

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

Matthew A. Combs<sup>1,8#</sup>, Danielle M. Tufts<sup>1,2</sup>, Ben Adams<sup>3</sup>, Yi-Pin Lin<sup>4,5</sup>, Sergios-Orestis Kolokotronis<sup>6,7,8\*</sup> Maria A. Diuk-Wasser<sup>1\*</sup>

<sup>1</sup> Department of Ecology, Evolution and Environmental Biology, Columbia University, New York, NY 10027

<sup>2</sup> Infectious Diseases and Microbiology Department, University of Pittsburgh, Pittsburgh, PA 15261

<sup>3</sup> Department of Mathematical Sciences, University of Bath, Bath, United Kingdom, BA27AY

<sup>4</sup> Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, NY 12201

<sup>5</sup> Department of Biomedical Sciences, State University of New York at Albany, Albany, NY 12203

<sup>6</sup> Department of Epidemiology and Biostatistics, School of Public Health, SUNY Downstate Health Sciences University, Brooklyn, NY 11203

<sup>7</sup> Division of Infectious Diseases, Department of Medicine, College of Medicine, SUNY Downstate Health Sciences University, Brooklyn, NY 11203

<sup>8</sup> Institute for Genomics in Health, SUNY Downstate Health Sciences University, Brooklyn, NY 11203

# Equal contribution

\* Equal role as senior author ([sok@downstate.edu](mailto:sok@downstate.edu), [mad2256@columbia.edu](mailto:mad2256@columbia.edu))

**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 (Sánchez *et al.* 2021). Recent meta-analyses and reviews have repeatedly identified host breadth, or the capacity to infect phylogenetically diverse species, as a key trait influencing spillover (Woolhouse & Gowtage-Sequeria 2005; Olival *et al.* 2017; Becker & Albery 2020). 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 (Elena & Lenski 2003; Sexton *et al.* 2017).

Wider pathogen host breadth is intrinsically linked to increased competence of potential hosts, defined as the capacity to acquire, maintain, and transmit infection (Stewart Merrill & Johnson 2020). Though host competence is increasingly used to model human disease risk, unexplained sources of variation can limit its utility (Downs *et al.* 2019; Becker *et al.* 2020; Stewart Merrill & Johnson 2020). In particular, intraspecific genetic variation in multi-strain pathogens may influence host-pathogen interactions, preventing accurate characterization of pathogen host breadth and host competence (Balmer & Tanner 2011; Kilpatrick *et al.* 2017).

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 (Elena & Lenski 2003; Hedrick 2006; Weedall & Conway 2010). 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 (Grenfell *et al.* 2004; Hedrick 2006; Gandon *et al.* 2008; Thrall *et al.* 2012; Koskella & Vos 2015). 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 (AVMA 2008; Wells & Clark 2019; Woods *et al.* 2019; Becker & Han 2021).

Vector-borne pathogens are increasingly responsible for emerging infectious diseases (Jones *et al.* 2008; Swei *et al.* 2020) 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 (Kugeler *et al.* 2021). LD is the most common vector-borne disease in the US and continues to increase in geographic range and case numbers (Radolf *et al.* 2020; CDC 2022), particularly in the Northeast and Midwest US, where *Bb* circulates primarily in small mammal and bird species via the generalist tick vector *Ixodes scapularis* (Halsey *et al.* 2018). 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 (Kilpatrick *et al.* 2017; Schwartz *et al.* 2021; Tsao *et al.* 2021).

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 (Qiu *et al.* 1997; Wang *et al.* 1999; Barbour & Travinsky 2010; Brisson *et al.* 2012). The *OspC* protein is required for host infection (Tilly *et al.* 2006) 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 (Lagal *et al.* 2006; Dykhuizen *et al.* 2008; Strle *et al.* 2011). The eco-evolutionary drivers maintaining *ospC* variation and implications for human disease risk remain a focus of active debate (O’Keeffe *et al.* 2020). 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 (Brisson & Dykhuizen 2004; Mechai *et al.* 2016). 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 (Swanson & Norris 2008; Haven *et al.* 2011).

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 ( $\alpha=1.77$ ) than in birds ( $\alpha=2.59$ ) and intermediate in nymphal ticks ( $\alpha=2.30$ ). 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 (Alberdi & Gilbert 2019) 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 (Margos *et al.* 2010). *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 association among the most strongly bird-adapted genotype (type E), which had merely a 0.99% probability of infecting a mouse (OR = 0.01), is stronger than that of the most mouse-adapted genotype (type K), which maintains a 16% probability of infecting a bird (OR = 0.20, when mouse is set as the reference). No *Bb* genotypes were found exclusively in a single host taxon (Table S1). The dominant genotype, accounting for the largest proportion of reads, in an individual was adapted to that host taxon in the majority of observations, occurring in 376 of 553 individual mice (68.0%) and 51 of 92 individual birds (55.4%).

Individual hosts or ticks are represented by their genotypic community, composed of one or more genotypes and their ranking within the individual, defined by relative sequencing depth. To understand similarity among individual genotypic communities among the populations of mice, birds, and ticks we used non-metric multidimensional scaling (NMDS). NMDS revealed moderate separation between genotypic communities typically found in each population, with birds and mice clustering further away from one another, and ticks forming an intermediate cluster with individual communities that share genotypic community characteristics with both host groups (Figure 1B). Separately, we used a pairwise similarity test of genotype overlap



among populations and found that both mammal genotype and bird genotype communities were more similar to nymph genotype communities (0.904 and 0.883, respectively), than they were to each other (0.672). We observed some variation in the frequency of genotype infection across *Bb* infected bird species (Table S3) but GLMs differentiating between commonly infected bird species (Carolina Wren, Common Yellowthroat, and American Robin) indicated that most genotypes exhibited parallel responses compared to white-footed mice (Figure S3). We thus combined bird-derived data in our analyses.

### Limited evidence of NFDS mechanisms in hosts

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

We used analysis of similarities (ANOSIM) to test for significant differences in the composition of individual communities within and between years for each taxonomic population, separating analyses across sampling sites for mice and nymphal ticks. We found significant differences in mouse genotype communities across years at only one of the three sampling sites (RH:  $p = 0.047$ , NR:  $p = 0.09$ , EI:  $p = 0.673$ ), but not among bird infection communities, which were sampled at multiple sites across the island ( $p = 0.85$ ). Pairwise post-hoc comparisons

revealed significant differences in mice at RH only between 2015 and 2016, 2018, and 2019. In contrast, questing nymphs exhibited significant yearly differences in their infection community across all three sites (RH:  $p = 0.001$ , NR:  $p = 0.001$ , EI:  $p = 0.004$ ). Graphic examination using NMDS plots revealed strong and consistent overlap for within-year variation in genotype communities across populations of mice, birds, and ticks (Figure 2B).

### *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

We examined evolutionary patterns at the *ospC* locus. The Neighbor Net phylogenetic network revealed long branches separating each *ospC* type, with pronounced reticulation at the center of the network (Figure 4A). The single exception was between type C and type J3, which are closely related (97.3% nucleotide similarity). Examination of genotype representative amino acid sequences revealed a K196Q mutation, located at the C-terminus of OspC's 5<sup>th</sup>  $\alpha$ -helix. All mouse-adapted genotypes and all but one of the generalist genotypes exhibit a lysine at this residue, while all but one bird-adapted genotypes exhibit a glutamine (Figure 4B).

Using a suite of recombination detection tools, we identified 11 intralocus recombination events across the 21 *ospC* types, each of which was detected by at least three independent methods (Table S7). Analysis of breakpoints suggested two major recombination hotspots around nucleotide positions 300 and 400, with lower confidence hotspots in positions 180 and the 500-550 region (Figure 4C).

## **Discussion**

A critical step in evaluating human disease risk from emerging zoonotic pathogens is understanding how interactions with endemic reservoir hosts shape a pathogen's genetic and

functional diversity, in particular its host breadth and ability to infect new species (Karesh *et al.* 2012; Olival *et al.* 2017). Here we demonstrate that *Borrelia burgdorferi* sensu stricto exhibits a spectrum of host-adapted phenotypes associated with the *ospC* alleles. Specifically, we identified several genotypes that show evidence of adaption to white-footed mice (Types C, E3, H, K) and passerine birds (Types E, F, G, I, M, N, O, T, Cj), while a group of genotypes (Types A, A3, B, D, J, U, J3) exhibit no significant association with either host (i.e. are generalists). Importantly, the prevalence of *Bb* genotypes in ticks, including human invasive strains (HIS), closely reflects genotypic frequency in the local host community. These results reveal cryptic intraspecific variation in pathogen host breadth, suggesting that host competence for *Bb*, as well as human disease risk and Lyme Disease severity, is dependent on the relative prevalence of locally circulating genotypes.

The evolutionary drivers maintaining *ospC* polymorphisms in nature have been debated in recent decades. Some field-based studies have suggested MNP drives *Bb* diversity through evidence of genotype-host associations defined by genotypes (Brisson & Dykhuizen 2004; Vuong *et al.* 2014) or multi-locus sequence typing (Mechai *et al.* 2016), though others found no such pattern (Hanincová *et al.* 2006). Lab-based studies have observed variable infection establishment, persistence, and transmission of *Bb* genotypes within and among host species (Wang *et al.* 2001, 2002; Derdákóvá *et al.* 2004; Hanincová *et al.* 2008; Lin *et al.* 2022). Yet, other short-term field studies (i.e. up to 3 years) observe shifting frequencies of *Bb* genotypes in tick vectors consistent with balancing selection mediated by NFDS (Qiu *et al.* 1997; Wang *et al.* 1999; States *et al.* 2014). Pervasive local recombination, revealed by recent *Bb* comparative genomics, appears to drive nucleotide and antigenic diversity at *ospC*, and NFDS is proposed as a parsimonious explanation for the maintenance of this variation, particularly if recombination

disrupts host-pathogen allele-level co-adaptation (Haven *et al.* 2011; Qiu & Martin 2014; Schwartz *et al.* 2021). Our results, spanning multiple host taxonomic groups and nymphal ticks over up to 7 years, provide strong evidence supporting MNP in structuring *Bb* genetic diversity. While we observed little evidence of temporal shifts in population-level genotype frequency, as predicted by NFDS, it is possible that antigenic exclusion in individual hosts accounts for minor variation among host-adapted genotypes.

While MNP appears to partially maintain *Bb* diversity through genotype-specific host adaptation, variable OspC proteins are also clearly under selection for distinct antigenic epitopes that limit cross-reactivity (Baum *et al.* 2013). Since processes governing initial establishment of infection through occur before the production of antibodies, additional selective forces must maintain diverse antigenic epitopes (Kurtenbach *et al.* 2006). 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 (Probert *et al.* 1997; Baum *et al.* 2012), 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 (Ogden *et al.* 2007; Haven *et al.* 2012). Indeed, variable tick phenology has been shown to be associated with infection persistence (Gatewood *et al.* 2009). 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 (Anderson & Norris 2006; Nadelman *et al.* 2012; Khatchikian *et al.* 2014), 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 (Andersson *et al.* 2013; Stewart Merrill & Johnson 2020). Experimental studies of bacterial evolution find that increasing environmental complexity (e.g. host diversity) results in overlapping niches and imperfect specialization (Barrett *et al.* 2005). 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 (Adams *et al.* 2021). These findings are suggestive of a role for host adaptation-related pathogen traits in maintaining pathogen fitness and persistence in infected hosts. The high genotype diversity observed in *Bb* appears to fit a ‘mass effect’ metacommunity model (Leibold *et al.* 2004; Alexander *et al.* 2012) where host-adapted genotypes are common in their respective hosts and strong propagule pressure (from infectious ticks) results in low frequency of non-adapted genotypes co-occurring with adapted ones.

# 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 (Seinost *et al.* 1999; Dykhuizen *et al.* 2008; Wormser *et al.* 2008). 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 (Ginsberg *et al.* 2006; Becker & Han 2021), previous work estimated that 27% of larval ticks feed on birds at our study location (Huang *et al.* 2019). 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 (Rataud *et al.* 2021). 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 (Ginsberg *et al.* 2006; Brinkerhoff *et al.* 2010, 2011; Vuong *et al.* 2014; Loss *et al.* 2016; Becker & Han 2021).

### 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 (Wywiał *et al.* 2009; Caine & Coburn 2016), though *ospC* has no close paralogs (Mongodin *et al.* 2013). 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 (Ramamoorthi *et al.* 2005; Lagal *et al.* 2006; Bonnot *et al.* 2010; Caine *et al.* 2017; Lin *et al.* 2020). 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 5<sup>th</sup>  $\alpha$ -helix, a region predicted to be important for genotype-specific antibody cross reactivity (Baum *et al.* 2013). 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 (Dykhuizen & Dean 2004). 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 (Wang *et al.* 1999; Lin *et al.* 2002; Brisson & Dykhuizen 2004). This bimodal distribution is expected to reflect antigenic interactions, whereby different genotypes display distinct epitopes that limit cross-reactivity of host antibodies (Probert *et al.* 1997; Earnhart *et al.* 2005). Interestingly, the genotype identified here as type C<sub>J</sub> does not conform to this widespread trend, exhibiting 2.7% nucleotide difference to type C. Our recombination analysis indicated that genotype type C<sub>J</sub> is the product of incorporation of the type J entire 4<sup>th</sup>  $\alpha$ -helix (residues 158-183) into type C. A similar variant (type C<sub>KR10</sub>) was previously reported circulating in tick populations in both the Northeast and Midwest US (Brisson *et al.* 2010). In our study, type C<sub>J</sub> 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 were collected on Block Island, RI, USA, between May and August, 2014 - 2020. Small mammals, mainly white-footed mice, were trapped in the same months from 2014 - 2019, and birds were sampled in the same months from 2013 - 2019. Grids were established for tick and small mammal sampling at three sites across the island at the Block Island National Wildlife Refuge (NR), a private property on the eastern part of the island (EI), and at Rodman's

Hollow (RH). Passerine birds were mist netted at two sites at the Ocean View Pavilion (OVP) and Bayrose cabin (BR) under the supervision of Kim Gaffett (U.S. Department of Interior Banding Permit #09636). Nymphal ticks were collected from each grid by a standard 1m<sup>2</sup> drag cloth method (Falco & Fish 1992) stopping every 10m to remove attached ticks which were immediately stored in 70% ethanol until DNA extraction. Small mammals were trapped biweekly at each site for seven sessions consisting of three consecutive trap nights each, using Sherman live traps in accordance with approved Columbia University IACUC protocol (AC-AAAS6470) and scientific collections permits from Rhode Island Department of Environmental Management and town of New Shoreham, RI. Block Island supports a low diversity small mammal community (Comings 2006) dominated by *Peromyscus leucopus* (white-footed mouse), with <1% other hosts (Huang *et al.* 2019). Upon capture morphological characteristics, such as body, tail, and foot measurements, weight, sex, age, and reproduction status were obtained. Each animal was carefully examined for ticks and other ectoparasites and marked with a uniquely numbered ear tag. Ear biopsy tissue was collected and stored in 100% ethanol until DNA extraction and pathogen analysis. Engorged *I. scapularis* larvae were removed from all bird species and stored in 70% ethanol until DNA extraction and subsequent pathogen analysis.

#### DNA extraction, amplification, and sequencing

Genomic DNA was extracted using the DNeasy Blood & Tissue kit optimized for QIAcube HT automation (Qiagen). Questing nymphal and engorged larval ticks from birds were dried, frozen with liquid nitrogen, and crushed with sterile pestles before extraction. Mouse ear tissue was processed using the manufacturer's optimized protocol. Individual samples were then screened using a duplex quantitative PCR for *Borrelia burgdorferi* and *Borrelia miyamotoi* in

duplicate on an ABI 7500 Fast Real-Time PCR System using TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific).

For *Bb* positive samples, we used a standard PCR to amplify a region centered around the *ospC* locus. Amplification was first attempted for a 1500bp region with the following primers GGGATCCAAAATCTAATACAA (forward) and CCCTTAACATACAATATCTCTTC (reverse). If samples failed to amplify due to DNA degradation, we attempted amplification of a smaller 750bp region that contained the entire *ospC* locus using the following primers GAGGCACAAATTAATGAAAAGAA (forward) and GACTTTATTTTCCAGTTACTTTTT (reverse). Both primer sets were designed to target conserved sequences identified across 18 *Bb* genotypes downloaded from GenBank (accession codes provided in Table S8). The PCR protocol included 30 s of denaturation at 98°C followed by 35 cycles of denaturation at 98°C for 10 s, annealing at 60°C for 10 s, and extension at 72°C for 1 min for the longer fragment or 30 s for the shorter fragment, followed by a final extension for 5 min at 72°C.

Each sample was amplified using a unique set of barcoded forward and reverse primers to enable pooling and downstream sample demultiplexing (Tables S9, S10). Amplification success was assessed by running 2 µl PCR product aliquots on a 1.5% agarose gel. Successful PCR products were cleaned with the ProNex Size-Selective Purification System (Promega) then quantified with a Qubit 3.0 Fluorometer using the dsDNA High Sensitivity Assay kit (Thermo Fisher Scientific) or Fragment Analyzer using the HS NGS Fragment kit (Agilent). Products were pooled at equimolar concentrations and libraries were constructed for sequencing using the single-molecule real-time (SMRTbell) sequencing Express Template Prep Kit 2.0 with an

extended DNA damage repair step, then sequenced with the Pacific Biosciences Sequel I platform at the Icahn School of Medicine at Mt. Sinai Genomics Core Facility (New York, NY).

