Another tick bites the dust: exploring the association of microbial composition with a broad transmission competence of tick vector species

Tiago F. Mota,[1] Eduardo R. Fukutani,[1] Kelsilandia A. Martins,[2] Vanessa R. Salgado,[3] Bruno B. Andrade,[1] Deborah B. M. Fraga,[1] Artur T. L. Queiroz[1]

AUTHOR AFFILIATIONS See affiliation list on p. 12.

ABSTRACT Ticks harbor and transmit many different pathogens. While some species are competent vectors of a single pathogen, mono-competent vectors (MCV), other species are proven to be competent vectors of several pathogens, pluri-competent vectors (PCV). Such a difference in vector competence might be related to the microbiome. To better comprehend its influence on the vector competence of ticks for one or several pathogens, a data-driven approach using publicly available databases was applied on bacterial 16S rRNA from MCV and PCV tick species. Alpha and beta diversity, co-occurrence networks, and functional profiles were analyzed. A differential analysis was performed to identify bacterial genera associated with PCV ticks. These tick species presented higher richness, and the bacterial composition showed a significant difference between MCV and PCV ticks. The bacteria genera of PCV ticks demonstrated fewer correlations within each other in comparison with MCV ticks. The differential analysis revealed 14 bacteria genera related to PCV tick species, such as *Rickettsia*, *Staphylococcus*, and *Corynebacterium*. Using 24 differently abundant genera, tick samples from another data set could be classified into either PCV or MCV with high accuracy and concordance. Moreover, pathway regulation related to reactive oxygen species detoxification, β-Lactam resistance, and dTDP-L-rhamnose biosynthesis could be participating in the competence of PCV ticks for several tick-borne pathogens. These findings enlighten our understanding of the bacteria community's role on some tick species' broad vector competence.

IMPORTANCE Some tick species are competent to transmit more than one pathogen while other species are, until now, known to be competent to transmit only one single or any pathogen. Such a difference in vector competence for one or more pathogens might be related to the microbiome, and understanding what differentiates these two groups of ticks could help us control several diseases aiming at the bacteria groups that contribute to such a broad vector competence. Using 16S rRNA from tick species that could be classified into these groups, genera such as Rickettsia and Staphylococcus seemed to be associated with such a broad vector competence. Our results highlight differences in tick species when they are divided based on the number of pathogens they are competent to transmit. These findings are the first step into understanding the relationship between one single tick species and the pathogens it transmits.

V̲ector-borne diseases are responsible for more than 700,000 deaths per year, and around 80% of the world's population are at risk of infection (1). Among the

arthropod vectors, ticks and mosquitoes are responsible for the majority of the vector-borne disease transmission (2). Ticks are obligatory hematophagous arthropods that have a broad host spectrum and are spread worldwide, being an important vector linking pathogens to vertebrates. These characteristics assign them as an important public health problem. Several tick species such as *Ixodes scapularis*, *Ixodes ricinus*, *Haemaphysalis longicornis*, and *Rhipicephalus microplus* are known to be vectors to many pathogens from distinct domains, i.e., virus, bacteria, and protozoa (3). Furthermore, they also impair the society's economy and public health (4, 5).

Usually, vector-pathogen studies on ticks focus on single-pathogen infections, despite co-infections cases being a commonly noted occurrence on field ticks (6). Although co-infected with many pathogens, these ticks should not be always considered responsible for their transmission. Vector competence represents the arthropod's ability to acquire a pathogen infection, maintain it with replication, and further transmission to a susceptible host leading to disease development (7). Thus, the incrimination as a vector should be empirically determined to avoid misclassification regarding pathogen transmission (8). Some species, such as *I. scapularis* and *I. ricinus*, were empirically described transmitting bacterial, protozoan, and viral pathogens (9–12). On the other hand, other tick species like *Ixodes holocyclus* and *Hyalomma dromedarii* were incriminated as vectors of many pathogens while having little empirical evidence for most of these pathogens. This classification regarding the number of transmitted pathogens was assumed in studies based on molecular detection, which does not prove vector competence (8). However, until now, empirical experiment-based studies suggested that they are competent in transmitting only one single pathogen (13–16).

The vector competence of a given tick species to transmit multiple pathogens appears to be related to its microbiome. It was demonstrated that pathogen-microbiome interaction could influence the infection establishment and transmission. The study of tick microbiome enlightened researchers on the tick-pathogen relationship's complexity, in which symbiont bacteria play an important role in tick fitness and vector competence (17). This type of discovery with the aid of high-throughput approaches enables the development of new control methods focused on the pathogen-microbiome-vector relationship (18). Moreover, a point of view from tick species that are competent to transmit several pathogens in comparison with those that have a narrower vector competence could bring us insights into control mechanisms to decrease such a broad competence. For instance, the introduction of modified symbiont *Sphingomonas* in tick microbiota has been demonstrated to reduce *Anaplasma phagocytophilum* infection in *I. scapularis* ticks (19). Along these lines, the first group of ticks could be assigned as being pluri-competent vectors (PVC) while the second group of ticks as being mono-competent vectors (MCV).

Thus, the aim of the present work was to explore the microbiome data from tick species identified as PCV in comparison with those that are, up to date, considered as MCV. This exploratory approach can help us investigate what tick microbiome can tell us about such a broad vector competence.

