Host tropism determination by convergent evolution of immunological evasion in the Lyme disease system

Thomas M. Hart[1,2], Alan P. Dupuis, II[1], Danielle M. Tufts[4], Anna M. Blom[5], Simon Starkey[1], Ryan O. M. Rego[6,7], Sanjay Ram[8], Peter Kraiczy[9], Laura D. Kramer[1,3], Maria A. Diuk-Wasser[4], Sergios-Orestis Kolokotronis[10,11,12]*, and Yi-Pin Lin[1,3]*

[1]Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, NY, USA, [2]Department of Biological Sciences and [3]Department of Biomedical Sciences, State University of New York at Albany, NY, USA, [4]Department of Ecology, Evolution, and Environmental Biology, Columbia University, New York, NY USA, [5]Division of Medical Protein Chemistry, Department of Translational Medicine, Lund University, Malmo, Sweden, [6] Institute of Parasitology, Czech Academy of Sciences, České Budějovice, Czech Republic, [7]Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic, [8]Division of Infectious Diseases and Immunology, University of Massachusetts Medical School, Worcester, MA, USA, [9]Institute of Medical Microbiology and Infection Control, University Hospital of Frankfurt, Frankfurt, Germany, [10]Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, Brooklyn, NY, USA, [11]Institute for Genomic Health, SUNY Downstate Health Sciences University, Brooklyn, NY, USA, [12]Division of Infectious Diseases, Department of Medicine, College of Medicine, SUNY Downstate Health Sciences University, Brooklyn, NY, USA..

Short title: Convergent evolution of a microparasite protein and host tropism

Key words: Lyme disease, CspA, Host tropism, *Borrelia*, Complement, Convergent evolution

*correspondence: Sergios-Orestis Kolokotronis, Ph.D.
Department of Epidemiology and Biostatistics. School of Public Health
SUNY Downstate Health Sciences University
450 Clarkson Ave, MSC43A, Brooklyn, NY 11203, USA
Telephone: +1 718-270-6741; Fax: +1 718-270-2533
Email: SOK@downstate.edu

*correspondence: Yi-Pin Lin, Ph.D.
Division of Infectious Disease, Wadsworth Center, New York State Department of Health
120 New Scotland Ave, Albany, NY 12047, USA
Telephone: +1 518-402-2233; Fax: +1 518-473-1326
Email: Yi-Pin.Lin@health.ny.gov

ABSTRACT (150 words)

Microparasites selectively adapt in some hosts, known as host tropism. Transmitted through ticks and carried mainly by mammals and birds, the Lyme disease (LD) bacterium is a well-suited model to study such tropism. LD bacteria species vary in host ranges through mechanisms eluding characterization. By feeding ticks infected with different LD bacteria species, utilizing feeding chambers and live mice and quail, we found species-level differences of bacterial transmission. These differences localize on the tick blood meal, and complement, a defense in vertebrate blood, and a bacterial polymorphic protein, CspA, which inactivates complement by binding to a host complement inhibitor, FH. CspA selectively confers bacterial transmission to vertebrates that produce FH capable of allele-specific recognition. Phylogenetic analyses revealed convergent evolution as the driver of such findings, which likely emerged during the last glacial maximum. Our results identify LD bacterial determinants of host tropism, defining an evolutionary mechanism that shapes host-microparasite associations.

INTRODUCTION

The interactions between hosts and microparasites (e.g. bacteria, viruses, and protozoa) have arisen through numerous evolutionary events (1, 2), often resulting in generalist microparasites, which adapt to most host environments, or specialists, which selectively survive in particular host species. The association between microparasites and their respective hosts is defined as "host tropism (or host specialization)" (3). For vector-borne microparasites, such a host tropism can be dictated by not only host factors but also host constituents in the vectors (e.g. blood meals) (4). Because many bacteria varying in host specificity are involved in the infection cycle, the Lyme disease bacterium is one of the models regularly applied to investigate the host-

59 microparasite interactions (4). Carried by *Ixodes* ticks, this disease is the most common vector-
60 borne disease in the northern hemisphere (5). The causative agent of Lyme disease is a
61 genospecies complex of the spirochete *Borrelia burgdorferi* sensu lato (also known as
62 *Borreliella burgdorferi* sensu lato (NCBI taxid: 139), Lyme borreliae) (6). Among these
63 genospecies, the most frequently isolated spirochetes from both ticks and vertebrate hosts are *B.*
64 *afzelii*, *B. garinii*, and *B. burgdorferi* sensu stricto (hereafter referred to as *B. burgdorferi*) (7).
65 Following the tick bite, spirochetes need to survive in tick blood meals, which permit
66 transmission to the bite site of the host skin. Survival in host bloodstream then is a prerequisite
67 for hematogenous dissemination and colonization of distant tissues, resulting in varied disease
68 manifestations involving different organs in incidental hosts such as humans (8). In nature, Lyme
69 borreliae can invade vertebrate reservoirs (mainly birds and rodents) but varying in their ability
70 to infect these animals in a host-specific manner (9, 10). For example, some Lyme borreliae
71 species, such as *B. burgdorferi*, can invade a wide range of hosts whereas others are selectively
72 infectious in few host taxa (e.g. *B. afzelii* for rodents and *B. garinii* for birds). However, the
73 molecular basis for such host tropism is largely unclear (9, 10).

74 The elimination of microparasites by host immune responses is a major bottleneck of
75 infectivity, limiting the breadth of host competence (9, 11, 12). Complement is one of the first
76 lines of host defenses in vertebrate animals and can be activated through three canonical routes:
77 the classical, lectin, and alternative pathways (13). The activation of complement on the
78 microparasite surface results in the formation of a protein complex called C3 convertase. C3
79 convertase is essential for complement function because it serves as a protease to cleave C3
80 protein to its active fragments. Further complement activation leads to the lysis of microparasites
81 due to the formation of membrane attack complex pores (C5b-9) on microparasite surface. C3

convertase formed by alternative pathway activation, called C3bBb (composed of C3b and Bb proteins), activates more C3 molecules and results in a positive feedback of the C3 amplification loop, a unique characteristic of the alternative pathway. To avoid tissue damage from unwanted complement activation in the absence of microparasites, hosts possess complement inhibitors (13, 14). One of these inhibitors is factor H (FH), which binds to C3b and blocks further activation of complement (13, 15). Like other microparasites, Lyme borreliae exploit such host self-regulatory mechanisms by recruiting complement inhibitors on their surface (10, 16-18), which allows spirochetes to evade complement-mediated killing in the host bloodstream or tick blood meals (10, 16-18). In fact, spirochetes bind to FH through the production of several bacterial FH-binding proteins, called Complement Regulators Acquiring Surface Proteins (CRASPs): CspA, CspZ, and OspE-related proteins (19, 20). Among these CRASPs, CspA is uniquely required for bacterial transmission from nymphal ticks to vertebrate animals by binding to FH, resulting in complement evasion in those feeding ticks (21-23).

Interestingly, the ability of CspA to bind to FH from different vertebrate hosts varies by alleles and dictates the specificity of a Lyme borreliae species to survive in the sera of various animals (21, 24, 25). Specifically, the ability of CspA variants to bind to mammalian FH and survive in homologous sera correlates with the ability of these variants to promote tick-to-mouse transmission (21). These findings raise a possibility that CspA-mediated, FH binding-dependent complement evasion drives Lyme borreliae host tropism (9, 10). However, that possibility could not be fully demonstrated until we could elucidate the roles of CspA variants that do not promote tick-to-mouse transmission to confer tickborne transmission to other hosts. Additionally, if CspA is one of the spirochete determinants of host tropism, what evolutionary mechanisms give rise to the host-spirochete associations mediated by this protein?

105 In this study, we used Lyme disease spirochetes, avian and mammalian hosts, and the blood of
106 mammals and birds as models to examine the role of complement in driving host tropism of
107 microparasites. We further identified CspA as a molecular determinant for such tropism and
108 elucidated the evolutionary mechanisms resulting in the allelically specific roles of this protein in
109 conferring maintenance of microparasites in diverse hosts during the infection cycle.
110
111 **RESULTS**
112 **Lyme borreliae genospecies differ in their levels of transmission to wild-type but not**
113 **complement-deficient mice and quail.** To examine tick-to-host transmission among Lyme
114 borreliae species, we intradermally injected wild-type BALB/c (wild-type; WT) mice with *B.*
115 *burgdorferi* B31-5A4, *B. garinii* ZQ1, or *B. afzelii* CB43. The tissues from B31-5A4- or CB43-
116 infected mice had significantly greater spirochete burdens than those from uninfected mice (Fig.
117 S1A to D). In contrast, except for the bladders from two mice, bacterial burdens in the tissues of
118 ZQ1-infected mice were below detection limits (10 bacteria per 100ng total DNA) (Fig. S1A to
119 D). To generate ticks harboring equal loads of spirochetes, we intradermally injected each of
120 these strains into C3-deficient BALB/c mice ($C3^{-/-}$) mice, which do not have functional
121 complement. After allowing *I. scapularis* larvae to feed on these mice, we found similar burdens
122 of these strains in all tested tissues, fed larvae, and post molting flat nymphs (Fig. S1 E to J).
123 We then permitted the nymphs carrying each of these strains to feed on WT mice and
124 determined the spirochete burdens in the replete nymphs, the skin at the tick bite site, and blood
125 from these animals at 7 days post feeding (dpf), and uninfected nymphs were included as control.
126 Strains B31-5A4 and CB43 survived at these sites (~10^3 spirochetes per tick or 10^2 to 10^3
127 spirochetes per 100ng DNA of tissues, Fig. 1A to C). Two out of six fed nymphs had

undetectable loads of ZQ1 whereas the other four ticks have bacterial loads ranging from 39 to 111 spirochetes per tick (Fig. 1A). These low, variable values were not significantly different compared to uninfected nymphs (Fig. 1A, $p = 0.52$). Strain ZQ1 was also largely undetectable in tick bite sites and blood (three and five out of five bite sites and blood samples, respectively; Fig. 1B and C). Additionally, we fed nymphs carrying B31-5A4, CB43, or ZQ1 on $C3^{-/-}$ mice and found that these strains were detected in fed nymphs, bite sites, and blood at similar levels (Fig. 1D to F). These results indicate that ZQ1 is less capable of surviving in fed nymphs and establishing infection than B31-5A4 and CB43 after ticks fed only on WT mice, but not on the mice lacking C3, suggesting that mouse complement dictates spirochete transmission. We further studied the tickborne transmission of these strains in a similar fashion using *Coturnix* quail, the avian model of Lyme disease (26, 27). We detected B31-5A4 and ZQ1 in fed nymphs, tick bite sites, and bloodstream (~10^4 spirochetes per tick or 10^2 to 10^3 Lyme borreliae per 100ng DNA of tissues, Fig. 1G to I). Conversely, strain CB43 was not detected above the detection limit in these ticks and tissue samples (Fig. 1G to I). When the nymphs carrying each of these strains were allowed to feed on quail treated with *O. moubata* complement inhibitor (OmCI), which blocks quail complement at the level of activation of C5 (Fig. S2)(28), we found similar levels of spirochetes in fed nymphs, tick bite sites and blood (Fig. 1J to L). These data showed less efficient tick-to-quail transmission of CB43 than that of B31-5A4 and ZQ1, and as was the case in mice, quail complement dictates transmission efficiency among those spirochete strains.

The different abilities of Lyme borreliae genospecies to evade complement in tick blood meals determine mammalian- or avian-specific spirochete transmission. To examine the role that the source of the blood meals plays in determining spirochete transmission, nymphs infected

with B31-5A4, ZQ1, or CB43 were allowed to feed on artificial feeding chambers with human blood, the mammalian blood representative (29)(Fig. 2A). Uninfected nymphs were included as controls. We found that B31-5A4 and CB43 survive in fed nymphs and blood (~10^3 spirochetes per tick and 10^2 spirochetes per 100 ng DNA of blood, Fig. 2B and C). In contrast, ZQ1 was undetectable in the majority of human blood-fed nymphs (six out of ten ticks) (Fig. 2B). Similarly, none of the blood samples fed on by ZQ1-infected nymphs had spirochete burdens significantly greater than uninfected blood (Fig. 2C). We also allowed nymphs carrying B31-5A4, ZQ1, or CB43 to feed on human blood treated with Cobra Venom Factor (CVF), which depletes human complement cascade from the level of C3 (30). We found similar burdens of these strains in the fed nymphs and the blood samples (~10^3 spirochetes per tick and ~10^2 spirochetes per 100 ng DNA of blood, Fig. 2D and E). These results indicate that ZQ1 is less competent than B31-5A4 or CB43 to survive in the tick blood meals from humans during transmission, and active human complement in the blood meals drives survival differences. We also performed similar work using quail blood and detected the strains B31-5A4 and ZQ1 in fed nymphs and blood (~10^3 spirochetes per tick and ~10^2 spirochetes per 100ng total DNA of blood, Fig. 2F and G). However, the strain CB43 was not detected in the majority of fed nymphs (7 out of 13 nymphs, Fig. 2F; $P > 0.05$ compared to uninfected nymphs). Further, spirochete burden values in blood samples fed on by nymphs carrying CB43 were indistinguishable from those of uninfected blood samples (Fig. 2G). Similar levels of each strain were observed when nymphs carrying each of these strains were fed on OmCI-treated quail blood (~10^3 spirochetes per tick and 10^2 spirochetes per 100 ng DNA of blood, Fig. 2H and I). These findings indicate that quail complement limits CB43 survival when ticks feed on quail blood.