# Sequence clustering and identification of genotypes

High fidelity (HiFi) reads were generated from subreads of each zero-mode waveguide (ZMW) using the circular consensus sequencing (*ccs*) tool (<https://github.com/PacificBiosciences/ccs>) and demultiplexed using *Lima* v.2.2.0 (<https://github.com/PacificBiosciences/barcoding>) with the following parameters: `--ccs --dump-removed --dump-clips --guess 80 --min-score 80 --split --output-handles 600`. To identify the distribution of *Bb* genotypes among samples we first created a FASTA containing all demultiplexed reads and used *PROKKA* gene annotation software (Seemann 2014) to isolate full-length gene sequences. We used *vsearch* v2.18.0 toolkit (Rognes *et al.* 2016), a free and open-source version of the popular *usearch* OTU clustering software, to cluster all sequenced *ospC* genes and 32 reference sequences with known *ospC* types (Table S11), using `--id 0.98`. Of 21 read clusters with a minimum of 200 reads (all other clusters contained <5 reads), 18 clustered with known *ospC* types. We subjected the centroid sequences representing 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<sub>J</sub>). 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 (Inouye *et al.* 2014) to assign reads from each demultiplexed sample represented by a minimum

of 20 HiFi reads to one of the 21 identified *ospC* types using the following parameters: --  
min\_depth 1 --prob\_err 0.005 --max\_divergence 2.5.

### Statistical analyses

Statistical analyses and data filtering were conducted in R v4.1.1 unless otherwise noted (R Core Team 2021). For each sampled host or nymphal tick, we first filtered any genotype represented by less than three individual HiFi reads. We then rank transformed genotype abundances for each sample, as PCR-based amplification preferentially amplifies the most common substrate, skewing relative differences between amplified products. We evaluated diversity profiles across mammal, bird, and nymph genotype communities using Hill numbers and further characterized the extent of genotype overlap between genotype communities with the Sorensen index, both implemented with the SpadeR R package (Chao *et al.* 2016). We also examined correlations between sequencing depth on genotype richness across samples using linear models.

Because the *ospC* locus is heavily influenced by recombination, individual phylogenetic trees do not accurately represent the evolutionary relationships among genotypes. Thus, we built a phylogenetic network using Neighbor-Net (Bryant & Moulton 2004) after first aligning representative full length *ospC* sequences from each genotype (i.e. centroids of USEARCH clusters) as translated amino acids with *MAFFT* (Katoh & Standley 2013).

We investigated patterns of co-occurrence between all pairwise genotype combinations across individual mice, birds, and ticks, separately using the *cooccur* R package, which evaluates the predicted and observed probability of genotype co-occurrence within individuals given their abundances in the dataset (Griffith *et al.* 2016). To better interpret patterns of co-occurrence

among genotypes we built linear models in which either nucleotide dissimilarity, as calculated by Clustal-Omega (Sievers & Higgins 2014), or phylogenetic network distance, was used to predict the effect size for each pair of genotypes.

We tested for evidence of recombination among genotypes using the RDP5 analysis suite across 30bp sliding windows using the following methods: RDP, GENECONV, BOOTSCAN, MAXCHI, CHIMERA, SISCAN. Putative breakpoints were accepted if detected across three or more methods independently (Martin *et al.* 2021).

To test for evidence of host-adapted associations between specific genotypes and mammalian or avian hosts we used binomial (logarithmic) generalized linear models (GLMs), specifying infection with an individual genotype as the response and sample type (i.e. mammal, bird, or nymph) as the predictor, setting “bird” as the reference to highlight differences between birds and mice. To account for detection bias introduced at low sequencing depth we included an offset term for samples with <100 reads (otherwise set to 1; Figure S1). A genotype was considered host-adapted if it displayed a significant association ( $\alpha = 0.05$ ) with mammalian or avian hosts. We visualized differences in genotype communities among hosts and ticks using nonmetric multidimensional scaling (NMDS) with the *metaMDS* function from the *vegan* R package (Oksanen *et al.* 2020). To examine differences among bird species, we conducted binomial GLMs as described above, but restricted bird samples to species with sample size greater than 10 (Carolina wren:  $n = 28$ , Common yellowthroat:  $n = 20$ , American robin:  $n = 16$ , all others  $\leq 8$ ) and included species as the predictor variable, setting “mouse” as the reference.

To test for evidence of temporal variation in genotype frequency, as expected under NFDS, we first calculated the frequency of each genotype within each sample type for each year, then used boxplots to visualize the distributions. We also plotted changes in frequency over time

for each genotype at each site. To assess changes in community structure we used Analysis of Similarities (ANOSIM) tests from the *vegan* R package (Oksanen *et al.* 2020). For each sample type, communities across each year were compared using Bray-Curtis dissimilarity matrices and run with 999 permutations. If this global test was significant, we tested pairwise yearly combinations for post-hoc identification of significant pairs. We visualized differences in genotype communities by year using NMDS as described above.

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, implemented with the *msm* R package (Jackson 2011). We established three infection states reflecting whether the mouse exhibited no *Bb* infection, *Bb* infection dominated (i.e. read depth rank = 1) by a mouse-associated genotype (type C, E3, H, K), or *Bb* infection dominated by a non-mouse-associated genotype. We estimated the transition rates between state pairs by maximum likelihood using ‘nlm’ optimization. We then extracted transition probabilities over a 4-week period, the mean sojourn time (i.e. persistence) for each state, and the probability of the next state given each starting state. To understand what factors influence genotype community changes we also used a binomial generalized linear mixed model (GLMM) to model the probability of a change in identity for a mouse’s dominant genotype as a function of time (days since first sampling), genotype richness, sex, maturity, and tick burden, using the site and year of capture as random variables.



## 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

- Adams B, Walter KS, Diuk-Wasser MA (2021) Host Specialisation, Immune Cross-Reaction and the Composition of Communities of Co-circulating *Borrelia* Strains. *Bulletin of Mathematical Biology*, **83**, 1–24.
- Alberdi A, Gilbert MTP (2019) A guide to the application of Hill numbers to DNA-based diversity analyses. *Molecular Ecology Resources*, **19**, 804–817.
- Alexander HM, Foster BL, Ballantyne F *et al.* (2012) Metapopulations and metacommunities: Combining spatial and temporal perspectives in plant ecology. *Journal of Ecology*, **100**, 88–103.
- Anderson JM, Norris DE (2006) Genetic diversity of *Borrelia burgdorferi* sensu stricto in *Peromyscus leucopus*, the primary reservoir of lyme disease in a region of endemicity in Southern Maryland. *Applied and Environmental Microbiology*, **72**, 5331–5341.
- Andersson M, Scherman K, Råberg L (2013) Multiple-Strain Infections of *Borrelia afzelii*: A Role for Within-Host Interactions in the Maintenance of Antigenic Diversity? *The American Naturalist*, **181**, 545–554.
- Association AVM (2008) *One health: A new professional imperative*.
- Balmer O, Tanner M (2011) Prevalence and implications of multiple-strain infections. *The Lancet Infectious Diseases*, **11**, 868–878.
- Barbour AG, Travinsky B (2010) Evolution and Distribution of the ospC Gene, a Transferable Serotype Determinant of *Borrelia burgdorferi*. *mBio*, **1**, 1–11.
- Barrett RDH, MacLean RC, Bell G (2005) Experimental evolution of *Pseudomonas fluorescens* in simple and complex environments. *American Naturalist*, **166**, 470–480.
- Baum E, Hue F, Barbour AG (2012) Experimental infections of the reservoir species *peromyscus leucopus* with diverse strains of *borrelia burgdorferi*, a Lyme disease agent. *mBio*, **3**, 1–11.
- Baum E, Randall AZ, Zeller M, Barbour AG (2013) 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**.
- Becker DJ, Albery GF (2020) Expanding host specificity and pathogen sharing beyond viruses. *Molecular ecology*, **29**, 3170–3172.
- Becker DJ, Han BA (2021) The macroecology and evolution of avian competence for *Borrelia burgdorferi*. *Global Ecology and Biogeography*, **30**, 710–724.
- Becker DJ, Seifert SN, Carlson CJ (2020) Beyond Infection: Integrating Competence into Reservoir Host Prediction. *Trends in Ecology and Evolution*, **35**, 1062–1065.
- Bonnot N, Gaillard JM, Coulon A *et al.* (2010) No difference between the sexes in fine-scale spatial genetic structure of roe deer. *PLoS ONE*, **5**.
- Brinkerhoff RJ, Bent SJ, Folsom-O’Keefe CM *et al.* (2010) Genotypic diversity of *Borrelia burgdorferi* strains detected in ixodes scapularis larvae collected from north American songbirds. *Applied and Environmental Microbiology*, **76**, 8265–8268.
- Brinkerhoff RJ, Folsom-O’Keefe CM, Tsao K, Diuk-Wasser MA (2011) Do birds affect Lyme disease risk? Range expansion of the vector-borne pathogen *Borrelia burgdorferi*. *Frontiers in Ecology and the Environment*, **9**, 103–110.
- Brisson D, Drecktrah D, Eggers CH, Samuels DS (2012) Genetics of *Borrelia burgdorferi*. *Annual Review of Genetics*, **46**, 515–536.
- Brisson D, Dykhuizen DE (2004) ospC diversity in *Borrelia burgdorferi*: Different hosts are different niches. *Genetics*, **168**, 713–722.

- Brisson D, Vandermause MF, Meece JK, Reed KD, Dykhuizen DE (2010) Evolution of northeastern and midwestern *Borrelia burgdorferi*, United States. *Emerging Infectious Diseases*, **16**, 911–917.
- Bryant D, Moulton V (2004) Neighbor-Net: An Agglomerative Method for the Construction of Phylogenetic Networks. *Molecular Biology and Evolution*, **21**, 255–265.
- Caine JA, Coburn J (2016) Multifunctional and redundant roles of *Borrelia burgdorferi* outer surface proteins in tissue adhesion, colonization, and complement evasion. *Frontiers in Immunology*, **7**, 1–11.
- Caine JA, Lin YP, Kessler JR *et al.* (2017) *Borrelia burgdorferi* outer surface protein C (OspC) binds complement component C4b and confers bloodstream survival. *Cellular Microbiology*, **19**, 1–14.
- Chao A, Ma KH, Hsieh TC, Chiu C-H (2016) SpadeR: species-richness prediction and diversity estimation with R. *CRAN*.
- Comings SB (2006) *The Nature of Block Island*. Royal Bruce Inc., New Shoreham, RI.
- Derdáková M, Dudiák V, Brei B *et al.* (2004) Interaction and transmission of two *Borrelia burgdorferi* sensu stricto strains in a tick-rodent maintenance system. *Applied and Environmental Microbiology*, **70**, 6783–6788.
- Downs CJ, Schoenle LA, Han BA, Harrison JF, Martin LB (2019) Scaling of Host Competence. *Trends in Parasitology*, **35**, 182–192.
- Dykhuizen DE, Brisson D, Sandigursky S *et al.* (2008) The Propensity of Different *Borrelia burgdorferi* sensu stricto Genotypes to Cause Disseminated Infections in Humans. *American Journal of Tropical Medicine and Hygiene*, **78**, 806–810.
- Dykhuizen DE, Dean AM (2004) Evolution of specialists in an experimental microcosm. *Genetics*, **167**, 2015–2026.
- Earnhart CG, Buckles EL, Dumler JS, Marconi RT (2005) Demonstration of OspC type diversity in invasive human Lyme disease isolates and identification of previously uncharacterized epitopes that define the specificity of the OspC murine antibody response. *Infection and Immunity*, **73**, 7869–7877.
- Elena SF, Lenski RE (2003) Evolution experiments with microorganisms: The dynamics and genetic bases of adaptation. *Nature Reviews Genetics*, **4**, 457–469.
- Falco RC, Fish D (1992) A comparison of methods for sampling the deer tick, *Ixodes dammini*, in a Lyme disease endemic area. *Experimental & Applied Acarology*, **14**, 165–173.
- Gandon S, Buckling A, Decaestecker E, Day T (2008) Host-parasite coevolution and patterns of adaptation across time and space. *Journal of Evolutionary Biology*, **21**, 1861–1866.
- Gatewood AG, Liebman KA, Gwenaél Vourc’h *et al.* (2009) Climate and tick seasonality are predictors of *Borrelia burgdorferi* genotype distribution. *Applied and Environmental Microbiology*, **75**, 2476–2483.
- Ginsberg HS, Buckley PA, Balmforth MG *et al.* (2006) Reservoir Competence of Native North American Birds for the Lyme Disease Spirochete, *Borrelia burgdorferi*. *Journal of Medical Entomology*, **42**, 445–449.
- Grenfell BT, Pybus OG, Gog JR *et al.* (2004) Unifying the Epidemiological and Evolutionary Dynamics of Pathogens. *Science*, **303**, 327–332.
- Griffith DM, Veech JA, Marsh CJ (2016) Cooccur: Probabilistic species co-occurrence analysis in R. *Journal of Statistical Software*, **69**, 1–17.
- Halsey SJ, Allan BF, Miller JR (2018) The role of *Ixodes scapularis*, *Borrelia burgdorferi* and wildlife hosts in Lyme disease prevalence: A quantitative review. *Ticks and Tick-borne*

- Diseases*, **9**, 1103–1114.
- Hanincová K, Kurtenbach K, Diuk-Wasser M, Brei B, Fish D (2006) Epidemic spread of Lyme borreliosis, northeastern United States. *Emerging Infectious Diseases*, **12**, 604–611.
- Hanincová K, Ogden NH, Diuk-Wasser M *et al.* (2008) Fitness variation of *Borrelia burgdorferi* sensu stricto strains in Mice. *Applied and Environmental Microbiology*, **74**, 153–157.
- Haven J, Magori K, Park AW (2012) Ecological and inhost factors promoting distinct parasite life-history strategies in Lyme borreliosis. *Epidemics*, **4**, 152–157.
- Haven J, Vargas LC, Mongodin EF *et al.* (2011) Pervasive recombination and sympatric genome diversification driven by frequency-dependent selection in *borrelia burgdorferi*, the lyme disease bacterium. *Genetics*, **189**, 951–966.
- Hedrick PW (2006) Genetic polymorphism in heterogeneous environments: The age of genomics. *Annual Review of Ecology, Evolution, and Systematics*, **37**, 67–93.
- Huang CI, Kay SC, Davis S *et al.* (2019) High burdens of *Ixodes scapularis* larval ticks on white-tailed deer may limit Lyme disease risk in a low biodiversity setting. *Ticks and Tick-borne Diseases*, **10**, 258–268.
- Inouye M, Dashnow H, Raven LA *et al.* (2014) SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Medicine*, **6**, 1–16.
- Jackson CH (2011) Multi-state models for panel data: The msm package for R. *Journal of Statistical Software*, **38**, 1–28.
- Jones KE, Patel NG, Levy MA *et al.* (2008) Global trends in emerging infectious diseases. *Nature*, **451**, 990–993.
- Karesh WB, Dobson A, Lloyd-Smith JO *et al.* (2012) Ecology of zoonoses: Natural and unnatural histories. *The Lancet*, **380**, 1936–1945.
- Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, **30**, 772–780.
- Khatchikian CE, Nadelman RB, Nowakowski J *et al.* (2014) Evidence for strain-specific immunity in patients treated for early lyme disease. *Infection and Immunity*, **82**, 1408–1413.
- Kilpatrick AM, Dobson ADM, Levi T *et al.* (2017) Lyme disease ecology in a changing world: consensus, uncertainty and critical gaps for improving control. *Proceedings of the Royal Society B*, **372**.
- Koskella B, Vos M (2015) Adaptation in Natural Microbial Populations. *Annual Review of Ecology, Evolution, and Systematics*, **46**, 503–522.
- Kugeler KJ, Schwartz AM, Delorey MJ, Mead PS, Hinckley AF (2021) Estimating the frequency of lyme disease diagnoses, United States, 2010–2018. *Emerging Infectious Diseases*, **27**, 616–619.
- Kurtenbach K, Hanincová K, Tsao JI *et al.* (2006) Fundamental processes in the evolutionary ecology of Lyme borreliosis. *Nature Reviews Microbiology*, **4**, 660–669.
- Lagal V, Portnoi D, Faure G, Postic D, Baranton G (2006) *Borrelia burgdorferi* sensu stricto invasiveness is correlated with OspC-plasminogen affinity. *Microbes and Infection*, **8**, 645–652.
- Leibold MA, Holyoak M, Mouquet N *et al.* (2004) The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, **7**, 601–613.
- Lin T, Oliver JH, Gao L (2002) Genetic diversity of the outer surface protein C gene of southern *Borrelia* isolates and its possible epidemiological, clinical, and pathogenetic implications. *Journal of Clinical Microbiology*, **40**, 2572–2583.