RESULTS

For the literature review on vector competence of ticks that have 16S data deposited in Sequence Read Archive (SRA), 2,651 papers have been screened and 226 empirical vector competence-related studies were fully read, Fig. 1. The majority of these tick species with 16S data in SRA, 32 out of 58, could not be classified into MCV (competent to transmit one or any pathogen) or PCV (competent to transmit more than one pathogen) as only one or no papers have been found that have evaluated its vector competence for any pathogen. Among the 26 species that could be classified based on 226 studies, 17 species had samples with the inclusion criteria for the microbiome analysis herein proposed, i.e., 16S V3-V4 or V4 region, paired-end sequences, 250 to 300 bases, field collected, washed with ethanol, whole body tissue, and free of pathogens. Among the 1,480 samples meeting the inclusion criteria, 352 samples were divided into two data
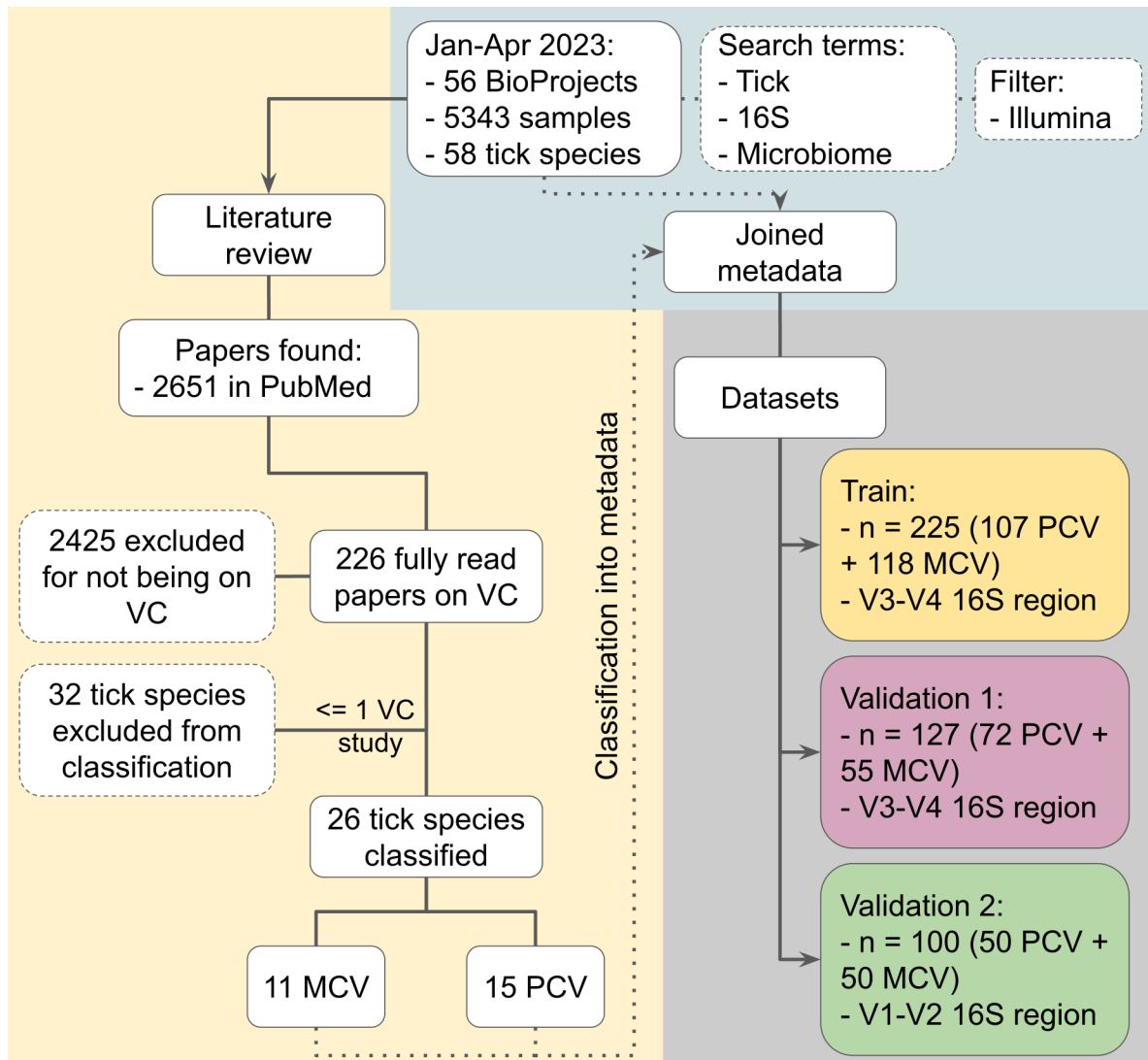
FIG 1 Flowchart of microbiome data search in NCBI's Sequence Read Archive followed by literature in PubMed review on tick species containing 16S rRNA data. The light-blue area represents the search in the SRA database. The light-yellow area represents the literature review and classification into MCV and PCV tick species. The purple area represents the data sets prepared for the microbiome analysis. VC, vector competence; MCV, mono-competent vector; PCV, pluri-competent vector.

sets. This sample number was selected in order to balance the number of samples of each tick species within each group as some species were overrepresented.

The training data set was used to characterize the microbiota and identify differentially abundant bacteria while the validation data set provided samples to assess the accuracy of the bacteria associated with any of the groups to distinguish them. For the training data set, 225 samples of seven species have been chosen, being the MCV group composed of *Dermacentor silvarum*, *Dasymutilla occidentalis*, and *I. holocyclus* while the PCV group of *Amblyomma cajennense*, *Ha. longicornis*, *I. ricinus*, and *Rhipicephalus sanguineus*. The remaining samples, 127, were used for the validation data set being the MCV group composed of *D. occidentalis*, *D. silvarum*, *Haemaphysalis leporispalustris*, *I. holocyclus*, and *Ixodes persulcatus*. The validation data set's PCV group was composed of *Dermacentor marginatus*, *Dermacentor reticulatus*, *Dermacentor variabilis*, *Ha. longicornis*, *Haemaphysalis punctata*, *Ixodes pacificus*, *I. ricinus*, and *R. sanguineus*. A second validation data set, 16S V1-V2 region, was composed of 100 samples from four tick species, being *Haemaphysalis humerosa* and *I. holocyclus* assigned to the MCV group while *I. ricinus* and
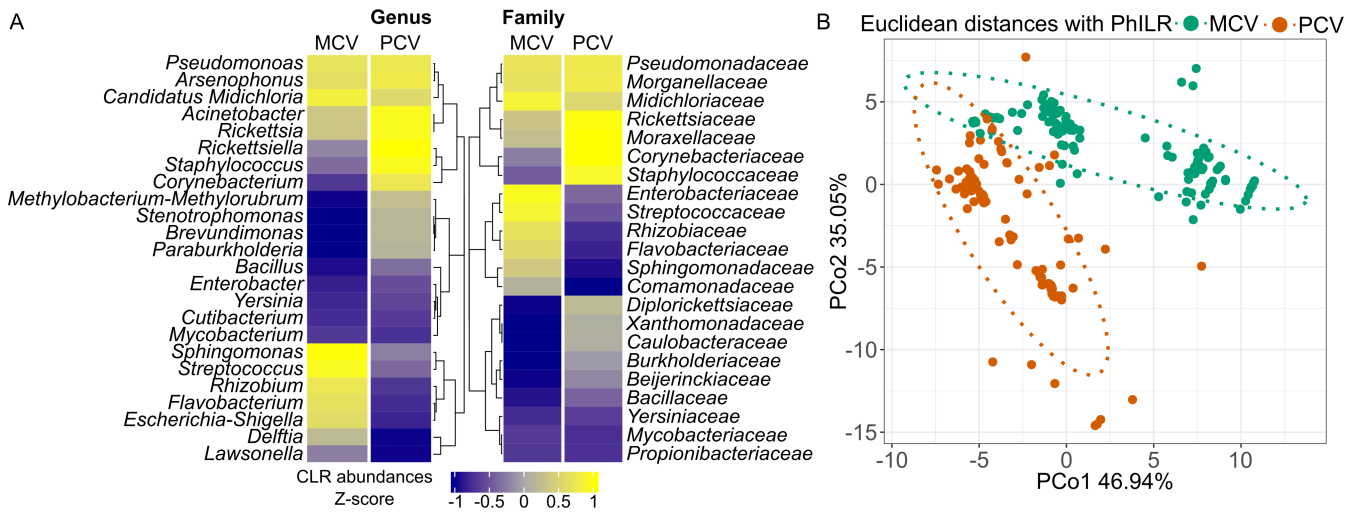
FIG 2 Microbial composition and beta diversity highlighting differences in the bacterial community of mono-competent vectors (MCV) and pluri-competent vectors (PCV). In A, a heatmap using mean CLR transformed abundances by group is represented. Z-score has been applied in order to reduce color scale range. In B, a PCoA from an Euclidean matrix of PhILR transformed counts between groups is represented.

