Host adaptation drives genetic diversity in a vector-borne disease system

Matthew A. Combs[a,h1], Danielle M. Tufts[a,b1], Ben Adams[c], Yi-Pin Lin[d,e], Sergios-Orestis Kolokotronis[f,g,h*] Maria A. Diuk-Wasser[a*]

[a] Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY 10027
[b] Infectious Diseases and Microbiology Department, University of Pittsburgh, Pittsburgh, PA 15261
[c] Department of Mathematical Sciences, University of Bath, Bath, United Kingdom, BA27AY
[d] Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, NY 12201
[e] Department of Biomedical Sciences, State University of New York at Albany, Albany, NY 12203
[f] Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, Brooklyn, NY 11203
[g] Division of Infectious Diseases, Department of Medicine, College of Medicine, SUNY Downstate Health Sciences University, Brooklyn, NY 11203
[h] Institute for Genomics in Health, SUNY Downstate Health Sciences University, Brooklyn, NY 11203

[1] Equal contribution

* Co-corresponding authors:
Maria Diuk-Wasser, Department of Ecology, Evolution and Environmental Biology, Columbia University, 1200 Amsterdam Ave, Schermerhorn Extension 10th Floor, New York, NY, 10027, mad2256@columbia.edu
Sergios-Orestis Kolokotronis, Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, 450 Clarkson Ave, MSC43, Brooklyn, NY 11203, sok@downstate.edu

Classification: Biological Sciences (Major), Evolution (Minor)

Keywords: Zoonotic, Borrelia, Host, Vector, Pathogen, Adaptation, Multiple niche polymorphism, Frequency dependent selection

Abstract The range of hosts a pathogen can infect is a key trait influencing human disease risk and reservoir host infection dynamics. *Borrelia burgdorferi* sensu stricto (*Bb*), an emerging zoonotic pathogen, causes Lyme disease and is widely considered a host generalist, commonly infecting mammals and birds. Yet the extent of intraspecific variation in *Bb* host breadth, its role in determining host competence and potential implications to human infection remain unclear. We conducted a long-term study of *Bb* diversity, defined by the polymorphic *ospC* locus, across white-footed mice, passerine birds, and tick vectors leveraging long-read amplicon sequencing. Our results reveal strong variation in host breadth across *Bb* genotypes, exposing a spectrum of genotype-specific host-adapted phenotypes. We found support for multiple niche polymorphism maintaining *Bb* diversity in nature and little evidence of temporal shifts in genotype dominance, as would be expected under negative frequency-dependent selection. Passerine birds support the circulation of several human invasive strains in the local tick population and harbor greater *Bb* genotypic diversity compared to white-footed mice. Mouse-adapted *Bb* genotypes exhibited longer persistence in individual mice compared to non-adapted genotypes and infection communities infecting individual mice preferentially became dominated by mouse-adapted genotypes over time. We posit that intraspecific variation in *Bb* host breadth and specificity helps maintain overall species fitness in response to transmission by a generalist vector. Because pathogen genotypes vary in host breadth and result in diverse human disease manifestations, our findings indicate that a more nuanced definition of 'host competence' incorporating local genotype frequency is warranted.

Significance Lyme disease is the most common vector-borne disease in the US with a causative agent (*Borrelia burgdorferi*) exhibiting high genetic diversity that partially correlates with human disease manifestations. Understanding the extent of host specificity in pathogens is critical for evaluating disease risk, but host specificity and mechanisms maintaining genetic diversity in Bb are unknown. We show that Bb genotypes exhibit variable host adaptation to white-footed mice and passerine birds, two common reservoir hosts, which appears to promote high intraspecific pathogen diversity. Conversely, we find limited evidence of negative frequency-dependent selection, an alternative mechanism for diversity maintenance. Our results reveal cryptic intraspecies host breadth variation and suggest that evaluating host competence depends on the frequency of host-adapted genotypes in local environments.

Introduction

Evaluating human disease risk from zoonotic pathogens, those shared between wildlife and humans, requires characterization of pathogen traits influencing their environmental distribution and spillover potential (1). Recent meta-analyses and reviews have repeatedly identified host breadth, or the capacity to infect phylogenetically diverse species, as a key trait influencing spillover (2–4). Greater host breadth provides pathogens more opportunities for population persistence across heterogeneous environments (e.g. variable host availability), but often invokes a fitness tradeoff as pathogens must adapt to diverse immunological selection pressures (5).

Wider pathogen host breadth is intrinsically linked to increased competence of potential hosts, defined as the capacity to acquire, maintain, and transmit infection (6). Though host competence is increasingly used to model human disease risk, unexplained sources of variation can limit its utility (6, 7). In particular, intraspecific genetic variation in multi-strain pathogens may influence host-pathogen interactions, preventing accurate characterization of pathogen host breadth and host competence (8).

To understand the consequences of genotypic diversity on pathogen phenotypes and human disease risk, one must also characterize the evolutionary drivers that maintain diversity and the ecological context in which variation is observed. Balancing selection can maintain diverse genotypes across a bacterial population when heterogeneous immunological and physiological host environments select for different traits (9). While observing evidence of balancing selection can be straightforward, identifying the eco-evolutionary processes responsible requires careful experimental design or intensive and targeted population sampling (9–11). Thus, to evaluate diversity patterns and evolutionary drivers of pathogen traits, it is

crucial to sample pathogens across endemic hosts and time scales relevant to detect natural selection (12, 13).