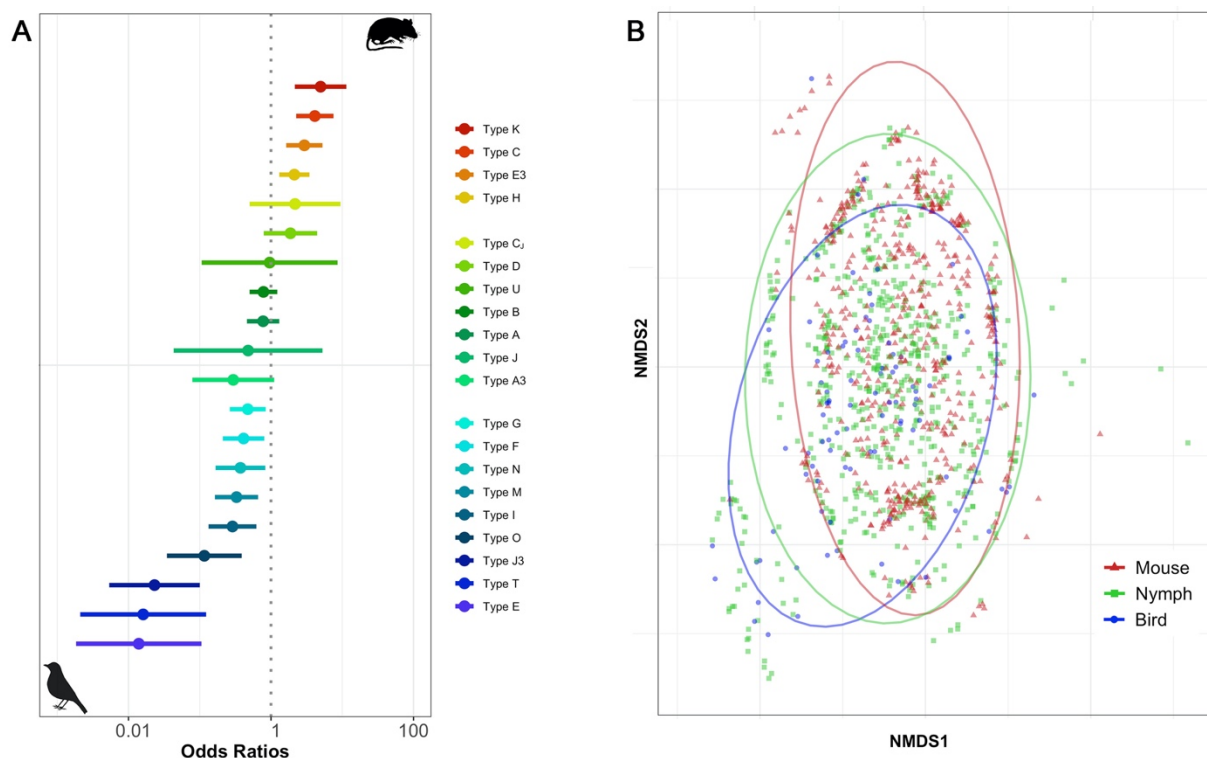
- Lin YP, Tan X, Caine JA *et al.* (2020) Strain-specific joint invasion and colonization by Lyme disease spirochetes is promoted by outer surface protein C. *PLoS Pathogens*, **16**, 1–29.
- Lin Y, Tufts DM, Combs M *et al.* (2022) Cellular and immunological mechanisms influence host-adapted phenotypes in a vector-borne microparasite. *Proceedings of the Royal Society B*.
- Loss SR, Noden BH, Hamer GL, Hamer SA (2016) A quantitative synthesis of the role of birds in carrying ticks and tick-borne pathogens in North America. *Oecologia*, **182**, 947–959.
- Margos G, Hojgaard A, Lane RS *et al.* (2010) Multilocus sequence analysis of *Borrelia bissettii* strains from North America reveals a new *Borrelia* species, *Borrelia kurtenbachii*. *Ticks and Tick-borne Diseases*, **1**, 151–158.
- Martin DP, Varsani A, Roumagnac P *et al.* (2021) RDP5: A computer program for analyzing recombination in, and removing signals of recombination from, nucleotide sequence datasets. *Virus Evolution*, **7**, 5–7.
- Mechai S, Margos G, Feil EJ *et al.* (2016) Evidence for host-genotype associations of *Borrelia burgdorferi* sensu stricto. *PLoS ONE*, **11**, 1–25.
- Mongodin EF, Casjens SR, Bruno JF *et al.* (2013) Inter- and intra-specific pan-genomes of *Borrelia burgdorferi* sensu lato: genome stability and adaptive radiation. *BMC Genomics*, **14**.
- Nadelman RB, Hanincová K, Mukherjee P *et al.* (2012) Differentiation of Reinfection from Relapse in Recurrent Lyme Disease. *New England Journal of Medicine*, **367**, 1883–1890.
- O’Keeffe KR, Oppler ZJ, Brisson D (2020) Evolutionary ecology of Lyme *Borrelia*. *Infection, Genetics and Evolution*, **85**, 104570.
- Ogden NH, Bigras-Poulin M, O’Callaghan CJ *et al.* (2007) Vector seasonality, host infection dynamics and fitness of pathogens transmitted by the tick *Ixodes scapularis*. *Parasitology*, **134**, 209–227.
- Oksanen AJ, Blanchet FG, Friendly M *et al.* (2020) Vegan: Community Ecology Package. *CRAN*.
- Olival KJ, Hosseini PR, Zambrana-Torrel C *et al.* (2017) Host and viral traits predict zoonotic spillover from mammals. *Nature*, **546**, 646–650.
- Prevention C for DC and (2022) Lyme Disease.
- Probert WS, Crawford M, Cadiz RB, LeFebvre RB (1997) Immunization with outer surface protein (Osp) A, but not OspC, provides cross-protection of mice challenged with North American isolates of *Borrelia burgdorferi*. *Journal of Infectious Diseases*, **175**, 400–405.
- Qiu WG, Bosler EM, Campbell JR *et al.* (1997) 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.
- Qiu WG, Martin CL (2014) Evolutionary Genomics of *Borrelia burgdorferi* sensu lato: Findings, Hypotheses, and the Rise of Hybrids. *Infection, Genetics and Evolution*, **23**, 1–7.
- Radolf JD, Strle K, Lemieux JE, Strle F (2020) Lyme disease in humans. *Current Issues in Molecular Biology*, **42**, 333–384.
- Ramamoorthi N, Narasimhan S, Pal U *et al.* (2005) The Lyme disease agent exploits a tick protein to infect the mammalian host Nandhini. *Nature*, **436**, 573–577.
- Rataud A, Henry PY, Moutailler S, Marsot M (2021) Research effort on birds’ reservoir host potential for Lyme borreliosis: A systematic review and perspectives. *Transboundary and Emerging Diseases*, 1–11.
- Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source



- tool for metagenomics. *PeerJ*, **2016**, 1–22.
- Sánchez CA, Venkatachalam-Vaz J, Drake JM (2021) Spillover of zoonotic pathogens: A review of reviews. *Zoonoses and Public Health*, **68**, 563–577.
- Schwartz I, Margos G, Casjens SR, Qiu WG, Eggers CH (2021) Multipartite genome of lyme disease borrelia: Structure, variation and prophages. *Current Issues in Molecular Biology*, **42**, 409–454.
- Seemann T (2014) Prokka: Rapid prokaryotic genome annotation. *Bioinformatics*, **30**, 2068–2069.
- Seinost G, Golde WT, Berger BW *et al.* (1999) Infection With Multiple Strains of *Borrelia burgdorferi* Sensu Stricto in Patients With Lyme Disease. *Archives of Dermatology*, **135**, 1329–1333.
- Sexton JP, Montiel J, Shay JE, Stephens MR, Slatyer RA (2017) Evolution of Ecological Niche Breadth. *Annual Review of Ecology, Evolution, and Systematics*, **48**, 183–206.
- Sievers F, Higgins DG (2014) Clustal Omega. *Current Protocols in Bioinformatics*, **2014**, 3.13.1–3.13.16.
- States SL, Brinkerhoff RJ, Carpi G *et al.* (2014) Lyme disease risk not amplified in a species-poor vertebrate community: similar *Borrelia burgdorferi* tick infection prevalence and *OspC* genotype frequencies. *Infection, Genetics and Evolution*, **27**, 266–575.
- Stewart Merrill TE, Johnson PTJ (2020) Towards a mechanistic understanding of competence: A missing link in diversity-disease research. *Parasitology*, **147**, 1159–1170.
- Strle K, Jones KL, Drouin EE, Li X, Steere AC (2011) *Borrelia burgdorferi* RST1 (*OspC* type A) genotype is associated with greater inflammation and more severe lyme disease. *American Journal of Pathology*, **178**, 2726–2739.
- Swanson KI, Norris DE (2008) Presence of Multiple Variants of *Borrelia burgdorferi* in the Natural Reservoir *Peromyscus leucopus* Throughout a Transmission Season. *Vector-Borne and Zoonotic Diseases*, **8**, 397–406.
- Swei A, Couper LI, Coffey LL, Kapan D, Bennett S (2020) Patterns, Drivers, and Challenges of Vector-Borne Disease Emergence. *Vector-Borne and Zoonotic Diseases*, **20**, 159–170.
- Team RC (2021) R: A Language and environment for statistical computing.
- Thrall PH, Laine AL, Ravensdale M *et al.* (2012) Rapid genetic change underpins antagonistic coevolution in a natural host-pathogen metapopulation. *Ecology Letters*, **15**, 425–435.
- Tilly K, Krum JG, Bestor A *et al.* (2006) *Borrelia burgdorferi* *OspC* protein required exclusively in a crucial early stage of mammalian infection. *Infection and Immunity*, **74**, 3554–3564.
- Tsao JI, Hamer SA, Han S, Sidge JL, Hickling GJ (2021) The Contribution of Wildlife Hosts to the Rise of Ticks and Tick-Borne Diseases in North America. *Journal of Medical Entomology*, **58**, 1565–1587.
- Vuong HB, Canham CD, Fonseca DM *et al.* (2014) Occurrence and transmission efficiencies of *Borrelia burgdorferi* *ospC* types in avian and mammalian wildlife. *Infection, Genetics and Evolution*, **27**, 594–600.
- Wang IN, Dykhuizen DE, Qiu W *et al.* (1999) Genetic diversity of *ospC* in a local population of *Borrelia burgdorferi* sensu stricto. *Genetics*, **151**, 15–30.
- Wang G, Ojaimi C, Iyer R *et al.* (2001) Impact of genotypic variation of *Borrelia burgdorferi* sensu stricto on kinetics of dissemination and severity of disease in C3H/HeJ mice. *Infection and Immunity*, **69**, 4303–4312.
- Wang G, Ojaimi C, Wu H *et al.* (2002) Disease Severity in a Murine Model of Lyme Borreliosis Is Associated with the Genotype of the Infecting *Borrelia burgdorferi* Sensu Stricto Strain .

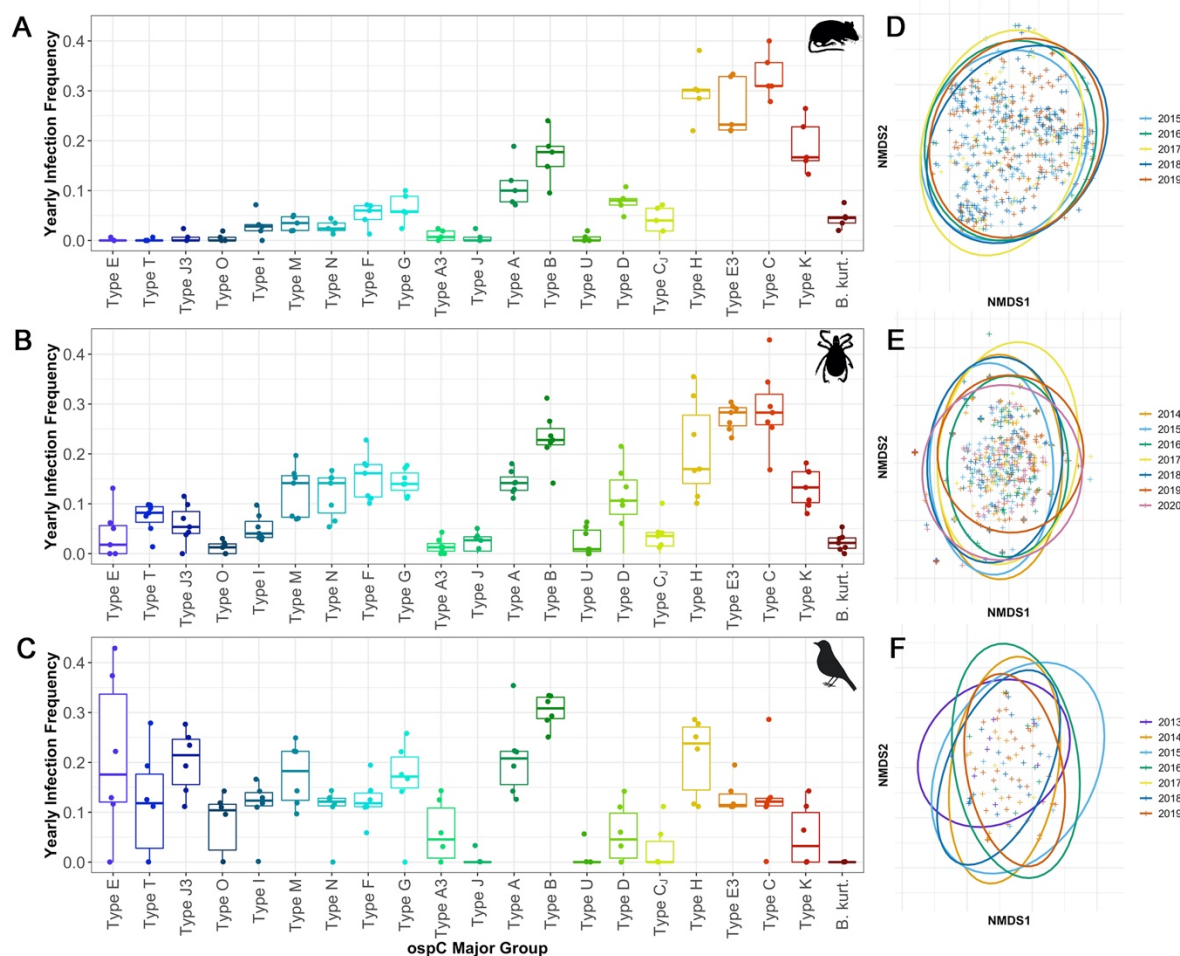
- The Journal of Infectious Diseases*, **186**, 782–791.
- Weedall GD, Conway DJ (2010) Detecting signatures of balancing selection to identify targets of anti-parasite immunity. *Trends in Parasitology*, **26**, 363–369.
- Wells K, Clark NJ (2019) Host Specificity in Variable Environments. *Trends in Parasitology*, **35**, 452–465.
- Woods R, Reiss A, Cox-Witton K, Grillo T, Peters A (2019) The importance of wildlife disease monitoring as part of global surveillance for zoonotic diseases: The role of Australia. *Tropical Medicine and Infectious Disease*, **4**.
- Woolhouse MEJ, Gowtage-Sequeria S (2005) Host range and emerging and reemerging pathogens. *Emerging Infectious Diseases*, **11**, 1842–1847.
- Wormser GP, Brisson D, Liveris D *et al.* (2008) *Borrelia burgdorferi* Genotype Predicts the Capacity for Hematogenous Dissemination during Early Lyme Disease . *The Journal of Infectious Diseases*, **198**, 1358–1364.
- Wywiał E, Haven J, Casjens SR *et al.* (2009) Fast, adaptive evolution at a bacterial host-resistance locus: The PFam54 gene array in *Borrelia burgdorferi*. *Gene*, **445**, 26–37.

**Figure 1** Host associations among *ospC* genotypes. (A) Log odds of infection with each genotype in mice compared to bird hosts. Significant associations exhibit confidence intervals that do not overlap the dotted line. (B) NMDS plot of individual mouse, nymph, and bird genotype communities with 95% confidence ellipses.