CspA-mediated quail FH-binding activity promotes tick-to-quail transmission of spirochetes by complement evasion. We previously showed that a *cspA*-deficient mutant *B. burgdorferi* producing a spirochete outer surface protein, CspA, from the *B. burgdorferi* B31 ($CspA_{B31}$) or *B. afzelii* PKo ($CspA_{PKo}$) but not from *B. garinii* ZQ1 ($CspA_{ZQ1}$), facilitates tick-to-mouse transmission by surviving in fed nymphs (21). This isogenic strain-specific transmission is dependent on the mouse FH-binding activity of these CspA variants to evade complement (21) (Table 1) and recapitulates the mouse-specific tickborne transmissibility of *B. burgdorferi* B31, *B. afzelii* CB43, and *B. garinii* ZQ1 (Fig. 1). (Note that CspA from *B. afzelii* CB43 ($CspA_{CB43}$) shares 99% amino acid identity with $CspA_{PKo}$, making these CspA variants likely confer similar FH-binding and transmission phenotypes (Fig. S3A)). To extend that isogenic strain-specific phenotype to other small mammals, we allowed nymphs carrying the that *cspA*-deficient *B. burgdorferi* producing $CspA_{B31}$, $CspA_{PKo}$, or $CspA_{ZQ1}$ to feed on a rodent reservoir of Lyme borreliae, *Peromyscus leucopus*. We found that expression of $CspA_{B31}$ and $CspA_{PKo}$, but not $CspA_{ZQ1}$, permitted spirochete transmission to that rodent species (Fig. S4). Further, our previous findings of $CspA_{ZQ1}$ (and $CspA_{B31}$) binding to quail FH and promoting survival in quail serum raised the hypothesis that CspA variants that bound FH drives tick-to-quail transmission (21) (Table 1). To test this hypothesis, quail were fed on by the nymphs carrying WT *B. burgdorferi* B31-5A15 (5A15), *B. burgdorferi* B31-5A4NP1Δ*cspA* harboring an empty vector (Δ*cspA*/Vector), or the *cspA*-deficient strain carrying plasmids to express $CspA_{B31}$, $CspA_{ZQ1}$, or $CspA_{PKo}$ in the background of B31-5A4NP1Δ*cspA*. We also included an isogenic strain producing $CspA_{B31}$-L246D, a $CspA_{B31}$ mutant selectively devoid of quail FH-binding activity (21) (Table 1). Uninfected ticks were included as control. The strain 5A15 but not Δ*cspA*/Vector or $cspA_{B31}$-L246D-complemented strain, had burdens above detection limits in the fed nymphs,

tick bite sites and blood (Fig. 3A to C). Strains producing $CspA_{B31}$ or $CspA_{ZQ1}$ but not $CspA_{PKo}$, had detectable burdens in the fed nymphs or tick bite sites and blood (~10^4 spirochetes per tick and ~10^2 spirochetes per 100 ng DNA of tissues or blood, Fig. 3A to C). In contrast, when nymphs carrying each of these spirochete strains were permitted to feed on OmCI-treated quail, all strains showed comparable burdens in fed nymphs, tick bite sites, and blood (Fig. 3D to F). These findings suggest that the $CspA_{ZQ1}$ and $CspA_{B31}$ as quail FH binders promote tick-to-quail transmission by evading complement, and CspA-mediated quail FH-binding activity dictates such a transmission.

Allelically variable, CspA-mediated FH-binding activity confers spirochete complement evasion in tick blood meals and transmissibility in a mammalian and avian blood-specific manner. We sought to examine whether CspA-mediated FH-binding activity facilitates spirochete evasion of complement in tick blood meals, and if that ability determines tickborne transmission in a host-specific manner. Human blood was allowed to be ingested by the nymphs carrying 5A15, Δ*cspA*/Vector, or this strain producing $CspA_{B31}$, $CspA_{ZQ1}$, $CspA_{PKo}$, or $CspA_{B31}$-L246D using feeding chambers. We detected 5A15, but not Δ*cspA*/Vector or $CspA_{B31}$-L246D-producing strains, in fed nymphs and blood (Fig. 4A and B). The $CspA_{B31}$- or $CspA_{PKo}$-producing strains were found in fed nymphs and in human blood (~10^4 spirochetes per tick (Fig. 4A) and more than 10 spirochetes per 100 ng total DNA of blood (Fig. 4B)). Conversely, the $CspA_{ZQ1}$-producing strain was not detectable in these samples (Fig. 4A and B). When we allowed nymphs carrying the same strains to feed on CVF-treated human blood, all strains were detected at similar burdens in fed nymphs and blood (Fig. 4C and D). These results suggest that $CspA_{B31}$ and $CspA_{PKo}$, but not $CspA_{ZQ1}$, permitted transmission to human blood by facilitating

human FH-binding mediated complement evasion in tick blood meals. We also permitted the nymphs carrying the above-mentioned strains to feed on quail blood in the same fashion and found B31-5A15, but not Δ*cspA*/Vector or $CspA_{B31}$-L246D, in the fed nymphs and blood had detectable burdens in both nymphs and blood (Fig. 4E and F). The $CspA_{B31}$- or $CspA_{ZQ1}$-producing strain was readily detected in fed nymphs and blood (Fig. 4E and F). Though three and two ticks carrying the $CspA_{PKo}$ and $CspA_{B31}$-L246D-producing strain, respectively, had burdens greater than detection limits, we did not detect any spirochetes in the remaining 10 nymphs (Fig. 4E). Additionally, the burdens of these strains in the blood were statistically indistinguishable from uninfected blood samples (Fig. 4F). When we performed similar experiments using OmCI-treated quail blood, all strains were found in nymphs and blood at comparable levels (Fig. 4G and H). These results show the contribution of quail FH-binding dependent complement evasion in tick blood meals in promoting transmission to quail blood, and $CspA_{B31}$ and $CspA_{ZQ1}$, but not $CspA_{PKo}$, conferred these activities.

CspA homologs showed discontinuous sequence variation and genospecies-specific polymorphisms. Given the finding that homologous CspA proteins from single strains of *B. burgdorferi*, *B. afzelii*, or *B. garinii* confer distinct host tropism, we examined CspA variation in publicly available sequences. CspA is nested in the fourth clade of a protein family encoded on the linear plasmid 54, lp54 (PFam54-IV) (41-81% nucleotide identity; Fig. 5A) (31, 32). However, the homology of all PFam54-IV proteins makes it difficult to easily identify CspA variants, leading to inaccurate annotations and misidentification (31). We thus compiled publicly available gene sequences encoding PFam54-IV available in GenBank from *B. burgdorferi*, *B. afzelii*, and *B. garinii*, and compared the pairwise nucleotide identities of codon alignments for

these genes from each species. Within PFam54-IV-encoding genes of any one particular *B. burgdorferi* strain, we identified one-to-one orthologous genes based on sequence conservation (>95% identity, green in Fig. S5). In contrast, we found moderate conservation among PFam54-IV homologs lacking such one-to-one orthology (<81% identity, red and yellow in Fig. S5). Sequence divergence patterns (inlets in Fig. S5-S7) allowed us to identify genes encoding $CspA_{B31}$, $CspA_{PKo}$, and $CspA_{ZQ1}$ as CspA orthologs in *B. burgdorferi*, *B. afzelii*, and *B. garinii*, respectively. Among these CspA orthologs, intraspecific diversity (i.e. within genospecies) exceeded 93% identity, while interspecific diversity (i.e. between genospecies) varied from 67 to 72% (Fig. 5B and C). These results suggest a genospecies-specific polymorphism among CspA variants, whereby variants of the same genospecies share notably high identity, while variants of different genospecies share relatively lower identities.

Host-specific FH-binding activity of CspA variants arose through convergent evolution. An average identity of 74% among the genes encoding CspA and other Pfam54-IV proteins raises the possibility that non-CspA members of Pfam54-IV share FH-binding functions. We thus examined the mouse (*Mus musculus*) and quail FH-binding ability of Pfam54-IV from *B. burgdorferi* B31-5A4, *B. afzelii* MMS, and *B. garinii* ZQ1 using ELISA. Note that PFam54-IV members from *B. afzelii* MMS, PKo, and CB43 are nearly identical (>99% identity) (Fig. S3A and S6). We used Pfam54-IV of MMS to represent these proteins of *B. afzelii* given that the recombinant version of these proteins from MMS had been generated in our previous work (24). As expected, PFam54-IV from *B. burgdorferi* B31-5A4, *B. afzelii* MMS, and *B. garinii* ZQ1 did not bind to BSA (Fig. 6A). We found that $CspA_{B31}$ and $CspA_{MMS}$ bound to mouse FH at levels greater than a negative control spirochete protein, DbpA (21) (Fig. 6B). Despite a high

concentration (2 µM) of other recombinant Pfam54-IV used, none bound to mouse FH over baseline levels seen with DbpA (Fig. 6B). Furthermore, we observed that $CspA_{B31}$ and $CspA_{ZQ1}$, but none of other tested Pfam54-IV, bound to quail FH (Fig. 6C). These results indicate that the host-specific and allelically variable FH-binding activity of Pfam54-IV is CspA-dependent.

To further study the evolutionary mechanisms leading to host-specific and allelically variable FH-binding activity of CspA, we estimated phylogenetic relationships among gene sequences encoding PFam54-IV from *B. burgdorferi* B31, *B. afzelii* MMS, and *B. garinii* ZQ1. We found that those sequences of the same genospecies do not form monophyletic assemblages, but CspA variants grouped in separate clades with moderate to high internode branch support (Bayesian posterior probabilities (PP) of 0.81 at $CspA_{B31}$ and $CspA_{MMS}$ nodes and 0.86 at $CspA_{MMS}$ and $CspA_{ZQ1}$ nodes) (Fig. 6D). Similar branching patterns were seen when phylogenetic relationships were estimated among genes encoding PFam54-IV available on GenBank from *B. burgdorferi*, *B. afzelii*, and *B. garinii* (SH-aLRT/ultrafast bootstrap supports of 98.3/100% at $CspA_{B.\ burgdorferi}$ and $CspA_{B.\ afzelii}$ nodes and 84.4/74% at $CspA_{B.\ afzelii}$ $CspA_{B.\ garinii}$ nodes Fig. S8). We then tested the plausibility of this evolutionary scenario by placing CspA variants in the same clades (due to the same FH-binding functions, Fig. S9A, left panel) or cladding PFam54-IV variants from the same genospecies together (Fig. S9A, right panel). The results supported neither alternative phylogeny (Fig. S9B), in agreement with the phylogeny placing CspA variants in separated clades but not with every PFam54-IV protein from the same genospecies. Further, the supported phylogeny raises the possibility that CspA-mediated FH-binding activities arose from 1) a common FH-binding ancestor or 2) the convergent evolution of PFam54-IV (Fig. 6D). However, our results from maximum likelihood and parsimony-based tree-building methods, rejected the former possibility (Fig. 6D and E), indicating that the allelically variable, host-specific FH-

binding activity of CspA is a result of convergent evolution within PFam54-IV. Based on a chromosome mutation rate estimated in a previous study (33), such an evolution event likely occurred approximately 15,000-55,000 years before present, coinciding with the end of the last glacial maximum (Fig. 6D).

DISCUSSION

The constant interaction of microparasites and hosts allows the microparasites to adapt to each of the host environments, by which they can evolve to become specialists (4, 34). However, the fact that generalists are present in nature suggests that generalization of host ranges for those microparasites also confers fitness advantages (4, 34). For vector-borne microparasites, the process leading to host tropism can be driven by host-derived components (i.e. immune molecules or nutrients) either in the hosts or acquired by vectors (4). The molecular determinants and evolutionary mechanisms by which microparasites specialize or generalize to be associated with hosts are largely unclear. Reflected by the variable host tropism of different spirochete genospecies transmitted through *Ixodes* ticks, the Lyme disease bacterium is a well-suited model to study host-microparasite interactions (9, 10). *I. scapularis* ticks were shown in laboratory infections to carry *B. burgdorferi*, *B. afzelii*, and *B. garinii* at similar levels (35, 36), suggesting the use of this tick to represent *Ixodes* vectors of Lyme disease. Using a single tick species carrying each of the tested spirochete species allows for attribution of the observations solely to host and/or pathogen determinants, the emphasis of this study. Nonetheless, *B. afzelii* and *B. garinii* are not endemic to North America where *I. scapularis* are found, and are thus isolated from other *Ixodes* ticks (i.e. *I. ricinus* and *I. persulcatus*) in the field. *B. burgdorferi* is the only Lyme borrelia genospecies in this study that is circulated in *I. scapularis* in nature (5). Thus,

utilizing *I. scapularis* as a vector representative may not completely address the role of vector competence in modulating host tropism of spirochetes (37, 38), which warrants further investigations. Additionally, Lyme borreliae-infected mice have commonly been used to generate nymphs harboring spirochetes through blood feeding by naïve larvae (39). However, rearing nymphs carrying similar burdens of each spirochete species in this fashion is difficult because wild-type mice do not maintain equal loads of these spirochetes (40, 41). Using complement-deficient mice ($C3^{-/-}$ mice), we found similar burdens of *B. burgdorferi* B31-5A4, *B. afzelii* CB43, and *B. garinii* ZQ1 in fed larvae, post-molting flat nymphs, and the tissues derived from spirochete-infected mice. This result provides a strategy to overcome the difficulty in tick-rearing and infection, and supports the concept that complement controls the spirochete infectivity during infection (21, 40, 42).

No definitive studies have been performed to test the long-held model that complement evasion by spirochetes determines Lyme borreliae host tropism (43). A hurdle for such an investigation is the inability to easily maintain and/or persistently infect non-mammalian hosts, such as birds (44-54). Though some wild-birds have been brought into laboratories to study spirochete infectivity (44-54), molecular mechanisms have not been elucidated because of the lack of avian-specific reagents. We and others have intradermally inoculated Lyme borreliae into *Coturnix* quail as this domestic bird can sustain detectable spirochete burdens for more than eight weeks (26, 27). We thus allowed ticks carrying spirochetes to feed on quail, similar to previous work performed in this species and other domestic aves (55-57). We found that *B. garinii* ZQ1 and *B. burgdorferi* B31-5A4 survive in fed ticks and are transmitted to quail whereas *B. afzelii* CB43 did not. These results demonstrate that the genospecies variation of spirochete transmissibility to birds, in agreement with prior studies [reviewed in (9)], supporting the use of quail as an avian

host representative. In contrast, when nymph feeding was performed on wild-type mice, *B. afzelii* CB43 and *B. burgdorferi* B31-5A4 survived in fed ticks and migrated to these animals while *B. garinii* ZQ1 did not. All three species survived in fed ticks and are transmitted to complement-deficient quail or mice. In support of previous *in vitro* evidence (43, 58), this study establishes complement evasion by spirochetes as a driver of Lyme borreliae host tropism. As the presence of complement in vertebrate blood, our findings raise a possibility that spirochetes must evade host complement specifically in the blood meal. However, tick feeding on live animals may introduce confounding factors of blood meal-independent, complement-mediated clearance (59). Using different sources of blood in "artificial feeding chambers" without the involvement of animals allows us to demonstrate that spirochete evasion to complement in blood meals dictates Lyme borreliae host tropism (29).

Both *Ixodes* ticks and Lyme borreliae produce complement-inactivating proteins to facilitate feeding and pathogen transmission (8, 60-62). Supported by many Lyme borreliae proteins are polymorphic, one attractive hypothesis is that polymorphisms in these proteins contribute to host tropism. Regrettably, using wild-type spirochete strains may not delineate the contribution these proteins individually because these spirochetes generate multiple polymorphic complement-inactivating proteins during transmission (22, 63-66). Thus, identical spirochete background strains have been used to express genes or alleles (also known as "isogenic strains") to define bacterial determinants of particular phenotypes (21) (67, 68). We have previously shown that isogenic strains in a *cspA*-deficient background producing CspA from *B. burgdorferi* B31 or *B. afzelii* PKo, but not *B. garinii* ZQ1, promote tick-to-mouse transmission. The ability of ticks to transmit Lyme borreliae is contingent on the ability of the spirochete to survive in fed ticks, consistent with the phenotypes observed using wild-type genospecies (21). Such differential

358 transmissibility depends on the ability of CspA to bind to mammal FH in the presence of mouse
359 complement, leading to the possibility that allelically-variable, CspA-mediated FH-binding
360 activity dictates host tropism (21, 24, 25). Nonetheless, that concept requires identification of the
361 animals that are susceptible to Lyme borreliae that express ZQ1-derived CspA. In agreement
362 with our previous work showing CspA from ZQ1 and B31 but not PKo binds to quail FH (21),
363 we found the ZQ1- and B31- (but not PKo-) derived CspA facilitates tick-to-quail transmission
364 of spirochetes. We further showed that such an allele-dependent, quail-specific transmission is
365 determined by the presence of quail complement and the ability of CspA to bind to this species'
366 FH. Our results using feeding chambers, isogenic bacterial strains and complement-intact or -
367 deficient human or quail blood demonstrated that CspA is a determinant of Lyme borreliae host
368 tropism by promoting host-specific, FH-binding-dependent complement evasion in tick blood
369 meals.