Vector-borne pathogens are increasingly responsible for emerging infectious diseases (14) and may provide unique insight into the links between the evolution of host breadth and ecology of human disease risk. *Borrelia burgdorferi* sensu stricto (hereafter *Bb*) is a spirochete bacterium that causes 476,000 human cases of Lyme disease (LD) in the United States annually (15). LD is the most common vector-borne disease in the US and continues to increase in geographic range and case numbers, particularly in the Northeast and Midwest US, where *Bb* circulates primarily in small mammal and bird species via the generalist tick vector *Ixodes scapularis* (Radolf *et al.* 2020; CDC 2022). Despite the increasing toll of tick-borne pathogens on human health, fundamental questions about their basic biological traits, including host breadth, and the role of immune environments in structuring underlying selection mechanisms, remain unanswered (18, 19).

Genetic studies of *Bb* regularly identify elevated genetic diversity and signatures of balancing selection at the pathogen's outer surface protein C (*ospC*) locus in natural populations of infected ticks, with more than 25 known alleles (20–22). The OspC protein is required for host infection (23) and *ospC* alleles are often used to distinguish among strains with variable phenotypes, including a subset deemed human-invasive strains (HIS) that exhibit more severe pathology in humans (24–26). The eco-evolutionary drivers maintaining *ospC* variation and implications for human disease risk remain a focus of active debate (27). Two competing hypotheses exist, centered around distinct host-pathogen interactions. The first suggests *Bb* diversity is maintained via multiple niche polymorphism (MNP) in which genotypes exhibiting host-adapted phenotypes segregate across different host species (28, 29). The second predicts

that negative frequency-dependent selection (NFDS) occurs via host antibodies, which iteratively induce fitness costs on common genotypes within the local population, driving temporal fluctuations in genotype frequency (30, 31).

Here we present the most comprehensive long-term study of *Bb ospC* diversity to date examined across two divergent reservoir host types, passerine birds and white-footed mice, as well as tick vectors in the United States to date. By leveraging long-read amplicon sequencing of the *ospC* locus, we reveal intraspecific variation in host adaptation phenotypes that yield strong support to the MNP hypothesis, while the absence of strain dominance shifts indicates a minor role for NFDS in maintaining strain diversity. Our results shed light on the drivers of balancing selection and the evolution of host breadth for this emerging zoonotic pathogen.

Results

Genotypic community structure

To understand the distribution and diversity of *ospC* major groups (oMGs, hereafter described as genotypes), we sequenced the *ospC* gene from 553 white-footed mice, 92 passerine birds (11 species, Table S1), and 628 individual nymphal *Ixodes scapularis* ticks. In total, we assigned 696,453 HiFi reads to 21 genotypes (Supplementary File 1) and sequencing depth per sample varied between 20 and 9874 reads. Coinfection with multiple genotypes was common across all samples (52.8%). Individual genotype richness ranged from 1 – 16, and was lower on average in mice (α=1.77) than in birds (α=2.59) and intermediate in nymphal ticks (α=2.30). Sequencing depth showed a correlation with genotype richness only at shallower coverage below 100 reads ($p < 0.001$; Figure S1). We used Hill numbers to evaluate diversity profiles (32) and

found that *Bb* diversity and evenness in mice are lower compared to those of birds and nymphs, despite similar richness across populations (Figure S2).

We identified three novel genotypes. One was designated subtype Cj and shares 97.3% sequence identity with *Bb* type C, but contains a 75bp region that is fully identical with *Bb* type J – evidence of a recent recombination event (see below). Another genotype exhibited >8% sequence dissimilarity with all known oMG genotypes and was designated type J3, which was found almost exclusively in birds (present in 0.2% of *Bb* infected mice compared to 17.4% of *Bb* infected birds). A third genotype was quite divergent with type T being its closest relative (11.2% dissimilarity), but matched *Borrelia kurtenbachii* (*B.kurt.*) *ospC* at 99.8%. *B.kurt.* is a recently described genospecies and close relative of *Bb* that infects mammals exclusively (33). *B.kurt.* was the only genotype never identified in birds (Table S1) and exhibited a distinct lack of co-occurrence with other genotypes in mice (Figure 3), indicating a low frequency of mixed infection between *Bb* and *B.kurt* among mice (4.8% of *Bb* infected mice).

Evidence supporting MNP mechanisms in hosts

If genotypic diversity in *Bb* communities is maintained by MNP, we expect significant associations between genotypes and hosts, with nymphal ticks harboring all circulating genotypes. We identified strong patterns of host association among the 20 *Bb* genotypes found in this study (Figure 1A, Table 1). Four genotypes (Types C, E3, H, K) were significantly more likely to be found in rodent compared to avian hosts, whereas nine genotypes (Types E, F, G, I, M, N, O, T, CJ) followed the opposite taxonomic trend. The remaining seven genotypes (Types A, A3, B, D, J, U, J3) were not associated with either host taxon. We hereafter refer to these genotypes as mouse-adapted, bird-adapted, and generalist, respectively. The signal of host

163 association among the most strongly bird-adapted genotype (type E), which had merely a 0.99%
164 probability of infecting a mouse (OR = 0.01), is stronger than that of the most mouse-adapted
165 genotype (type K), which maintains a 16% probability of infecting a bird (OR = 0.20, when
166 mouse is set as the reference). No *Bb* genotypes were found exclusively in a single host taxon
167 (Table S1). The dominant genotype, accounting for the largest proportion of reads, in an
168 individual was adapted to that host taxon in the majority of observations, occurring in 376 of 553
169 individual mice (68.0%) and 51 of 92 individual birds (55.4%).