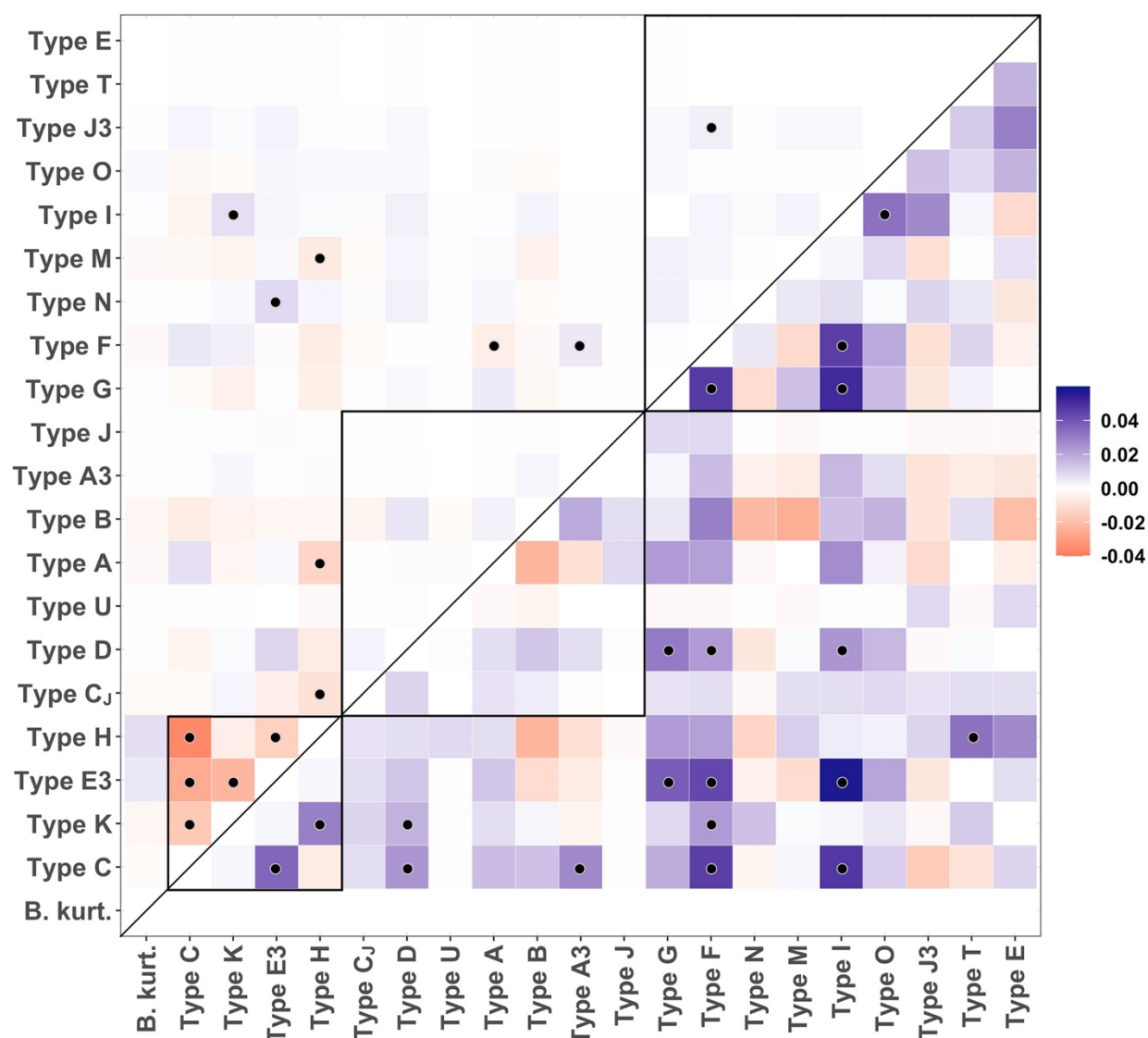




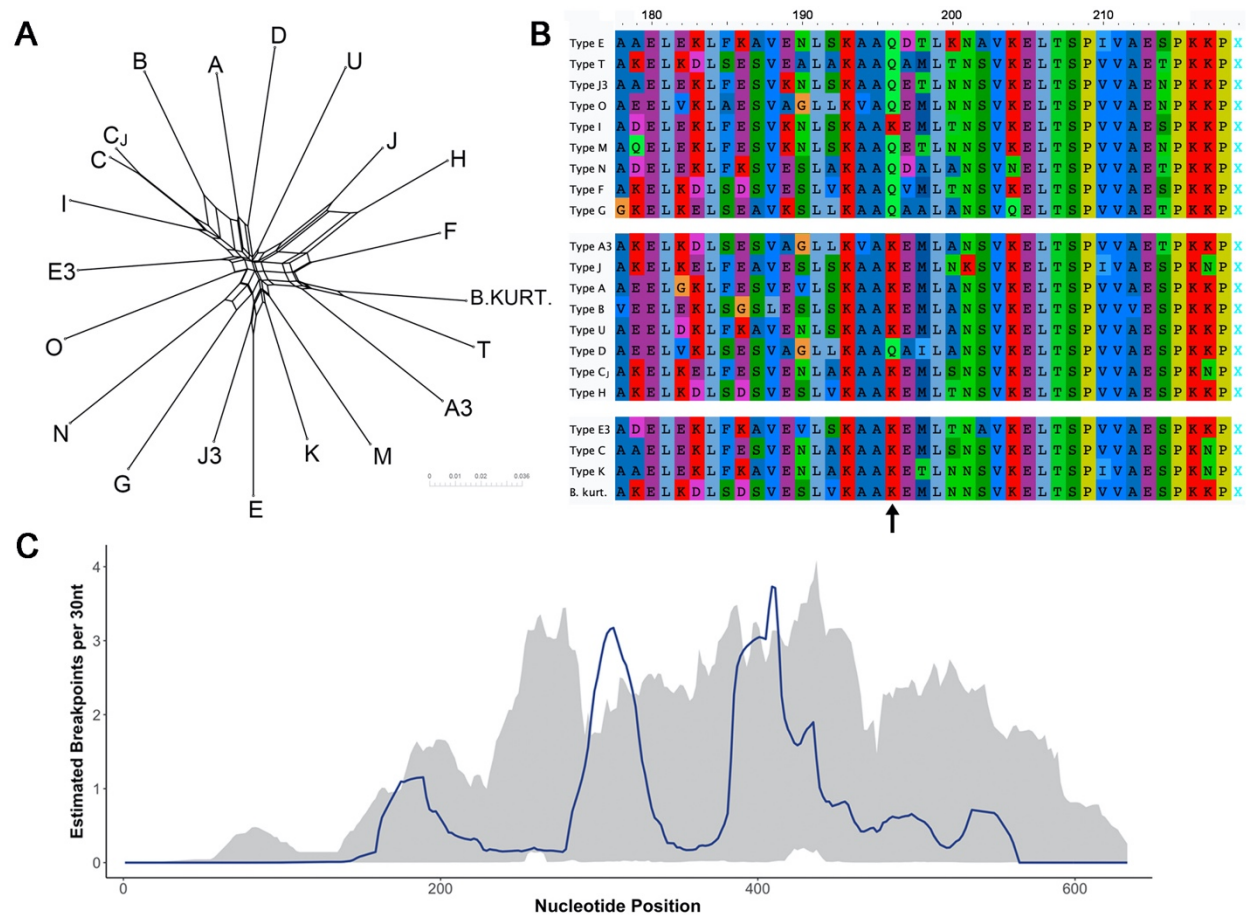
**Figure 2** Yearly variation in genotype frequency and individual community structure. Boxplots of ospC genotype frequency distributions observed in mice (A), nymphs (B), and birds (C). NMDS plots of individual communities separated by year among mice (D), nymphs (E), and birds (F) with 95% confidence ellipses. Boxplot colors reflect strength of host-association shown in Figure 1.



**Figure 3** Heatmap of effect size for cooccurrence probability estimates for ospC genotypes found in individual mammal (top triangle) or bird hosts (bottom triangle). Color indicates the strength of the effect (blue = positive, red = negative), and black dots indicate significant associations ( $p < 0.05$ ). Black boxes within the heatmap surround the putative mouse-adapted genotypes (bottom), generalist genotypes (middle), and bird-adapted genotypes (top).



**Figure 4** Evolutionary patterns at the *ospC* locus. (A) Phylogenetic network of 21 observed genotypes. (B) Amino acid alignment of 5' end of *ospC* gene with arrow identifying the K196Q mutation. Genotypes are separated by phenotypic groups: bird-adapted (top), generalist (middle), mouse-adapted (bottom). (C) Recombination breakpoint estimation (blue line) across *ospC*. Shaded area represents 95% confidence intervals for local hot/coldspot test. Areas where observed counts exceed confidence interval are inferred recombination hotspots.



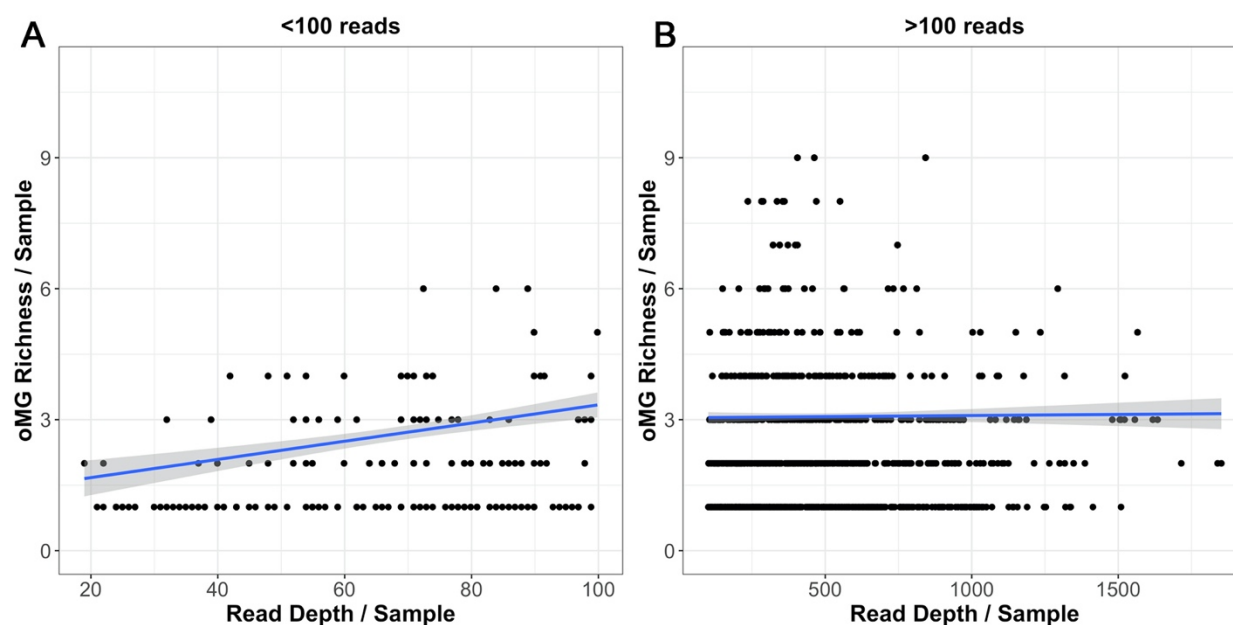
**Table 1** Log odds ratio of strain-specific infection in mouse host and tick vector relative to bird hosts. Significant odds ratios are in bold and 95% confidence intervals provided in parentheses.

<i>ospC</i> Genotype	Mouse		Tick	
	Odds Ratio	p-value	Odds Ratio	p-value
K	<b>4.97 (2.16 – 11.43)</b>	<0.001	<b>2.32 (1.00 – 5.37)</b>	0.050
C	<b>4.12 (2.25 – 7.53)</b>	<0.001	<b>2.74 (1.50 – 4.99)</b>	0.001
E3	<b>2.94 (1.63 – 5.28)</b>	<0.001	<b>2.31 (1.29 – 4.14)</b>	0.005
H	<b>2.13 (1.31 – 3.46)</b>	0.002	1.10 (0.67 – 1.80)	0.705
CJ	2.17 (0.50 – 9.44)	0.300	1.98 (0.47 – 8.45)	0.354
D	1.88 (0.79 – 4.44)	0.153	2.02 (0.87 – 4.70)	0.103
U	0.96 (0.11 – 8.60)	0.968	0.16 (0.06 – 0.43)	<0.001
B	0.78 (0.50 – 1.23)	0.285	0.84 (0.55 – 1.29)	0.423
A	0.77 (0.46 – 1.31)	0.338	0.70 (0.42 – 1.17)	0.173
J	0.48 (0.04 – 5.28)	0.546	2.47 (0.33 – 18.82)	0.382
G	<b>0.47 (0.26 – 0.84)</b>	0.011	0.84 (0.50 – 1.41)	0.505
F	<b>0.41 (0.21 – 0.81)</b>	0.010	1.12 (0.63 – 2.00)	0.700
N	<b>0.37 (0.17 – 0.83)</b>	0.016	1.24 (0.63 – 2.44)	0.529
M	<b>0.33 (0.16 – 0.66)</b>	0.002	0.88 (0.49 – 1.59)	0.680
A3	<b>0.30 (0.08 – 1.11)</b>	0.071	0.41 (0.13 – 1.30)	0.130
I	<b>0.29 (0.13 – 0.62)</b>	0.002	<b>0.45 (0.23 – 0.89)</b>	0.021
O	<b>0.12 (0.03 – 0.39)</b>	<0.001	<b>0.16 (0.06 – 0.43)</b>	<0.001
J3	<b>0.02 (0.01 – 0.10)</b>	<0.001	<b>0.32 (0.18 – 0.56)</b>	<0.001
T	<b>0.02 (0.00 – 0.12)</b>	<0.001	<b>0.16 (0.06 – 0.43)</b>	<0.001
E	<b>0.01 (0.00 – 0.11)</b>	<0.001	<b>0.21 (0.11 – 0.41)</b>	<0.001

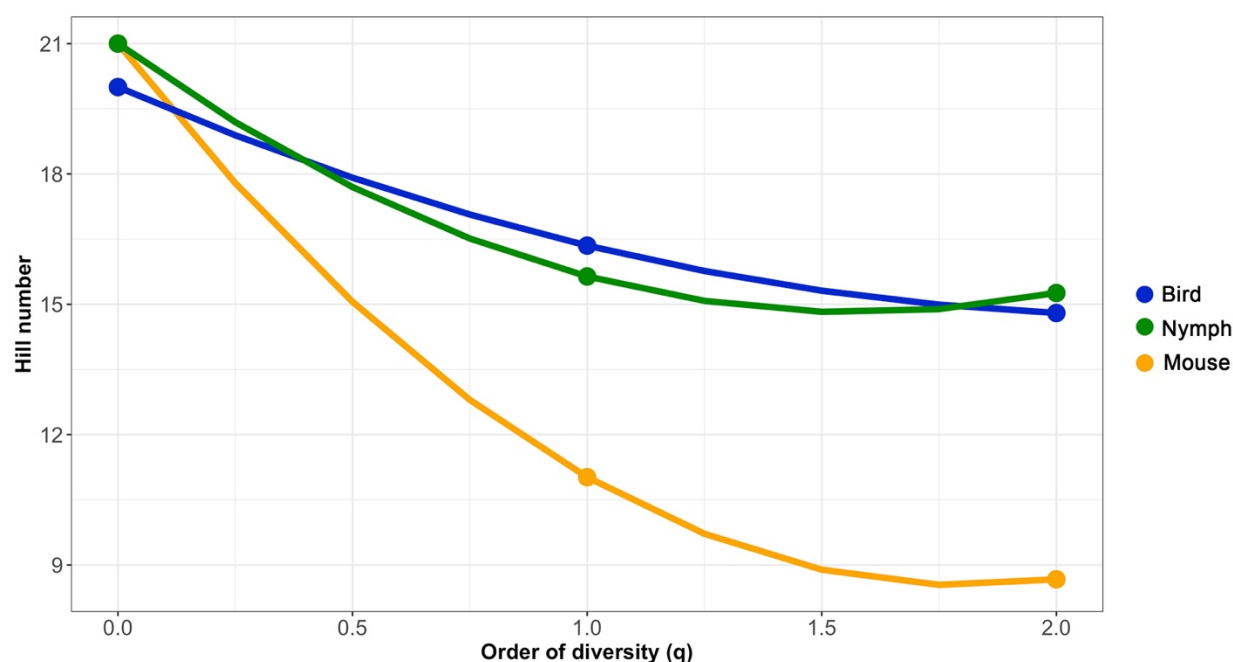
**Table 2** Transition probabilities between infection states in white-footed mice estimated over a 4-week period with the multistate Markov model. Columns and rows represent starting and transition states, respectively. 95% confidence intervals are provided in parentheses.

	<b>Uninfected</b>	<b>Adapted</b>	<b>Non-Adapted</b>
Uninfected	0.722 (0.664 - 0.765)	0.346 (0.277 - 0.258)	0.167 (0.121 - 0.237)
Adapted	0.164 (0.131 - 0.212)	0.501 (0.395 - 0.580)	0.336 (0.252 - 0.42.7)
Non-Adapted	0.114 (0.088 - 0.146)	0.152 (0.111 - 0.205)	0.497 (0.401 - 0.592)

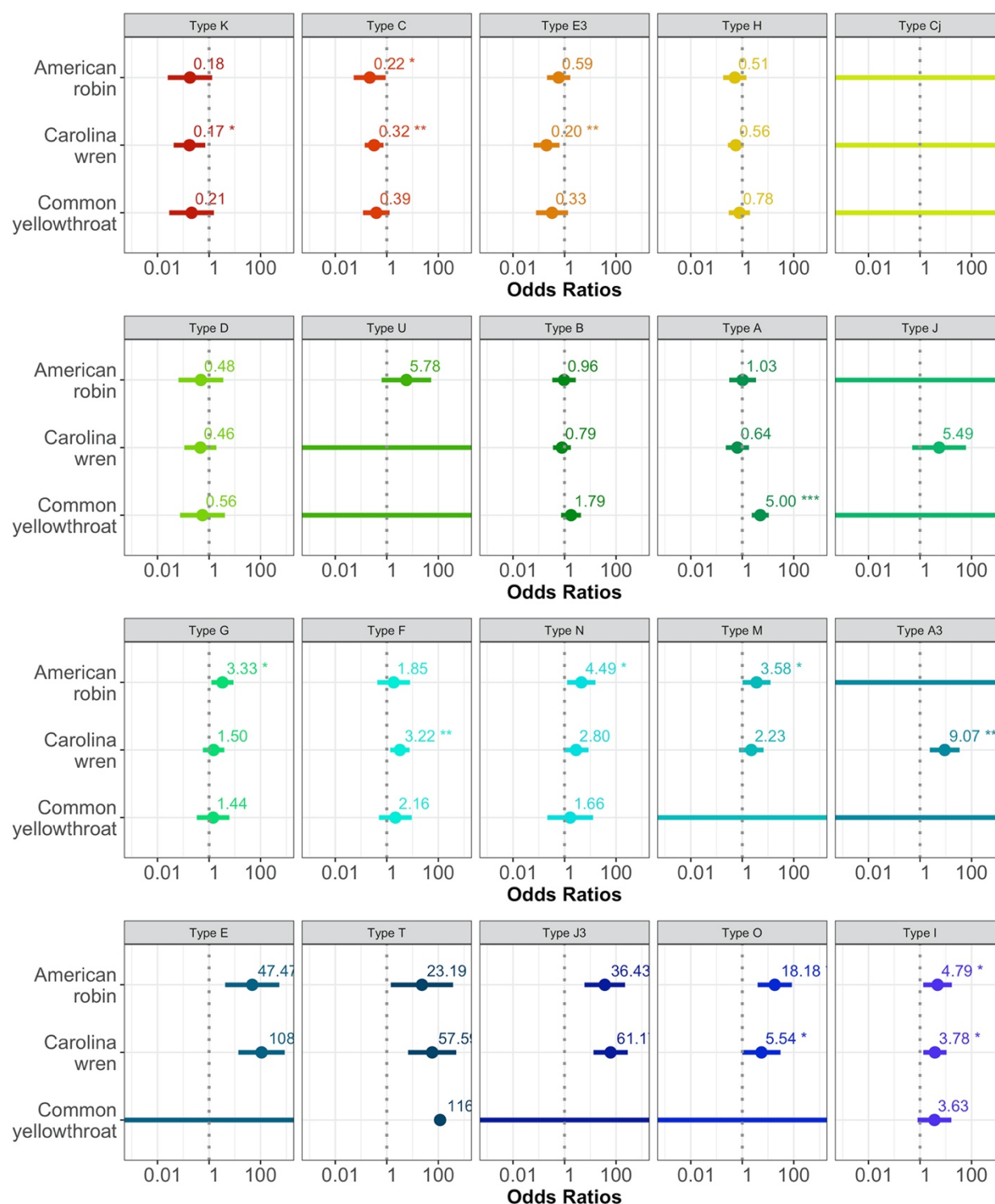
**Figure S1** Sequencing depth bias observed for samples represented by fewer than 100 reads (A) but absent for samples represented by more than 100 reads (B).



