

106. Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol 33:1870–1874.

107. Trifinopoulos J, Nguyen L-T, von Haeseler A, Quang Minh B. 2016. W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44.

108. Rozas J, Ferrer-Mata A, Carlos anchez-DelBarrio JS, Guirao-Rico S, Librado P, Ramos-Onsins SE, Sanchez-Gracia A. 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Mol Biol Evol 34:3299–3302.

109. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754–1760.

110. Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

111. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin A V., Sirotkin A V., Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol 19:455–477.

112. Li R, Zhu H, Ruan J, Qian W, Fang X, Shi Z, Li Y, Li S, Shan G, Kristiansen K, Li S, Yang H, Wang J, Wang J. 2010. De novo assembly of human genomes with massively parallel short read sequencing. Genome Res 20:265–272.

113. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829.

114. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

115. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:1–9.

116. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST Server: rapid annotations using subsystems technology. BMC Genomics 9:75.

117. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res 42:D206–D214.

118. Paradis E. 2010. Pegas: An R package for population genetics with an integrated-modular approach. Bioinformatics 26:419–420.

119. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. Mol Biol Evol 30:772–780.

120. Steere AC, Coburn J, Glickstein L. 2004. The emergence of Lyme disease. J Clin Invest 113:1093–1101.

121. Walter KS, Carpi G, Caccone A, Diuk-Wasser MA. 2017. Genomic insights into the ancient spread of Lyme disease across North America. Nat Ecol Evol 1:1569–1576.

122. Eisen L. 2020. Vector competence studies with hard ticks and *Borrelia burgdorferi* sensu lato spirochetes: A review. Ticks Tick Borne Dis 11:101359.

123. Margos G, Vollmer SA, Ogden NH, Fish D. 2011. Population genetics, taxonomy, phylogeny and evolution of *Borrelia burgdorferi* sensu lato. Infect Genet Evol 11:1545–1563.

124. Comstedt P, Bergstrom S, Olsen B, Garpmo U, Marjavaara L, Mejlon H, Barbour AG, Bunikis J. 2006. Migratory passerine birds as reservoirs of Lyme borreliosis in Europe. Emerg Infect Dis 12:1087–1095.

125. Preac-Mursic V, Wilske B, Schierz G. 1986. European *Borrelia burgdorferi* Isolated from Humans and Ticks Culture Conditions and Antibiotic Susceptibility. Zentralblatt für Bakteriol Mikrobiol und Hyg 263:112–118.

126. Pollack RJ, Telford SR, Spielman A. 1993. Standardization of medium for culturing Lyme disease spirochetes. J Clin Microbiol 31:1251–1255.

127. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

128. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biol 5:R 12.

129. Delcher AL, Phillippy A, Carlton J, Salzberg SL. 2002. Fast algorithms for large-scale genome alignment and comparison. Nucleic Acids Res 30:2478–2483.

130. Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics 111:147–164.

131. Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics 17:754–755.

132. Ronquist F, Teslenko M, Van Der Mark P, Ayres DL, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP. 2012. MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. Syst Biol 61:539–542.

133. Tavaré S. 1986. Some Probabilistic and Statistial Problems in the Analysis of DNA Sequences. Lect Math teh Life Sci 17:57–86.

134. Rambaut A, Drummond AJ, Xie D, Baele G, Suchard MA. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst Biol 67:901–904.

135. Casjens S, Huang WM. 1993. Linear chromosomal physical and genetic map of *Borrelia burgdorferi*, the Lyme disease agent. Mol Microbiol 8:967–980.

136. R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

137. Nei M. 1987. Molecular Evolutionary Genetics. Columbia University Press, New York.

138. Tajima F. 1989. Statistical Method for Testing the Neutral Mutation Hypothesis by DNA Polymorphism. Genetics 123:585–595.

139. Excoffier L, Smouse PE, Quattro JM. 1992. Analysis of Molecular Variance Inferred From Metric Distances Among DNA Haplotypes: Application to Human Mitochondrial DNA Restriction Data Laurent. Genetics 131:479–491.

140. Kamvar ZN, Brooks JC, Grünwald NJ. 2015. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. Front Genet 6:208.

141. Hudson RR, Slatkin M, Maddison WP. 1992. Estimation of levels of gene flow from DNA sequence data. Genetics 132:583–589.

142. Pfeifer B, Wittelsbürger U, Ramos-Onsins SE, Lercher MJ. 2014. PopGenome: An efficient swiss army knife for population genomic analyses in R. Mol Biol Evol 31:1929–1936.

143. Bates D, Maechler M, Bolker B, Walker S. 2015. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 67.

144. Gelman A, Su Y-S. 2016. arm: Data Analysis Using Regression and Multilevel/Hierarchical Models. R package verstion 1. 10-1.

145. Nakagawa S, Schielzeth H. 2010. Repeatability for Gaussian and non-Gaussian data: A practical guide for biologists. Biol Rev 85:935–956.