170 Individual hosts or ticks are represented by their genotypic community, composed of one
171 or more genotypes and their ranking within the individual, defined by relative sequencing depth.
172 To understand similarity among individual genotypic communities among the populations of
173 mice, birds, and ticks we used non-metric multidimensional scaling (NMDS). NMDS revealed
174 moderate separation between genotypic communities typically found in each population, with
175 birds and mice clustering further away from one another, and ticks forming an intermediate
176 cluster with individual communities that share genotypic community characteristics with both
177 host groups (Figure 1B). Separately, we used a pairwise similarity test of genotype overlap
178 among populations and found that both mammal genotype and bird genotype communities were
179 more similar to nymph genotype communities (0.904 and 0.883, respectively), than they were to
180 each other (0.672). We observed some variation in the frequency of genotype infection across *Bb*
181 infected bird species (Table S3) but GLMs differentiating between commonly infected bird
182 species (Carolina Wren, Common Yellowthroat, and American Robin) indicated that most
183 genotypes exhibited parallel responses compared to white-footed mice (Figure S3). We thus
184 combined bird-derived data in our analyses.
185

186 *Limited evidence of NFDS mechanisms in hosts*

187 If genotypic diversity in *Bb* is maintained by NFDS, we would expect significant
188 variation in the dominant genotype and shifting genotypic frequencies within host populations
189 over time. We first examined genotype frequency and similarity of individual genotype
190 communities within populations across years (Figure 2). For mice, where sample sizes among
191 years were consistently high, the distributions of annual genotype frequencies exhibited low
192 dispersion over time (Figure 2A, Figures S4). Similar patterns were observed among nymphal
193 ticks and bird hosts, though smaller sample sizes and variation among bird species led to more
194 variation in frequency distributions (Figure 2A, Figure S5, S6).

195 We used analysis of similarities (ANOSIM) to test for significant differences in the
196 composition of individual communities within and between years for each taxonomic population,
197 separating analyses across sampling sites for mice and nymphal ticks. We found significant
198 differences in mouse genotype communities across years at only one of the three sampling sites
199 (RH: $p = 0.047$, NR: $p = 0.09$, EI: $p = 0.673$), but not among bird infection communities, which
200 were sampled at multiple sites across the island ($p = 0.85$). Pairwise post-hoc comparisons
201 revealed significant differences in mice at RH only between 2015 and 2016, 2018, and 2019. In
202 contrast, questing nymphs exhibited significant yearly differences in their infection community
203 across all three sites (RH: $p = 0.001$, NR: $p = 0.001$, EI: $p = 0.004$). Graphic examination using
204 NMDS plots revealed strong and consistent overlap for within-year variation in genotype
205 communities across populations of mice, birds, and ticks (Figure 2B).

206

207 *Divergent genotype co-occurrence patterns among hosts*

Competitive or facilitative interactions among genotypes within hosts may influence their probability of co-occurrence within the same individual host community. For each population of mice, birds, and ticks we compared the observed pairwise co-occurrence of genotypes within individual infection communities to the random distributions predicted by the total number of genotype occurrences in that respective population. We found contrasting patterns in the number and directionality of significant genotype pairs across mice and birds (Figure 3). Among mice, we observed 13 significant pairwise correlations, including multiple negative correlations among the four mouse-adapted genotypes (i.e. pairs of genotypes were observed together less frequently than expected). Among birds, we observed a greater total number of significant correlations (n =19) than in mice, all of which were positive (i.e. pairs of genotypes were observed together more frequently than expected), with no consistent relationship among bird-adapted genotypes. Among nymphs, we observed both positive and negative significant correlations among genotypes, with strong negative interactions between those genotypes with the strongest signals of adaptation to birds and mice, suggesting ticks rarely acquire both types of genotypes through a larval bloodmeal (Figure S7).

We examined whether genetic dissimilarity or phylogenetic network distance among genotype pairs predict the observed co-occurrence effect size using linear models. No significant relationship was found between either predictor variable and co-occurrence in mice ($p = 0.097$ and $p = 0.096$) or nymphs ($p = 0.087$ and $p = 0.082$), but a significant negative relationship between both predictors and co-occurrence was found in birds ($p = 0.007$ and $p = 0.008$, Figure S8), indicating that co-occurrence was less frequent when genotype pairs were more genetically divergent.

Individual mouse genotype community dynamics

The genotype community of individual hosts may change over time due to the host's ability to clear infections by specific genotypes and due to sequential introductions through multiple nymphal tick bites. To understand how patterns of genotype host adaptation influence individual-level infection dynamics, we used a multistate Markov model to dissect mice genotype infection in mice sampled multiple times within a single year. We characterized the state of infection based on the identity of the dominant genotype (i.e. greatest sequencing depth) at each sampling time, specifying three possible states: 1) uninfected, 2) infected with a mouse-adapted genotype, 3) infected with a non-mouse-adapted genotype (i.e. generalist or bird-adapted). We found that mice in any initial state are more likely to transition to a mouse-adapted than a non-mouse-adapted genotype infection (Table 2). We also found that the mean persistence (i.e. mean sojourn time) of infections was nearly three times greater in mouse-adapted than that in non-mouse-adapted genotypes in mice (27.6 days vs. 9.5 days; Table S4). Yet, when the state of an uninfected mouse changes, we found it was twice as likely to become infected with a non-mouse-adapted genotype than a mouse-adapted genotype (Table S5). The dominant genotype was significantly more likely to change with increasing time since initial capture (Figure S9) but was not influenced by individual characteristics or tick burden (Table S6). Together, our model suggests that while infections with non-mouse adapted genotypes are common in mice, these genotypes exhibit weak persistence and are often replaced with infections dominated by mouse-adapted genotypes.