370 CspA, like other Pfam54-IV proteins, experienced numerous events of duplications and
371 deletions, resulting in moderate identity (~40-80%) among the variants in this protein family
372 (31, 32, 69). These observations suggest rapid evolution, potentially indicating novel functions in
373 this region (31, 32, 70). This notion is supported by some important functions (e.g. complement
374 evasion, cell adhesion, plasminogen binding, and tissue colonization) identified for CspA and
375 other Pfam54 proteins (21, 71-74). However, the moderate sequence similarity from limited
376 number of strains in those studies makes it difficult to definitively determine the one-to-one
377 orthologs of each Pfam54-IV member (31, 32), creating hurdles to further investigate the
378 evolutionary mechanisms giving rise to those functions. By comparing the sequences that encode
379 PFam54-IV from different strains within each genospecies, we found that some comparisons
380 showed moderate identity (<80%), while others shared a high degree of identity (~90% or more).

These results allow us to define these highly similar sequences as paralogs. These findings also suggest that the phenotypes conferred by one Pfam54-IV protein (e.g. host-specific FH-binding activity/transmissibility) are likely shared by its orthologs from different strains within the same genospecies (75). Additionally, we observed notably less identity (<79%) when comparing genes encoding Pfam54-IV among different spirochete genospecies, indicating genospecies-specific polymorphisms. These results, combined with the fact that CspA variants confer variable complement evasion and host tropism (21, 24, 25), are similar to findings of allelically-variable phenotypes in other Lyme borreliae polymorphic proteins (65, 67, 68). Further, these data suggest that any unidentified functions of CspA or other Pfam54-IV may vary among the strains from different genospecies.

Some phenotypes are shared among multiple Pfam54-IV proteins from the same Lyme borreliae genospecies (e.g. C7 and C9 binding-mediated complement evasion by Bga66 and Bga72 from *B. bavariensis*) (71). This finding raises the possibility that FH-binding is a common feature for non-CspA Pfam54-IV proteins within the same genospecies. However, our finding showing no detectable mouse or quail FH-binding to non-CspA Pfam54-IV indicates that the FH-binding activity is unique to CspA, in agreement with the results of human FH-binding activity of Pfam54-IV (24). Further, our data from phylogenetic reconstructions rejected the possibility of functional cladding of CspA or a common FH-binding ancestor, but supported convergent evolution as the mechanism leading to such an allelically variable, CspA-mediated FH-binding activity (76). With the fact that such an activity dictates Lyme borreliae host tropism, allelically variable, CspA-mediated FH-binding activity would effectively isolate different populations of Lyme borreliae in their respective hosts in nature, resulting in adaptive radiation of spirochetes. An intriguing question is whether such adaptation is the result of allopatric

ecological speciation of the Lyme borreliae (77-82). However, the phylogenetic reconstruction of Pfam54-IV suggests that CspA of *B. burgdorferi* evolved approximately 15,000 to 55,000 years ago. The emergence of the CspA variants thus likely occurred after ancient Lyme borreliae speciation as the earliest common ancestor of *B. burgdorferi* in North America was dated at ~ 60,000 years ago (33). Rather, that timeline of CspA emergence (~15,000-55,000 years ago) coincides with the latter half of the last glacial maximum in North America and Europe (83, 84). An attractive possibility can be considered that massive climatic changes triggered ecological shifts necessitating new strategies to be maintained in the enzootic cycle, such as allelically variable, CspA-mediated FH-binding activity. Using the Lyme disease bacterium as a model, this work is a pioneering study defining the mechanisms that dictate host tropism of microparasites, identifying the molecular determinants, and elucidating the evolutionary drivers of such host-microparasite associations. These findings will provide significant impacts into the origin of a vector-borne enzootic cycle and establish the groundwork for future studies to investigate the mechanisms in shaping host-microparasite interaction.

MATERIALS AND METHODS

Ethics statement. All mouse and quail 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 protocol was approved by the Institutional Animal Care and Use Committee (IACUC) of Wadsworth Center, New York State Department of Health (Protocol docket number 19-451) and Columbia University (Protocol docket number AC-AAAO4551). All efforts were made to minimize animal suffering.

Mouse, quail, tick, bacterial strains, OmCI, and FH. BALB/c mice were purchased from Taconic (Hudson, NY). $C3^{-/-}$ mice in BALB/c background were generated from the $C3^{-/-}$ (C57BL/6) purchased from Jackson Laboratory (Bar Harbor, ME) as described in our previous study (21). *P. leucopus* mice were ordered from *Peromyscu*s genetic stock center at University of South Carolina (Columbia, SC). *Ixodes scapularis* tick larvae were purchased from National Tick Research and Education Center, Oklahoma State University (Stillwater, OK) or obtained from BEI Resources (Manassas, VA). Lyme borreliae-infected nymphs were generated as described in the section "Generation of ticks carrying Lyme borreliae." The *Borrelia* and *Escherichia coli* strains used in this study are described in Table S1. *E. coli* strains DH5α, M15, and derivatives were grown in Luria-Bertani (BD Bioscience) broth or agar, supplemented with kanamycin (50 µg/ml), ampicillin (100 µg/ml), or no antibiotics as appropriate. All *B. burgdorferi*, *B. afzelii*, and *B. garinii* strains were grown in BSK-II completed medium supplemented with kanamycin (200 µg/mL), streptomycin (50 µg/mL), gentamicin (50 µg/mL), or no antibiotics (see Table S1). Mouse FH was purchased from MyBiosource. Quail FH and recombinant OmCI proteins were generated as described previously (21, 26, 28).

Mouse infection using needle inoculation. Four-week-old female BALB/c or $C3^{-/-}$ mice in BALB/c background were used for experiments involved in needle infection of Lyme borreliae strains. Mice were infected by intradermal injection as previously described (68) with 10^6 of *B. afzelii* CB43, *B. garinii* ZQ1, or *B. burgdorferi* B31-5A4. The plasmid profile of the strain B31-5A4 was verified prior to infection as described to ensure the stability of the vector and no loss of plasmids (85, 86). As the information of genome is not available for plasmid profiling,

the strains CB43 and ZQ1 used in this study were less than ten passages. Mice were sacrificed at 21 days post-infection, the inoculation site of the skin, the ankle joints, ears, and bladder were collected to quantitatively evaluate levels of colonization during infection as described in the section "Determination of Lyme borreliae burdens in infected ticks, tissues and blood samples.".

Generation of ticks carrying Lyme borreliae. The procedure of the tick infection has been described previously (21, 87). Basically, four-week-old male and female $C3^{-/-}$ mice in BALB/c background were infected with 10^6 of *B. afzelii* CB43, *B. garinii* ZQ1, or *B. burgdorferi* B31-5A4, B31-5A15, B31-5A4NP1Δ*cspA*-V or this *cspA* mutant strain producing $CspA_{B31}$, $CspA_{PKo}$, $CspA_{ZQ1}$, or $CspA_{B31}$L246D by intradermal injection as described above. The ear punches from those mice were collected at 13 days post infection, and DNA was extracted to perform qPCR using *Borrelia* 16S rRNA primers as previously described (26) (Table S2) to confirm the infection (See section "Determination of Lyme borreliae burdens in infected ticks, tissues and blood samples."). At 14 days post infection, the uninfected larvae were allowed to feed to repletion on those spirochete-infected mice as described previously (21, 87). Approximately 100 to 200 larvae were allowed to feed on each mouse. The engorged larvae were collected and allowed to molt into nymphs in a desiccator at room temperature and 95% relative humidity in a room with light dark control (light to dark, 16:8 h). (21, 85).

Serum resistance assays. *Coturnix* quail were subcutaneously injected with OmCI (1 mg/kg of quail) or PBS buffer, and the sera were collected at 6, 24, 48, 72, and 96 h post injection. A serum sensitive, high passaged *B. burgdorferi* strain B313 was cultivated to mid-log phase,

followed by being diluted to a final concentration of 5×10^6 bacteria/ml in BSKII medium without rabbit serum. These bacteria were then incubated with each of these quail serum samples (final concentration: 40% of serum). We also included the bacteria mixed with heat inactivated serum samples, which have been incubated at 56ºC for 2 h prior to being mixed with spirochetes. An aliquot was taken from each reaction at 0 and 4 h post injection to determine the number of motile bacteria under a Nikon Eclipse E600 darkfield microscope, as previously described (21). The survival percentage for those motile spirochetes was calculated using the number of mobile spirochetes at 4 h post incubation normalized to that at the very beginning of incubation with serum.

Mouse, quail, and *P. leucopus* infection by ticks. The flat nymphs were placed in a chamber on four- to six-week old male and female BALB/c or $C3^{-/-}$ mice in BALB/c background, and the engorged nymphs were collected from the chambers at seven days post nymph feeding as described (88). For ticks feeding on quail, the feathers located on the back of quail's neck were plucked to expose approximately 2 to 3 cm^2 of skin, close to the back of its head. 1.2 mL screw-top 'cryo' microcentrifuge vials (ThermoFisher Scientific) were cut to be used as mini chambers. The top of the caps from the chambers was pierced with a 25-gauge needle to create air holes, and sand papers were used to smooth any sharp edges along the cut surface edge of these chambers. Vetbond Tissue Adhesive (3M) was used to attach the chambers onto the exposed quail skin followed by manually restraining quail while the surgical glue dries (1-2 min). Ten nymphs were placed into the chambers on mice or quail, which those ticks to feed on. For OmCI-treated quail, the quail were subcutaneously injected with OmCI (1 mg/kg of quail) a day prior to the nymph feeding. The engorged nymphs were obtained from the chambers. The

mice and quail were placed into a small cage, which then placed above the moat in a larger cage (for mice) or plastic bin (for quail). Ticks feeding on *Peromyscus leucopus* mice have been described previously (89, 90). Ten nymphs were placed in the ears of each mouse, five nymphs per ear and were allowed to feed until repletion. *P. leucopus* mice were separately placed in water cages which consisted of the cage being filled with approximately 2.5 cm of water along the bottom, a wire rack to keep the mouse out of the water, and then being placed in a larger hamster cage with water to prevent ticks from escaping. The engorged nymphs were recovered from the water cage beginning five days post nymph feeding. Blood and tick placement site from quail, mice, and *P. leucopus* mice were collected at seven days post nymph feeding.

Feeding chamber assays by ticks. Artificial feeding chambers were prepared as described in our previous study (21). In short, the silicone rubber-saturated rayon membrane was generated as described (21), with the exception of adhering fiberglass mesh (3-mm pore; Lowe's Inc.) to the membrane before attaching it to the rest of the chamber. Such membrane was attached to one side of a 2-cm length of polycarbonate tubing (hereafter called the chamber; inner diameter: 2.5 cm; outer diameter: 3.2 cm; (Amazon Inc.), as described, with the exception of using a rubber band to hold the chamber in place instead of a rubber O-ring (21). Feeding stimuli including hair and hair extract from white-tailed deer (*Odocoileus virginianus*) and a plastic tile spacer (Lowe's Inc.) were added as described with the exception of using 3 stainless steel bearings (Amazon Inc.) instead of a nickel coin (21). *I. scapularis* nymphs carrying *B. burgdorferi* B31-5A4, *B. afzelii* CB43, *B. garinii* ZQ1; or *B. burgdorferi* B31-5A4, B31-5A15, B31-5A4NP1Δ*cspA*-V or this *cspA* mutant strain producing $CspA_{B31}$, $CspA_{PKo}$, $CspA_{ZQ1}$, or

519 $CspA_{B31}$L246D were then added onto the chamber (5-8 ticks/chamber). Chamber feedings were
520 carried out as previously described using human (BioIVT, Westbury, NY) or quail blood
521 (Canola Poultry Market, Brooklyn, NY) (Fig. 2A). Chambers in blood were placed in a sealed
522 Styrofoam cooler with wet paper towels to maintain humidity at approximately 87 to 95%.
523 Depending on the experiments, blood was treated with CVF (ComTech) or OmCI to a final
524 concentration of 17 µg/ml. Blood was changed daily, and was collected along with ticks after 5
525 days of feeding. SYBR-based qPCR was used to determine bacterial burdens in the ticks and
526 blood using *Borrelia* 16S rRNA gene primers (Table S2).
527
528 **Determination of Lyme borreliae burdens in infected ticks, tissues and blood samples.** The
529 ticks fed on quail, mice, *P. leucopus* mice, or feeding chambers were homogenized by hand in
530 a 1.5 ml Eppendorf tube (Eppendorf) with a plastic pestle (ThermoFisher Scientific). The DNA
531 from tissues or blood or homogenized ticks was extracted using the EZ-10 Genomic DNA kit
532 (Biobasic). The quantity and quality of DNA was assessed using a Nanodrop 1000 UV/Vis
533 spectrophotometer (ThermoFisher Scientific). The 280:260 ratio was between 1.75 and 1.85,
534 indicating the lack of contaminating RNA or proteins. qPCR was then performed to quantitate
535 bacterial loads. Spirochete genomic equivalents were calculated using an ABI 7500 Real-Time
536 PCR System (ThermoFisher Scientific) in conjunction with PowerUp SYBR Green Master Mix
537 (ThermoFisher Scientific), based on amplification of the Lyme borreliae 16S rRNA gene using
538 primers 16SrRNAfp and 16SrRNArp (Table S2), as described previously (23, 68). Cycling
539 parameters for SYBR green-based reactions were 50°C for 2 min, 95°C for 10 min, and 45
540 cycles of 95°C for 15 s, 55°C for 30 s, and 60°C for 1 min. The number of 16S rRNA copies
541 was calculated by establishing a threshold cycle (Cq) standard curve of a known number of 16S