146. Dulebohn DP, Bestor A, Rosa PA. 2013. *Borrelia burgdorferi* linear plasmid 28-3 confers a selective advantage in an experimental mouse-tick infection model. Infect Immun 81:2986–

2996.

147. Ellis TC, Jain S, Linowski AK, Rike K, Bestor A, Rosa PA, Halpern M, Kurhanewicz S, Jewett MW. 2013. In Vivo Expression Technology Identifies a Novel Virulence Factor Critical for *Borrelia burgdorferi* Persistence in Mice. PLoS Pathog 9.

148. Embers ME, Alvarez X, Ooms T, Philipp MT. 2008. The failure of immune response evasion by linear plasmid 28-1-deficient *Borrelia burgdorferi* is attributable to persistent expression of an outer surface protein. Infect Immun 76:3984–3991.

149. Grimm D, Eggers CH, Caimano MJ, Tilly K, Stewart PE, Elias AF, Radolf JD, Rosa PA. 2004. Experimental assessment of the roles of linear plasmids lp25 and lp28-1 of *Borrelia burgdorferi* throughout the infectious cycle. Infect Immun 72:5938–5946.

150. Grimm D, Tilly K, Bueschel DM, Fisher MA, Policastro PF, Gherardini FC, Schwan TG, Rosa PA. 2005. Defining plasmids required by *Borrelia burgdorferi* for colonization of tick vector *Ixodes scapularis* (Acari: Ixodidae). J Med Entomol 42:676–684.

151. Margos G, Gatewood AG, Aanensen DM, Ra Hanincová K, Terekhova D, Vollmer SA, Cornet M, Piesman J, Donaghy M, Bormane A, Hurn MA, Feil EJ, Fish D, Casjens S, Wormser GP, Schwartz I, Kurtenbach K, Hanincová K, Terekhova D, Vollmer SA, Cornet M, Piesman J, Donaghy M, Bormane A, Hurn MA, Feil EJ, Fish D, Casjens S, Wormser GP, Schwartz I, Kurtenbach K, Ra Hanincová K, Terekhova D, Vollmer SA, Cornet M, Piesman J, Donaghy M, Bormane A, Hurn MA, Feil EJ, Fish D, Casjens S, Wormser GP, Schwartz I, Kurtenbach K. 2008. MLST of housekeeping genes captures geographic population structure and suggests a European origin of *Borrelia burgdorferi*. Proc Natl Acad Sci U S A 105:8730–8735.

152. Takano A, Nakao M, Masuzawa T, Takada N, Yano Y, Ishiguro F, Fujita H, Ito T, Ma X, Oikawa Y, Kawamori F, Kumagai K, Mikami T, Hanaoka N, Ando S, Honda N, Taylor K, Tsubota T, Konnai S, Watanabe H, Ohnishi M, Kawabata H. 2011. Multilocus sequence typing implicates rodents as the main reservoir host of human-pathogenic *Borrelia garinii* in Japan. J Clin Microbiol 49:2035–2039.

153. Kovalev SY, Mukhacheva TA. 2014. Tick-borne encephalitis virus subtypes emerged through rapid vector switches rather than gradual evolution. Ecol Evol 4:4307–4316.

154. Comstedt P, Bergström S, Olsen B, Garpmo U, Marjavaara L, Mejlon H, Barbour AG, Bunikis J. 2006. Migratory passerine birds as reservoirs of Lyme borreliosis in Europe. Emerg Infect Dis 12:1087–1095.

155. Gómez-Díaz E, Boulinier T, Sertour N, Cornet M, Ferquel E, McCoy KD. 2011. Genetic structure of marine *Borrelia garinii* and population admixture with the terrestrial cycle of Lyme borreliosis. Environ Microbiol 13:2453–2467.

156. Gallais F, De Martino SJ, Sauleau EA, Hansmann Y, Lipsker D, Lenormand C, Talagrand-Reboul E, Boyer PH, Boulanger N, Jaulhac B, Schramm F. 2018. Multilocus sequence typing of clinical *Borreliella afzelii* strains: Population structure and differential ability to disseminate in humans. Parasites and Vectors 11:1–13.

157. Margos G, Wilske B, Sing A, Hizo-Teufel C, Cao WC, Chu C, Scholz H, Straubinger RK, Fingerle V. 2013. *Borrelia bavariensis* sp. nov. is widely distributed in Europe and Asia. Int J Syst Evol Microbiol 63:4284–4288.

158. Munro HJ, Ogden NH, Lindsay LR, Robertson GJ, Whitney H, Lang S. 2017. Evidence for *Borrelia bavariensis* Infections of *Ixodes uriae* within Seabird Colonies of the North Atlantic Ocean 83:1–9.

159. Heylen D, Matthysen E, Fonville M, Sprong H. 2014. Songbirds as general transmitters but selective amplifiers of *Borrelia burgdorferi* sensu lato genotypes in *Ixodes rinicus* ticks. Environ Microbiol 16:2859–2868.