Evolutionary patterns among genotypes

253 We examined evolutionary patterns at the *ospC* locus. The Neighbor Net phylogenetic
254 network revealed long branches separating each *ospC* type, with pronounced reticulation at the
255 center of the network (Figure 4A). The single exception was between type C and type J3, which
256 are closely related (97.3% nucleotide similarity). Examination of genotype representative amino
257 acid sequences revealed a K196Q mutation, located at the C-terminus of OspC's 5^{th} α-helix. All
258 mouse-adapted genotypes and all but one of the generalist genotypes exhibit a lysine at this
259 residue, while all but one bird-adapted genotypes exhibit a glutamine (Figure 4B).
260 Using a suite of recombination detection tools, we identified 11 intralocus recombination
261 events across the 21 *ospC* types, each of which was detected by at least three independent
262 methods (Table S7). Analysis of breakpoints suggested two major recombination hotspots
263 around nucleotide positions 300 and 400, with lower confidence hotspots in positions 180 and
264 the 500-550 region (Figure 4C).
265

266 **Discussion**

267 A critical step in evaluating human disease risk from emerging zoonotic pathogens is
268 understanding how interactions with endemic reservoir hosts shape a pathogen's genetic and
269 functional diversity, in particular its host breadth and ability to infect new species (2). Here we
270 demonstrate that *Borrelia burgdorferi* sensu stricto exhibits a spectrum of host-adapted
271 phenotypes associated with the *ospC* alleles. Specifically, we identified several genotypes that
272 show evidence of adaption to white-footed mice (Types C, E3, H, K) and passerine birds (Types
273 E, F, G, I, M, N, O, T, Cj), while a group of genotypes (Types A, A3, B, D, J, U, J3) exhibit no
274 significant association with either host (i.e. are generalists). Importantly, the prevalence of *Bb*
275 genotypes in ticks, including human invasive strains (HIS), closely reflects genotypic frequency

276 in the local host community. These results reveal cryptic intraspecific variation in pathogen host
277 breadth, suggesting that host competence for *Bb*, as well as human disease risk and Lyme
278 Disease severity, is dependent on the relative prevalence of locally circulating genotypes.
279 The evolutionary drivers maintaining *ospC* polymorphisms in nature have been debated
280 in recent decades. Some field-based studies have suggested MNP drives *Bb* diversity through
281 evidence of genotype-host associations defined by genotypes (28, 34) or multi-locus sequence
282 typing (29), though others found no such pattern (35). Lab-based studies have observed variable
283 infection establishment, persistence, and transmission of *Bb* genotypes within and among host
284 species (36–38). Yet, other short-term field studies (i.e. up to 3 years) observe shifting
285 frequencies of *Bb* genotypes in tick vectors consistent with balancing selection mediated by
286 NFDS (20, 39, 40). Pervasive local recombination, revealed by recent *Bb* comparative genomics,
287 appears to drive nucleotide and antigenic diversity at *ospC*, and NFDS is proposed as a
288 parsimonious explanation for the maintenance of this variation, particularly if recombination
289 disrupts host-pathogen allele-level co-adaptation (30, 41). Our results, spanning multiple host
290 taxonomic groups and nymphal ticks over up to 7 years, provide strong evidence supporting
291 MNP in structuring *Bb* genetic diversity. While we observed little evidence of temporal shifts in
292 population-level genotype frequency, as predicted by NFDS, it is possible that antigenic
293 exclusion in individual hosts accounts for minor variation among host-adapted genotypes.
294 While MNP appears to partially maintain *Bb* diversity through genotype-specific host
295 adaptation, variable OspC proteins are also clearly under selection for distinct antigenic epitopes
296 that limit cross-reactivity (42). Since processes governing initial establishment of infection
297 through occur before the production of antibodies, additional selective forces must maintain
298 diverse antigenic epitopes (43). We propose that limited antibody cross-reactivity serves to allow

MNP processes by enabling serial susceptibility of hosts. Because hosts produce antibodies targeting specific genotypic epitopes (44), subsequent infections by different genotypes encounter essentially naïve hosts. Thus, short-lived infections by non-adapted genotypes should not restrict the future success of more host-adapted genotypes, improving the likelihood of transmission despite variable infection success. In the Northeast US, nymphal ticks are abundant for 6-8 weeks, several weeks before the emergence of larvae, allowing hosts to experience multiple *Bb* introductions from independent tick bites. Longer duration phenotypes then have a higher likelihood of infecting the next tick cohort's larvae (45). Indeed, variable tick phenology has been shown to be associated with infection persistence (46). In our study, not only did mouse-adapted genotypes exhibit much longer persistence times in mice sampled multiple times than non-mouse-adapted genotypes, but mouse infection communities preferentially shifted mouse-adapted genotype dominance over time. While temporal shifts in *Bb* infections have previously been observed in mice and humans (47, 48), our results highlight how the combination of limited antibody cross-reactivity and variable host-adapted pathogen genotypes may provide a fitness advantage for *Bb* when hosts experience multiple independent infections.