**Figure S2** Diversity profiles using Hill numbers for genotype communities observed within mice (orange), birds (blue), or nymphs (green). Flatter profiles represent greater evenness and steeper slopes represent increasing unevenness.  $q = 0$  is equivalent to community richness, while a  $q = 1$  is equivalent to the Shannon index of diversity, and a  $q = 2$  is equivalent to the multiplicative inverse of the Simpson diversity index.

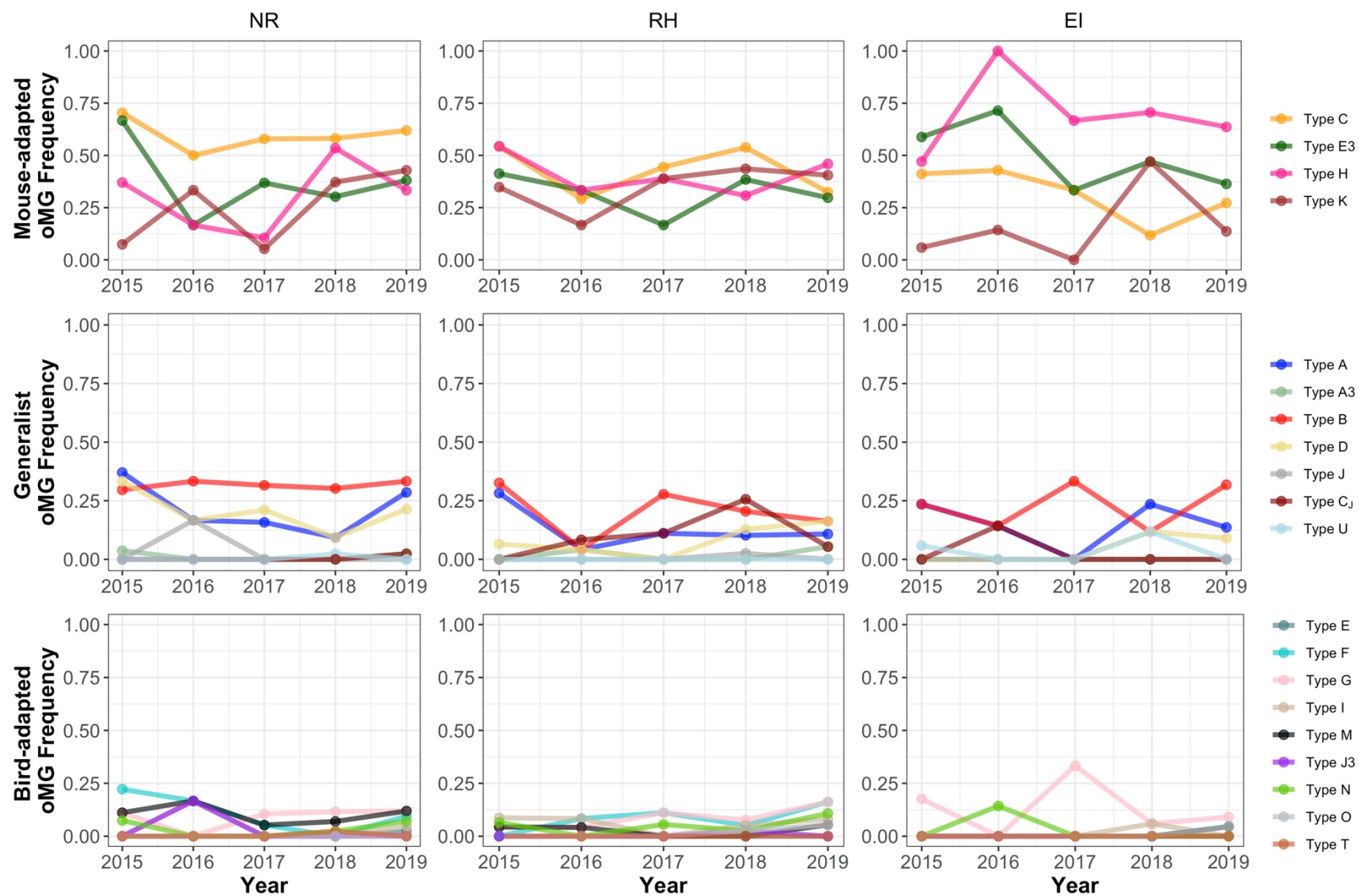


**Figure S3** Bird species specific binomial GLMs for each *Bb* genotype. Odds ratios are relative to genotype infection in mice. Colors represent strength of host association observed when bird species were treated as a single taxa group, as labeled in Figure 1.





**Figure S4** Yearly frequency of individual genotypes in mice across sites.

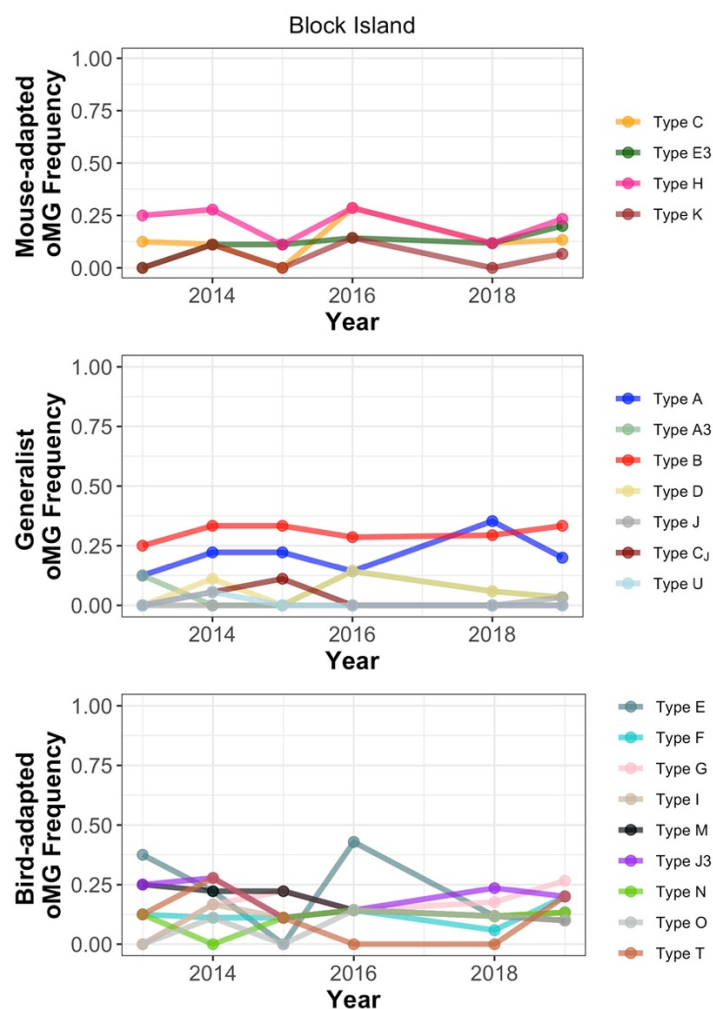




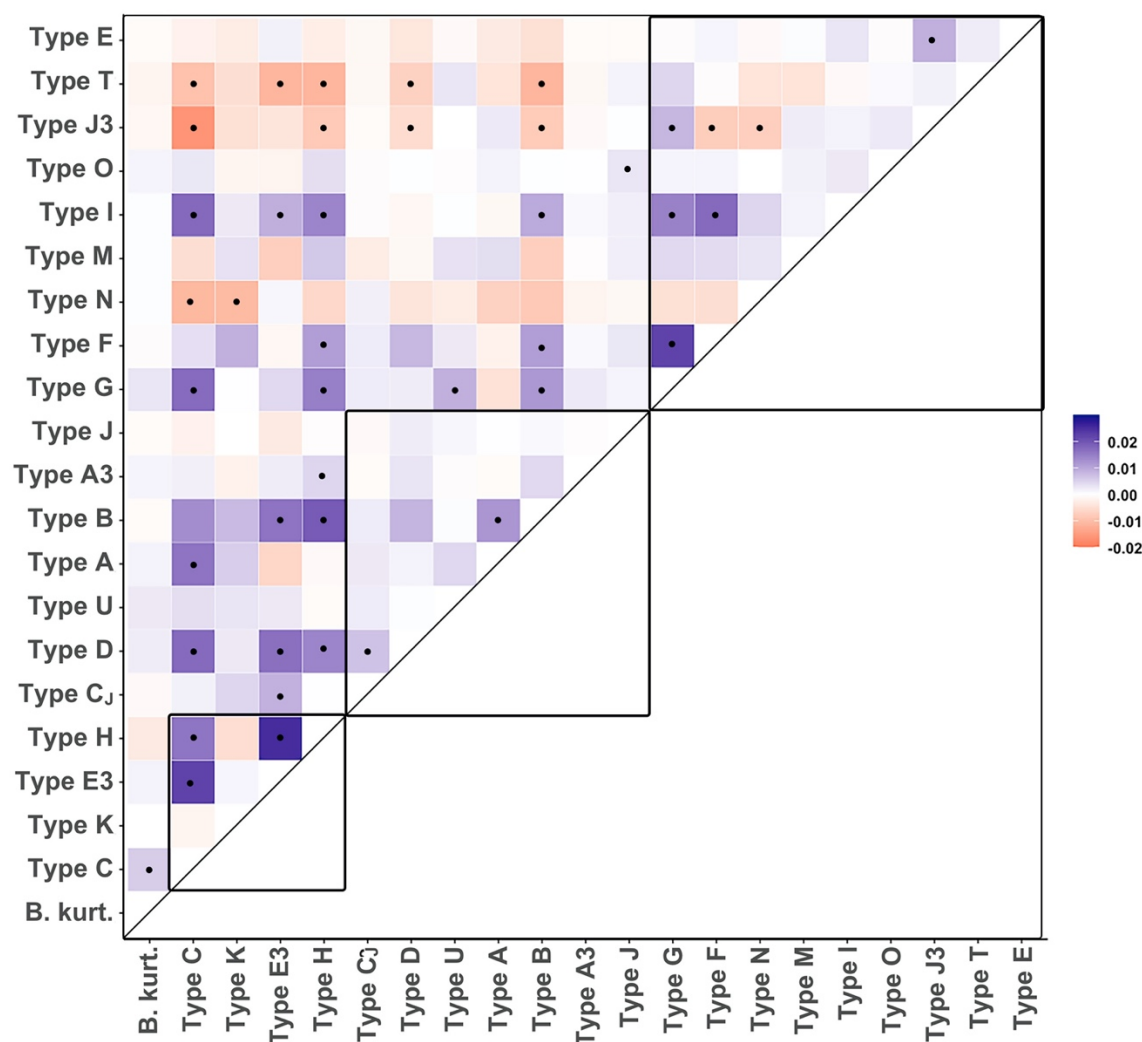
**Figure S5** Yearly frequency of individual genotypes in nymphal ticks across sites.



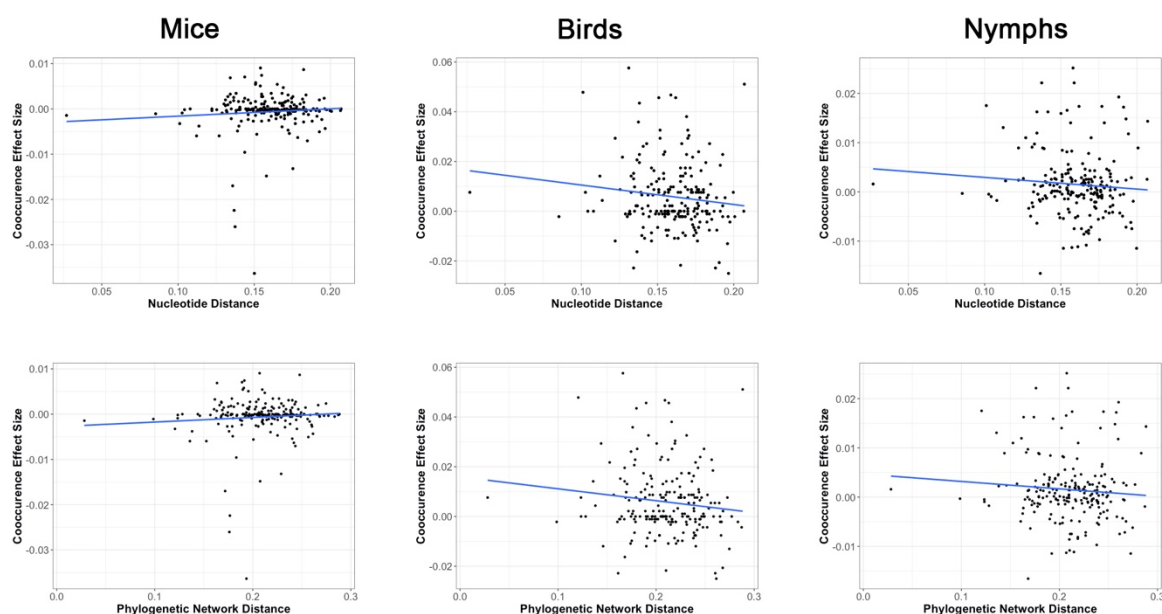
**Figure S6** Yearly frequency of individual genotypes in bird hosts, excluding 2017 where n = 2.



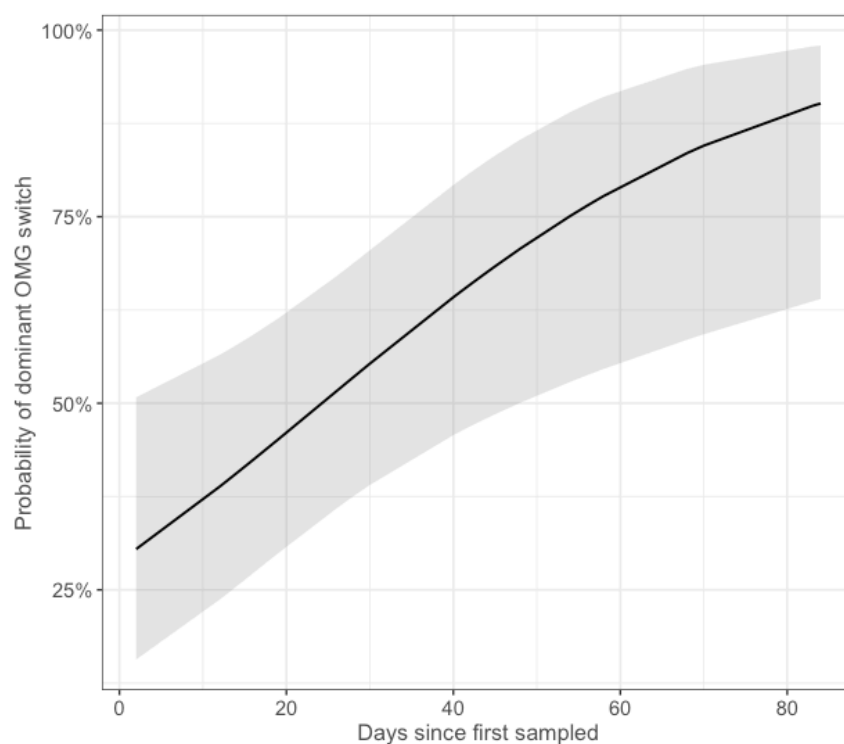
**Figure S7** Heatmap of genotype co-occurrence probability effect sizes for nymphal ticks. Color indicates the strength of the effect (blue = positive, red = negative), and black dots indicate significant associations ( $p < 0.05$ ). Black boxes within the heatmap surround the putative mouse-adapted genotypes (bottom), generalist genotypes (middle), and bird-adapted genotypes (top).



**Figure S8** Correlations between pairwise co-occurrence effect sizes among genotypes and nucleotide distance (top row) or phylogenetic network distance (bottom row) among mice (left column), birds (middle column), and nymphal ticks (right column).



**Figure S9** Predicted relationship between time (days) and the probability of turnover in the dominant genotype within individual mammal communities.



**Table S1** Sequencing of genotypes across *Bb* infected tick nymphs, white footed mice, and passerine birds. Details for individual bird species are provided in Table S3.

Source	Nymphs		Mice		Birds	
n	553		628		92	
Richness	2.6		1.6		2.6	
oMG	Count	Freq.	Count	Freq.	Count	Freq.
A	88	0.140	66	0.119	20	0.217
B	146	0.232	94	0.170	28	0.304
C	184	0.293	178	0.322	12	0.130
Cj	24	0.038	18	0.033	2	0.022
D	72	0.115	46	0.083	6	0.065
E	22	0.035	1	0.002	16	0.174
F	95	0.151	25	0.045	14	0.152
G	93	0.148	37	0.067	18	0.196
H	133	0.212	162	0.293	20	0.217
I	34	0.054	15	0.027	12	0.130
J	15	0.024	2	0.004	1	0.011
K	82	0.131	113	0.204	6	0.065
M	76	0.121	20	0.036	14	0.152
N	75	0.119	16	0.029	10	0.109
O	8	0.013	4	0.007	8	0.087
T	48	0.076	1	0.002	14	0.152
U	16	0.025	4	0.007	1	0.011
A3	10	0.016	5	0.009	4	0.043
E3	171	0.272	144	0.260	13	0.141
J3	39	0.062	2	0.004	19	0.207
B.kurt.	15	0.024	27	0.049	0	0.000

**Table S2** Pairwise genotype community similarity among populations of mammals, birds, and nymphs calculated with the local (Sørensen-type) species overlap measure.

	Mammals	Birds	Nymphs
Mammals	1.000	0.672	0.904
Birds		1.000	0.883
Nymphs			1.000

**Table S3** Sequencing of genotypes across individual species of *Bb* passerine birds.

[illegible]

**Table S4** Mean sojourn time for each infection state as determined by the MSM model. Values represent the expected duration of a mouse infection in days. 95% confidence intervals are provided in parentheses.