160. Masuzawa T. 2004. Terrestrial distribution of the Lyme borreliosis agent *Borrelia burgdorferi* sensu lato in East Asia. Jpn J Infect Dis 57:229–235.

161. Ishiguro F, Takada N, Masuzawa T. 2005. Molecular evidence of the dispersal of Lyme disease

Borrelia from the Asian continent to Japan via migratory birds. Jpn J Infect Dis 58:184–186.

162. Nakao M, Miyamoto K, Fukunaga M. 1994. Lyme Disease Spirochetes in Japan: Enzootic Transmission Cycles in Birds, Rodents, and *Ixodes persulcatus* Ticks. J Infect Dis 170:878–882.

163. Smith RP, Muzaffar S Bin, Lavers J, Lacombe EH, Cahill BK, Lubelczyk CB, Kinsler A, Mathers AJ, Rand PW. 2006. *Borrelia garinii* in seabird ticks (*Ixodes uriae*), Atlantic Coast, North America. Emerg Infect Dis 12:1909–1912.

164. Chaconas G, Kobryn K. 2010. Structure, function, and evolution of linear replicons in *Borrelia*. Annu Rev Microbiol. Annual Reviews.

165. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, Bertoni A, Swerdlow HP, Gu Y. 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics 13:341.

166. Venter JC. 2010. Multiple personal genomes await. Nature. Nature Publishing Group.

167. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, Holden MTG, Dougan G, Bentley SD, Parkhill J, Peacock SJ. 2012. Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. PLoS Pathog 8:e1002824.

168. Jacquot M, Gonnet M, Ferquel E, Abrial D, Claude A, Gasqui P, Choumet V, Charras-Garrido M, Garnier M, Faure B, Sertour N, Dorr N, De Goër J, Vourc'H G, Bailly X. 2014. Comparative population genomics of the *Borrelia burgdorferi* species complex reveals high degree of genetic isolation among species and underscores benefits and constraints to studying intra-specific epidemiological processes. PLoS One 9.

169. Rollins RE, Yeyin Z, Wyczanska M, Alig N, Hepner S, Fingerle V, Margos G, Becker NS. 2021. Spatial variability in prevalence and genospecies distributions of *Borrelia burgdorferi* sensu lato from ixodid ticks collected in southern Germany. Ticks Tick Borne Dis 12.

170. Pérez-Eid C. 2007. Les tiques: identification, biologie, importance médicalvétérinaire. Lavoisier, Paris.

171. Fingerle V, Schulte-Spechtel UC, Ruzic-Sabljic E, Leonhard S, Hofmann H, Weber K, Pfister K, Strle F, Wilske B. 2008. Epidemiological aspects and molecular characterization of *Borrelia burgdorferi* s.l. from southern Germany with special respect to the new species *Borrelia spielmanii* sp. nov. Int J Med Microbiol 298:279–290.

172. Eiffert H, Ohlenbusch A, Christen HJ, Thomssen R, Spielman A, Matuschka FR. 1995. Nondifferentiation between Lyme disease spirochetes from vector ticks and human cerebrospinal fluid. J Infect Dis 171:476–479.

173. Rizzoli A, Hauffe HC, Carpi G, Vourc'h GI, Neteler M, Rosà R. 2011. Lyme borreliosis in Europe. Eurosurveillance 16:1–8.

174. Müller I, Freitag MH, Poggensee G, Scharnetzky E, Straube E, Schoerner C, Hlobil H, Hagedorn H-J, Stanek G, Schubert-Unkmeir A, Norris DE, Gensichen J, Hunfeld K-P. 2012. Evaluating Frequency, Diagnostic Quality, and Cost of Lyme Borreliosis Testing in Germany: A Retrospective Model Analysis. Clin Dev Immunol 2012:595427.

175. Schwartz AM, Hinckley AF, Mead PS, Hook SA, Kugeler KJ. 2017. Surveillance for Lyme Disease - United States, 2008-2015. Morb Mortal Wkly report Surveill Summ (Washington, DC 2002) 66:1–12.

176. Kugeler KJ, Eisen RJ. 2020. Challenges in Predicting Lyme Disease Risk. JAMA Netw Open 3:e200328–e200328.

177. Kugeler K, Schwartz A, Delorey M, Mead P, Hinckley A. 2021. Estimating the Frequency of Lyme Disease Diagnoses, United States, 2010–2018. Emerg Infect Dis J 27:616.

178. Rauer S, Kastenbauer S, Hofmann H, Fingerle V, Huppertz H-I, Hunfeld K-P, Krause A, Ruf B, Dersch R, group C. 2020. Guidelines for diagnosis and treatment in neurology - Lyme neuroborreliosis. Ger Med Sci 18:Doc03–Doc03.

179. Gern L. 2008. *Borrelia burgdorferi* sensu lato, the agent of Lyme borreliosis: life in the wilds. Parasite 15:244–247.