A second mechanism maintaining genotype diversity is the lack of complete host specialization; even *Bb* genotypes with the strongest evidence for host adaptation were observed in their non-adapted hosts. This pattern may be an intrinsic aspect of the genotype phenotype, may be due to intraspecific variation in host competence, or result from facilitative interactions among genotypes that allow non-adapted genotypes to persist in the presence of host-adapted genotypes (6). Experimental studies of bacterial evolution find that increasing environmental complexity (e.g. host diversity) results in overlapping niches and imperfect specialization (49). A recent model examining the interaction between MNP and NFDS in *Bb* also suggested that more

diverse suites of differentially host-adapted phenotypes can coexist in a population when antigenic variation, a form of environmental heterogeneity, is high (50). These findings are suggestive of a role for host adaptation-related pathogen traits in maintaining pathogen fitness and persistence in infected hosts.

Host community composition drives patterns of Bb diversity in ticks and human disease risk

The frequency of specific *Bb* genotypes in nymphal ticks, which are the main source of human infections, matches the genotype prevalence observed in the local host community (Figure 2A). This trend is particularly important for understanding the local prevalence of human invasive strains (HIS), a subset of *Bb* genotypes including ospC types A, B, I K, and N that produce more severe human disease outcomes (24–26). Human disease risk appears to reflect both the presence of reservoir hosts as well as the host breadth and specificity phenotypes exhibited by local *Bb* genotypes. Further, suggests a more nuanced view of a host's pathogen competence and their role as amplification or dilution hosts, where the infectious phenotype of the host-adapted genotype should be considered in addition to the pathogen's prevalence in the host.

In our study, birds were the main source of ospC types I and N and harbor types A and B in high frequency, although not as high as mice. Though *Bb* infection prevalence varies among bird species (12), previous work estimated that 27% of larval ticks feed on birds at our study location (51). Higher genotype richness and evenness were observed in bird hosts compared to white-footed mice. Further, we found that individual bird infection communities were characterized by greater genotype co-occurrence, indicating facilitation, in contrast with mammal

communities that appear to exhibit only competitive interactions among genotypes. These findings may be partially explained by birds' longer lifespan, different immunity mechanisms, larger home ranges, or other life-history traits resulting in more opportunities for infection (52). Thus, our study emphasizes the role of passerine birds in maintaining *Bb* transmission and increasing human disease risk locally, in addition to their known roles as reservoir hosts and tick dispersers (12, 34, 53, 54).

Signatures of adaptation in ospC

Bb genomes are characterized by the presence of paralogous lipoprotein gene families, whose members provide multiple and redundant roles throughout different stages of infection (55), though o*spC* has no close paralogs (56). The OspC protein has known roles binding to multiple host ligands to promote evasion of host complement, binding to tick salivary proteins to prevent host antibody attack, and promoting host tissue colonization and dissemination (57–59). The *ospC* K196Q mutation we identified may contribute to host adaptation *Bb* phenotypes, as it was observed in 8 out of 9 bird-associated genotypes, but only in 1 out of 7 generalist genotypes, and was absent in mouse-adapted genotypes. This mutation localizes in the N-terminus of OspC's 5th α-helix, a region predicted to be important for genotype-specific antibody cross reactivity (42). Though *ospC* variation clearly impacts antigenic recognition, the elevated role of MNP identified here suggests it may also influence host specific adaptation. If this gene is involved in host specific complement evasion or tissue dissemination in addition to influencing antibody cross-reactivity, it would suggest that competing or complementary evolutionary pressures govern standing variation. Importantly, we posit that genotype identity alone does not dictates host adaptation phenotypes, as other loci in linkage disequilibrium with *ospC* are likely

involved in conferring host-adapted phenotypes via multiple interrelated mechanisms (60). Our results suggest genotype-host associations may be partially due to variation at *ospC* directly impacting fitness across host immune environments and should be a target of further investigation.

Studies of *ospC* variation consistently report <2% and >8% nucleotide variation within and among all oMG genotypes, respectively (28, 39). This bimodal distribution is expected to reflect antigenic interactions, whereby different genotypes display distinct epitopes that limit cross-reactivity of host antibodies (42). Interestingly, the genotype identified here as type C_J does not conform to this widespread trend, exhibiting 2.7% nucleotide difference to type C. Our recombination analysis indicated that genotype type C_J is the product of incorporation of the type J entire 4th α-helix (residues 158-183) into type C. A similar variant (type C_{KR10}) was previously reported circulating in tick populations in both the Northeast and Midwest US (61). In our study, type C_J exhibits decreased association with white-footed mice compared to genotype type C. Together, these patterns suggest that selective pressures otherwise maintaining coinfection with divergent genotypes is relaxed for this genotype, though the mechanism involved and the potential role of differential host adaptation require further study.

Methods

Study site and sample collection

Nymphal ticks, white-footed mouse tissue, and engorged bird-fed tick larvae were collected on Block Island, RI, USA between May and August 2013-2020 and stored for DNA

extraction and subsequent pathogen analysis. Additional details are described in SI Materials and Methods.