R. sanguineus to the PCV group. For each data set, samples were randomly selected in order to balance the number of samples per tick species within each group. Differences of the sample number among groups were not statistically significant when compared with the other data sets by χ^2 test. The whole classification results and list of tick species and pathogen pairs that have been evaluated can be seen in Table S1.

The bacterial community of ticks from each studied group (MCV and PCV) was composed of similar families and genera, however, with different centered log-ratio (CLR) abundances among the groups (Fig. 2A). For instance, *Rickettsiaceae*, *Moraxellaceae*, *Staphylococcaceae*, and *Corynebacteriaceae* families were more abundant in the PCV group. On the other hand, *Enterobacteriaceae*, *Streptococcaceae*, *Rhizobiaceae*, *Flavobacteriaceae*, *Sphingomonadaceae*, and *Comamonadaceae* were more present in the MCV group when comparing both groups in the heatmap, Fig. 2A. In the beta diversity analysis using Phylogenetic Isometric Log-Ratio Transformation (PhILR) followed by an Euclidean distance matrix, MCV and PCV groups presented different bacterial communities as seen in the principal coordinate analysis (PCoA), Fig. 2B. This separation between groups was statistically significant (PERMANOVA, $F = 93.04$, $R^2 = 0.294$, $P < 0.001$). According to the PERMANOVA test, 29.4% (R^2) of the distance between groups could be explained by their vector competence grouping (MCV and PCV). There was no statistically significant difference in multivariate dispersion in the comparison between groups, $F = 0.18$ and $P > 0.674$. Thus, P-value inflation in the PERMANOVA test should not be expected.

In order to evaluate how bacteria genera correlated with each other within groups, networks using Sparse Correlations for Compositional data (SparCC) have been built for each group separately. The co-occurrence network of PCV tick species was less modular, with lower overall centrality and density as shown in Fig. 3. Due to this result, the alpha diversity has been evaluated using rarefied counts to see if the smaller and less-dense network seen in PCV tick species was related to a lower diversity of bacteria. Contrasting this previous idea, alpha diversity was higher in PCV ticks as measured by both Shannon and inverse Simpson metrics with Mann-Whitney $P < 0.0001$, Fig. 3D.

Among the 408 pathways predicted with PICRUSt2's functional analysis, 197 were associated to any of the groups, MCV or PCV, with $\log_2$ fold change > 1 and BH adjusted $P < 0.05$, according to linear models for differential abundance (LinDA). The pathways predicted to be more active in PCV tick species were related to the biosynthesis of dTDP-L-rhamnose, mycothiol, peptidoglycan, and methyl ketone. On the other hand, the
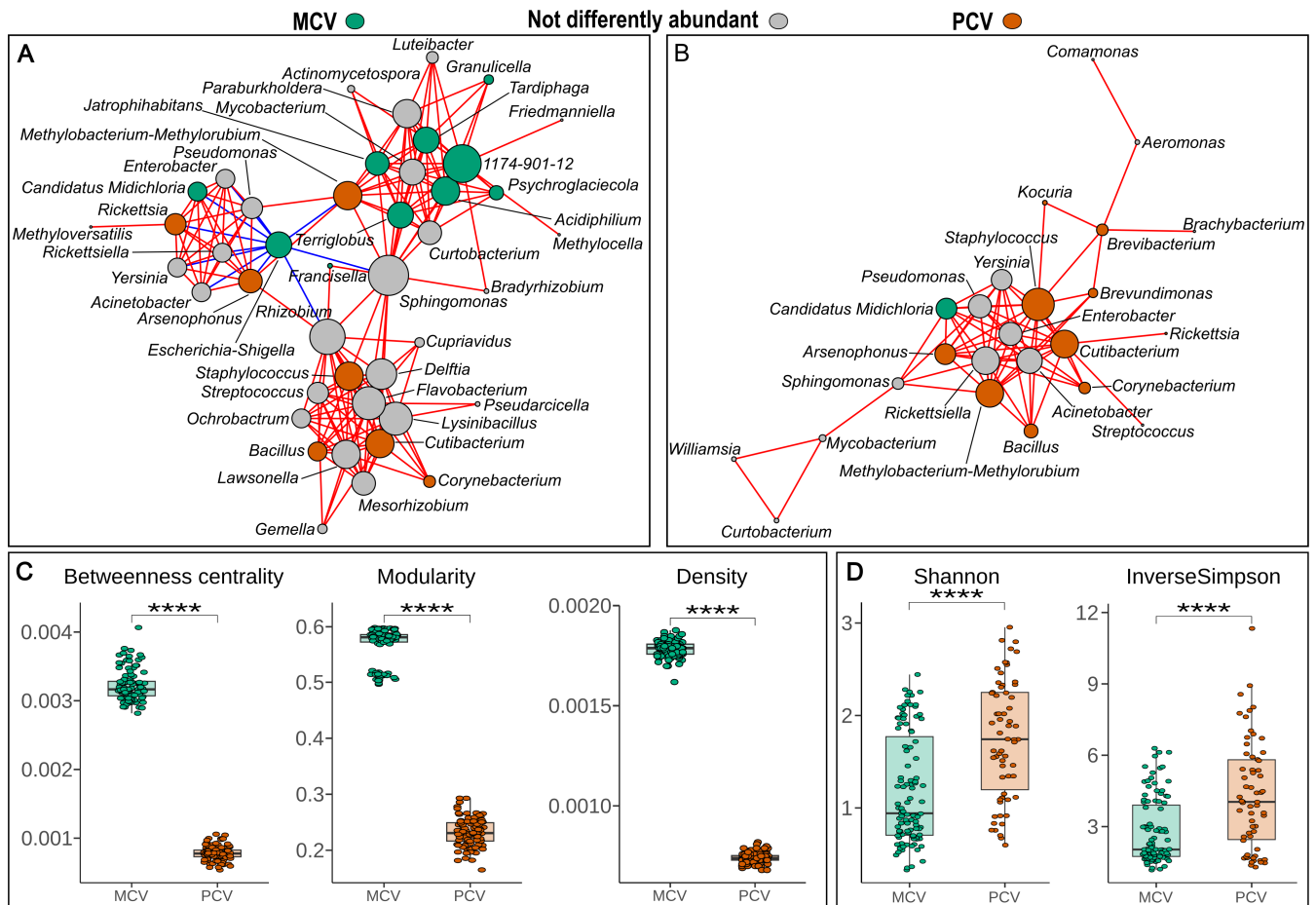
FIG 3 Co-occurrence networks of mono-competent vectors (MCV) (A) and pluri-competent vectors (PCV) (B) tick species. The node diameter was measured by each node degree. The color of nodes is related to the genera that were differently associated with each of the groups in the linear models for differential abundance analysis. Color of edges represents (red) positive and (blue) negative correlations > 0.4 between genera. In C and D, boxplots of network performance metrics and alpha diversity metrics are shown. Comparisons between groups were done using Mann-Whitney test with $P < 0.0001$.

pathways which were predicted to be upregulated in MCV tick species were L-tryptophan biosynthesis, sulfoglycolysis, and degradation of D-galactarate and D-glucarate. A complete table of all differentially regulated pathways can be seen in Table S2.