	Mean Sojourn Time
Uninfected	53.9 (41.3 - 71.2)
Adapted	27.6 (19.6 - 38.7)
Non-Adapted	9.5 (6.2 - 14.5)

**Table S5** Estimated probabilities of a mouse's next different infection state (rows), given the current state (columns), determined by the MSM model. 95% confidence intervals are provided in parentheses.

	Uninfected	Adapted	Non-Adapted
Uninfected	NA	0.37 (0.16 - 0.65)	0.51 (0.35 - 0.70)
Adapted	0.32 (0.14 - 0.55)	NA	0.49 (0.30 - 0.65)
Non-Adapted	0.68 (0.45 - 0.86)	0.63 (0.35 - 0.84)	NA

**Table S6** Model results for the GLMM testing influence of time and individual mouse characteristics on the probability of observing a switch in the dominant genotype.

Predictors	Odds Ratios	CI	p
(Intercept)	0.51	0.20 – 1.30	0.164
Time since first sample (days)	1.03	1.01 – 1.06	<b>0.006</b>
oMG richness	1.13	0.84 – 1.53	0.43
Sex (Male)	0.92	0.50 – 1.71	0.796
Age (non-adult)	1.52	0.28 – 11.41	0.643
Nmph1 Burden	0.99	0.94 – 1.06	0.862



**Table S7** Recombination analysis results.

In Alignment		In Recombinant Sequence					Detection Method						
Begin	End	Begin	End	Recombinant Sequence(s)	Minor Parental Sequence(s)	Major Parental Sequence(s)	RDP	GENECONV	Bootscan	Maxchi	Chimaera	SiScan	3Seq
303	413	297	398	B	Cj, C	J, H	5.15E-05	2.21E-04	2.18E-06	1.62E-07	1.37E-07	1.62E-09	1.12E-07
459	554	441	530	B.kurt	J	C	NS	2.84E-07	1.28E-08	2.56E-04	2.56E-04	1.63E-05	8.66E-08
350	408	335	393	I	D	Cj, C	3.84E-07	2.78E-06	1.53E-06	4.60E-03	7.96E-04	2.64E-05	2.38E-05
303	493	297	469	J3	A	M	NS	2.38E-03	9.61E-04	5.91E-06	9.63E-03	3.89E-07	6.14E-07
142	418	139	403	A	G	C, Cj	NS	NS	NS	5.17E-05	4.68E-06	4.50E-03	4.55E-04
320	526	317	508	B.kurt	E	H	NS	NS	NS	1.91E-05	4.38E-02	NS	1.03E-02
628	186	604	183	H	O	B.kurt	NS	2.22E-03	1.96E-03	1.88E-02	1.43E-02	NS	2.22E-03
518	198	497	195	E3	U	G	NS	NS	8.77E-03	2.09E-04	5.25E-05	4.67E-04	NS
608	294	587	291	K	J	N	NS	NS	8.03E-04	NS	5.73E-03	NS	NS
250	309	247	303	J3	E	A	NS	NS	NS	2.51E-03	NS	1.49E-03	1.66E-03
340	418	326	400	O	C, Cj	J, H	NS	NS	2.19E-03	4.52E-03	1.66E-03	2.85E-05	4.28E-02

**Table S8** Accession numbers for strains used for alignment during primer design.

Accession	Strain
<b>CP001550</b>	29805
<b>CP002239</b>	N40
<b>CP001446</b>	WI91-23
<b>CP017202</b>	B331
<b>CP001535</b>	118a
<b>CP001375</b>	72a
<b>CP001484</b>	CA-11.2A
<b>CP001493</b>	94a
<b>CP031398</b>	MM1
<b>CP002316</b>	JD1
<b>CP001271</b>	156a
<b>CP002268</b>	297
<b>CP001422</b>	64b
<b>CP001212</b>	ZS7
<b>AE000792</b>	B31
<b>CP019917</b>	Pabe
<b>CP019755</b>	B31_NRZ
<b>CP001568</b>	Bol26

**Table S9** Barcoded *ospC* primer set 1 used in this study to amplify 1500bp region centered around the *ospC* locus. Barcodes selected from a set of 384 sequences provided by Pacific Biosciences.

Primer Name	Barcoded Primer Sequence
<b>ospC_16.6_F_1</b>	CACATATCAGAGTGCGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_2</b>	ACACACAGACTGTGAGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_3</b>	ACACATCTCGTGAGAGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_4</b>	CACGCACACACGCGCGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_5</b>	CACTCGACTCTCGCGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_6</b>	CATATATATCAGCTGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_7</b>	TCTGTATCTCTATGTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_8</b>	ACAGTCGAGCGCTGCGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_9</b>	ACACACGCGAGACAGAGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_10</b>	ACGCGCTATCTCAGAGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_11</b>	CTATACGTATATCTATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_12</b>	ACACTAGATCGCGTGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_13</b>	CTCTCGCATACGCGAGGGGATCCAAAATCTAATACAA

<b>ospC_16.6_F_14</b>	CTCACTACGCGCGCGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_15</b>	CGCATGACACGTGTGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_16</b>	CATAGAGAGATAGTATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_17</b>	CACACGCGCGCTATATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_18</b>	TCACGTGCTCACTGTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_19</b>	ACACACTCTATCAGATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_20</b>	CACGACACGACGATGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_21</b>	CTATACATAGTGATGTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_22</b>	CACTCACGTGTGATATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_23</b>	CAGAGAGATATCTCTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_24</b>	CATGTAGAGCAGAGAGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_25</b>	CGCGACACGCTCGCGCGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_26</b>	CACAGAGACACGCACAGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_27</b>	CTCACACTCTCTCACAGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_28</b>	CTCTGCTCTGACTCTCGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_29</b>	TATATATGTCTATAGAGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_30</b>	TCTCTCTATCGCGCTCGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_31</b>	GATGTCTGAGTGTGTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_32</b>	GAGACTAGAGATAGTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_33</b>	TCTCGTCGCAGTCTCTGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_34</b>	ATGTGTATATAGATATGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_35</b>	GCGCGCGCACTCTCTGGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_36</b>	GAGACACGTCGCACACGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_37</b>	ACACATATCGCACTACGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_38</b>	GTGTGTCTCGATGCGCGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_39</b>	CGCACACATAGATACAGGGATCCAAAATCTAATACAA
<b>ospC_16.6_F_40</b>	TGTCATATGAGAGTGTGGGATCCAAAATCTAATACAA
<b>ospC-18.1_R_1</b>	CGCACTCTGATATGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_2</b>	CTCACAGTCTGTGTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_3</b>	CTCTCACGAGATGTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_4</b>	CGCGCGTGTGTGCGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_5</b>	ACGCGAGAGTCGAGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_6</b>	ACAGCTGATATATATGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_7</b>	CACATAGAGATACAGACCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_8</b>	CGCAGCGCTCGACTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_9</b>	TCTGTCTCGCGTGTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_10</b>	CTCTGAGATAGCGCGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_11</b>	ATAGATATACGTATAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_12</b>	ACACGCGATCTAGTGTCCCTTAACATACAATATCTCTTC

<b>ospC-18.1_R_13</b>	CTCGCGTATGCGAGAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_14</b>	ACGCGCGCGTAGTGAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_15</b>	ACACACGTGTCATGCGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_16</b>	ATACTATCTCTCTATGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_17</b>	ATATAGCGCGCGTGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_18</b>	CACAGTGAGCACGTGACCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_19</b>	ATCTGATAGAGTGTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_20</b>	ACATCGTCGTGTCGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_21</b>	ACATCACTATGTATAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_22</b>	ATATCACACGTGAGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_23</b>	CAGAGATATCTCTCTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_24</b>	CTCTCTGCTCTACATGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_25</b>	GCGCGAGCGTGTGCGGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_26</b>	TGTGCGTGTCTCTGTGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_27</b>	TGTGAGAGAGTGTGAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_28</b>	GAGAGTCAGAGCAGAGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_29</b>	TCTATAGACATATATACCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_30</b>	GAGCGCGATAGAGAGACCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_31</b>	CACACACTCAGACATCCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_32</b>	CACTATCTCTAGTCTCCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_33</b>	AGAGACTGCGACGAGACCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_34</b>	ATATCTATATACACATCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_35</b>	CAGAGAGTGCGCGCGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_36</b>	GTGTGCGACGTGTCTCCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_37</b>	GTAGTGCGATATGTGTCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_38</b>	GCGCATCGAGACACACCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_39</b>	TGTATCTATGTGTGCGCCCTTAACATACAATATCTCTTC
<b>ospC-18.1_R_40</b>	ACACTCTCATATGACACCCTTAACATACAATATCTCTTC

**Table S10** Barcoded *ospC* primer set 2 used in this study to amplify 750bp region encompassing the *ospC* locus. Barcodes selected from a set of 384 sequences provided by Pacific Biosciences.

Primer Name	Barcoded Primer Sequence
BL-1F-BC1	CACATATCAGAGTGCGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC2	ACACACAGACTGTGAGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC3	ACACATCTCGTGAGAGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC4	CACGCACACACGCGCGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC5	CACTCGACTCTCGCGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC6	CATATATATCAGCTGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC7	TCTGTATCTCTATGTGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC8	ACAGTCGAGCGCTGCGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC9	ACACACGCGAGACAGAGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC10	ACGCGCTATCTCAGAGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC11	CTATACGTATATCTATGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC12	ACACTAGATCGCGTGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC13	CTCTCGCATACGCGAGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC14	CTCACTACGCGCGCGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC15	CGCATGACACGTGTGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC16	CATAGAGAGATAGTATGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC17	CACACGCGCGCTATATGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC18	TCACGTGCTCACTGTGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC19	ACACACTCTATCAGATGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC20	CACGACACGACGATGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC21	CTATACATAGTGATGTGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC22	CACTCACGTGTGATATGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC23	CAGAGAGATATCTCTGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC24	CATGTAGAGCAGAGAGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC25	CGCGACACGCTCGCGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC26	CACAGAGACACGCACAGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC27	CTCACACTCTCTCACAGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC28	CTCTGCTCTGACTCTCGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC29	TATATATGTCTATAGAGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC30	TCTCTCTATCGCGCTCGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC31	GATGTCTGAGTGTGTGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC32	GAGACTAGAGATAGTGGAGGCACAAATTAATGAAAAAGAA
BL-1F-BC33	TCTCGTCGAGTCTCTGAGGCACAAATTAATGAAAAAGAA

<b>BL-1F-BC34</b>	ATGTGTATATAGATATGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC35</b>	GCGCGCGCACTCTCTGGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC36</b>	GAGACACGTGCGACACGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC37</b>	ACACATATCGCACTACGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC38</b>	GTGTGTCTCGATGCGCGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC39</b>	CGCACACATAGATACAGAGGCACAAATTAATGAAAAAGAA
<b>BL-1F-BC40</b>	TGTCATATGAGAGTGTGAGGCACAAATTAATGAAAAAGAA
<b>BL-2R-BC1</b>	CGCACTCTGATATGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC2</b>	CTCACAGTCTGTGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC3</b>	CTCTCACGAGATGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC4</b>	CGCGCGTGTGTGCGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC5</b>	CACATAGAGATACAGAGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC6</b>	CGCAGCGCTCGACTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC7</b>	TCTGTCTCGCGTGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC8</b>	CTCTGAGATAGCGCGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC9</b>	TCTGTCTCGCGTGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC10</b>	CTCTGAGATAGCGCGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC11</b>	ATAGATATACGTATAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC12</b>	ACACGCGATCTAGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC13</b>	CTCGCGTATGCGAGAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC14</b>	ACGCGCGCGTAGTGAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC15</b>	ACACACGTGTCATGCGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC16</b>	ATACTATCTCTCTATGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC17</b>	ATATAGCGCGCGTGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC18</b>	CACAGTGAGCACGTGAGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC19</b>	ATCTGATAGAGTGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC20</b>	ACATCGTCGTGTCGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC21</b>	ACATCACTATGTATAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC22</b>	ATATCACACGTGAGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC23</b>	CAGAGATATCTCTCTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC24</b>	CTCTCTGCTCTACATGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC25</b>	GCGCGAGCGTGTGCGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC26</b>	TGTGCGTGTCTCTGTGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC27</b>	TGTGAGAGAGTGTGAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC28</b>	GAGAGTCAGAGCAGAGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC29</b>	TCTATAGACATATATAGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC30</b>	GAGCGCGATAGAGAGAGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC31</b>	CACACACTCAGACATCGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC32</b>	CACTATCTCTAGTCTCGACTTTATTTTTCCAGTTACTTTTT

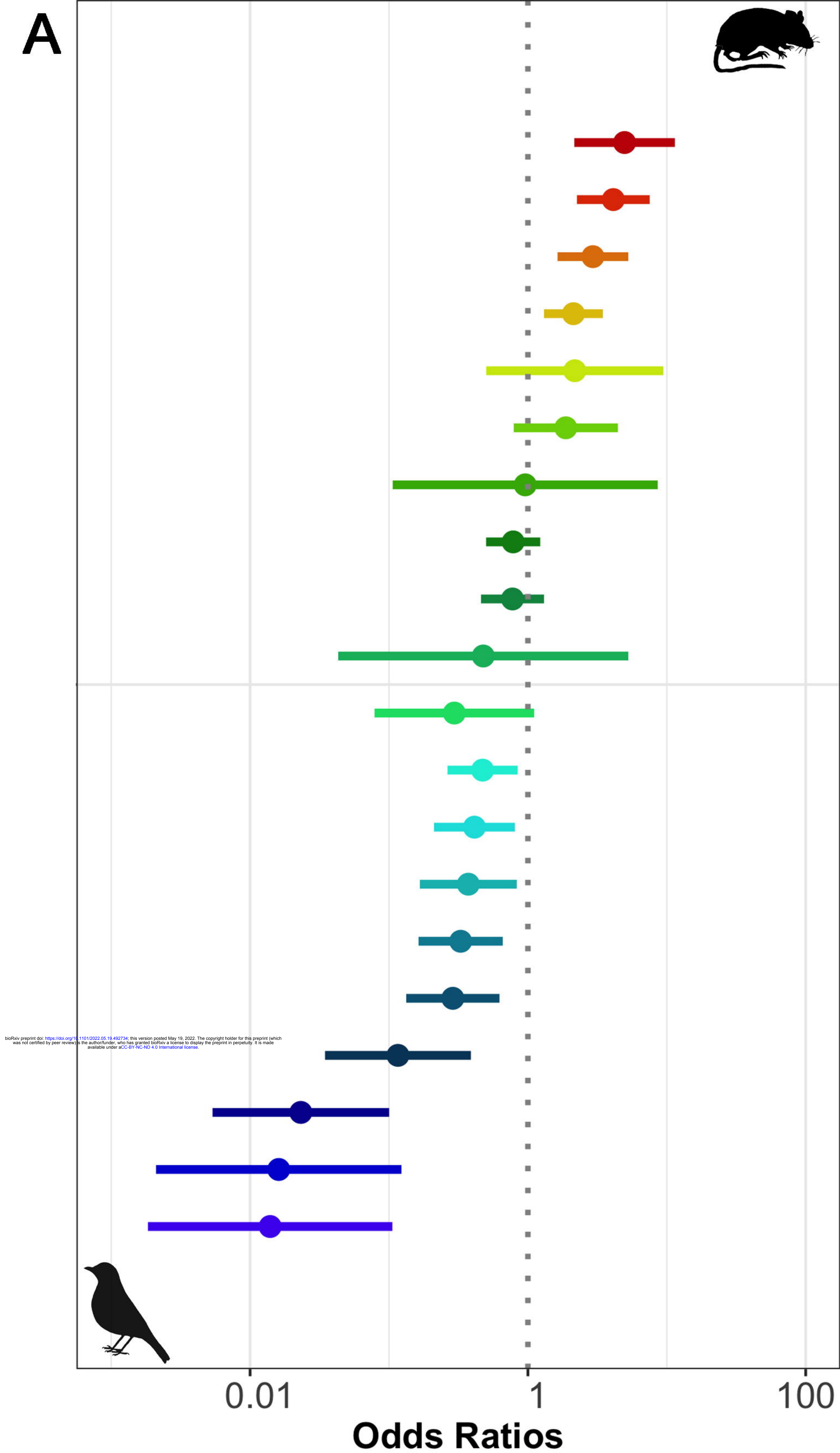
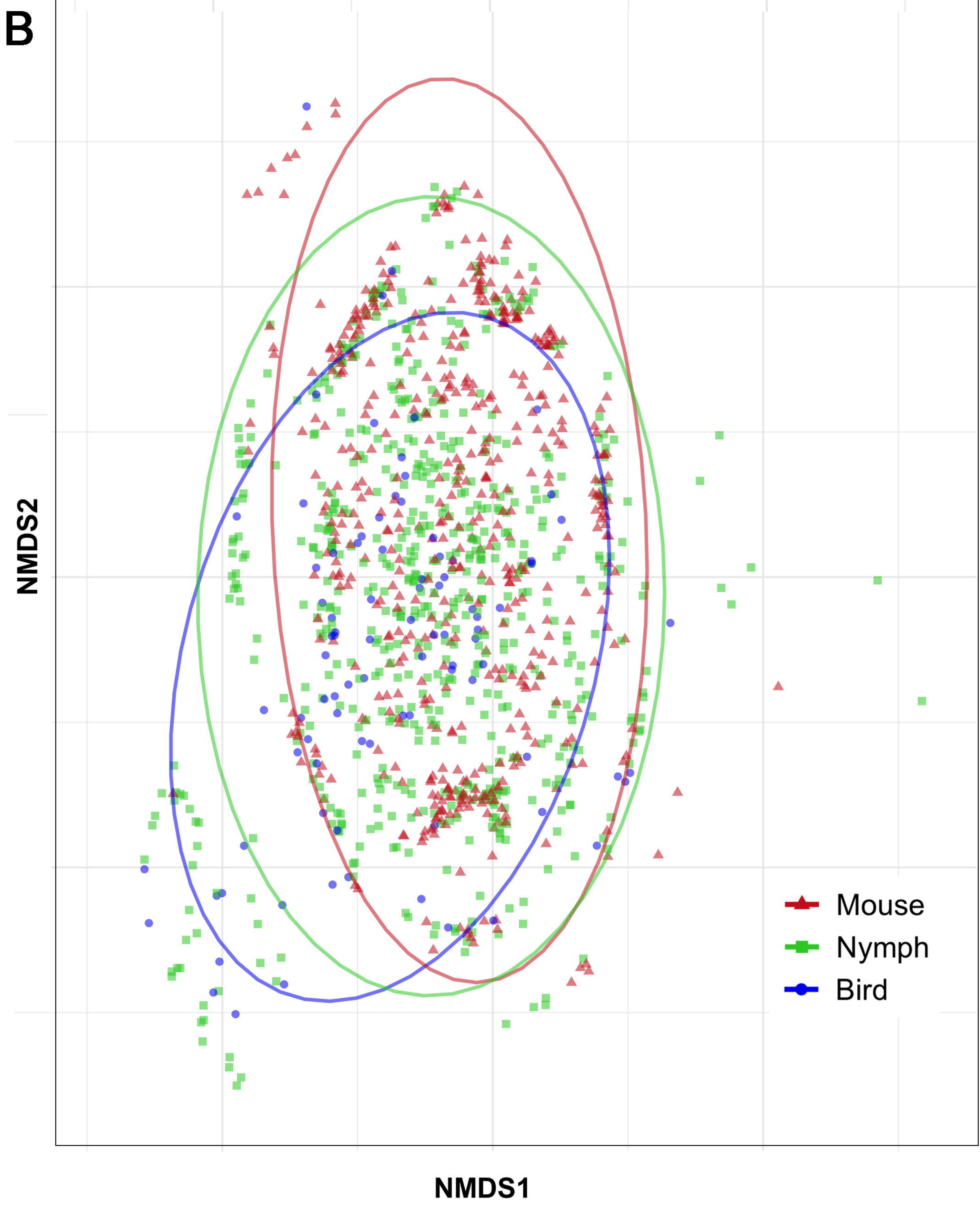
<b>BL-2R-BC33</b>	AGAGACTGCGACGAGAGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC34</b>	ATATCTATATACACATGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC35</b>	CAGAGAGTGCGCGCGCGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC36</b>	GTGTGCGACGTGTCTCGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC37</b>	GTAGTGCGATATGTGTGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC38</b>	GCGCATCGAGACACACGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC39</b>	TGTATCTATGTGTGCGGACTTTATTTTTCCAGTTACTTTTT
<b>BL-2R-BC40</b>	ACACTCTCATATGACAGACTTTATTTTTCCAGTTACTTTTT