DNA extraction, amplification, and sequencing

Genomic DNA was extracted and individual samples were screened using a duplex quantitative PCR for *Borrelia burgdorferi* and *Borrelia miyamotoi* infection. For each *Bb* positive sample, we amplified the entire *ospC* locus using standard PCR with a unique set of barcoded primers. Amplicon products were then sequenced with the Pacific Biosciences Sequel I platform. Additional details describing DNA extraction, amplification, and sequencing are provided in SI Materials and Methods.

Sequence clustering and identification of genotypes

Sequencing reads were demultiplexed and clustered with 32 reference sequences representing known ospC types (Table S11). We subjected any genotypes without known matches to BLASTN against the full Genbank nucleotide database for further identification. Any novel genotype with >8% nucleotide dissimilarity to known genotypes was named according to the sequential list of all known genotypes (i.e. starting with J3), while those with <8% nucleotide dissimilarity were given a subtype designation according to its closest match (i.e. type C_J). A novel genotype most closely matching the *ospC* of *Borrelia kurtenbachii* was named simply "B.kurt." but not given a designated type name. Finally, we used SRST2 (62) to assign reads from each demultiplexed sample. Additional details are provided in SI Materials and Methods.

Statistical analyses

For each sampled host or nymphal tick, we first filtered any genotype represented by less than three individual HiFi reads. We evaluated diversity profiles using Hill numbers and examined correlations between sequencing depth on genotype richness. We built a phylogenetic network to represent the evolutionary relationships among genotypes. We investigated patterns of co-occurrence between all pairwise genotype combinations across individual mice, birds, and ticks, separately using the *cooccur* R package. We also tested for evidence of recombination among genotypes using the RDP5 analysis suite. Additional details for each procedure are provided in SI Materials and Methods.

To test for evidence of host-adapted associations between specific genotypes and mammalian or avian hosts we used binomial (logarithmic) generalized linear models (GLMs). We visualized differences in genotype communities among hosts and ticks using nonmetric multidimensional scaling (NMDS). To understand the extent of temporal variation in genotype frequency, we plotted the distribution of infection frequency of each genotype across years for each sample type. We also assessed changes in community structure among years using Analysis of Similarities and NMDS. To evaluate the dynamics of genotype community transitions and persistence at the individual host scale we examined mice that were sampled multiple times within a single year (n = 383) using a multi-state Markov (MSM) model. Additional details for each procedure are provided in SI Materials and Methods.

Acknowledgements

The authors thank all the Block Island field crews who collected samples between 2013 - 2020 and Ella Steiger for laboratory work. We thank Kim Gaffett and the Nature Conservancy staff on Block Island for their continued support. We also thank Mathilde Cordellier, Karina Garcia, and Nina Skinner for their contributed silhouettes to phylopic.org. This study was supported by the National Institute of General Medical Sciences, National Institutes of Health, Ecology and Evolution of Infectious Disease Program (R01 GM105246; United States), and National Science Foundation (IOS 174995, 1755286, 1755370; United States).

References

1. C. A. Sánchez, J. Venkatachalam-Vaz, J. M. Drake, Spillover of zoonotic pathogens: A review of reviews. *Zoonoses Public Health* **68**, 563–577 (2021).
2. K. J. Olival, *et al.*, Host and viral traits predict zoonotic spillover from mammals. *Nature* **546**, 646–650 (2017).
3. D. J. Becker, G. F. Albery, Expanding host specificity and pathogen sharing beyond viruses. *Mol. Ecol.* **29**, 3170–3172 (2020).
4. M. E. J. Woolhouse, S. Gowtage-Sequeria, Host range and emerging and reemerging pathogens. *Emerg. Infect. Dis.* **11**, 1842–1847 (2005).
5. J. P. Sexton, J. Montiel, J. E. Shay, M. R. Stephens, R. A. Slatyer, Evolution of ecological niche breadth. *Annu. Rev. Ecol. Evol. Syst.* **48**, 183–206 (2017).
6. T. E. Stewart Merrill, P. T. J. Johnson, Towards a mechanistic understanding of competence: A missing link in diversity-disease research. *Parasitology* **147**, 1159–1170 (2020).
7. D. J. Becker, S. N. Seifert, C. J. Carlson, Beyond infection: Integrating competence into reservoir host prediction. *Trends Ecol. Evol.* **35**, 1062–1065 (2020).
8. O. Balmer, M. Tanner, Prevalence and implications of multiple-strain infections. *Lancet Infect. Dis.* **11**, 868–878 (2011).
9. P. W. Hedrick, Genetic polymorphism in heterogeneous environments: The age of genomics. *Annu. Rev. Ecol. Evol. Syst.* **37**, 67–93 (2006).
10. S. Gandon, A. Buckling, E. Decaestecker, T. Day, Host-parasite coevolution and patterns of adaptation across time and space. *J. Evol. Biol.* **21**, 1861–1866 (2008).
11. B. Koskella, M. Vos, Adaptation in natural microbial populations. *Annu. Rev. Ecol. Evol. Syst.* **46**, 503–522 (2015).
12. D. J. Becker, B. A. Han, The macroecology and evolution of avian competence for *Borrelia burgdorferi*. *Glob. Ecol. Biogeogr.* **30**, 710–724 (2021).
13. R. Woods, A. Reiss, K. Cox-Witton, T. Grillo, A. Peters, The importance of wildlife disease monitoring as part of global surveillance for zoonotic diseases: The role of Australia. *Trop. Med. Infect. Dis.* **4** (2019).
14. A. Swei, L. I. Couper, L. L. Coffey, D. Kapan, S. Bennett, Patterns, drivers, and challenges of vector-borne disease emergence. *Vector-Borne Zoonotic Dis.* **20**, 159–170 (2020).
15. K. J. Kugeler, A. M. Schwartz, M. J. Delorey, P. S. Mead, A. F. Hinckley, Estimating the frequency of Lyme disease diagnoses, United States, 2010-2018. *Emerg. Infect. Dis.* **27**, 616–619 (2021).
16. Centers for Disease Control and Prevention, Lyme Disease (2022).
17. J. D. Radolf, K. Strle, J. E. Lemieux, F. Strle, Lyme disease in humans. *Curr. Issues Mol. Biol.* **42**, 333–384 (2020).
18. A. M. Kilpatrick, *et al.*, Lyme disease ecology in a changing world: consensus, uncertainty and critical gaps for improving control. *Proc. R. Soc. B* **372** (2017).
19. J. I. Tsao, S. A. Hamer, S. Han, J. L. Sidge, G. J. Hickling, The contribution of wildlife hosts to the rise of ticks and tick-borne diseases in North America. *J. Med. Entomol.* **58**, 1565–1587 (2021).
20. W. G. Qiu, *et al.*, A population genetic study of *Borrelia burgdorferi* sensu stricto from eastern Long Island, New York, suggested frequency-dependent selection, gene flow and