542 rRNA gene extracted from *B. burgdorferi* strain B31-5A4, then comparing the Cq values of the
543 experimental samples.
544
545 **Sequence analysis of PFam54-IV.** Nucleotide sequences of the PFam54-IV genes, including
546 *bba68* ($cspA_{B31}$), and *bba69* from *B. burgdorferi* B31 (AE000790), *pkoa0062*, *pkoa0063*,
547 *pkoa0064*, *pkoa0065*, *pkoa0066*, and *pkoa0067* ($cspA_{PKo}$), from *B. afzelii* PKo (CP002950),
548 and *zqa67*, *zqa68* ($cspA_{ZQ1}$), *zsa69*, *zsa70*, *zsa71*, and *zsa72* from *B. garinii* ZQ1 (AJ786369),
549 were used as queries against the NCBI GenBank nr database using BLASTN(91). Sequences
550 from organisms other than *B. burgdorferi*, *B. afzelii*, and *B. garinii* and duplicate sequences
551 were discarded. The remaining sequences were then adjusted to include the full open reading
552 frames with removal of any sequences containing premature stop codons. Sequences were
553 grouped by species, and the sequences of the genes encoding the third and fifth clade of
554 PFam54 (PFam54-III and Pfam54-V, respectively) from *B. burgdorferi* B31, *B. afzelii* PKo,
555 and *B. garinii* ZQ1 were added to the sequence set (Genbank accession codes are found on on
556 Fig. S5 to S7). Codon alignments were generated using T-Coffee on the TranslatorX server
557 (92, 93), followed by calculation of pairwise identity in Clustal Omega (94). Only Pfam54-IV
558 best hits were kept. The remaining sequences were then realigned by codon and pairwise
559 identity was calculated as above. The percent identity between each of these genes and each of
560 the identified PFam54-IV alleles is shown in Fig. S5 to S7.
561
562 **Generation of recombinant PFam54 proteins.** The pQE30Xa vectors encoding the open
563 reading frames lacking the putative signal sequences of *bba68* ($cspA_{B31}$), *bba69* from *B.*
564 *burgdorferi* strain B31-5A4, *mmsa67*, *mmsa68*, *mmsa69*, or *mmsa70*, and *mmsa71* ($cspA_{MMS}$)

565 from *B. afzelii* strain MMS, or *zqa67*, *zqa68* (*cspA*$_{ZQ1}$), *zsa69*, *zsa70*, *zsa71*, or *zsa72* from *B.*
566 *garinii* strain ZQ1 were obtained previously to generate recombinant histidine-tagged proteins
567 (24). The plasmids were transformed into *E. coli* strain M15, and the plasmid inserts were
568 sequenced using Sanger sequencing on an ABI 3730xl DNA Analyzer (ThermoFisher
569 Scientific) at the NYSDOH Wadsworth Center ATGC Core Facility. The resulting M15
570 derived strains were used to produce respective recombinant PFam54-IV (Table S1). The
571 histidine-tagged PFam54-IV were produced and purified by nickel affinity chromatography
572 with Ni-NTA agarose according to the manufacturer's instructions (Qiagen, Valencia, CA).
573
574 **FH binding assay by qualitative ELISA.** Qualitative ELISA for FH binding by PFam54-IV
575 was performed as described (21, 95). One microgram of BSA (negative control; Sigma-Aldrich,
576 St. Louis, MO) or FH from mouse or quail was coated onto microtiter plate wells by incubating
577 the plate for overnight at 4ºC. Then, 100 µl 2 µM of histidine-tagged DbpA from *B.*
578 *burgdorferi* strain B31 (negative control) (96) or each of the PFam54-IV was added to the wells.
579 Mouse anti-histidine tag 1:200× (Sigma-Aldrich, St. Louis, MO) and HRP-conjugated goat
580 anti-mouse IgG 1:1,000× (Seracare Life Sci., Inc, Milford, MA) were used as primary and
581 secondary antibodies, respectively, to detect the binding of histidine-tagged proteins. The plates
582 were washed three times with PBST (0.05% Tween 20 in PBS), and 100 µl of tetramethyl
583 benzidine (TMB) solution (ThermoFisher Scientific) was added to each well and incubated for
584 5 min. The reaction was stopped by adding 100 µl of 0.5% hydrosulfuric acid to each well.
585 Plates were read at 405 nm using a Tecan Sunrise Microplate reader at five minutes after the
586 incubation (Tecan Life science, Männedorf, Switzerland).
587

588 **Phylogenetic reconstruction.** The PFam54-IV codon alignment from *B. burgdorferi* B31, *B.*
589 *afzelii* MMS, and *B. garinii* ZQ1 was used to generate a Bayesian phylogenetic reconstruction
590 was carried out in BEAST v1.8.4 with a relaxed lognormal clock, an estimated mutation rate of
591 4.75×10^{-6} substitutions/site/year and a coalescent Bayesian skyline model (33, 97). A Markov
592 chain Monte Carlo chain length of 10,000,000 steps was used with a 100,000-step thinning,
593 resulting in effective sample sizes greater than 200, an indication of an adequate chain mixing.
594 The resulting maximum clade credibility tree was visualized in FigTree v1.4.4 (98). We
595 evaluated alternative, competing evolutionary scenarios for PFam54-IV based on species-
596 specific divergence, or clustering by FH-binding activity shown in Fig. S10A (99). A battery of
597 statistical phylogenetic tests was deployed in IQ-TREE (Kishino-Hasegawa, Shimodaira-
598 Hasegawa, Expected Likelihood Weight, and Approximately Unbiased tests)(101-104).
599
600 **Statistical analysis.** Significant differences between samples were assessed using the Mann-
601 Whitney *U* test or the Kruskal-Wallis test with the two-stage step-up method of Benjamini,
602 Krieger, and Yekutieli. A P-value < 0.05 (*) or ($^{\#}$) was considered to be significant (105).
603
604 **ACKNOWLEDGEMENTS**
605 We thank Frank Blaisdell, and Dierdre Torrisi from the Wadsworth vet sciences facility and
606 Ashley Marcinkiewicz and Patricia Lederman for animal husbandry, Levi Poirier and Ing-Nang
607 Wang for assistance with gene sequencing and phylogenetic reconstructions, Ashley
608 Marcinkiewicz for critical reading of the manuscript, Patricia Rosa for sharing the unpublished
609 observation of *B. afzelii* strain PKo, and Roxie Giradin for assistance with SDS-PAGE. We
610 thank Wadsworth ATGC core for plasmid sequencing, Leslie Eisele of Wadsworth

611 Biochemistry and Immunology Core for HPLC performance, and Karen Chave of the
612 Wadsworth Protein Expression Core for purifying factor H. This work was supported by NIH-
613 U01CK000509 (DMT and MDW), NSF-IOS1755286 (DMT, MDW, SOK, ADII, LDK, TMH,
614 and YL), DoD-TB170111, NIH-R21AI144891, NIH-R21AI146381, New York State
615 Department of Health Wadsworth Center Start-Up Grant (TMH and YL), the Czech Science
616 Foundation grant No. 17-21244S (ROMR), NIH R01AI121401 (PK), and the LOEWE Center
617 DRUID Novel Drug Targets against Poverty-Related and Neglected Tropical Infectious
618 Diseases, project C3 (PK). The funders had no role in study design, data collection and
619 analysis, decision to publish, or preparation of the manuscript. The authors declare that the
620 research was conducted in the absence of any commercial or financial relationships that could
621 be construed as a potential conflict of interest.

622

623 **REFERENCES**

624 1. Kilpatrick AM, Altizer, S. Disease Ecology. Nature Education Knowledge.
625 2010;3(10):55.
626 2. Killilea ME, Swei A, Lane RS, Briggs CJ, Ostfeld RS. Spatial dynamics of Lyme
627 disease: a review. Ecohealth. 2008;5(2):167-95.
628 3. Douam F, Gaska JM, Winer BY, Ding Q, von Schaewen M, Ploss A. Genetic
629 Dissection of the Host Tropism of Human-Tropic Pathogens. Annual review of genetics.
630 2015;49:21-45.
631 4. Kurtenbach K, Hanincova K, Tsao JI, Margos G, Fish D, Ogden NH. Fundamental
632 processes in the evolutionary ecology of Lyme borreliosis. Nature Reviews Microbiology.
633 2006;4(9):660-9.

5. Steere AC, Strle F, Wormser GP, Hu LT, Branda JA, Hovius JW, et al. Lyme borreliosis. Nature Reviews Disease primers. 2016;2:16090.

6. Radolf JD, Caimano MJ, Stevenson B, Hu LT. Of ticks, mice and men: understanding the dual-host lifestyle of Lyme disease spirochaetes. Nature Reviews Microbiology. 2012;10(2):87-99.

7. Brisson D, Drecktrah D, Eggers CH, Samuels DS. Genetics of *Borrelia burgdorferi*. Annual review of genetics. 2012;46:515-36.

8. Kurokawa C, Lynn GE, Pedra JHF, Pal U, Narasimhan S, Fikrig E. Interactions between *Borrelia burgdorferi* and ticks. Nature Reviews Microbiology. 2020;18(10):587-600.

9. Tufts DM, Hart TM, Chen GF, Kolokotronis SO, Diuk-Wasser MA, Lin YP. Outer surface protein polymorphisms linked to host-spirochete association in Lyme borreliae. Molecular microbiology. 2019.

10. Lin YP, Diuk-Wasser MA, Stevenson B, Kraiczy P. Complement Evasion Contributes to Lyme Borreliae-Host Associations. Trends in parasitology. 2020;36(7):634-45.

11. Zhou W, Brisson D. Interactions between host immune response and antigenic variation that control *Borrelia burgdorferi* population dynamics. Microbiology (Reading). 2017;163(8):1179-88.

12. Gomez-Chamorro A, Battilotti F, Cayol C, Mappes T, Koskela E, Boulanger N, et al. Susceptibility to infection with *Borrelia afzelii* and TLR2 polymorphism in a wild reservoir host. Scientific reports. 2019;9(1):6711.

13. Zipfel PF, Skerka C. Complement regulators and inhibitory proteins. Nature Reviews Immunology. 2009;9(10):729-40.

14. Blom AM. The role of complement inhibitors beyond controlling inflammation. Journal of internal medicine. 2017;282(2):116-28.

15. Jozsi M, Zipfel PF. Factor H family proteins and human diseases. Trends in immunology. 2008;29(8):380-7.

16. Ermert D, Ram S, Laabei M. The hijackers guide to escaping complement: Lessons learned from pathogens. Molecular immunology. 2019;114:49-61.

17. Dulipati V, Meri S, Panelius J. Complement evasion strategies of *Borrelia burgdorferi* sensu lato. FEBS Letters. 2020;594(16):2645-56.

18. Skare JT, Garcia BL. Complement Evasion by Lyme Disease Spirochetes. Trends in Microbiology. 2020.

19. Lin YP, Frye AM, Nowak TA, Kraiczy P. New Insights Into CRASP-Mediated Complement Evasion in the Lyme Disease Enzootic Cycle. Frontiers in Cellular and Infection Microbiology. 2020;10:1.

20. Kraiczy P, Stevenson B. Complement regulator-acquiring surface proteins of *Borrelia burgdorferi*: Structure, function and regulation of gene expression. Ticks and Tick-borne Diseases. 2013;4(1-2):26-34.

21. Hart T, Nguyen NTT, Nowak NA, Zhang F, Linhardt RJ, Diuk-Wasser M, et al. Polymorphic factor H-binding activity of CspA protects Lyme borreliae from the host complement in feeding ticks to facilitate tick-to-host transmission. PLoS Pathogens. 2018;14(5):e1007106.

22. Kraiczy P, Hellwage J, Skerka C, Becker H, Kirschfink M, Simon MM, et al. Complement resistance of *Borrelia burgdorferi* correlates with the expression of BbCRASP-1,

a novel linear plasmid-encoded surface protein that interacts with human factor H and FHL-1 and is unrelated to Erp proteins. The Journal of Biological Chemistry. 2004;279(4):2421-9.

23. Bykowski T, Woodman ME, Cooley AE, Brissette CA, Brade V, Wallich R, et al. Coordinated expression of *Borrelia burgdorferi* complement regulator-acquiring surface proteins during the Lyme disease spirochete's mammal-tick infection cycle. Infection and Immunity. 2007;75(9):4227-36.

24. Wallich R, Pattathu J, Kitiratschky V, Brenner C, Zipfel PF, Brade V, et al. Identification and functional characterization of complement regulator-acquiring surface protein 1 of the Lyme disease spirochetes *Borrelia afzelii* and *Borrelia garinii*. Infection and Immunity. 2005;73(4):2351-9.

25. Hammerschmidt C, Koenigs A, Siegel C, Hallstrom T, Skerka C, Wallich R, et al. Versatile roles of CspA orthologs in complement inactivation of serum-resistant Lyme disease spirochetes. Infection and Immunity. 2014;82(1):380-92.

26. Marcinkiewicz AL, Dupuis AP, 2nd, Zamba-Campero M, Nowak N, Kraiczy P, Ram S, et al. Blood treatment of Lyme borreliae demonstrates the mechanism of CspZ-mediated complement evasion to promote systemic infection in vertebrate hosts. Cellular Microbiology. 2019;21(2):e12998.

27. Isogai E, Tanaka S, Braga IS, 3rd, Itakura C, Isogai H, Kimura K, et al. Experimental *Borrelia garinii* infection of Japanese quail. Infection and Immunity. 1994;62(8):3580-2.

28. Frye AM, Hart TM, Tufts DM, Ram S, Diuk-Wasser MA, Kraiczy P, et al. A soft tick *Ornithodoros moubata* salivary protein OmCI is a potent inhibitor to prevent avian complement activation. Ticks and Tick-borne Diseases. 2020;11(2):101354.

700 29. Hart T, Yang X, Pal U, Lin YP. Identification of Lyme borreliae proteins promoting
701 vertebrate host blood-specific spirochete survival in *Ixodes scapularis* nymphs using artificial
702 feeding chambers. Ticks and Tick-borne Diseases. 2018.
703 30. Finnie JA, Stewart RB, Aston WP. A comparison of cobra venom factor-induced
704 depletion of serum C3 in eight different strains of mice. Dev Comp Immunol. 1981;5(4):697-
705 701.
706 31. Wywial E, Haven J, Casjens SR, Hernandez YA, Singh S, Mongodin EF, et al. Fast,
707 adaptive evolution at a bacterial host-resistance locus: the PFam54 gene array in *Borrelia*
708 *burgdorferi*. Gene. 2009;445(1-2):26-37.
709 32. Qiu WG, Martin CL. Evolutionary genomics of *Borrelia burgdorferi* sensu lato:
710 findings, hypotheses, and the rise of hybrids. Infection, genetics and evolution : Journal of
711 Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases. 2014;27:576-93.
712 33. Walter KS, Carpi G, Caccone A, Diuk-Wasser MA. Genomic insights into the ancient
713 spread of Lyme disease across North America. Nat Ecol Evol. 2017;1(10):1569-76.
714 34. O'Keeffe KR, Oppler ZJ, Brisson D. Evolutionary ecology of Lyme *Borrelia*. Infection,
715 genetics and evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in
716 Infectious Diseases. 2020:104570.
717 35. Dolan MC, Piesman J, Mbow ML, Maupin GO, Peter O, Brossard M, et al. Vector
718 competence of *Ixodes scapularis* and *Ixodes ricinus* (Acari: Ixodidae) for three genospecies of
719 Borrelia burgdorferi. Journal of Medical Entomology. 1998;35(4):465-70.
720 36. Eisen L. Vector competence studies with hard ticks and *Borrelia burgdorferi* sensu lato
721 spirochetes: A review. Ticks and Tick-borne Diseases. 2020;11(3):101359.