Several bacteria genera were associated with each of the studied tick groups, presenting LinDA's $\log_2$ fold change > 0.5 and BH adjusted $P < 0.05$ (Fig. 4A). Genera such as *Candidatus Midichloria*, *Escherichia-Shigella*, *Francisella*, and *Acidiphilum* were related to the MCV group. To the same degree, *Rickettsia*, *Staphylococcus*, *Methylobacterium-Methylorubrum*, and *Corynebacterium* were associated with the PCV group. In order to validate LinDA analysis results, a PCoA with PhILR and Euclidean distance matrix was used to assess if separation was better using only the genera associated with each group. As seen in Fig. 4B, the clustering separation between the groups was greater than that observed before the filter. In both situations, there was a statistically significant difference in centroid positions from both groups (PERMANOVA, $P < 0.001$), while there was an increased R^2 value in the abovementioned validation data set from 0.294 to 0.445. In both cases, the homogeneity dispersion of samples within each group showed no statistically significant differences, $F = 0.18$ and $P > 0.674$; $F = 3.60$ and $P = 0.057$, respectively. Using the train data set, i.e., the same data set used for the differential analysis, ticks were classified using a principal component regression followed by a receiver operating curve (ROC) which showed an area under the curve (AUC) of 0.859. When another data set with new samples that sequenced the same 16S variable region,
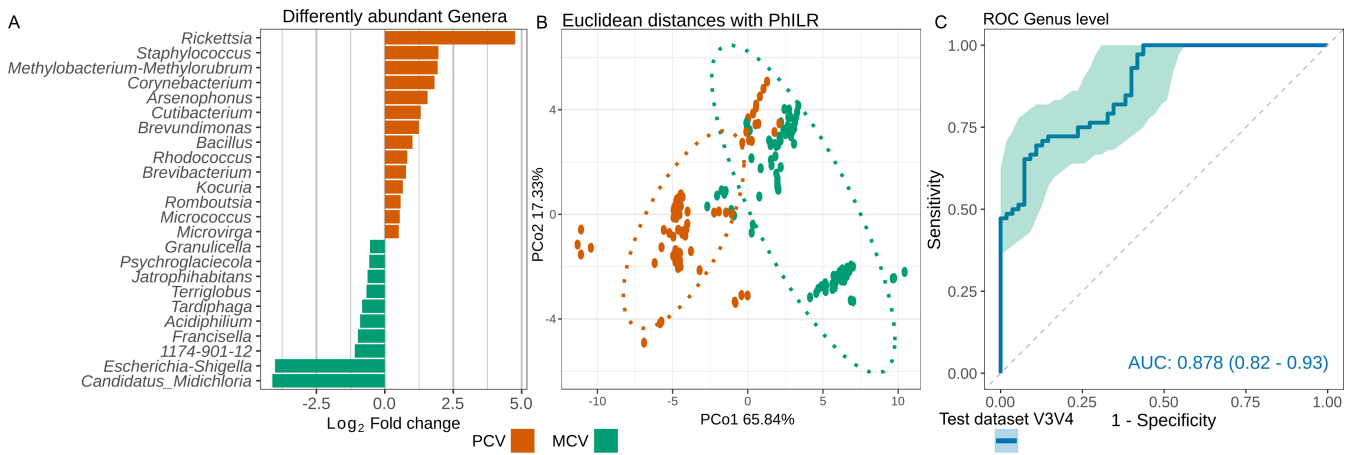
FIG 4 Differential analysis to identify to which group bacteria genera were associated with and validation of using these genera to measure the accuracy of using them to separate samples from another data set into MCV or PCV ticks. In A, LinDA analysis of the microbiota from both mono-competent vectors (MCV) and pluri-competent vectors (PCV) groups. Only taxa with log$_2$ fold change > 0.5 and Benjamini-Hochberg adjusted P < 0.05 are shown. In B, principal coordinate analysis of the beta diversity, with PhILR and Euclidean distance, when only this group of bacteria was kept. In C, receiver operating curves (ROC) and the area under the curve (AUC) in the train and validation (V3-V4) data sets based on principal component regression.

V3-V4, was used, a similar AUC of 0.878 was obtained. Such comparison could not be performed for the second validation data set using samples that have sequenced the 16S V1-V2 region as the 24 differently associated bacteria genera by LinDA analysis were not present in this data set. The model using the 24 genera most associated with MCV or PCV ticks demonstrated a high accuracy of 0.95 (95%CI 0.90–0.98) with a non-informative rate of 0.57 and high concordance, kappa coefficient of 0.90.

DISCUSSION

The tick species herein evaluated were selected based on the presence of laboratory confirmation and its vector competence being classified as MCV and PCV. Only species assessed under experimental conditions were used. This evaluation was performed by a comprehensive literature review of vector competence studies which highlighted a publication bias in such kind of studies. For instance, some tick species such as *I. ricinus*, *I. scapularis*, and *H. longicornis,* have more studies on their vector competence for any pathogen. These species are most likely to be part of the PCV group as more information regarding their competence to transmit a broader range of pathogens is available in the literature. Regardless of such limitations, several tick species with microbiome data deposited in NCBI's SRA could be classified as being MCV or PCV allowing us to further evaluate differences in their microbiota.

The alpha diversity results showed a higher richness in PCV tick species. Such observation contrasts what has been seen under empirical condition where *I. scapularis* co-infected with three pathogens, *Borrelia burgdorferi*, *A. phagocytophilum*, and *Babesia microti*, did not show differences in alpha diversity's Shannon index when compared with pathogen-free ticks (20). Considering our network analysis, the richer PCV group had a network with less correlations among the genera. Such contrast demonstrates that although the microbiota composition of PCV ticks being richer, less correlated microbiota might be important before co-infection for multiple pathogen vector competence, since the samples herein evaluated were free of infection. The microbiota's beta diversity of MCV and PCV ticks differed in terms of phylogenetic distance as shown in the beta diversity evaluation, PERMANOVA P < 0.001, and as multivariate dispersion has not been observed, the P-value tends not be over-inflated. Such a difference in beta diversity suggests unique microbial composition related to competence in transmitting several pathogens which was further evaluated with LinDA.

Considering the bacteria genera with LinDA's $\log_2 > 0.5$ and BH adjusted $P < 0.05$, the *Francisella* genus was associated with the MCV group, suggesting that bacteria belonging to this genus should direct tick species toward transmitting less pathogens. This genus is composed of pathogenic species as well as endosymbionts that have been identified in many tick species (21). In particular, the latter which are detected in all life stages both in colony-reared and in field-collected ticks (21). High prevalence of *Francisella* was associated with decreased *Rickettsia* genus, which impairs tick borne pathogen transmission by outcompeting other bacterial genera (22, 23). In our analysis, no direct correlation was observed between these two genera.