180. Piesman J, Gern L. 2004. Lyme borreliosis in Europe and North America. Parasitology 129:S191–S220.

181. Wilske B, Busch U, Eiffert H, Fingerle V, Pfister H-W, Rössler D, Preac-Mursic V. 1996. Diversity of OspA and OspC among cerebrospinal fluid isolates of *Borrelia burgdorferi* sensu lato from patients with neuroborreliosis in Germany. Med Microbiol Immunol 184:195–201.

182. Caine JA, Lin Y-P, Kessler JR, Sato H, Leong JM, Coburn J. 2017. *Borrelia burgdorferi* outer surface protein C (OspC) binds complement component C4b and confers bloodstream survival. Cell Microbiol 19:e12786.

183. Xie J, Zhi H, Garrigues RJ, Keightley A, Garcia BL, Skare JT. 2019. Structural determination of the complement inhibitory domain of *Borrelia burgdorferi* BBK32 provides insight into classical pathway complement evasion by Lyme disease spirochetes. PLOS Pathog 15:e1007659.

184. Lin YP, Tan X, Caine JA, Castellanos M, Chaconas G, Coburn J, Leong JM. 2020. Strain-specific joint invasion and colonization by Lyme disease spirochetes is promoted by outer surface protein C. PLoS Pathog 16:1–29.

185. Schmit VL, Patton TG, Gilmore RDJ, Patton VLS, Gilmore RDJ. 2011. Analysis of *Borrelia burgdorferi* surface proteins as determinants in establishing host cell interactions. Front Microbiol 2:1–8.

186. Tkáčová Z, Pulzová LB, Mochnáčová E, Jiménez-Munguía I, Bhide K, Mertinková P, Majerová P, Kulkarni A, Kováč A, Bhide M. 2020. Identification of the proteins of *Borrelia garinii* interacting with human brain microvascular endothelial cells. Ticks Tick Borne Dis 11:e101451.

187. Wywial E, Haven J, Casjens SR, Hernandez YA, Singh S, Monogodin EF, Fraser-Ligget CM, Luft BJ, Schutzer SE, Qiu WG. 2009. Fast, Adaptive Evolution at a Bacterial Host-Resistance Locus: the PFam54 Gene Array in *Borrelia burgdorferi*. Gene 445:26–37.

188. Van Burgel ND, Kraiczy P, Schuijt TJ, Zipfel PF, Van Dam AP. 2010. Identification and functional characterisation of Complement Regulator Acquiring Surface Protein-1 of serum resistant *Borrelia garinii* OspA serotype 4. BMC Microbiol 10:1–12.

189. Patton TG, Dietrich G, Dolan MC, Piesman J, Carroll JA, Gilmore Jr. RD. 2011. Functional Analysis of the *Borrelia burgdorferi* bba64 Gene Product in Murine Infection via Tick Infestation. PLoS One 6:e19536.

190. Patton TG, Brandt KS, Nolder C, Clifton DR, Carroll JA, Gilmore RD. 2013. *Borrelia burgdorferi* bba66 gene inactivation results in attenuated mouse infection by tick transmission. Infect Immun 81:2488–2498.

191. Hammerschmidt C, Koenigs A, Siegel C, Hallström T, Skerka C, Wallich R, Zipfel PF, Kraiczy P. 2014. Versatile roles of CspA orthologs in complement inactivation of serum-resistant Lyme disease spirochetes. Infect Immun 82:380–392.

192. Walter L, Sürth V, Röttgerding F, Zipfel PF, Fritz-Wolf K, Kraiczy P. 2019. Elucidating the Immune Evasion Mechanisms of *Borrelia mayonii*, the Causative Agent of Lyme Disease. Front Immunol 10:1–18.

193. Zimmermann L, Stephens A, Nam S-Z, Rau D, Kübler J, Lozajic M, Gabler F, Söding J, Lupas AN, Alva V. 2018. A Completely Reimplemented MPI Bioinformatics Toolkit with a New HHpred Server at its Core. J Mol Biol 430:2237–2243.

194. Gabler F, Nam S-Z, Till S, Mirdita M, Steinegger M, Söding J, Lupas AN, Alva V. 2020. Protein Sequence Analysis Using the MPI Bioinformatics Toolkit. Curr Protoc Bioinforma 72:e108.

195. Herzberger P, Siegel C, Skerka C, Fingerle V, Schulte-Spechtel U, Van Dam A, Wilske B, Brade V, Zipfel PF, Wallich R, Kraiczy P. 2007. Human pathogenic *Borrelia spielmanii* sp. nov. resists complement-mediated killing by direct binding of immune regulators factor H and factor H-like protein 1. Infect Immun 75:4817–4825.

196. Siegel C, Schreiber J, Haupt K, Skerka C, Brade V, Simon MM, Stevenson B, Wallich R, Zipfel PF, Kraiczy P. 2008. Deciphering the ligand-binding sites in the *Borrelia burgdorferi* complement regulator-acquiring surface protein 2 required for interactions with the human immune regulators factor H and factor H-like protein 1. J Biol Chem 283:34855–34863.