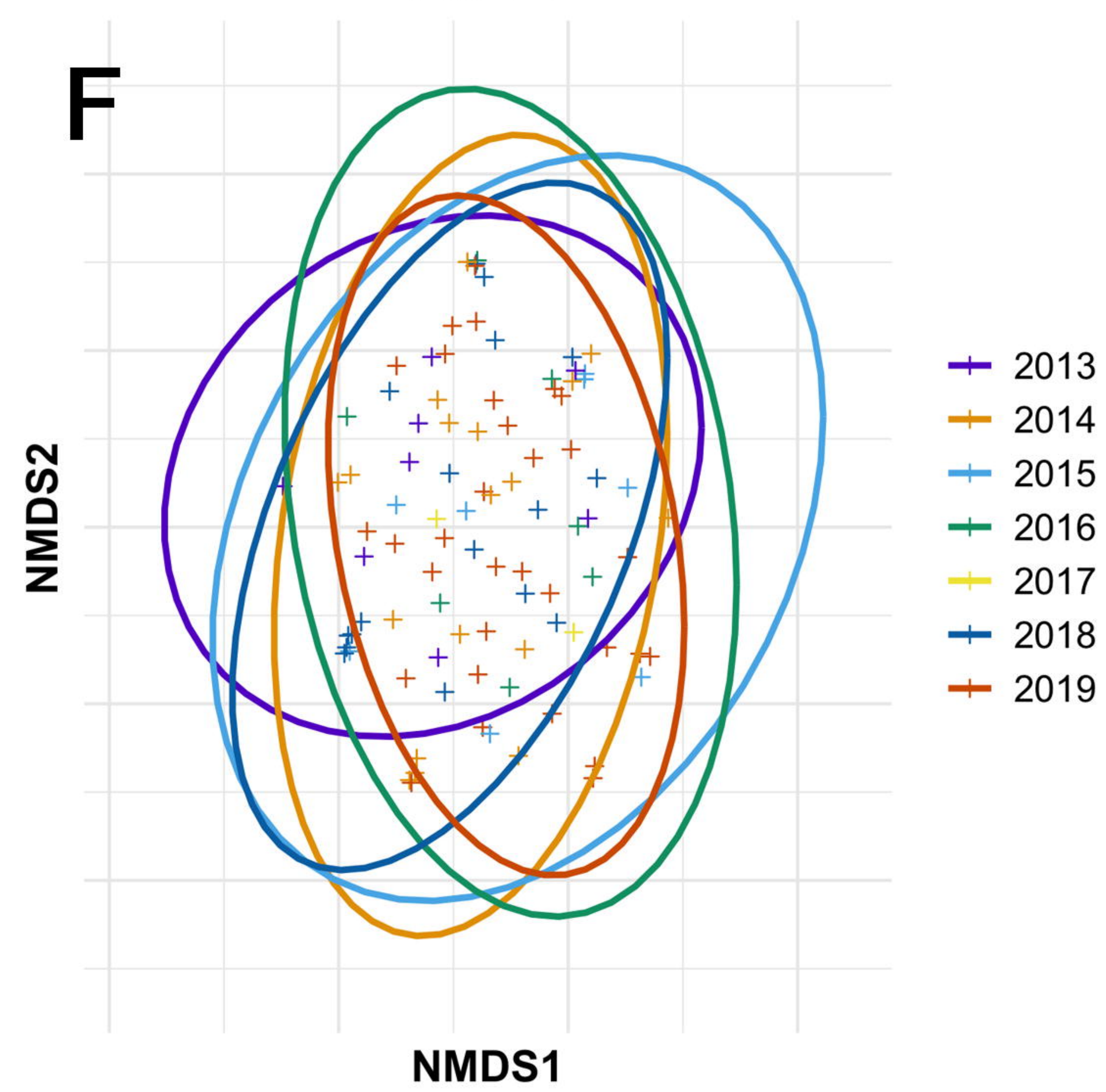
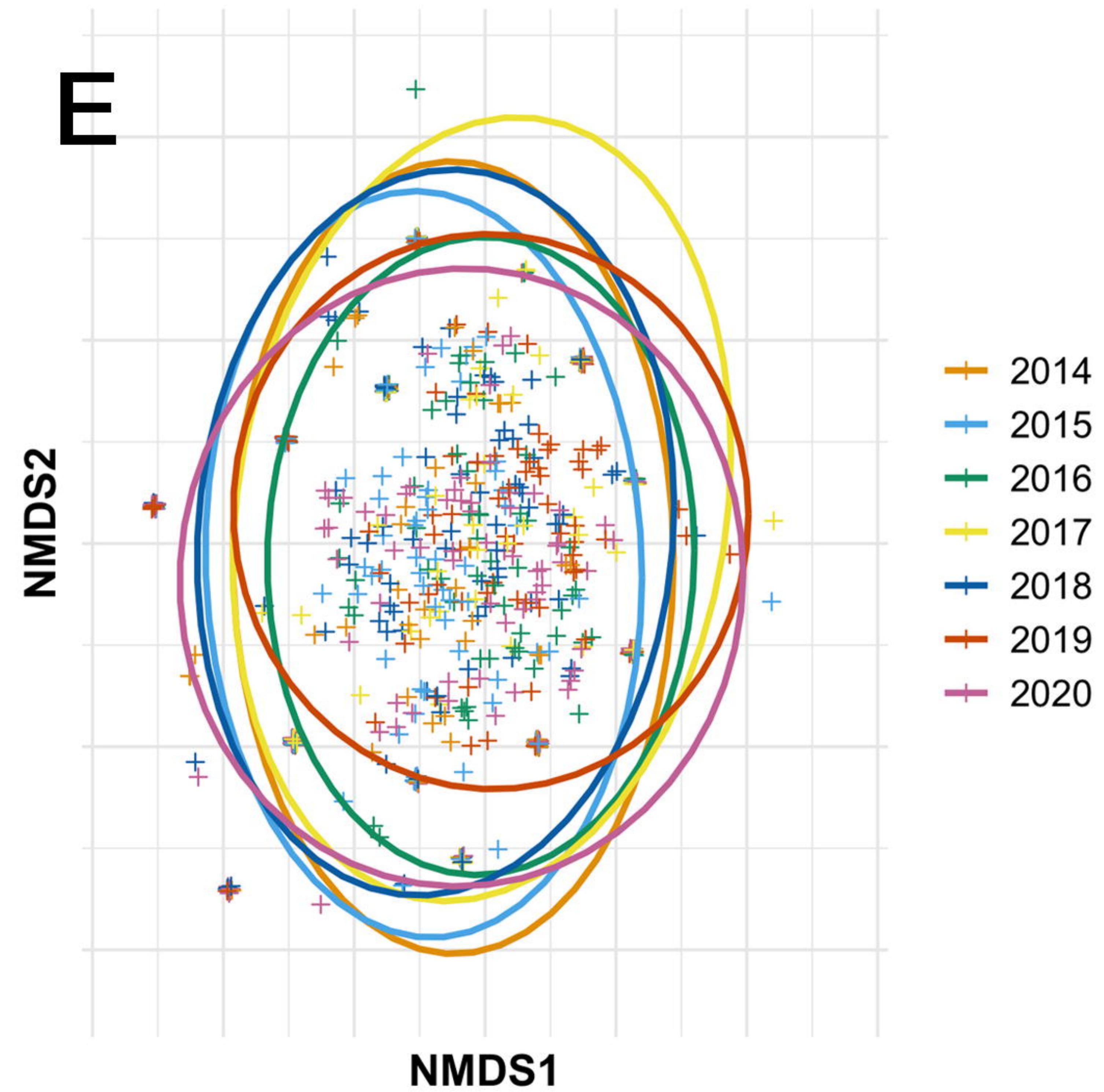
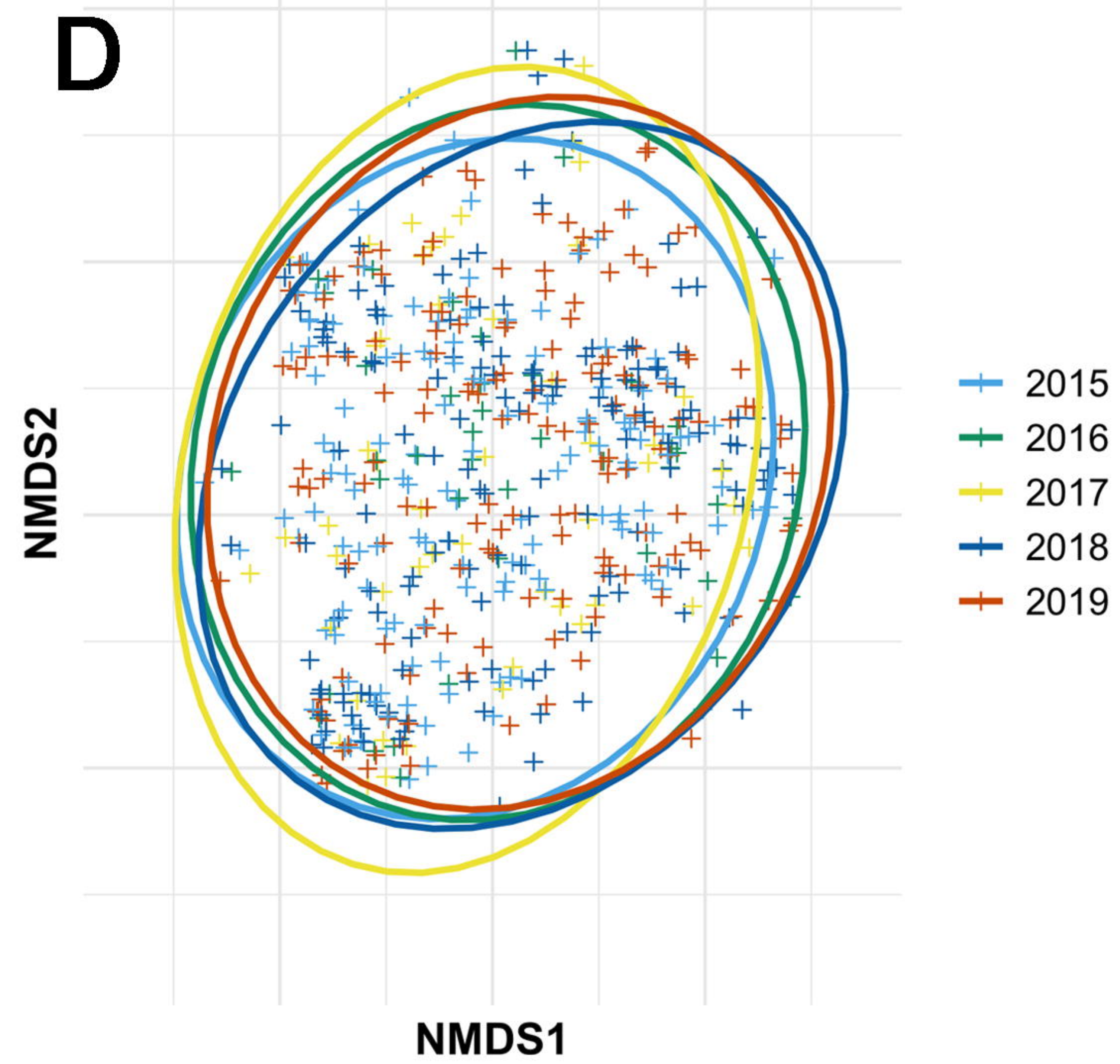
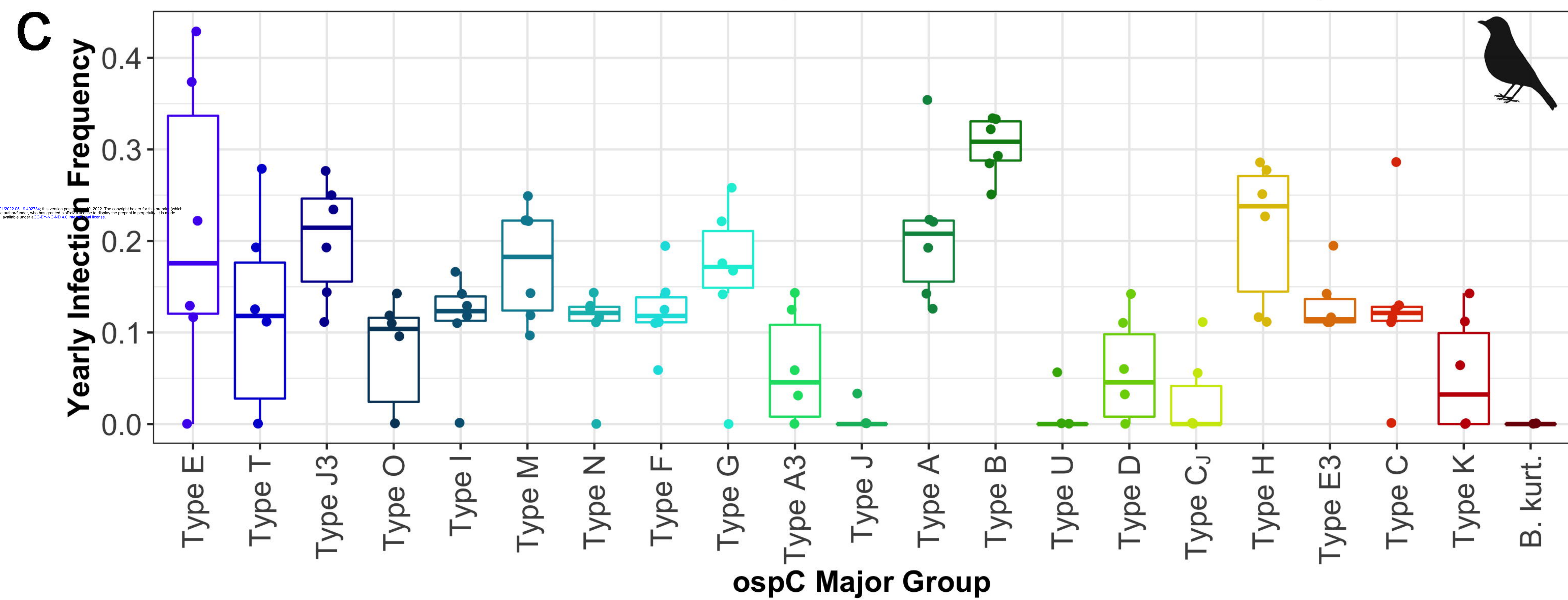
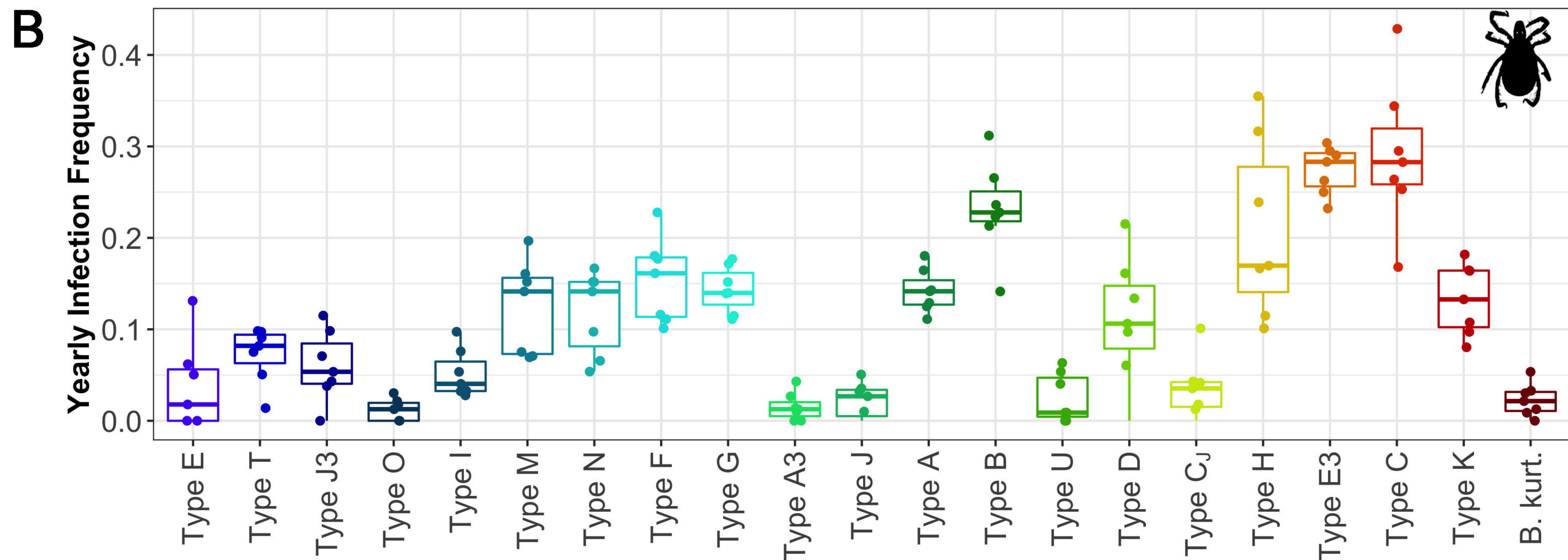
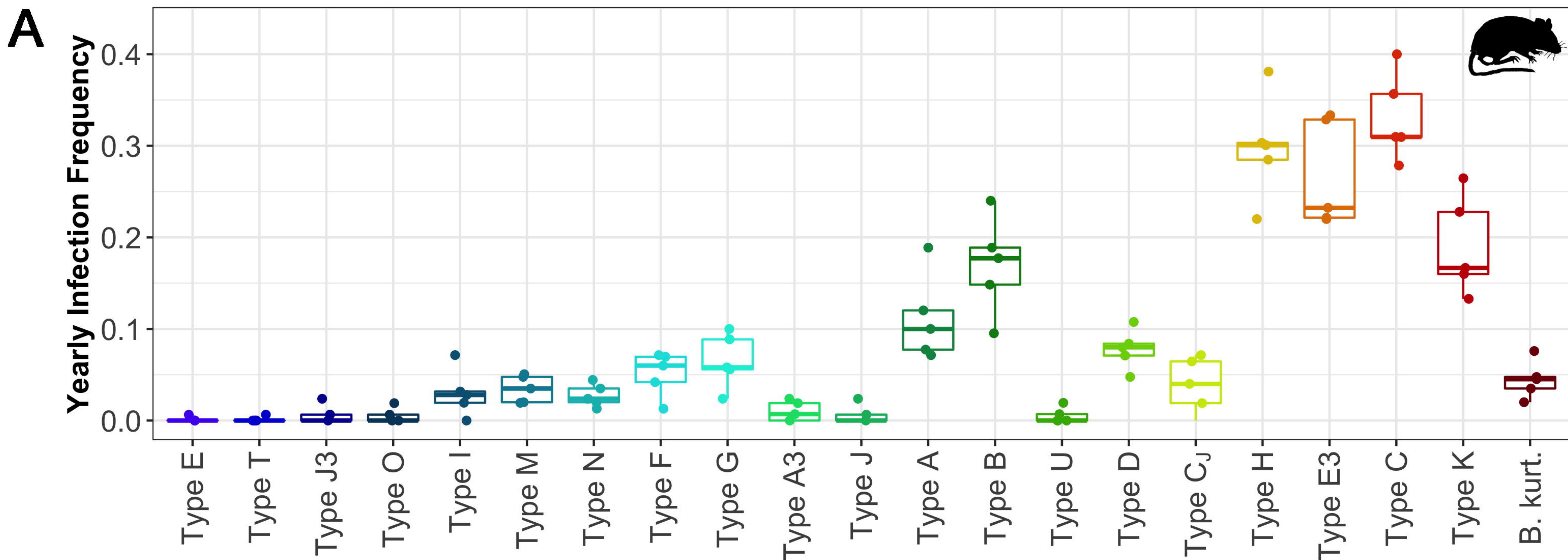
**Table S11** Accession numbers of reference *ospC* sequences used during sequence clustering.