host adaptation. *Hereditas* **127**, 203–216 (1997).
21. D. Brisson, D. Drecktrah, C. H. Eggers, D. S. Samuels, Genetics of *Borrelia burgdorferi*. *Annu. Rev. Genet.* **46**, 515–536 (2012).
22. A. G. Barbour, B. Travinsky, Evolution and distribution of the *ospC* Gene, a transferable serotype determinant of *Borrelia burgdorferi*. *MBio* **1**, 1–11 (2010).
23. K. Tilly, *et al.*, *Borrelia burgdorferi* OspC protein required exclusively in a crucial early stage of mammalian infection. *Infect. Immun.* **74**, 3554–3564 (2006).
24. G. Seinost, *et al.*, Infection with multiple strains of *Borrelia burgdorferi* Sensu Stricto in patients with Lyme disease. *Arch. Dermatol.* **135**, 1329–1333 (1999).
25. G. P. Wormser, *et al.*, *Borrelia burgdorferi* genotype predicts the capacity for hematogenous dissemination during early Lyme disease . *J. Infect. Dis.* **198**, 1358–1364 (2008).
26. D. E. Dykhuizen, *et al.*, The propensity of different *Borrelia burgdorferi* sensu stricto genotypes to cause disseminated infections in humans. *Am. J. Trop. Med. Hyg.* **78**, 806–810 (2008).
27. K. R. O'Keeffe, Z. J. Oppler, D. Brisson, Evolutionary ecology of Lyme Borrelia. *Infect. Genet. Evol.* **85**, 104570 (2020).
28. D. Brisson, D. E. Dykhuizen, ospC diversity in *Borrelia burgdorferi*: Different hosts are different niches. *Genetics* **168**, 713–722 (2004).
29. S. Mechai, *et al.*, Evidence for host-genotype associations of *Borrelia burgdorferi* sensu stricto. *PLoS One* **11**, 1–25 (2016).
30. J. Haven, *et al.*, Pervasive recombination and sympatric genome diversification driven by frequency-dependent selection in *Borrelia burgdorferi*, the lyme disease bacterium. *Genetics* **189**, 951–966 (2011).
31. K. I. Swanson, D. E. Norris, Presence of multiple variants of *Borrelia burgdorferi* in the natural reservoir *Peromyscus leucopus* throughout a transmission season. *Vector-Borne Zoonotic Dis.* **8**, 397–406 (2008).
32. A. Alberdi, M. T. P. Gilbert, A guide to the application of Hill numbers to DNA-based diversity analyses. *Mol. Ecol. Resour.* **19**, 804–817 (2019).
33. G. Margos, *et al.*, Multilocus sequence analysis of *Borrelia bissettii* strains from North America reveals a new Borrelia species, *Borrelia kurtenbachii*. *Ticks Tick. Borne. Dis.* **1**, 151–158 (2010).
34. H. B. Vuong, *et al.*, Occurrence and transmission efficiencies of *Borrelia burgdorferi* ospC types in avian and mammalian wildlife. *Infect. Genet. Evol.* **27**, 594–600 (2014).
35. K. Hanincová, K. Kurtenbach, M. Diuk-Wasser, B. Brei, D. Fish, Epidemic spread of Lyme borreliosis, northeastern United States. *Emerg. Infect. Dis.* **12**, 604–611 (2006).
36. M. Derdáková, *et al.*, Interaction and transmission of two *Borrelia burgdorferi* sensu stricto strains in a tick-rodent maintenance system. *Appl. Environ. Microbiol.* **70**, 6783–6788 (2004).
37. K. Hanincová, *et al.*, Fitness variation of *Borrelia burgdorferi* sensu stricto strains in Mice. *Appl. Environ. Microbiol.* **74**, 153–157 (2008).
38. Y. Lin, *et al.*, Cellular and immunological mechanisms influence host-adapted phenotypes in a vector-borne microparasite. *Proc. R. Soc. B* (2022).
39. I. N. Wang, *et al.*, Genetic diversity of ospC in a local population of *Borrelia burgdorferi* sensu stricto. *Genetics* **151**, 15–30 (1999).
40. S. L. States, *et al.*, Lyme disease risk not amplified in a species-poor vertebrate