Candidatus Midichloria was also associated with the MCV group. This bacterial genus is closely related to symbionts such as *Midichloria mitochondrii*. This species has been recently demonstrated to be positively associated with *B. burgdorferi* and *Neoehrlichia mikurensis* while being negatively associated with *A. phagocytophilum* (24). However, a method has been used which modeled co-occurrence of bacteria based on the chance of these bacteria being present together with *M. mitochondrii*. As it is a symbiont bacteria present in many ixodid tick species (25), co-occurrence with pathogens does not mean that it plays a role in infection or transmission. In our analysis the genus *Ca. Midichloria* was positively correlated with environmental genera such as *Staphylococcus*, *Enterobacter*, and *Acinetobacter* (26) as well as some symbiont bacteria such as *Arsenophonus sp.* and *Rickettsia sp.* (27). Such correlation with environmental and symbiont bacteria might be indicating a bridge role of *Ca. Midichloria* as environmental bacteria are acquired from habitat and blood meal shaping tick microbiota (26). Some of these genera were also associated in the present work with PCV tick species indicating that *Ca. Midichloria* should be important for both evaluated tick groups as it showed positive correlations in both networks.

Recently, *M. mitochondrii*'s pathways and tissue tropism have been studied showing that it may play a role helping the tick by supplying nutrients, increasing its survival as well as maintaining homeostasis and antioxidant defenses (28). Thus, this symbiont might be beneficial for the survival of many bacteria that favor correlations of genera. Altogether, it should be promoting an environment associated with both mono- and pluri-competence as suggested by the results of the present work. The other bacteria genera associated with the MCV group have not been related to vectorial competence or any kind of influence on microbial composition yet. Nonetheless, as demonstrated by the results of the present work, they might be somehow affecting bacterial composition leaning vectorial competence for several pathogens.

The *Staphylococcus* and *Corynebacterium* genera are known environmental bacteria that can be acquired by ticks and might interact with tick's microbiota (26, 29). These genera were associated with uninfected PCV ticks and might, directly or indirectly, favor the colonization of several pathogens in one single tick. The genus *Bacillus* was also associated with PCV ticks and correlated positively with other bacteria related to this same group such as *Staphylococcus* and *Cutibacterium*. In recent work, *Bacillus* has been negatively associated with ticks infected by the pathogen *R. helvetica* while being positively associated with the genus *Rickettsia* (30). In the present work, such a positive correlation with *Rickettsia* was not observed, at least not directly in the PCV network. However, its positive correlation with other genera also related to PCV ticks indicates that it somehow should be helping to promote a favorable environment for the competence to transmit some pathogens.

Although the genus *Rickettsia* has been associated with PCV ticks, it showed more correlations with other genera in the MCV network demonstrating positive correlations with other bacteria related to both groups. Since pathogen-infected samples were removed from the analysis and most of the species belonging to *Rickettsia* are non-pathogenic (31), we believe that this genus detected in our analysis refers to non-pathogenic symbionts. Rickettsial endosymbionts like *Rickettsia peacockii* and *Rickettsia buchneri* have been strongly related to rickettsial pathogens such as *Rickettsia rickettsii* and *Rickettsia monacensis*, respectively (32, 33). However, little is known about the

influence of these endosymbionts in tick vector competence (34). Based on the results of the present work, these bacteria should be somehow related to the transmission of several pathogens by one single tick species.

Although geographically distinct populations of a given tick species have microbiomes with different composition, the functional profile is conserved and redundant within tick species (35). The results from the functional analysis showed a higher predicted regulation of mycothiol biosynthesis in PCV ticks which is associated with reactive oxygen species (ROS) detoxification as well as protection against electrophilic compounds (36). Such a defensive mechanism in PCV ticks against ROS and electrophilic compounds might be contributing with the broader vector competence of these tick species as it facilitates pathogen replication and transmission (37).

Biosynthesis of dTDP-L-rhamnose is an important pathway for viability or virulence of many bacterial pathogens such as Group B *Streptococcus*, *Enterococcus faecalis*, and *Pseudomonas spp*. (38). This pathway was associated with PCV ticks, and since it is absent in mammals (39), it might be an interesting aim to modulate vector competence of this tick group in order to make it less competent to transmit many pathogens.

It has been shown that *Rickettsia* and *Corynebacterium* genera were related to penicillin degradation pathways (40). Thus, these genera might be related to the higher β-lactam resistance predicted by PICRUSt2 analysis in PCV ticks, i.e., higher peptidoglycan biosynthesis V β-lactam (MetaCyc). Both of these genera have been associated with PCV ticks and together with the higher β-lactam resistance might represent an important mechanism to promote many pathogen development.

Most of the studies evaluating functional differences in the host-pathogen relationship test one single pathogen versus one single host. Our results alongside these one-to-one host-pathogen interactions from previous studies have shown important shared pathways that contribute with higher β-lactam resistance regulation and tick's redox balance mechanisms. Altogether, these pathways might be related to the competence to transmit several pathogens.

After selecting the bacterial genera associated with any of the tested groups, beta diversity results differed in terms of R^2 being about 1.5 times higher after picking taxa with LinDA's $log_2 > 0.5$ and BH adjusted $P < 0.05$. This demonstrates that these bacteria are better able to distinguish MCV and PCV tick groups by PhILR and Euclidean distance. However, this performance should be applicable only to other tick samples that had sequenced the 16S rRNA V3-V4 region as it had high accuracy, 0.95 (95% confidence interval 0.90–0.98) when validated with another set of samples from this same variable region. Validation with V1-V2 data set could not be tested as none of the 24 genera indicated by LinDA analysis were present in this data set. This difference in bacterial composition should be associated with the lower comparability between data sets from different 16S variable regions due to specificities in bacteria classification of each region (41).

Several bacteria groups related to PCV tick species are commonly found in the environment which might be related to the sequenced tick tissue and pre-washing step. Samples herein analyzed had the tick's whole body sequenced and were surface washed with ethanol. Such a cleaning method is known to be the least effective for surface decontamination when compared with other methods, such as 5% sodium hypochlorite (42). However, if samples that had been previously washed with 5% sodium hypochlorite were chosen as selection criteria for the analysis, most of the tick samples would have been Australian ticks which sequenced the V1-V2 16S rRNA region. This criterion would have decreased the influence of environmental bacteria in the analysis; nevertheless, it would have inserted a geographical bias as well as diminished the number of tick species from both groups.

It is noteworthy that the relationship between environmental bacteria and tick microbiota has been proposed affecting its composition (43, 44). Thus, these bacteria may play a role in the tick microbiota indirectly affecting the tick's vectorial competence for many pathogens. Silva's 16S rRNA database presents limitations to accurately identify

bacteria to the species level (45). Therefore, the analysis of the present study was limited to the bacteria family and genera level. Such accuracy level bewilders the separation of species that could be actually part of the environment from those that had already established a co-evolutionary relationship in the tick microbiome and thus play a role in the vector competence of PCV tick species. Although environmental bacteria could not be removed from our analysis, assuming the influence of such bacteria in microbial composition of ticks, changes in the environment could help us modulate the tick microbiome (46). Thus, the bacteria genera herein appointed as associated with the MCV group, such as *Escherichia-Shigella* and *Francisella*, could help us, until a certain point, diminish the vectorial competence for several pathogens of a PCV tick species. As an example, such an approach can be performed as recently proposed by Mazuecos and collaborators (19) where modified symbiont *Sphingomonas* promoted a reduction of *A. phagocytophilum* in *I. scapularis* ticks. Thus, changing the tick's microbiota using the abovementioned bacteria associated with mono-competent vector ticks could be a method to reduce such a broad vector competence. Naturally, due to the abovementioned limitations regarding the samples, predictions, and mechanisms, the results of the present work should be empirically validated in the future for better comprehension of the mechanisms related to PCVs.