197. Marcinkiewicz AL, Lieknina I, Kotelovica S, Yang X, Kraiczy P, Pal U, Lin Y-PP, Tars K. 2018. Eliminating factor H-Binding activity of Borrelia burgdorferi CspZ combined with virus-like particle conjugation enhances its efficacy as a Lyme disease vaccine. Front Immunol 9:1–11.

198. Williams SK, Weiner ZP, Gilmore RD. 2018. Human neuroglial cells internalize *Borrelia burgdorferi* by coiling phagocytosis mediated by Daam1. PLoS One 13:e0197413.

199. Koenigs A, Hammerschmidt C, Jutras BL, Pogoryelov D, Barthel D, Skerka C, Kugelstadt D, Wallich R, Stevenson B, Zipfel PF, Kraiczy P. 2013. BBA70 of *Borrelia burgdorferi* is a novel plasminogen-binding protein. J Biol Chem 288:25229–25243.

200. Hallström T, Haupt K, Kraiczy P, Hortschansky P, Wallich R, Skerka C, Zipfel PF. 2010. Complement regulator-acquiring surface protein 1 of *Borrelia burgdorferi* binds to human bone morphogenic protein 2, several extracellular matrix proteins, and plasminogen. J Infect Dis 202:490–498.

201. Lin Y-P, Frye AM, Nowak TA, Kraiczy P. 2020. New Insights Into CRASP-Mediated Complement Evasion in the Lyme Disease Enzootic Cycle. Front Cell Infect Microbiol 10:1.

202. Schwan TG, Burgdorfer W, Garon CF. 1988. Changes in infectivity and plasmid profile of the Lyme disease spirochete, *Borrelia burgdorferi*, as a result of in vitro cultivation. Infect Immun 56:1831–1836.

203. Dorothee G, F. EA, Kit T, A. RP. 2003. Plasmid Stability during In Vitro Propagation of *Borrelia burgdorferi* Assessed at a Clonal Level. Infect Immun 71:3138–3145.

204. Carroll JA, Cordova RM, Garon CF. 2000. Identification of 11 pH-regulated genes in *Borrelia burgdorferi* localizing to linear plasmids. Infect Immun 68:6677–6684.

205. Casjens S, Palmer N, van Vugt R, Huang WM, Stevenson B, Rosa P, Lathigra R, Sutton G, Peterson J, Dodson RJ, Haft D, Hickey E, Gwinn M, White O, Fraser CM. 2000. A bacterial genome in flux: the twelve linear and nine circular extrachromosomal DNAs in an infectious isolate of the Lyme disease spirochete *Borrelia burgdorferi*. Mol Microbiol 35:490–516.

206. Woodman ME, Cooley AE, Miller JC, Lazarus JJ, Tucker K, Bykowski T, Botto M, Hellwage J, Wooten RM, Stevenson B. 2007. *Borrelia burgdorferi* binding of host complement regulator factor H is not required for efficient mammalian infection. Infect Immun 75:3131–3139.

207. Lawrenz MB, Wooten RM, Zachary JF, Drouin SM, Weis JJ, Wetsel RA, Norris SJ. 2003. Effect of complement component C3 deficiency on experimental Lyme borreliosis in mice. Infect Immun 71:4432–4440.

208. Sertour N, Cotté V, Garnier M, Malandrin L, Ferquel E, Choumet V. 2018. Infection Kinetics and Tropism of *Borrelia burgdorferi* sensu lato in Mouse After Natural (via Ticks) or Artificial (Needle) Infection Depends on the Bacterial Strain. Front Microbiol 9.

209. Genné D, Rossel M, Sarr A, Battilotti F, Rais O, Rego ROM, Voordouw MJ. 2021. Competition between strains of Borrelia afzelii in the host tissues and consequences for transmission to ticks. ISME J 15:2390–2400.

210. Norte AC, Lopes de Carvalho I, Núncio MS, Araújo PM, Matthysen E, Albino Ramos J, Sprong H, Heylen D. 2020. Getting under the birds' skin: tissue tropism of *Borrelia burgdorferi* s.l. in naturally and experimentally infected avian hosts. Microb Ecol 79:756–769.

211. Brisson D, Baxamusa N, Schwartz I, Wormser GP. 2011. Biodiversity of *Borrelia burgdorferi* strains in tissues of Lyme disease patients. PLoS One 6:e22926.

212. Hanincová K, Ogden NH, Diuk-Wasser M, Pappas CJ, Iyer R, Fish D, Schwartz I, Kurtenbach K. 2008. Fitness variation of *Borrelia burgdorferi* sensu stricto strains in Mice. Appl Environ Microbiol 74:153–157.