community: similar *Borrelia burgdorferi* tick infection prevalence and OspC genotype frequencies. *Infect. Genet. Evol.* **27**, 266–575 (2014).
41. I. Schwartz, G. Margos, S. R. Casjens, W. G. Qiu, C. H. Eggers, Multipartite genome of lyme disease borrelia: Structure, variation and prophages. *Curr. Issues Mol. Biol.* **42**, 409–454 (2021).
42. E. Baum, A. Z. Randall, M. Zeller, A. G. Barbour, Inferring epitopes of a polymorphic antigen amidst broadly cross-reactive antibodies using protein microarrays: A study of OspC proteins of *Borrelia burgdorferi*. *PLoS One* **8** (2013).
43. K. Kurtenbach, *et al.*, Fundamental processes in the evolutionary ecology of Lyme borreliosis. *Nat. Rev. Microbiol.* **4**, 660–669 (2006).
44. E. Baum, F. Hue, A. G. Barbour, Experimental infections of the reservoir species *Peromyscus leucopus* with diverse strains of *Borrelia burgdorferi*, a Lyme disease agent. *MBio* **3**, 1–11 (2012).
45. J. Haven, K. Magori, A. W. Park, Ecological and inhost factors promoting distinct parasite life-history strategies in Lyme borreliosis. *Epidemics* **4**, 152–157 (2012).
46. A. G. Gatewood, *et al.*, Climate and tick seasonality are predictors of *Borrelia burgdorferi* genotype distribution. *Appl. Environ. Microbiol.* **75**, 2476–2483 (2009).
47. J. M. Anderson, D. E. Norris, Genetic diversity of *Borrelia burgdorferi* sensu stricto in *Peromyscus leucopus*, the primary reservoir of lyme disease in a region of endemicity in Southern Maryland. *Appl. Environ. Microbiol.* **72**, 5331–5341 (2006).
48. C. E. Khatchikian, *et al.*, Evidence for strain-specific immunity in patients treated for early Lyme disease. *Infect. Immun.* **82**, 1408–1413 (2014).
49. R. D. H. Barrett, R. C. MacLean, G. Bell, Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. *Am. Nat.* **166**, 470–480 (2005).
50. B. Adams, K. S. Walter, M. A. Diuk-Wasser, Host specialisation, immune cross-reaction and the composition of communities of co-circulating *Borrelia* strains. *Bull. Math. Biol.* **83**, 1–24 (2021).
51. C. I. Huang, *et al.*, High burdens of *Ixodes scapularis* larval ticks on white-tailed deer may limit Lyme disease risk in a low biodiversity setting. *Ticks Tick. Borne. Dis.* **10**, 258–268 (2019).
52. A. Rataud, P. Y. Henry, S. Moutailler, M. Marsot, Research effort on birds' reservoir host potential for Lyme borreliosis: A systematic review and perspectives. *Transbound. Emerg. Dis.*, 1–11 (2021).
53. H. S. Ginsberg, *et al.*, Reservoir competence of native North American birds for the Lyme Disease spirochete, *Borrelia burgdorferi*. *J. Med. Entomol.* **42**, 445–449 (2006).
54. R. J. Brinkerhoff, C. M. Folsom-O'Keefe, K. Tsao, M. A. Diuk-Wasser, Do birds affect Lyme disease risk? Range expansion of the vector-borne pathogen *Borrelia burgdorferi*. *Front. Ecol. Environ.* **9**, 103–110 (2011).
55. J. A. Caine, J. Coburn, Multifunctional and redundant roles of *Borrelia burgdorferi* outer surface proteins in tissue adhesion, colonization, and complement evasion. *Front. Immunol.* **7**, 1–11 (2016).
56. E. F. Mongodin, *et al.*, Inter- and intra-specific pan-genomes of *Borrelia burgdorferi* sensu lato: genome stability and adaptive radiation. *BMC Genomics* **14** (2013).
57. V. Lagal, D. Portnoï, G. Faure, D. Postic, G. Baranton, *Borrelia burgdorferi* sensu stricto invasiveness is correlated with OspC-plasminogen affinity. *Microbes Infect.* **8**, 645–652 (2006).

58. J. A. Caine, *et al.*, *Borrelia burgdorferi* outer surface protein C (OspC) binds complement component C4b and confers bloodstream survival. *Cell. Microbiol.* **19**, 1–14 (2017).
59. Y. P. Lin, *et al.*, Strain-specific joint invasion and colonization by Lyme disease spirochetes is promoted by outer surface protein C. *PLoS Pathog.* **16**, 1–29 (2020).
60. D. E. Dykhuizen, A. M. Dean, Evolution of specialists in an experimental microcosm. *Genetics* **167**, 2015–2026 (2004).
61. D. Brisson, M. F. Vandermause, J. K. Meece, K. D. Reed, D. E. Dykhuizen, Evolution of northeastern and midwestern *Borrelia burgdorferi*, United States. *Emerg. Infect. Dis.* **16**, 911–917 (2010).
62. M. Inouye, *et al.*, SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med.* **6**, 1–16 (2014).