Conclusion

Based on our exploratory approach on the microbiome of uninfected PCV versus MCV ticks, the bacteria genera *Rickettsia*, *Staphylococcus*, and *Corynebacterium* and *Arsenophonus* among others are related to a broader vector competence. Additionally, pathways related to ROS detoxification, dTDP-L-rhamnose biosynthesis, and β-lactam resistance regulation could be participating in the infection of several tick-borne pathogens in ticks. These findings could be used in the future to regulate the permissiveness of PCV ticks in order to reduce their vector competence or even block the transmission of several pathogens by one single tick species. As we used and integrated publicly available data, it was not possible to determine the species level for all amplicon sequence variants (ASVs). Due to this limitation, the genus level was used for the major analysis. Deeper insights would be found evaluating each bacteria species on those ticks. Despite these limitations, we were able to provide understanding on the bacteria community role on the vector competence of tick species studied.

MATERIALS AND METHODS

Vector competence literature review

For this analysis, tick species that have 16S rRNA gene sequencing data in the SRA and NCBI's BioProject, between January and April 2023, were selected. Following, a literature review was performed, assessing studies which evaluated the tick species' vector competence for one or more pathogens. Thus, all 58 tick species with 16S rRNA microbiota data deposited in SRA were reviewed for their vector competence for pathogens of medical and veterinary importance. Reviews were carried out for each species through systematic searches, using the species name and terms related to vector competence (i.e., "Vectorial competence" OR "Vector competence" OR "Vector capability") as well as pathogen transmission pathways (vertical OR horizontal OR transstadial OR transovarial AND transmi*). The searches were performed on the NCBI's PubMed. When less than 15 studies were found, only the species name was used for the search and all result titles and abstracts were filtered to find works that evaluated the vector competence of the species in a laboratory setting.

To assess the vector competence of each species, tick species and pathogen pairs were classified according to the degree of evidence presented, which used adaptations of the criteria previously outlined in the literature (8). Thus, the criteria to determine vector competence were experimental demonstration of the following: (i) acquisition of

pathogens when feeding uninfected ticks on experimental infective hosts, (ii) maintenance of pathogens during the seedlings of experimentally infected tick life stages (vertical transmission and/or transstadial passage), and (iii) transmission of pathogens to a susceptible host in a subsequent blood meal.

Each study was evaluated, focused on demonstration of any of these criteria, and a tick/pathogen pair was classified according to the level of evidence presented by the applied method. The criteria were as follows:

- Grade 1 when there was only detection of the pathogen in the tick collected in endemic areas;

- Grade 2 when ticks collected from the field were evaluated in the laboratory for transmission of pathogens to susceptible hosts or to other life stages (vertical or transstadial transmission) as well as laboratory experiments using artificial infections;

- Grade 3 when there was experimental demonstration of at least one of the aforementioned criteria using colony-reared ticks;

- Grade 4 when all vector competence criteria have been demonstrated in the laboratory.

Species were classified as competent to transmit a given pathogen only when there were studies that evaluated vector competence with evidence degrees 3 or 4 and empirically demonstrated such competence. Despite existing works declaring vector competence for certain pathogens, if they were based only on pathogen detection in ticks or on citations of another published work without enough vector competence evidence, the tick species was still classified as an incompetent vector for such pathogens. When tick species presented vector competence for more than one pathogen with a grade 3 or 4 evidence degree, it was classified as being PCV. On the other hand, if it showed vector competence for only one or any pathogen with grade 3 or 4 evidence degree, it was classified as MCV. A table summarizing the findings and classification of the review can be found in Table S1.

Data selection and tick classification

From the data on each tick sample with bacterial 16S-rRNA sequencing deposited at the NCBI, a metadata was assembled containing information such as life stage, pre-sequencing washing method, variable region of the 16S gene, tick collection location, and tick species. A new variable was added regarding the number of pathogens that each tick species is competent to transmit as the aforementioned classification. In order to reduce the heterogeneity between samples and retain the highest number of tick species, only samples from data sets containing field-collected ticks by flagging, ticks' whole body, ticks without previous detection of infection of any pathogen, and ticks with prewash using alcohol or alcohol plus some other reagent in the methodology were included. In addition, the sequencing was filtered for V3-V4 or V4 regions with a paired-end layout and 250 or 300 bases as a higher number of samples was attained in comparison with other variable regions.

Data processing

All samples were downloaded using SRA toolkit's fastq-dump (github.com/ncbi/sra-tools) and quality checked by Trimmomatic v0.32 (47) to remove the sequencing adapter, small and unspecific reads, and those with Phred quality score lower than 30. After performing the data quality control, all forward and reverse reads have been joined and processed on the QIIME2 2022.11 (48) pipeline, removing chimeric sequences and building ASVs using the Deblur algorithm (49). Then, all ASVs have been classified using the SILVA v138.1 16S database (50) assuming 97% identity. Before usage, the

database was filtered to remove sequences that could be associated with fungi, filtering out Eukaryota, plants, removing Mitochondria and Chloroplast, and sequences with unspecified species classification such as *unkown*, *uncultured*, NA, *metagenome*, and *unidentified*. After classification, a phylogenetic tree was built using MAFFT (multiple alignment using fast Fourier transform (51) and fasttree (52) methods. For the analysis, all of these data generated on QIIME2, i.e., ASV table, ASV taxonomic classification, and phylogenetic tree, have been imported in R environment v4.2.3 using Qiime2R package (53). This package creates a phyloseq object that can be further analyzed by the Phyloseq package (54) to measure diversity indices.

Batch effect correction

Before any analysis, a batch correction has been performed to reduce the technical bias variation that could have been caused by batch, e.g., each project being sequenced and performed by different research groups and sequencing machines. Such correction was performed by the ConQuR_libsize function (55), from the ConQuR package v2.0. After batch correction, the corrected ASV table was incorporated in a phyloseq object. For future analysis, samples with less than 100 reads and taxa with less than 10 counts have been removed.

Compositional and diversity analysis

To visualize the microbial composition of MCV and PCV tick species, heatmaps using the batch-corrected mean counts transformed by centered log-ratio were created. CLR transformed counts were scaled using Z-scores in order to reduce the range of values for better visualization of differences between groups. As the alpha diversity metric could not be measured using CLR transformed counts, the rarefied batch-corrected counts have been used for such analysis, rarefying for 1,000 counts and using Shannon and Inverse Simpson indices. Beta diversity was assessed by means of ASV count transformation using PhILR (56), followed by a Euclidean distance matrix. Such an approach takes into account the data's compositionality characteristic together with the phylogenetic distance from taxa present in samples. Principal coordinate analysis was applied to evaluate group distances in terms of beta diversity followed by PERMANOVA test (57), in order to determine statistically significant differences in each group centroid according to the Euclidean distance matrix. The multivariate dispersion has been tested through the vegan's package v2.6.4 betadisper function to assess the homogeneity of each sample distance in relation to its group centroid. If groups are equally homogeneous, then the *P*-value from the PERMANOVA test is not inflated.