722 37. Pospisilova T, Urbanova V, Hes O, Kopacek P, Hajdusek O, Sima R. Tracking of
723 *Borrelia afzelii* Transmission from Infected Ixodes ricinus Nymphs to Mice. Infection and
724 Immunity. 2019;87(6).
725 38. Sertour N, Cotte V, Garnier M, Malandrin L, Ferquel E, Choumet V. Infection Kinetics
726 and Tropism of *Borrelia burgdorferi* sensu lato in Mouse After Natural (via Ticks) or Artificial
727 (Needle) Infection Depends on the Bacterial Strain. Frontiers in Microbiology. 2018;9:1722.
728 39. Rosa PA, Tilly K, Stewart PE. The burgeoning molecular genetics of the Lyme disease
729 spirochaete. Nature Reviews Microbiology. 2005;3(2):129-43.
730 40. Lawrenz MB, Wooten RM, Zachary JF, Drouin SM, Weis JJ, Wetsel RA, et al. Effect
731 of complement component C3 deficiency on experimental Lyme borreliosis in mice. Infection
732 and Immunity. 2003;71(8):4432-40.
733 41. van Burgel ND, Balmus NC, Fikrig E, van Dam AP. Infectivity of *Borrelia burgdorferi*
734 sensu lato is unaltered in C3-deficient mice. Ticks and Tick-borne Diseases. 2011;2(1):20-6.
735 42. Woodman ME, Cooley AE, Miller JC, Lazarus JJ, Tucker K, Bykowski T, et al.
736 *Borrelia burgdorferi* binding of host complement regulator factor H is not required for efficient
737 mammalian infection. Infection and Immunity. 2007;75(6):3131-9.
738 43. Kurtenbach K, De Michelis S, Etti S, Schafer SM, Sewell HS, Brade V, et al. Host
739 association of *Borrelia burgdorferi* sensu lato--the key role of host complement. Trends in
740 Microbiology. 2002;10(2):74-9.
741 44. Ginsberg HS, Buckley PA, Balmforth MG, Zhioua E, Mitra S, Buckley FG. Reservoir
742 competence of native North American birds for the lyme disease spirochete, *Borrelia*
743 *burgdorfieri*. Journal of Medical Entomology. 2005;42(3):445-9.

744 45. Heylen D, Adriaensen F, Van Dongen S, Sprong H, Matthysen E. Ecological factors
745 that determine *Ixodes ricinus* tick burdens in the great tit (*Parus major*), an avian reservoir of
746 *Borrelia burgdorferi* s.l. International Journal for Parasitology. 2013;43(8):603-11.
747 46. Heylen DJA, Muller W, Vermeulen A, Sprong H, Matthysen E. Virulence of recurrent
748 infestations with *Borrelia*-infected ticks in a *Borrelia*-amplifying bird. Scientific Reports.
749 2015;5:16150.
750 47. Heylen D, Fonville M, van Leeuwen AD, Sprong H. Co-infections and transmission
751 dynamics in a tick-borne bacterium community exposed to songbirds. Environ Microbiol.
752 2016;18(3):988-96.
753 48. Heylen DJ, Sprong H, Krawczyk A, Van Houtte N, Genne D, Gomez-Chamorro A, et
754 al. Inefficient co-feeding transmission of *Borrelia afzelii* in two common European songbirds.
755 Scientific Reports. 2017;7:39596.
756 49. Norte AC, Lopes de Carvalho I, Nuncio MS, Araujo PM, Matthysen E, Albino Ramos
757 J, et al. Getting under the birds' skin: tissue tropism of *Borrelia burgdorferi* s.l. in naturally and
758 experimentally infected avian hosts. Microb Ecol. 2020;79(3):756-69.
759 50. Norte AC, Costantini D, Araujo PM, Eens M, Ramos JA, Heylen D. Experimental
760 infection by microparasites affects the oxidative balance in their avian reservoir host the
761 blackbird *Turdus merula*. Ticks and Tick-borne Diseases. 2018;9(3):720-9.
762 51. Richter D, Spielman A, Komar N, Matuschka FR. Competence of American robins as
763 reservoir hosts for Lyme disease spirochetes. Emerging Infectious Diseases. 2000;6(2):133-8.
764 52. Piesman J, Dolan MC, Schriefer ME, Burkot TR. Ability of experimentally infected
765 chickens to infect ticks with the Lyme disease spirochete, *Borrelia burgdorferi*. The American
766 Journal of Tropical Medicine and Hygiene. 1996;54(3):294-8.

767 53. Olsen B, Gylfe A, Bergstrom S. Canary finches (*Serinus canaria*) as an avian infection
768 model for Lyme borreliosis. Microbial Pathogenesis. 1996;20(6):319-24.
769 54. Burgess EC. Experimental inoculation of mallard ducks (*Anas platyrhynchos*
770 *platyrhynchos*) with Borrelia burgdorferi. Journal of Wildlife Diseases. 1989;25(1):99-102.
771 55. Kurtenbach K, Peacey M, Rijpkema SG, Hoodless AN, Nuttall PA, Randolph SE.
772 Differential transmission of the genospecies of *Borrelia burgdorferi* sensu lato by game birds
773 and small rodents in England. Applied and Environmental Microbiology. 1998;64(4):1169-74.
774 56. Moraru GM, Goddard J, Paddock CD, Varela-Stokes A. Experimental infection of
775 cotton rats and bobwhite quail with *Rickettsia parkeri*. Parasites & Vectors. 2013;6:70.
776 57. Kurtenbach K, Carey D, Hoodless AN, Nuttall PA, Randolph SE. Competence of
777 pheasants as reservoirs for Lyme disease spirochetes. Journal of Medical Entomology.
778 1998;35(1):77-81.
779 58. Kurtenbach K, Sewell HS, Ogden NH, Randolph SE, Nuttall PA. Serum complement
780 sensitivity as a key factor in Lyme disease ecology. Infection and Immunity. 1998;66(3):1248-
781 51.
782 59. Bernard Q, Grillon A, Lenormand C, Ehret-Sabatier L, Boulanger N. Skin Interface, a
783 Key Player for *Borrelia* Multiplication and Persistence in Lyme Borreliosis. Trends in
784 Parasitology. 2020;36(3):304-14.
785 60. Tyson K, Elkins C, Patterson H, Fikrig E, de Silva A. Biochemical and functional
786 characterization of Salp20, an *Ixodes scapularis* tick salivary protein that inhibits the
787 complement pathway. Insect Mol Biol. 2007;16(4):469-79.

788 61. Schuijt TJ, Hovius JW, van Burgel ND, Ramamoorthi N, Fikrig E, van Dam AP. The
789 tick salivary protein Salp15 inhibits the killing of serum-sensitive *Borrelia burgdorferi* sensu
790 lato isolates. Infection and Immunity. 2008;76(7):2888-94.
791 62. Schuijt TJ, Coumou J, Narasimhan S, Dai J, Deponte K, Wouters D, et al. A tick
792 mannose-binding lectin inhibitor interferes with the vertebrate complement cascade to enhance
793 transmission of the lyme disease agent. Cell Host & Microbe. 2011;10(2):136-46.
794 63. Kraiczy P, Skerka C, Brade V, Zipfel PF. Further characterization of complement
795 regulator-acquiring surface proteins of *Borrelia burgdorferi*. Infection and Immunity.
796 2001;69(12):7800-9.
797 64. Hellwage J, Meri T, Heikkila T, Alitalo A, Panelius J, Lahdenne P, et al. The
798 complement regulator factor H binds to the surface protein OspE of *Borrelia burgdorferi*. The
799 Journal of Biological Chemistry. 2001;276(11):8427-35.
800 65. Caine JA, Lin YP, Kessler JR, Sato H, Leong JM, Coburn J. *Borrelia burgdorferi* outer
801 surface protein C (OspC) binds complement component C4b and confers bloodstream survival.
802 Cellular Microbiology. 2017.
803 66. Garcia BL, Zhi H, Wager B, Hook M, Skare JT. *Borrelia burgdorferi* BBK32 Inhibits
804 the Classical Pathway by Blocking Activation of the C1 Complement Complex. PLoS
805 Pathogens. 2016;12(1):e1005404.
806 67. Lin YP, Tan X, Caine JA, Castellanos M, Chaconas G, Coburn J, et al. Strain-specific
807 joint invasion and colonization by Lyme disease spirochetes is promoted by outer surface
808 protein C. PLoS Pathogens. 2020;16(5):e1008516.

809 68. Lin YP, Benoit V, Yang X, Martinez-Herranz R, Pal U, Leong JM. Strain-specific
810 variation of the decorin-binding adhesin DbpA influences the tissue tropism of the lyme
811 disease spirochete. PLoS Pathogens. 2014;10(7):e1004238.
812 69. Brangulis K, Akopjana I, Petrovskis I, Kazaks A, Tars K. Structural analysis of the
813 outer surface proteins from *Borrelia burgdorferi* paralogous gene family 54 that are thought to
814 be the key players in the pathogenesis of Lyme disease. Journal of Structural Biology.
815 2020;210(2):107490.
816 70. Kondrashov FA, Rogozin IB, Wolf YI, Koonin EV. Selection in the evolution of gene
817 duplications. Genome Biology. 2002;3(2):RESEARCH0008.
818 71. Hammerschmidt C, Klevenhaus Y, Koenigs A, Hallstrom T, Fingerle V, Skerka C, et
819 al. BGA66 and BGA71 facilitate complement resistance of *Borrelia bavariensis* by inhibiting
820 assembly of the membrane attack complex. Molecular Microbiology. 2016;99(2):407-24.
821 72. Tkacova Z, Pulzova LB, Mochnacova E, Jimenez-Munguia I, Bhide K, Mertinkova P,
822 et al. Identification of the proteins of *Borrelia garinii* interacting with human brain
823 microvascular endothelial cells. Ticks and Tick-borne Diseases. 2020;11(4):101451.
824 73. Koenigs A, Hammerschmidt C, Jutras BL, Pogoryelov D, Barthel D, Skerka C, et al.
825 BBA70 of *Borrelia burgdorferi* is a novel plasminogen-binding protein. The Journal of
826 Biological Chemistry. 2013;288(35):25229-43.
827 74. Lin T, Gao L, Zhang C, Odeh E, Jacobs MB, Coutte L, et al. Analysis of an ordered,
828 comprehensive STM mutant library in infectious *Borrelia burgdorferi*: insights into the genes
829 required for mouse infectivity. PLoS One. 2012;7(10):e47532.
830 75. Tian W, Skolnick J. How well is enzyme function conserved as a function of pairwise
831 sequence identity? Journal of molecular biology. 2003;333(4):863-82.

832 76. Arendt J, Reznick D. Convergence and parallelism reconsidered: what have we learned
833 about the genetics of adaptation? Trends Ecol Evol. 2008;23(1):26-32.
834 77. Becker NS, Margos G, Blum H, Krebs S, Graf A, Lane RS, et al. Recurrent evolution of
835 host and vector association in bacteria of the *Borrelia burgdorferi* sensu lato species complex.
836 BMC Genomics. 2016;17(1):734.
837 78. Mechai S, Margos G, Feil EJ, Barairo N, Lindsay LR, Michel P, et al. Evidence for
838 Host-Genotype Associations of *Borrelia burgdorferi* Sensu Stricto. PLoS One.
839 2016;11(2):e0149345.
840 79. Ogden NH, Feil EJ, Leighton PA, Lindsay LR, Margos G, Mechai S, et al. Evolutionary
841 aspects of emerging Lyme disease in Canada. Applied and Environmental Microbiology.
842 2015;81(21):7350-9.
843 80. Ogden NH, Mechai S, Margos G. Changing geographic ranges of ticks and tick-borne
844 pathogens: drivers, mechanisms and consequences for pathogen diversity. Frontiers in Cellular
845 and Infection Microbiology. 2013;3:46.
846 81. Vollmer SA, Feil EJ, Chu CY, Raper SL, Cao WC, Kurtenbach K, et al. Spatial spread
847 and demographic expansion of Lyme borreliosis spirochaetes in Eurasia. Infection, genetics
848 and evolution : Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious
849 Diseases. 2013;14:147-55.
850 82. Schulze-Lefert P, Panstruga R. A molecular evolutionary concept connecting nonhost
851 resistance, pathogen host range, and pathogen speciation. Trends Plant Sci. 2011;16(3):117-25.
852 83. Yokoyama Y, Lambeck K, De Deckker P, Johnston P, Fifield LK. Timing of the Last
853 Glacial Maximum from observed sea-level minima. Nature. 2000;406(6797):713-6.

854 84. Clark PU, Dyke AS, Shakun JD, Carlson AE, Clark J, Wohlfarth B, et al. The Last
855 Glacial Maximum. Science. 2009;325(5941):710-4.
856 85. Purser JE, Norris SJ. Correlation between plasmid content and infectivity in *Borrelia*
857 *burgdorferi*. Proceedings of the National Academy of Sciences of the United States of
858 America. 2000;97(25):13865-70.
859 86. Bunikis I, Kutschan-Bunikis S, Bonde M, Bergstrom S. Multiplex PCR as a tool for
860 validating plasmid content of *Borrelia burgdorferi*. Journal of Microbiological Methods.
861 2011;86(2):243-7.
862 87. Kern A, Zhou CW, Jia F, Xu Q, Hu LT. Live-vaccinia virus encapsulation in pH-
863 sensitive polymer increases safety of a reservoir-targeted Lyme disease vaccine by targeting
864 gastrointestinal release. Vaccine. 2016;34(38):4507-13.
865 88. Coleman AS, Yang X, Kumar M, Zhang X, Promnares K, Shroder D, et al. *Borrelia*
866 *burgdorferi* complement regulator-acquiring surface protein 2 does not contribute to
867 complement resistance or host infectivity. PLoS One. 2008;3(8):3010e.
868 89. Derdakova M, Dudioak V, Brei B, Brownstein JS, Schwartz I, Fish D. Interaction and
869 transmission of two *Borrelia burgdorferi* sensu stricto strains in a tick-rodent maintenance
870 system. Applied and Environmental Microbiology. 2004;70(11):6783-8.
871 90. Rynkiewicz EC, Brown J, Tufts DM, Huang CI, Kampen H, Bent SJ, et al. Closely-
872 related *Borrelia burgdorferi* (sensu stricto) strains exhibit similar fitness in single infections
873 and asymmetric competition in multiple infections. Parasites & Vectors. 2017;10(1):64.
874 91. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search
875 tool. Journal of Molecular Biology. 1990;215(3):403-10.