213. Hanincova K, Mukherjee P, Ogden NH, Margos G, Wormser GP, Reed KD, Meece JK, Vandermause MF, Schwartz I. 2013. Multilocus sequence typing of *Borrelia burgdorferi* suggests existence of lineages with differential pathogenic properties in humans. PLoS One 8:e73066.

214. Drecktrah D, Samuels DS. 2018. Genetic Manipulation of *Borrelia* spp. Curr Top Microbiol Immunol 415:113–140.

215. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG. 2012. Primer3--new capabilities and interfaces. Nucleic Acids Res2012/06/22. 40:e115.

216. Kõressaar T, Lepamets M, Kaplinski L, Raime K, Andreson R, Remm M. 2018. Primer3_masker: integrating masking of template sequence with primer design software. Bioinformatics 34:1937–1938.

217. Koressaar T, Remm M. 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291.

218. Okonechnikov K, Golosova O, Fursov M, Varlamov A, Vaskin Y, Efremov I, German Grehov OG, Kandrov D, Rasputin K, Syabro M, Tleukenov T. 2012. Unipro UGENE: A unified bioinformatics toolkit. Bioinformatics 28:1166–1167.

219. Huang X, Madan A. 1999. CAP3: A DNA sequence assembly program. Genome Res 9:868–877.

220. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792–1797.

221. Edgar RC. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.

222. Larsson A. 2014. AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30:3276–3278.

223. Ostfeld RS, Canham CD, Oggenfuss K, Winchcombe RJ, Keesing F. 2006. Climate, deer, rodents, and acorns as determinants of variation in Lyme-disease risk. PLoS Biol 4:1058–1068.

224. Randolph SE. 2008. The impact of tick ecology on pathogen transmission dynamics, p. 40–72. *In* Ticks: Biology, Disease and Control. Cambridge University Press.

225. MacDougall AS, McCann KS, Gellner G, Turkington R. 2013. Diversity loss with persistent human disturbance increases vulnerability to ecosystem collapse. Nature 494:86–89.

226. Halos L, Bord S, Cotté V, Gasqui P, Abrial D, Barnouin J, Boulouis HJ, Vayssier-Taussat M, Vourc'H G. 2010. Ecological factors characterizing the prevalence of bacterial tick-borne pathogens in *Ixodes ricinus* ticks in pastures and woodlands. Appl Environ Microbiol 76:4413–4420.

227. Hewison AJM, Vincent JP, Joachim J, Angibault JM, Cargnelutti B, Cibien C. 2001. The effects of woodland fragmentation and human activity on roe deer distribution in agricultural landscapes. Can J Zool 79:679–689.

228. Diuk-Wasser MA, VanAcker MC, Fernandez MP. 2020. Impact of Land Use Changes and Habitat Fragmentation on the Eco-epidemiology of Tick-Borne Diseases. J Med Entomol 1–19.

229. Mwangangi JM, Mbogo CM, Orindi BO, Muturi EJ, Midega JT, Nzovu J, Gatakaa H, Githure J, Borgemeister C, Keating J, Beier JC. 2013. Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years. Malar J 12:1–9.

230. Rollins RE, Mouchet A, Margos G, Chitimia-Dobler L, Fingerle V, Becker NS, Dingemanse NJ. 2021. Repeatable differences in exploratory behaviour predict tick infestation probability in wild great tits. Behav Ecol Sociobiol 75.

231. Tack W, Madder M, Baeten L, Vanhellemont M, Gruwez R, Verheyen K. 2012. Local habitat and landscape affect *Ixodes ricinus* tick abundances in forests on poor, sandy soils. For Ecol Manage 265:30–36.

232. Ruyts SC, Ampoorter E, Coipan EC, Baeten L, Heylen D, Sprong H, Matthysen E, Verheyen K. 2016. Diversifying forest communities may change Lyme disease risk: Extra dimension to the dilution effect in Europe. Parasitology 143:1310–1319.

233. Ehrmann S, Ruyts SC, Scherer-Lorenzen M, Bauhus J, Brunet J, Cousins SAO, Deconchat M, Decocq G, De Frenne P, De Smedt P, Diekmann M, Gallet-Moron E, Gärtner S, Hansen K, Kolb A, Lenoir J, Lindgren J, Naaf T, Paal T, Panning M, Prinz M, Valdés A, Verheyen K, Wulf M, Liira J. 2018. Habitat properties are key drivers of *Borrelia burgdorferi* (s.l.) prevalence in *Ixodes ricinus* populations of deciduous forest fragments. Parasites and Vectors 11:1–15.

234. Rollins RE, Schaper S, Kahlhofer C, Frangoulidis D, Strauß AFT, Cardinale M, Springer A, Strube C, Bakkes DK, Becker NS, Chitimia-Dobler L. 2021. Ticks (Acari: Ixodidae) on birds migrating to the island of Ponza, Italy, and the tick-borne pathogens they carry. Ticks Tick Borne Dis 12:10–14.