Co-occurrence network

Co-occurrence networks have been built using Sparse Correlations for Compositional data between samples from each group, MCV and PCV. In order to compare networks from both groups, 100 bootstrap replicates of SparCC correlation matrices were made to extract network performance metrics such as betweenness centrality, modularity, and density. After collecting these metrics from 100 bootstrapped networks, they were compared between MCV and PCV groups with the Mann-Whitney test. The package ggClusterNet v0.1.0 has been used to measure SparCC correlation matrices, and networks were built using igraph v1.4.1.

Functional analysis

In order to further describe PCV ticks, a functional analysis was performed using PICRUSt2 (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States), predicting metagenome functions that are regulated by a given bacterial community based on 16S-rRNA sequencing (58). Predicted pathways were described according to the MetaCyc database and further hierarchically classified into superclasses as provided by the file2meco R package v0.5.1. Differently abundant pathways were

identified using Linear models for Differential Abundance (59) with a cut-off of 1 $\log_2$ fold change and Benjamini-Hochberg false discovery rate (BH) adjusted $P < 0.05$.

Differential analysis and validation of differentially abundant bacteria

To determine groups of bacteria with higher probability to explain differences among groups (MCV vs PCV), LinDA has been applied as it takes into account the data's compositionality. Differently abundant genera related to MCV or PCV were determined when $\log_2$ fold change was >0.5 with an BH adjusted $P < 0.05$. Tick life stages have been used in the LinDA model to adjust the vector competence variable.

The bacteria identified as differentially abundant between MCV and PCV ticks in LinDA were used in a principal component regression approach to assess if these bacterial genera accurately separate tick species as MCV or PCV. Thus, only bacterial genera with the abovementioned fold change and P-value criteria were used to run a PCA. The first two components were modeled in a logistic regression to determine if samples belonged to any of the studied tick groups. A validation was performed on a data set, composed of other tick samples which sequenced V1-V2 and V3-V4 16S rRNA regions and were unused in the differential analysis as well as samples from other tick species in both groups. The performance of the logistic model built from the train data set was measured using the abovementioned validation data sets containing only the differently abundant bacterial genera. Receiver operating characteristic curves, package pROC v1.18.0, and the confusion matrix method, package caret v6.0.94, have been used to measure the accuracy, non-informative rate, and concordance kappa between train and test data sets.

AUTHOR AFFILIATIONS

[1]Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Bahia, Brazil
[2]Division of Biomedical and Life Sciences, Lancaster University, Lancaster, United Kingdom
[3]Faculdade de Medicina Veterinária da União Metropolitana de Educação e Cultura (UNIME), Lauro de Freitas, Bahia, Brazil

AUTHOR ORCIDs

Tiago F. Mota http://orcid.org/0000-0001-6575-8208

AUTHOR CONTRIBUTIONS

Tiago F. Mota, Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft | Eduardo R. Fukutani, Methodology, Writing – review and editing | Kelsilandia A. Martins, Conceptualization, Data curation, Investigation, Methodology, Writing – review and editing | Vanessa R. Salgado, Data curation, Methodology, Writing – review and editing | Bruno B. Andrade, Writing – review and editing | Deborah B. M. Fraga, Conceptualization, Supervision, Writing – review and editing | Artur T. L. Queiroz, Conceptualization, Project administration, Supervision, Visualization, Writing – review and editing

DATA AVAILABILITY

Raw sequences were obtained from NCBI's SRA under the following BioProjects: PRJEB36903, PRJNA574713, PRJNA577275, PRJNA631062, PRJNA661974, PRJNA664219, PRJNA732915, PRJNA733831, PRJNA766341, PRJNA801881, PRJNA352452, PRJNA401547, PRJNA438789, PRJNA494526, PRJNA523509, PRJNA530927, PRJNA548395, PRJEB46006, PRJNA640465, and PRJNA559059. Metadata and code are available upon request.

ADDITIONAL FILES

The following material is available online.

Supplemental Material

Supplemental tables (Spectrum02156-23-s0001.xlsx). Results from the literature review and differential functional analysis.