876 92. Notredame C, Higgins DG, Heringa J. T-Coffee: A novel method for fast and accurate
877 multiple sequence alignment. Journal of Molecular Biology. 2000;302(1):205-17.
878 93. Abascal F, Zardoya R, Telford MJ. TranslatorX: multiple alignment of nucleotide
879 sequences guided by amino acid translations. Nucleic Acids Research. 2010;38(Web Server
880 issue):W7-13.
881 94. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, et al. The EMBL-EBI
882 search and sequence analysis tools APIs in 2019. Nucleic Acids Research.
883 2019;47(W1):W636-W41.
884 95. Lin YP, Greenwood A, Nicholson LK, Sharma Y, McDonough SP, Chang YF.
885 Fibronectin binds to and induces conformational change in a disordered region of leptospiral
886 immunoglobulin-like protein B. The Journal of Biological Chemistry. 2009;284(35):23547-57.
887 96. Benoit VM, Fischer JR, Lin YP, Parveen N, Leong JM. Allelic variation of the Lyme
888 disease spirochete adhesin DbpA influences spirochetal binding to decorin, dermatan sulfate,
889 and mammalian cells. Infection and Immunity. 2011;79(9):3501-9.
890 97. Suchard MA, Lemey P, Baele G, Ayres DL, Drummond AJ, Rambaut A. Bayesian
891 phylogenetic and phylodynamic data integration using BEAST 1.10. Virus Evol.
892 2018;4(1):vey016.
893 98. Rambaut A. FigTree GitHub repository [Available from:
894 http://tree.bio.ed.ac.uk/software/figtree/.
895 99. Maddison WP, Maddison, D.R. Mesquite: a modular system for evolutionary analysis.
896 Version 3.61 2019 [Available from: http://www.mesquiteproject.org.
897 100. Kishino H, Miyata, T., Hasegawa, M. Maximum likelihood inference of protein
898 phylogeny and the origin of chloroplasts. Journal of Molecular Evolution. 1990;31:151-60.

899 101. Kishino H, Hasegawa M. Evaluation of the maximum likelihood estimate of the
900 evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea.
901 J Mol Evol. 1989;29(2):170-9.
902 102. Shimodaira H, Hasegawa, M. Multiple Comparisons of Log-Likelihoods with
903 Applications to Phylogenetic Inference. Molecular Biology and Evolution. 1999;16(8):1114.
904 103. Strimmer K, Rambaut A. Inferring confidence sets of possibly misspecified gene trees.
905 Proc Biol Sci. 2002;269(1487):137-42.
906 104. Shimodaira H. An approximately unbiased test of phylogenetic tree selection. Syst Biol.
907 2002;51(3):492-508.
908 105. Benjamini YK, A. M. Yekutieli, D. . Adaptive linear step-up procedures that control the
909 false discovery rate. Biometrika. 2006;93:491-507.
910 106. Kraiczy P, Hartmann K, Hellwage J, Skerka C, Kirschfink M, Brade V, et al.
911 Immunological characterization of the complement regulator factor H-binding CRASP and Erp
912 proteins of *Borrelia burgdorferi*. International journal of medical microbiology : IJMM.
913 2004;293 Suppl 37:152-7.
914 107. Stepanova-Tresova G, Kopecky J, Kuthejlova M. Identification of *Borrelia burgdorferi*
915 sensu stricto, *Borrelia garinii* and *Borrelia afzelii* in *Ixodes ricinus* ticks from southern
916 Bohemia using monoclonal antibodies. Zentralblatt fur Bakteriologie : International Journal of
917 Medical Microbiology. 2000;289(8):797-806.
918 108. Kenedy MR, Vuppala SR, Siegel C, Kraiczy P, Akins DR. CspA-mediated binding of
919 human factor H inhibits complement deposition and confers serum resistance in *Borrelia*
920 *burgdorferi*. Infection and Immunity. 2009;77(7):2773-82.

921 109. Tilly K, Krum JG, Bestor A, Jewett MW, Grimm D, Bueschel D, et al. *Borrelia*
922 *burgdorferi* OspC protein required exclusively in a crucial early stage of mammalian infection.
923 Infection and Immunity. 2006;74(6):3554-64.

924

TABLES

Table 1. Summarized findings of allelically variable, host-specific transmissibility and FH-binding activity of *B. burgdorferi*, *B. afzelii*, and *B. garinii* and their derived PFam54-IV proteins in this study and the previous study (21).

Spirochete strains	Transmissibility by wild type strains[a]		PFam54-IV variants	Transmissibility by Δ*cspA* producing PFam54-IV[c]		FH-binding activity[f]	
	Mouse	Quail		Mouse	Quail	Mouse	Quail
***B. burgdorferi* B31-5A4**	+	+	BBA68	+	+	+	+
			BBA69	n.d.[d]	n.d.	-	-
***B. afzelii* CB43, PKo, or MMS**	+[b]	-[b]					
			MMSA67	n.d.	n.d.	-	-
			MMSA68	n.d.	n.d.	-	-
			MMSA69	n.d.	n.d.	-	-
			MMSA70	n.d.	n.d.	-	-
			PKoA71 ($CspA_{PKo}$) or MMSA71 ($CspA_{MMS}$)	+[e]	-[e]	+[g]	-[g]
***B. garinii* ZQ1**	-	+	ZQA67	n.d.	n.d.	-	-
			ZQA68 ($CspA_{ZQ1}$)	-	+	-	+
			ZSA69	n.d.	n.d.	-	-
			ZSA70	n.d.	n.d.	-	-
			ZSA71	n.d.	n.d.	-	-
			ZSA72	n.d.	n.d.	-	-

[a]Significant increase or no difference in spirochete burdens in indicated mouse or quail compared to uninfected respective animals is defined as "+" and "-", respectively. The results were shown in Figure 1A to C and G to I.

[b]Determined using *B. afzelii* CB43

[c]Significant increase or no difference in spirochete burdens in indicated mouse or quail compared to uninfected respective animals is defined as "+" and "-", respectively. The results are shown in Figure 3A to C and (21).

[d]Note determined.

[e]Determined using the *B. burgdorferi* strain Δ*cspA* harboring the plasmid producing PKoA71($CspA_{PKo}$).

[f]Determined by qualitative ELISA in Figure 6A to C and described in (21). Statistically increasing or no different levels of binding by indicated PFam54-IV compared to that by negative control DbpA proteins are defined as "+" and "-."

[g]Determined using recombinant version of MMSA71 ($CspA_{MMS}$).

[h]Dertermined by reducing and non-reducing SDS-PAGE shown in Figure S8.

[i]Unable to determine due to no detection of proteins on nonreducing SDS-PAGE.

946

947

FIGURE LEGENDS

Figure 1. Lyme borreliae display species-level variation of tickborne transmission to wildtype but not complement-deficient mice and quail. *Ixodes scapularis* nymphs infected with *B. burgdorferi* B31-5A4 ("*Bb* B31-5A4"), *B. garinii* ZQ1 ("*Bg* ZQ1"), or *B. afzelii* CB43 ("*Ba* CB43") fed on **(A-C)** BALB/c mice, **(D- F)** $C3^{-/-}$ BALB/c mice, **(G- I)** *Coturnix* quail, or **(J- L)** OmCI-treated *Coturnix* quail. Uninfected nymphs and mouse and quail tissues were included as control ("Uninfect."). Fed nymphs were collected upon repletion, and blood and the tick bite sites of skin were collected at 7 days post nymph feeding ("dpf"). Spirochete burdens in **(A, D, G, and J)** replete nymphs, **(B, E, H, and K)** tick bite site of skin ("Inoc. site"), and **(C, F, I, and L)** blood were determined by qPCR. For the burdens in tissue and blood samples, the resulting values were normalized to 100ng total DNA. Shown are the geometric means of bacterial loads ± 95% confidence interval of bacterial burdens in tissues and blood from 5 mice or quail per group or nymphs feeding on mice (7 nymphs carrying *Bb* B31-5A4, 6 nymphs carrying *Bg* ZQ1, or 9 nymphs carrying *Ba* CB43), $C3^{-/-}$ mice (15 nymphs carrying *Bb* B31-5A4 or *Bg* ZQ1, or 13 nymphs carrying *Ba* CB43), quail (10 nymphs carrying *Bb* B31-5A4, 13 nymphs carrying *Bg* ZQ1, or 17 nymphs carrying *Ba* CB43), or OmCI-treated quail (15 nymphs carrying *Bb* B31-5A4, 11 nymphs carrying *Bg* ZQ1, or 15 nymphs carrying *Ba* CB43). Significant differences ($P < 0.05$, Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli) in the spirochete burdens relative to uninfected ticks or tissues are indicated with an asterisk.

Figure 2. Complement in tick blood meals determines human or quail blood-specific spirochete transmission in feeding chamber. (A) (left panel) The schematic diagram showing

the artificial feeding chamber that is used to examine the tickborne spirochete transmission in this study. **(Right panel)** The picture showing the engorged *I. scapularis* nymphs (indicated by arrows) in the chamber feeding on OmCI-treated quail blood. **(B-I)** *I. scapularis* nymphs infected with *B. burgdorferi* B31-5A4 ("*Bb* B31-5A4"), *B. garinii* ZQ1 ("*Bg* ZQ1"), or *B. afzelii* CB43 ("*Ba* CB43"), were allowed to feed in artificial feeding chambers submerging into six well plates containing **(B and C)** untreated or **(D and E)** CVF-treated human blood, or **(F and G)** untreated or **(H and I)** OmCI-treated quail blood. Blood was changed every 24-h and collected along with ticks on the fifth day of feeding. Uninfected nymphs and blood were included as control ("Uninfect."). Spirochete burdens in **(B, D, F, and H)** fed nymphs and **(C, E, G, and I)** blood were determined by qPCR. The spirochete burdens in the blood were obtained by normalizing the resulting values to 100 ng total DNA. Shown are the geometric means of bacterial loads ± 95% confidence interval of bacterial burdens from the 3 human or quail blood samples or nymphs feeding on untreated human blood (15 nymphs carrying *Bb* B31-5A4, 10 nymphs carrying *Bg* ZQ1, or 13 nymphs carrying *Ba* CB43), CVF-treated human blood (11 nymphs carrying *Bb* B31-5A4, *Bg* ZQ1, or *Ba* CB43), quail blood (8 nymphs carrying *Bb* B31-5A4 or *Bg* ZQ1 or 13 nymphs carrying *Ba* CB43), or OmCI-treated quail blood (15 nymphs carrying *Bb* B31-5A4 or *Bg* ZQ1 or 16 nymphs carrying *Ba* CB43). Significant differences ($P < 0.05$, Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli) in the spirochete burdens relative to uninfected ticks or blood are indicated with an asterisk

Figure 3. Polymorphic quail FH-binding activity of CspA confer the distinct transmissibility of Lyme borreliae to quail. *Ixodes scapularis* nymphs infected with WT *B.*

994 *burgdorferi* B31-5A15 ("B31-5A15"), *B. burgdorferi* B31-5A4NP1*ΔcspA* transformed with an
995 empty shuttle vector ("Vector"), or this deletion strain producing a mutated variant of CspA from
996 *B. burgdorferi* B31 selectively devoid of quail FH binding activity ("L246D"), WT *B.*
997 *burgdorferi* B31 ("B31"), *B. garinii* ZQ1 ("ZQ1"), or *B. afzelii* PKo ("PKo") were allowed to
998 feed on **(A-C)** untreated or **(D-F)** OmCI-treated quail. Uninfected nymphs and quail tissues were
999 included as control ("Uninfect."). Fed nymphs were collected upon repletion, and blood and
1000 tissues were collected at 7 days post nymph feeding ("dpf"). Spirochete burdens in **(A and D)**
1001 replete nymphs, **(B and E)** tick bite sites of skin ("Inoc. site"), and **(C and F)** blood were
1002 determined by qPCR. For the burdens in tissue samples, the resulting values were normalized to
1003 100ng total DNA. Shown are the geometric means of bacterial loads ± 95% confidence interval
1004 of bacterial burdens from 5 quail tissues and blood or nymphs feeding on untreated quail (13
1005 nymphs carrying the strains B31-5A15 or pCspA-PKo, 15 nymphs carrying the strains "Vector",
1006 12 nymphs carrying the strain pCspA-B31, 21 nymphs carrying the strain pCspA-ZQ1, or 17
1007 nymphs carrying the strain pCspA-L246D), or the nymphs from OmCI-treated quail (15 nymphs
1008 carrying the strains B31-5A15, pCspA-B31, or pCspA-PKo, 8 nymphs carrying the strains
1009 "Vector", 9 nymphs carrying the strain pCspA-L246D, 8 nymphs carrying the strain pCspA-
1010 ZQ1). Significant differences ($P < 0.05$, Kruskal-Wallis test with the two-stage step-up method
1011 of Benjamini, Krieger, and Yekutieli) in the spirochete burdens relative to uninfected ticks or
1012 quail tissues are indicated with an asterisk.
1013
1014 **Figure 4. Allelically variable FH-binding activity of CspA dictates human- and quail blood-**
1015 **specific transmission in feeding chambers by evading complement in tick blood meals.** *I.*
1016 *scapularis* nymphs infected with WT *B. burgdorferi* B31-5A15 ("B31-5A15"), *B. burgdorferi*

B31-5A4NP1*ΔcspA* transformed with an empty shuttle vector ("Vector"), or this deletion strain transformed to produce a mutated variant of CspA from *B. burgdorferi* B31 selectively devoid of FH binding activity ("L246D"), WT *B. burgdorferi* B31 ("B31"), *B. garinii* ZQ1 ("ZQ1"), or *B. afzelii* PKo ("PKo") were allowed to feed in feeding chambers submerged into 6-well plates containing **(A and B)** untreated or **(C and D)** CVF-treated human blood, or **(E and F)** untreated or **(G and H)** OmCI-treated quail blood. Blood was changed every 24-h and was collected along with ticks on the fifth day of feeding. Uninfected nymphs and blood were included as control ("Uninfect."). Spirochete burdens from **(A, C, E and G)** fed nymphs and **(B, D, F, and H)** blood were determined by qPCR. For the burdens in tissue samples, the resulting values were normalized to 100ng total DNA. Shown are the geometric means of bacterial loads and 95% confidence interval of bacterial burdens from 3 human and quail blood samples, or nymphs feeding on untreated human blood (15 nymphs carrying the strains B31-5A15 or pCspA-B31, 14 nymphs carrying the strains "Vector" or pCspA-PKo, 17 nymphs carrying the strain pCspA-L246D, or 16 nymphs carrying the strain pCspA-ZQ1), or CVF-treated human blood (9 nymphs carrying the strain B31-5A15, 7 nymphs carrying the strains "Vector", pCspA-B31, or pCspA-ZQ1, or 11 nymphs carrying the strains pCspA-L246D or pCspA-PKo), untreated quail blood (11 nymphs carrying the strains B31-5A15 or "Vector", 8 nymphs carrying the strains pCspA-B31 or pCspA-PKo, or 7 nymphs carrying the strains pCspA-ZQ1 or pCspA-L246D) or OmCI-treated quail blood (15 nymphs carrying the strains B31-5A15, pCspA-PKo, or pCspA-ZQ1, 12 nymphs carrying the strains "Vector" or pCspA-B31, or 13 nymphs carrying the strain pCspA-L246D). Significant differences ($P < 0.05$, Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli) in the spirochete burdens relative to uninfected ticks or quail tissues are indicated with an asterisk.