235. Krebs JR. 1982. Territorial Defence in the Great Tit (Parus major): Do Residents Always Win? Behav Ecol Sociobiol 11:185–194.

236. Hogstad O. 1993. Structure and dynamics of a passerine bird community in a spruce-dominated boreal forest. A 12-year study. Ann Zool Fennici 30:43–54.

237. Rosà R, Andreo V, Tagliapietra V, Baráková I, Arnoldi D, Hauffe HC, Manica M, Rosso F, Blaňarová L, Bona M, Derdáková M, Hamšíková Z, Kazimírová M, Kraljik J, Kocianová E, Mahríková L, Minichová L, Mošanský L, Slovák M, Stanko M, Špitalská E, Ducheyne E, Neteler M, Hubálek Z, Rudolf I, Venclikova K, Silaghi C, Overzier E, Farkas R, Földvári G, Hornok S, Takács N, Rizzoli A. 2018. Effect of climate and land use on the spatio-temporal variability of tick-borne bacteria in Europe. Int J Environ Res Public Health 15:1–15.

238. Bogdziewicz M, Zwolak R, Crone EE. 2016. How do vertebrates respond to mast seeding? Oikos 125:300–307.

239. Pearse IS, Koenig WD, Kelly D. 2016. Mechanisms of mast seeding: resources, weather, cues, and selection. New Phytol 212:546–562.

240. Ascoli D, Maringer J, Hacket-Pain A, Conedera M, Drobyshev I, Motta R, Cirolli M, Kantorowicz W, Zang C, Schueler S, Croisé L, Piussi P, Berretti R, Palaghianu C, Westergren M, Lageard JGA, Burkart A, Gehrig Bichsel R, Thomas PA, Beudert B, Övergaard R, Vacchiano G. 2017. Two centuries of masting data for European beech and Norway spruce across the European continent. Ecology 98:1473.

241. Brisson D. 2018. Negative frequency-dependent selection is frequently confounding. Front Ecol Evol 6:1–9.

242. Haven J, Vargas LC, Mongodin EF, Xue V, Hernandez Y, Pagan P, Fraser-Liggett CM, Schutzer SE, Luft BJ, Casjens SR, Qiu WG. 2011. Pervasive recombination and sympatric genome diversification driven by frequency-dependent selection in Borrelia burgdorferi, the Lyme disease bacterium. Genetics 189:951–966.

243. Weaver SC, Barrett ADT. 2004. Transmission cycles, host range, evolution and emergence of arboviral disease. Nat Rev Microbiol 2:789–801.

244. De Vienne DM, Hood ME, Giraud T. 2009. Phylogenetic determinants of potential host shifts in fungal pathogens. J Evol Biol 22:2532–2541.

245. Woolhouse MEJ, Haydon DT, Antia R. 2005. Emerging pathogens: The epidemiology and evolution of species jumps. Trends Ecol Evol 20:238–244.

246. Bonneaud C, Longdon B. 2020. Emerging pathogen evolution. EMBO Rep 21:1–5.

247. De Fine Licht HH. 2018. Does pathogen plasticity facilitate host shifts? PLoS Pathog 14:1–9.

248. Livey I, Gibbs CP, Schuster R, Dorner F. 1995. Evidence for lateral transfer and recombination in OspC variation in Lyme disease *Borrelia*. Mol Microbiol 18:257–269.

249. Stevenson B, Casjens S, Rosa P. 1998. Evidence of past recombination events among the genes encoding the Erp antigens of *Borrelia burgdorferi*. Microbiology 144:1869–1879.

250. Cohan FM, Perry EB. 2007. A systematics for discovering the fundamental units of bacterial diversity. Curr Biol 17:R373-86.

251. Cohan FM. 2001. Bacterial species and speciation. Syst Biol 50:513–524.

252. Bendall ML, Stevens SLR, Chan L-K, Malfatti S, Schwientek P, Tremblay J, Schackwitz W, Martin J, Pati A, Bushnell B, Froula J, Kang D, Tringe SG, Bertilsson S, Moran MA, Shade A, Newton RJ, McMahon KD, Malmstrom RR. 2016. Genome-wide selective sweeps and gene-specific sweeps in natural bacterial populations. ISME J 10:1589–1601.

253. Masuzawa T, Takada N, Kudeken M, Fukui T, Yano Y, Ishiguro F, Kawamura Y, Imai Y, Ezaki T. 2001. *Borrelia sinica* sp. nov., a Lyme disease-related *Borrelia* species isolated in China. Int J Syst Evol Microbiol 51:1817–1824.

254. Margos G, Chu CY, Takano A, Jiang BG, Liu W, Kurtenbach K, Masuzawa T, Fingerle V, Cao WC, Kawabata H. 2015. *Borrelia yangtzensis* sp. nov., A rodent-associated species in Asia, is related to Borrelia valaisiana. Int J Syst Evol Microbiol 65:3836–3840.