<b><i>ospC</i> type</b>	<b>GenBank accession</b>	<b>Region (as applicable)</b>
<b>A</b>	AE000792	16903 - 17535
<b>B</b>	CP001422	16904 - 17539
<b>C</b>	CP002316	16909 - 17541
<b>D</b>	CP001484	16911 - 17546
<b>E</b>	CP002239	16885 - 17514
<b>F</b>	MH071432	
<b>G</b>	CP001375	16911 - 17543
<b>H</b>	CP001271	16904 - 17536
<b>I</b>	CP001446	16886 - 17518
<b>J</b>	CP001535	16910 - 17545
<b>K</b>	CP002268	16902 - 17537
<b>L</b>	X81524	
<b>M</b>	CP001550	16917 - 17555
<b>N</b>	MH071433	
<b>O</b>	MH071434	
<b>P</b>	U91796	
<b>Q</b>	U91790	
<b>R</b>	U91800	
<b>S</b>	U91793	
<b>T</b>	MH071436	
<b>U</b>	CP001493	16908 - 17549
<b>X</b>	HM047876	
<b>Y</b>	HM047875	
<b>A3</b>	EF592541	
<b>B3</b>	MH071430	
<b>C3</b>	EF592543	
<b>D3</b>	EF592544	
<b>E3</b>	EF592545	
<b>F3</b>	EF592547	
<b>H3</b>	FJ932733	
<b>I3</b>	FJ932734	
<b>C14</b>	MH071431	

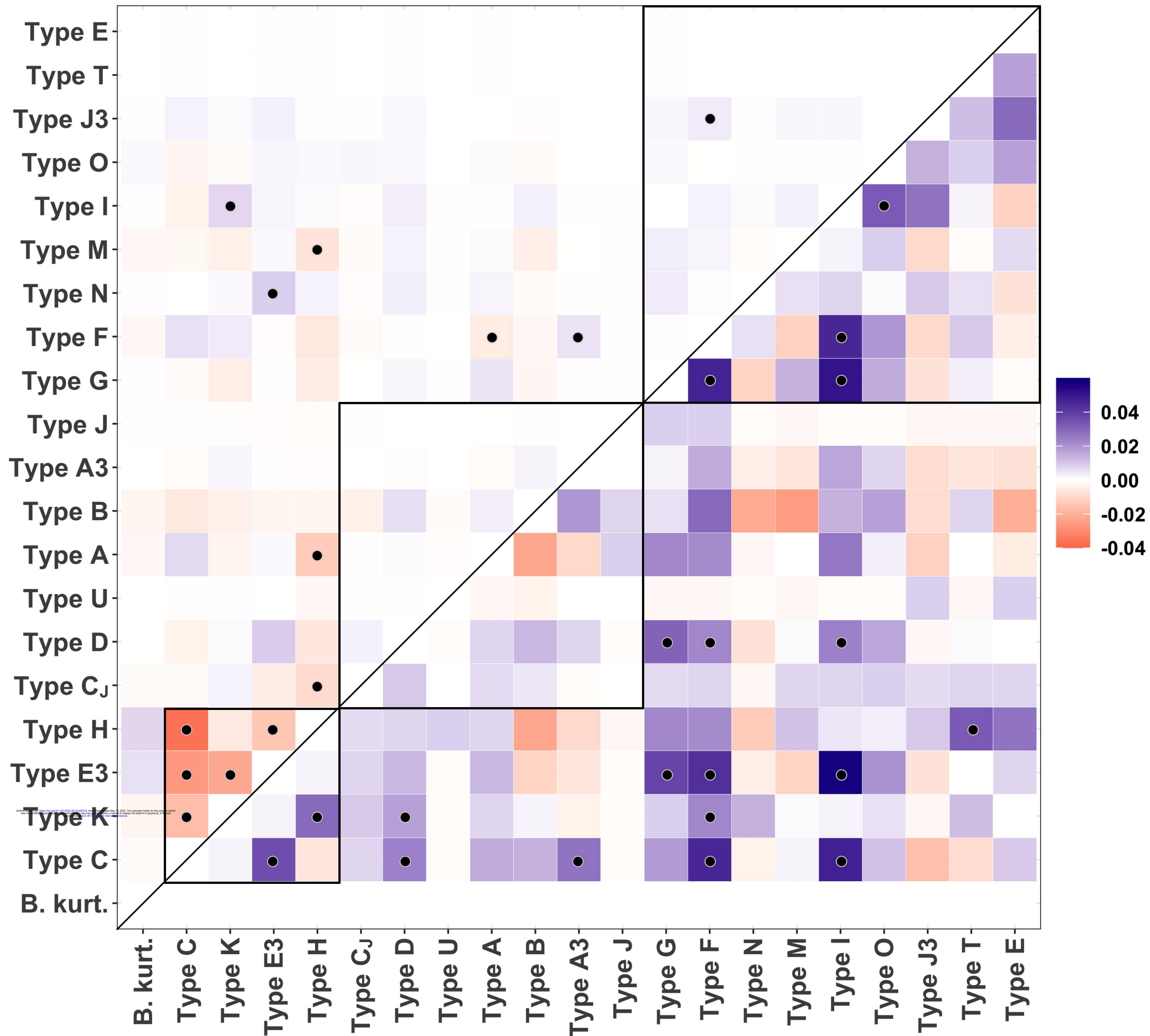


**A****B**

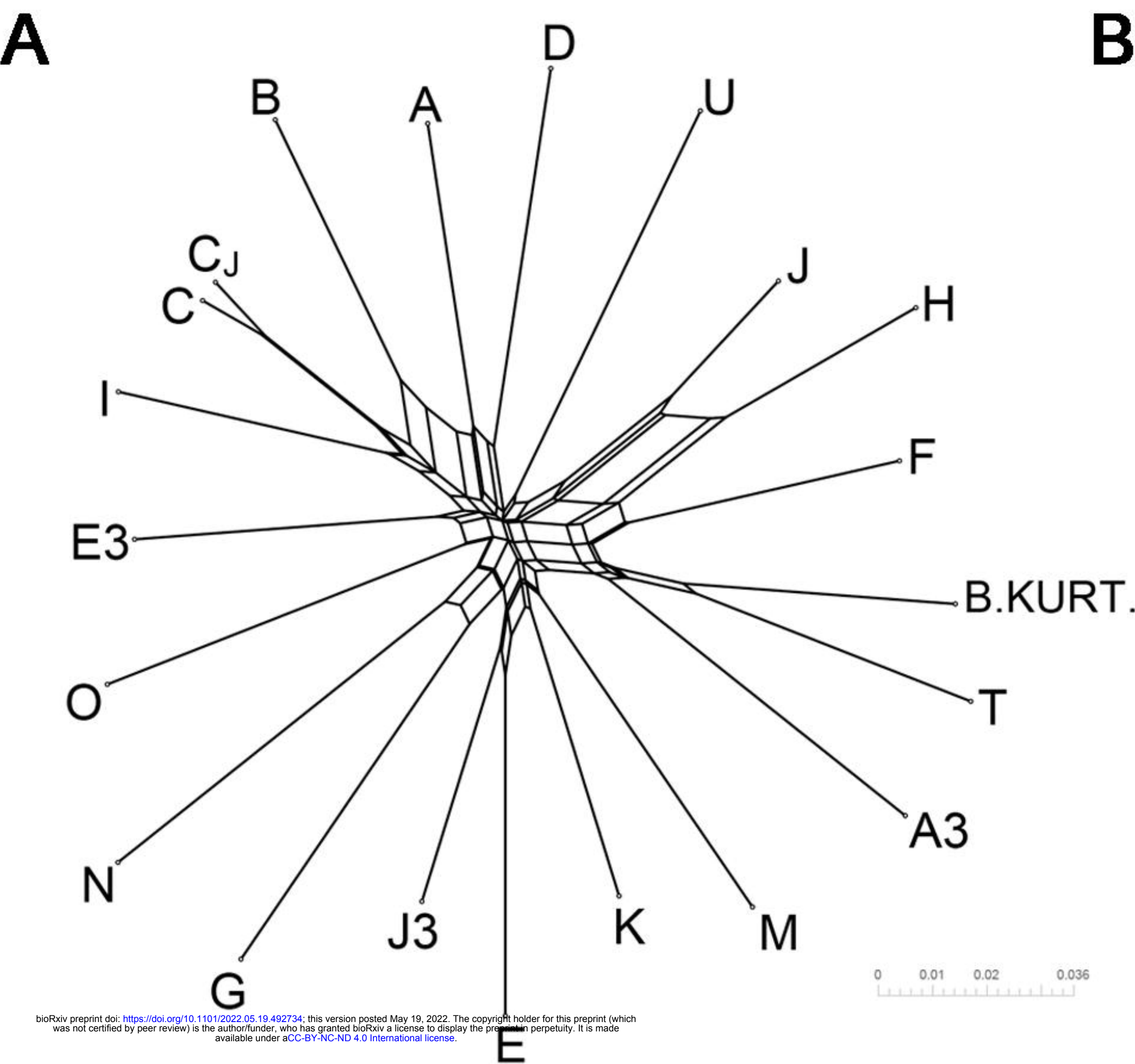












bioRxiv preprint doi: <https://doi.org/10.1101/2022.05.19.492734>; this version posted May 19, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**B**

Type E	A	A	E	L	E	K	L	F	K	A	V	E	N	L	S	K	A	A	Q	D	T	L	K	N	A	V	K	E	L	T	S	P	I	V	A	E	S	P	K	K	P	X
Type T	A	K	E	L	K	D	L	S	E	S	V	E	A	L	A	K	A	A	Q	A	M	L	T	N	S	V	K	E	L	T	S	P	V	V	A	E	T	P	K	K	P	X
Type J3	A	A	E	L	E	K	L	F	E	S	V	K	N	L	S	K	A	A	Q	E	T	L	N	N	S	V	K	E	L	T	S	P	V	V	A	E	N	P	K	K	P	X
Type O	A	E	E	L	V	K	L	A	E	S	V	A	G	L	L	K	V	A	Q	E	M	L	N	N	S	V	K	E	L	T	S	P	V	V	A	E	N	P	K	K	P	X
Type I	A	D	E	L	E	K	L	F	E	S	V	K	N	L	S	K	A	A	K	E	M	L	T	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type M	A	Q	E	L	E	K	L	F	E	S	V	K	N	L	S	K	A	A	Q	E	T	L	N	N	S	V	K	E	L	T	S	P	V	V	A	E	N	P	K	K	P	X
Type N	A	D	E	L	E	K	L	F	K	S	V	E	S	L	A	K	A	A	Q	D	A	L	A	N	S	V	N	E	L	T	S	P	V	V	A	E	T	P	K	K	P	X
Type F	A	K	E	L	K	D	L	S	D	S	V	E	S	L	V	K	A	A	Q	V	M	L	T	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type G	G	K	E	L	K	E	L	S	E	A	V	K	S	L	L	K	A	A	Q	A	A	L	A	N	S	V	Q	E	L	T	S	P	V	V	A	E	T	P	K	K	P	X
Type A3	A	K	E	L	K	D	L	S	E	S	V	A	G	L	L	K	V	A	K	E	M	L	A	N	S	V	K	E	L	T	S	P	V	V	A	E	T	P	K	K	P	X
Type J	A	K	E	L	K	E	L	F	E	A	V	E	S	L	S	K	A	A	K	E	M	L	N	K	S	V	K	E	L	T	S	P	I	V	A	E	S	P	K	N	P	X
Type A	A	E	E	L	G	K	L	F	E	S	V	E	V	L	S	K	A	A	K	E	M	L	A	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type B	V	E	E	L	E	K	L	S	G	S	L	E	S	L	S	K	A	A	K	E	M	L	A	N	S	V	K	E	L	T	S	P	V	V	V	E	S	P	K	K	P	X
Type U	A	E	E	L	D	K	L	F	K	A	V	E	N	L	S	K	A	A	K	E	M	L	A	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type D	A	E	E	L	V	K	L	S	E	S	V	A	G	L	L	K	A	A	Q	A	I	L	A	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type C <sub>J</sub>	A	K	E	L	K	E	L	F	E	S	V	E	N	L	A	K	A	A	K	E	M	L	S	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	N	P	X
Type H	A	K	E	L	K	D	L	S	D	S	V	E	S	L	V	K	A	A	K	E	M	L	T	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type E3	A	D	E	L	E	K	L	F	K	A	V	E	V	L	S	K	A	A	K	E	M	L	T	N	A	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X
Type C	A	A	E	L	E	K	L	F	E	S	V	E	N	L	A	K	A	A	K	E	M	L	S	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	N	P	X
Type K	A	A	E	L	E	K	L	F	K	A	V	E	N	L	A	K	A	A	K	E	T	L	N	N	S	V	K	E	L	T	S	P	I	V	A	E	S	P	K	N	P	X
B. kurt.	A	K	E	L	K	D	L	S	D	S	V	E	S	L	V	K	A	A	K	E	M	L	N	N	S	V	K	E	L	T	S	P	V	V	A	E	S	P	K	K	P	X