Figure 5. Pairwise comparisons reveal spirochete genospecies-specific CspA polymorphisms. (A) The synteny of Pfam54-IV proteins from *B. burgdorferi* B31, *B. afzelii* PKo, and *B. garinii* ZQ1 **(B)** CspA variants from *B. burgdorferi*, *B. afzelii*, and *B. garinii* identified in Fig. S5-S7 were aligned by codon using T-Coffee on the TranslatorX server. **(C)** The pairwise identity percentages were plotted versus the number of comparisons with those respective values, with bin widths of 0.25. Clear breaks in the pairwise sequence identity distribution were inspected to differentiate the highly identical (> 95%) from moderately divergent comparisons (< 80%). More than 95% identity in the comparisons of the variants between strains within spirochete genospecies but < 72% identity between genospecies indicate genospecies-specific *cspA* variation.

Figure 6. Allelically variable, host-specific FH-binding activity of CspA has emerged from convergent evolution. (A to C) 2 µM of histidine-tagged Pfam54-IV proteins from *B. burgdorferi* B31-5A4, *B. afzelii* PKo, or *B. garinii* ZQ1 or recombinant histidine-tagged DbpA from *B. burgdorferi* B31-5A4 (negative control) were added in triplicate to wells coated with of purified **(A)** BSA (negative control) or **(B)** mouse or **(C)** quail FH. The protein binding was measured by ELISA in three independent experiments. Each bar represents the geometric mean ± 95% confidence interval of three replicates in one representative experiment. Significant differences ($P < 0.05$, Kruskal-Wallis test with the two-stage step-up method of Benjamini, Krieger, and Yekutieli) in the levels of FH binding of indicated proteins relative to DbpA (‘‘*’’) are indicated. The ability of each protein in binding to factor H is summarized in Table 1. **(D)** A Bayesian phylogenetic reconstruction was generated based on the nucleotide sequences encoding

Pfam54-IV proteins from *B. burgdorferi* B31 (blue), *B. afzelii* PKo (green), *B. garinii* ZQ1 (orange). In brief, these sequences were aligned by codon using T-Coffee on the TranslatorX webserver. Phylogenetic reconstruction was generated based on the resulting nucleotide alignment using BEAST2 with a lognormal uncorrelated relaxed clock with an estimated mutation rate of 4.75×10^{-6} substitutions/site/year ("s/s/y") with a coalescent Bayesian skyline population. The resulting tree is drawn to scale, with branch lengths measured in the number of substitutions per site and rooted at the midpoint for clarity. The scale bar at the bottom represents an approximate timeline of evolution, in years before present ("YBP"), using the estimated substitution rate of 4.75×10^{-6} substitutions/site/year. Node bars represent the 95% highest posterior density of the node age. Node circles represent the posterior probability support. Branches are colored based on estimated the median substitution rate as per the legend to the left. **(E)** Maximum likelihood- and parsimony-based ancestral state reconstructions were used to predict FH-binding activities at ancestral nodes. FH-binding activities were not predicted at any node, and the likelihood of FH-binding activities at nodes immediately prior to CspA variants are indicated.