255. Fukunaga M, Hamase A, Okada K, Nakao M. 1996. *Borrelia tanukii* sp. nov. and *Borrelia turdae* sp. nov. found from ixodid ticks in Japan: Rapid species identification by 16S rRNA gene-targeted PCR analysis. Microbiol Immunol 40:877–881.

256. Antonovics J, Boots M, Ebert D, Koskella B, Poss M, Sadd BM. 2013. The origin of specificity by means of natural selection: Evolved and nonhost resistance in host-pathogen interactions. Evolution (N Y) 67:1–9.

257. Charrier NP, Hermouet A, Hervet C, Agoulon A, Barker SC, Heylen D, Toty C, McCoy KD, Plantard O, Rispe C. 2019. A transcriptome-based phylogenetic study of hard ticks (Ixodidae). Sci Rep 9:1–13.

258. Victorino LR, Margos G, Feil EJ, Collares-Pereira M, Zé-Zé L, Kurtenbach K. 2008. Fine-scale phylogeographic structure of *Borrelia lusitaniae* revealed by multilocus sequence typing. PLoS One 3.

259. Norte AC, Boyer PH, Castillo-Ramirez S, Chvostáč M, Brahami MO, Rollins RE, Woudenberg T, Didyk YM, Derdakova M, Núncio MS, de Carvalho IL, Margos G, Fingerle V. 2021. The population structure of *Borrelia lusitaniae* is reflected by a population division of its *Ixodes* vector. Microorganisms 9:1–22.

260. Gibbs JP. 2001. Demography versus habitat fragmentation as determinants of genetic variation in wild populations. Biol Conserv 100:15–20.

261. Nielsen R, Slatkin M. 2013. An Introdcution to Population Genetics: Theory and Applications, 6th ed. Sinauer Associates, Inc., Sunderland, MA.

262. Brisson D, Dykhuizen DE. 2004. ospC diversity in *Borrelia burgdorferi*: different hosts are different niches. Genetics 168:713–722.

263. Baranton G, Seinost G, Theodore G, Postic D, Dykhuizen D. 2001. Distinct levels of genetic diversity of *Borrelia burgdorferi* are associated with different aspects of pathogenicity. Res Microbiol 152:149–156.

264. Margos G, Fedorova N, Becker NS, Kleinjan JE, Marosevic D, Krebs S, Hui L, Fingerle V, Lane RS. 2020. *Borrelia maritima* sp. nov., a novel species of the *Borrelia burgdorferi* sensu lato complex, occupying a basal position to North American species. Int J Syst Evol Microbiol 70:849–856.

Acknowledgments

There are many people who have made this journey possible. I would first like to thank my supervisor, Dr. Noémie S. Becker, for providing me with this opportunity and guiding me through the last five years of projects and learning to become a scientist. Of course, I would like to thank the other members of my thesis advisory committee, Prof. Dr. Dirk Metzler, and PD Dr. Jürgen Lassak for their insights and guidance throughout my doctoral studies. I would also like to thank Dr. Ricardo J. Pereira for his guidance not only in terms of science but also in learning how to take ownership of my ideas and feel confident in my scientific ability. Thank you also to Hilde Lainer for her help in integrating lab work here at the LMU campus, and always being available for questions and advice.

To the members of the National Reference Center for *Borrelia* including Christl, Sylvia, Cecilia, and Wiltrud, thank you for all the help in isolating (or not isolating) *Borrelia*, running tests, learning new methods, and supporting me over the years. I would like to give a special thanks to Dr. Gabriele Margos and Dr. Volker Fingerle for giving me the opportunity to work on this fascinating system in their lab group and to the many nights where I received guidance over a glass of wine. This is true for Sabrina Hepner and Dr. Durdica Marosevic as well, who I would like to thank for her support in the good and hard times, always supporting that I would succeed.

I would like to also thank my family especially my sister, Tabby, and my parents, Noelle and Jason, for their constant support in my dream to become a scientist. Thank you for always supporting me in this endeavor and making a world possible where I could pursue this dream. Finally, to my partner, Fabi, I do not know if I could have done this without your constant support and love. Thank you for always believing in me and showing me that my dreams are worth fighting for. I am so excited for what the future will bring, and I know you will be my side for it all.

Additional publications

Hepner, S., Fingerle, V., Duscher, G.G., Felsberger, G., Marosevic, D., **Rollins, R.E.,** etal. (2020) Population structure of *Borrelia turcica* from Greece and Turkey. Infection, Genetics, & Evolution. 77: 104050.

Rollins, R.E. etal. (2021). Repeatable differences in exploratory behaviour predict tick infestation probability in wild great tits. Behavioral Ecology and Sociobiology. 75:48.

Norte, A.C., Boyer, P.H., Castillo-Ramierz, S., Chvostáč, M., Brahami, M.O., **Rollins, R.E.**, etal. (2021). The Population Structure of *Borrelia lusitaniae* Is Reflected by a Population Division of Its *Ixodes* Vector. Microorganisms. 9: 933.