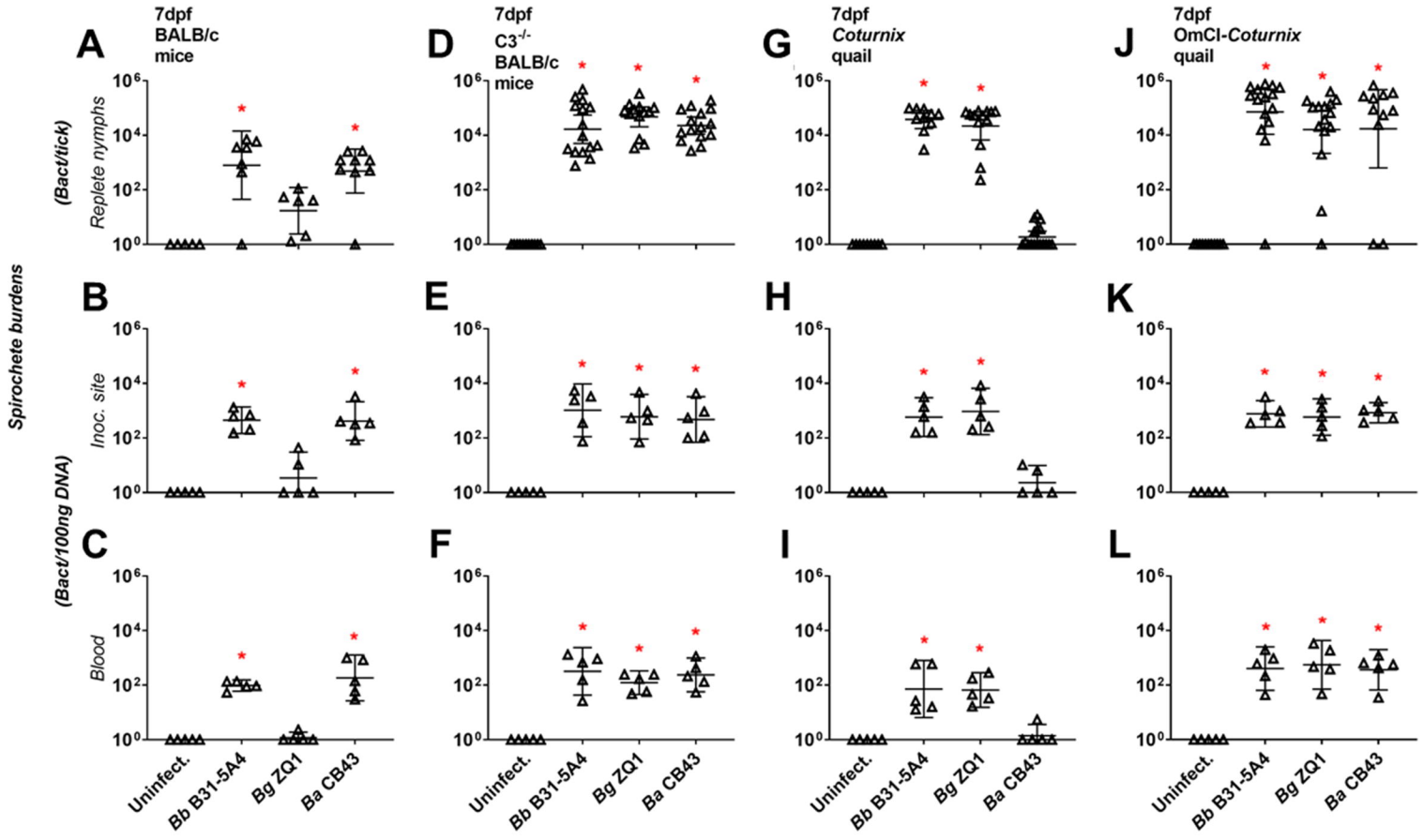
1086

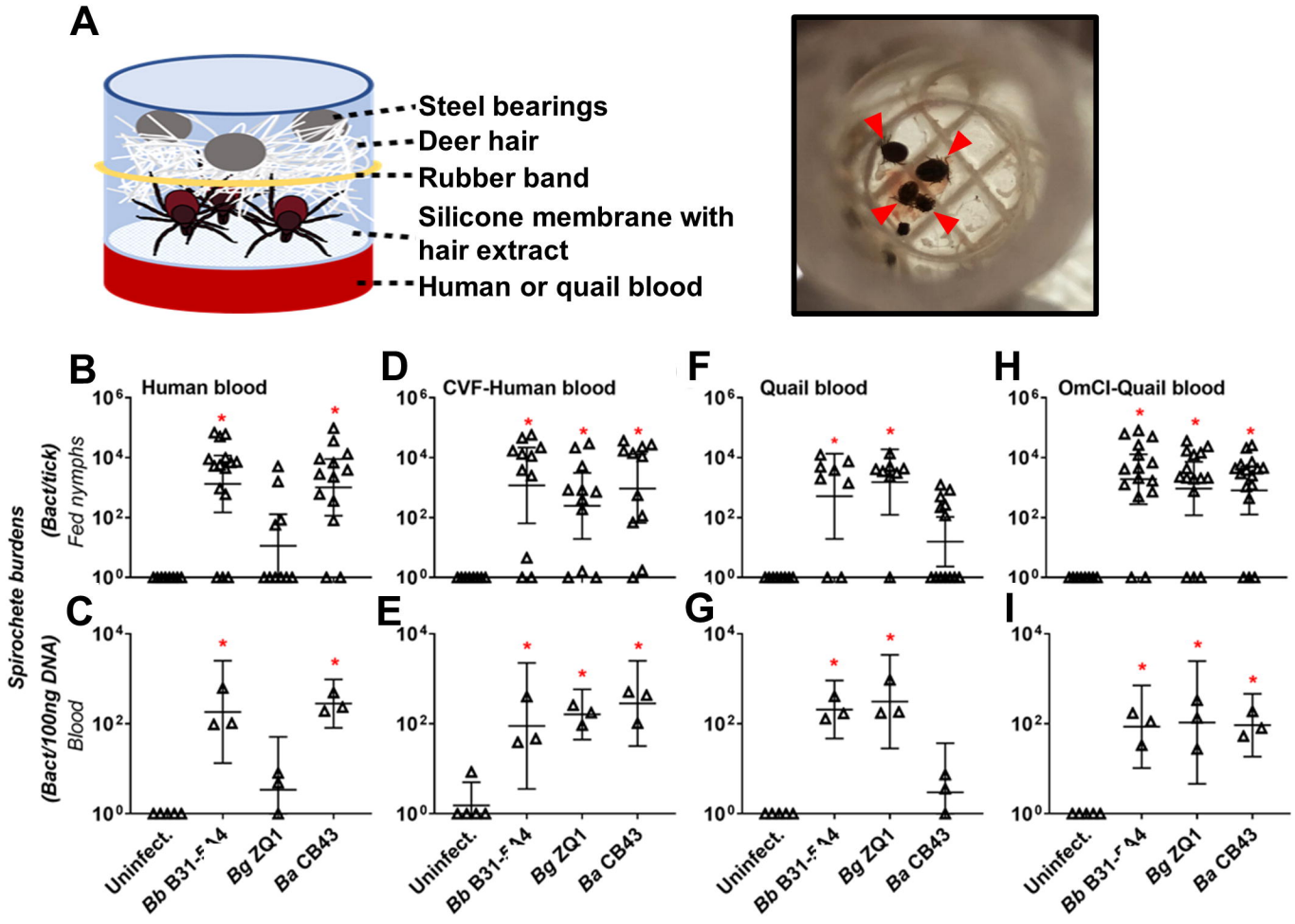
1087

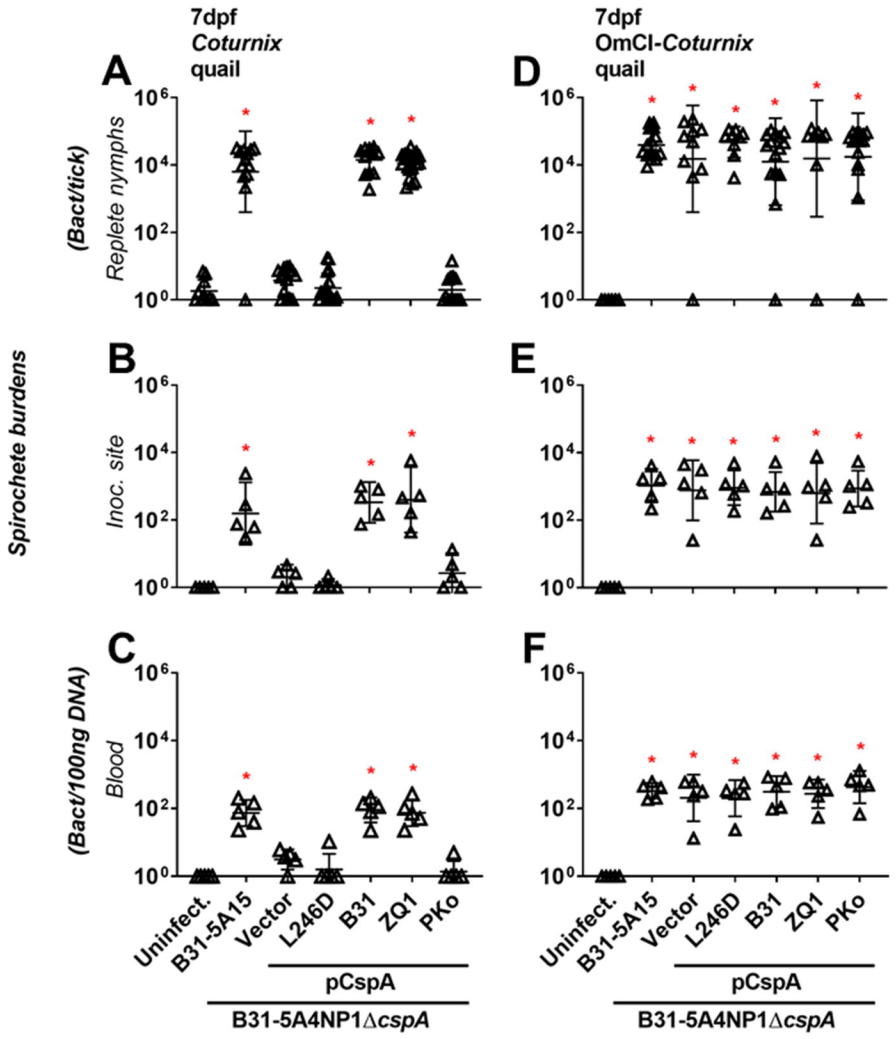
1088

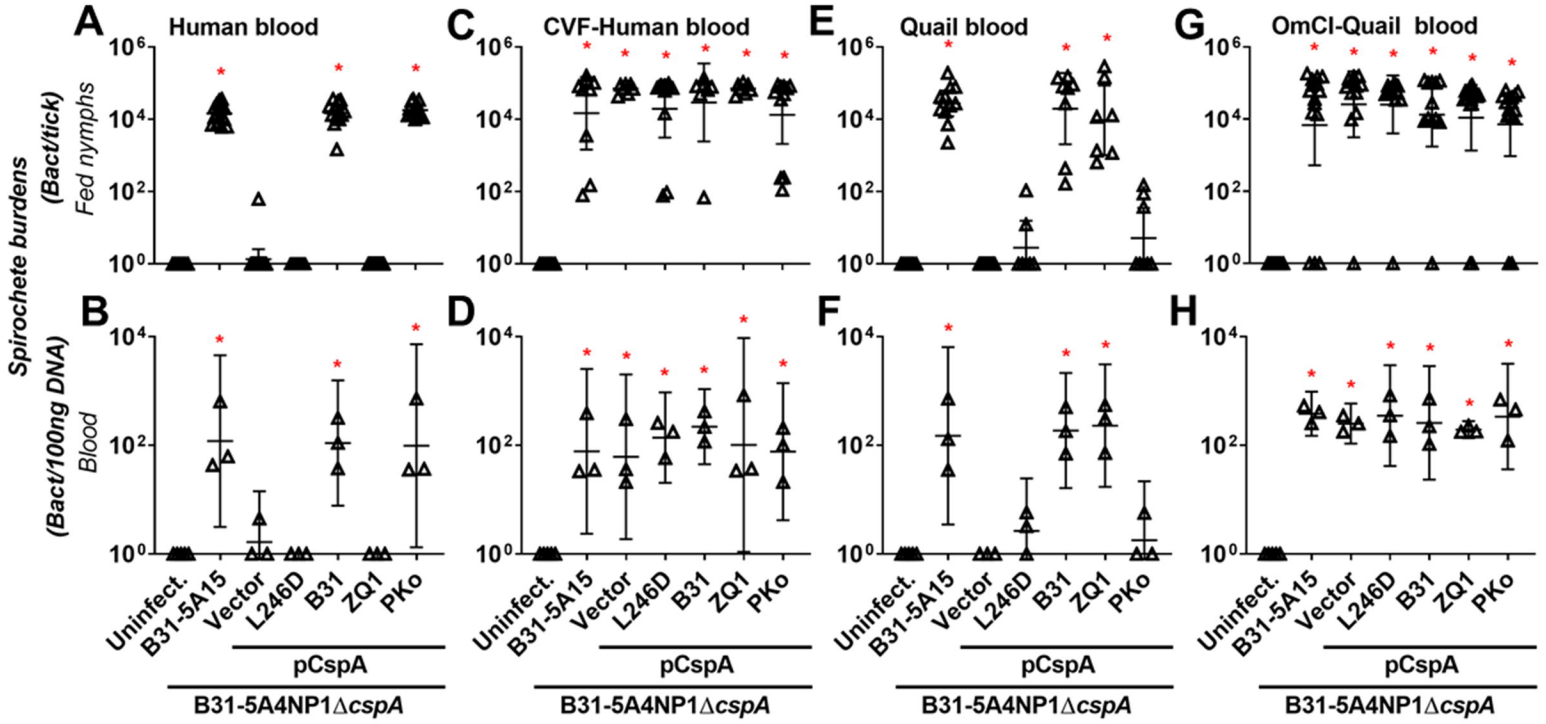
1089

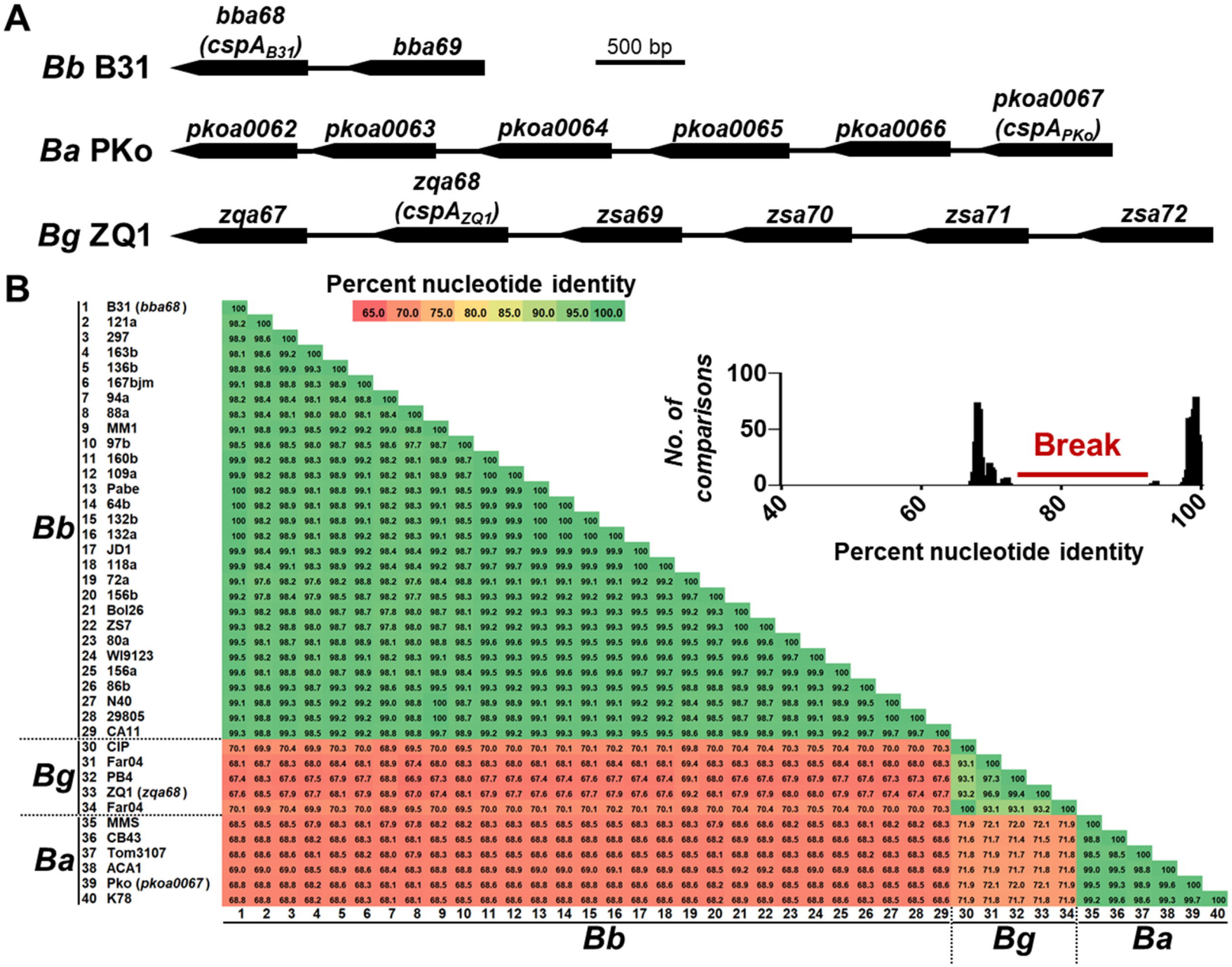
1090

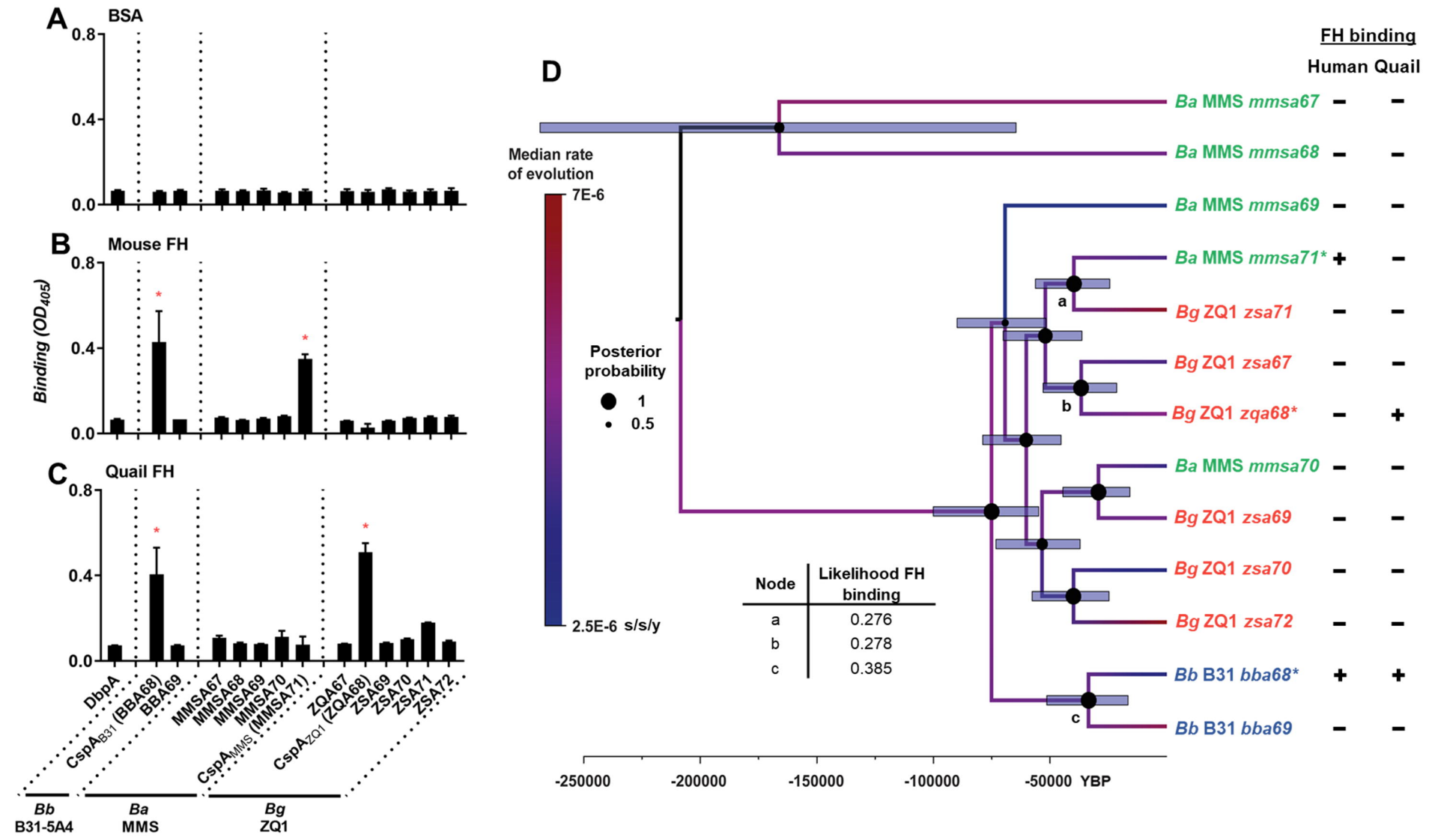

A
7dpf
BALB/c
mice
D
7dpf
C3-/-
BALB/c
mice
G
7dpf
Coturnix
quail
J
7dpf
OmCI-Coturnix
quail
B
E
H
K
C
F
I
L
Spirochete burdens
(Bact/tick)
Replete nymphs
(Bact/100ng DNA)
Inoc. site
Blood
10^6
10^4
10^2
10^0
Uninfect.
Bb B31-5A4
Bg ZQ1
Ba CB43


A
Steel bearings
Deer hair
Rubber band
Silicone membrane with hair extract
Human or quail blood
B
Human blood
D
CVF-Human blood
F
Quail blood
H
OmCI-Quail blood
Spirochete burdens
(Bact/tick)
Fed nymphs
(Bact/100ng DNA)
Blood
C
E
G
I
Uninfect.
Bb B31-5A4
Bg ZQ1
Ba CB43


7dpf
Coturnix
quail
7dpf
OmCI-Coturnix
quail
A
B
C
D
E
F
Spirochete burdens
(Bact/tick)
Replete nymphs
Inoc. site
(Bact/100ng DNA)
Blood
10^6
10^4
10^2
10^0
Uninfect.
B31-5A15
Vector
L246D
B31
ZQ1
PKo
pCspA
B31-5A4NP1Δcspa


A
Human blood
C
CVF-Human blood
E
Quail blood
G
OmCI-Quail blood
B
D
F
H
Spirochete burdens
(Bact/tick)
Fed nymphs
(Bact/100ng DNA)
Blood
10^6
10^4
10^2
10^0
Uninfect.
B31-5A15
Vector
L246D
B31
ZQ1
PKo
pCspA
B31-5A4NP1Δ*cspA*


A
bba68
(cspA B31)
bba69
500 bp
Bb B31
pkoa0062
pkoa0063
pkoa0064
pkoa0065
pkoa0066
pkoa0067
(cspA PKo)
Ba PKo
zqa68
(cspA ZQ1)
zqa67
zsa69
zsa70
zsa71
zsa72
Bg ZQ1
B
Percent nucleotide identity
65.0 70.0 75.0 80.0 85.0 90.0 95.0 100.0
1 B31 (bba68)
2 121a
3 297
4 163b
5 136b
6 167bjm
7 94a
8 88a
9 MM1
10 97b
11 160b
12 109a
13 Pabe
14 64b
15 132b
16 132a
17 JD1
18 118a
19 72a
20 156b
21 Bol26
22 ZS7
23 80a
24 WI9123
25 156a
26 86b
27 N40
28 29805
29 CA11
30 CIP
31 Far04
32 PB4
33 ZQ1 (zqa68)
34 Far04
35 MMS
36 CB43
37 Tom3107
38 ACA1
39 Pko (pkoa0067)
40 K78
Bb
Bg
Ba
No. of comparisons
100
50
0
Break
40
60
80
100
Percent nucleotide identity


A
BSA
B
Mouse FH
C
Quail FH
Binding (OD_{405})
0.8
0.4
0.0
DbpA
CspA$_{B31}$ (BBA68)
BBA69
MMSA67
MMSA68
MMSA69
MMSA70
CspA$_{MMS}$ (MMSA71)
ZQA67
CspA$_{ZQ1}$ (ZQA68)
ZSA69
ZSA70
ZSA71
ZSA72
Bb B31-5A4
Ba MMS
Bg ZQ1
D
FH binding
Human Quail
Median rate of evolution
7E-6
2.5E-6 s/s/y
Posterior probability
1
0.5
Node
Likelihood FH binding
a 0.276
b 0.278
c 0.385
Ba MMS mmsa67 − −
Ba MMS mmsa68 − −
Ba MMS mmsa69 − −
Ba MMS mmsa71* + −
Bg ZQ1 zsa71 − −
Bg ZQ1 zsa67 − −
Bg ZQ1 zqa68* − +
Ba MMS mmsa70 − −
Bg ZQ1 zsa69 − −
Bg ZQ1 zsa70 − −
Bg ZQ1 zsa72 − −
Bb B31 bba68* + +
Bb B31 bba69 − −
-250000
-200000
-150000
-100000
-50000 YBP