REFERENCES

1. WHO. 2017. Global vector control response 2017–2030. World Health Organization, Geneva.
2. Franklinos LHV, Jones KE, Redding DW, Abubakar I. 2019. The effect of global change on mosquito-borne disease. Lancet Infect Dis 19:e302–e312. https://doi.org/10.1016/S1473-3099(19)30161-6
3. Eisen L. 2022. Tick species infesting humans in the United States. Ticks Tick Borne Dis 13:102025. https://doi.org/10.1016/j.ttbdis.2022.102025
4. Singh K, Kumar S, Sharma AK, Jacob SS, RamVerma M, Singh NK, Shakya M, Sankar M, Ghosh S. 2022. Economic impact of predominant ticks and tick-borne diseases on Indian dairy production systems. Exp Parasitol 243:108408. https://doi.org/10.1016/j.exppara.2022.108408
5. Johnson N, Phipps LP, Hansford KM, Folly AJ, Fooks AR, Medlock JM, Mansfield KL. 2022. One health approach to tick and tick-borne disease surveillance in the United Kingdom. Int J Environ Res Public Health 19:5833. https://doi.org/10.3390/ijerph19105833
6. Moutailler S, Valiente Moro C, Vaumourin E, Michelet L, Tran FH, Devillers E, Cosson J-F, Gasqui P, Van VT, Mavingui P, Vourc'h G, Vayssier-Taussat M. 2016. Co-infection of ticks: the rule rather than the exception. PLoS Negl Trop Dis 10:e0004539. https://doi.org/10.1371/journal.pntd.0004539
7. Wu VY, Chen B, Christofferson R, Ebel G, Fagre AC, Gallichotte EN, Sweeny AR, Carlson CJ, Ryan SJ. 2022. A minimum data standard for vector competence experiments. Sci Data 9:634. https://doi.org/10.1038/s41597-022-01741-4
8. Eisen L. 2020. Vector competence studies with hard ticks and borrelia burgdorferi sensu lato spirochetes: A review. Ticks Tick Borne Dis 11:101359. https://doi.org/10.1016/j.ttbdis.2019.101359
9. Hajnická V, Kúdelová M, Štibrániová I, Slovák M, Bartíková P, Halásová Z, Pančík P, Belvončíková P, Vrbová M, Holíková V, Hails RS, Nuttall PA. 2017. Tick-borne transmission of murine gammaherpesvirus 68. Front Cell Infect Microbiol 7:458. https://doi.org/10.3389/fcimb.2017.00458
10. Król N, Militzer N, Stöbe E, Nijhof AM, Pfeffer M, Kempf VAJ, Obiegala A. 2021. Evaluating transmission paths for three different bartonella spp. in ixodes ricinus ticks using artificial feeding. Microorganisms 9. https://doi.org/10.3390/microorganisms9050901
11. Levin ML, Stanley HM, Hartzer K, Snellgrove AN. 2021. Incompetence of the Asian Longhorned tick (Acari: Ixodidae) in transmitting the agent of human granulocytic anaplasmosis in the United States. J Med Entomol 58:1419–1423. https://doi.org/10.1093/jme/tjab015
12. Sharma R, Cozens DW, Armstrong PM, Brackney DE. 2021. Vector competence of human-biting ticks Ixodes scapularis, amblyomma americanum and dermacentor variabilis for powassan virus. Parasit Vectors 14:466. https://doi.org/10.1186/s13071-021-04974-1
13. Smith DJW. 1942. Studies in the epidemiology of Q fever. Aust J Exp Biol Med Sci20:213–217. https://doi.org/10.1038/icb.1942.37
14. Piesman J, Stone BF. 1991. Vector competence of the australian paralysis tick, Ixodes holocyclus, for the lyme disease spirochete borrelia burgdorferi. Int J Parasitol 21:109–111. https://doi.org/10.1016/0020-7519(91)90127-s
15. Soares JF, Soares HS, Barbieri AM, Labruna MB. 2012. Experimental infection of the tick Amblyomma cajennense, cayenne tick, with Rickettsia rickettsii, the agent of rocky mountain spotted fever. Med Vet Entomol 26:139–151. https://doi.org/10.1111/j.1365-2915.2011.00982.x
16. Horta MC, Sabatini GS, Moraes-Filho J, Ogrzewalska M, Canal RB, Pacheco RC, Martins TF, Matushima ER, Labruna MB. 2010. Experimental infection of the opossum didelphis aurita by rickettsia felis, rickettsia bellii, and rickettsia parkeri and evaluation of the transmission of the infection to ticks amblyomma cajennense and amblyomma dubitatum. Vector Borne Zoonotic Dis 10:959–967. https://doi.org/10.1089/vbz.2009.0149
17. Hussain S, Perveen N, Hussain A, Song B, Aziz MU, Zeb J, Li J, George D, Cabezas-Cruz A, Sparagano O. 2022. The symbiotic continuum within ticks: opportunities for disease control. Front Microbiol 13:854803. https://doi.org/10.3389/fmicb.2022.854803
18. de la Fuente J, Antunes S, Bonnet S, Cabezas-Cruz A, Domingos AG, Estrada-Peña A, Johnson N, Kocan KM, Mansfield KL, Nijhof AM, Papa A, Rudenko N, Villar M, Alberdi P, Torina A, Ayllón N, Vancova M, Golovchenko M, Grubhoffer L, Caracappa S, Fooks AR, Gortazar C, Rego ROM. 2017. Tick-pathogen interactions and vector competence: identification of molecular drivers for tick-borne diseases. Front Cell Infect Microbiol 7:114. https://doi.org/10.3389/fcimb.2017.00114
19. Mazuecos L, Alberdi P, Hernández-Jarguín A, Contreras M, Villar M, Cabezas-Cruz A, Simo L, González-García A, Díaz-Sánchez S, Neelakanta G, Bonnet SI, Fikrig E, de la Fuente J. 2023. Frankenbacteriosis targeting interactions between pathogen and symbiont to control infection in the tick vector. iScience 26:106697. https://doi.org/10.1016/j.isci.2023.106697
20. Gil JC, Helal ZH, Risatti G, Hird SM. 2020. Microbiome correlates with life stage, not the presence of human pathogens. PeerJ 8. https://doi.org/10.7717/peerj.10424
21. Kumar D, Sharma SR, Adegoke A, Kennedy A, Tuten HC, Li AY, Karim S. 2022. Recently evolved -Like endosymbiont outcompetes an ancient and evolutionarily associated -Like endosymbiont in the lone Star tick () linked to the alpha-gal syndrome. Front Cell Infect Microbiol 12:787209. https://doi.org/10.3389/fcimb.2022.787209
22. Budachetri K, Browning RE, Adamson SW, Dowd SE, Chao CC, Ching WM, Karim S. 2014. An insight into the Microbiome of the Amblyomma maculatum (Acari: Ixodidae). J Med Entomol 51:119–129. https://doi.org/10.1603/me12223
23. Choubdar N, Karimian F, Koosha M, Oshaghi MA. 2021. An integrated overview of the bacterial flora composition of hyalomma anatolicum, the main vector of cchf. PLoS Negl Trop Dis 15:e0009480. https://doi.org/10.1371/journal.pntd.0009480
24. Krawczyk AI, Röttjers S, Coimbra-Dores MJ, Heylen D, Fonville M, Takken W, Faust K, Sprong H. 2022. Tick microbial associations at the crossroad

of horizontal and vertical transmission pathways. Parasit Vectors 15:380. https://doi.org/10.1186/s13071-022-05519-w

25. Epis S, Sassera D, Beninati T, Lo N, Beati L, Piesman J, Rinaldi L, McCoy KD, Torina A, Sacchi L, Clementi E, Genchi M, Magnino S, Bandi C. 2008. Midichloria Mitochondrii is widespread in hard ticks (Ixodidae) and resides in the mitochondria of phylogenetically diverse species. Parasitology 135:485–494. https://doi.org/10.1017/S0031182007004052

26. Li S-S, Zhang X-Y, Zhou X-J, Chen K-L, Masoudi A, Liu J-Z, Zhang Y-K. 2022. Bacterial microbiota analysis demonstrates that ticks can acquire bacteria from habitat and host blood meal. Exp Appl Acarol 87:81–95. https://doi.org/10.1007/s10493-022-00714-x

27. Duron O, Binetruy F, Noël V, Cremaschi J, McCoy KD, Arnathau C, Plantard O, Goolsby J, Pérez de León AA, Heylen DJA, Van Oosten AR, Gottlieb Y, Baneth G, Guglielmone AA, Estrada-Peña A, Opara MN, Zenner L, Vavre F, Chevillon C. 2017. Evolutionary changes in Symbiont community structure in ticks. Mol Ecol 26:2905–2921. https://doi.org/10.1111/mec.14094

28. Olivieri E, Epis S, Castelli M, Varotto Boccazzi I, Romeo C, Desirò A, Bazzocchi C, Bandi C, Sassera D. 2019. Tissue Tropism and metabolic pathways of Midichloria Mitochondrii suggest tissue-specific functions in the symbiosis with ixodes ricinus. Ticks Tick Borne Dis 10:1070–1077. https://doi.org/10.1016/j.ttbdis.2019.05.019

29. Nhan T-X, Parienti J-J, Badiou G, Leclercq R, Cattoir V. 2012. Microbiological investigation and clinical significance of Corynebacterium spp. In respiratory specimens. Diagn Microbiol Infect Dis 74:236–241. https://doi.org/10.1016/j.diagmicrobio.2012.07.001

30. Maitre A, Wu-Chuang A, Mateos-Hernández L, Foucault-Simonin A, Moutailler S, Paoli J-C, Falchi A, Díaz-Sánchez AA, Banović P, Obregón D, Cabezas-Cruz A. 2022. Rickettsia Helvetica infection is associated with microbiome modulation in Ixodes ricinus collected from humans in serbia. Sci Rep 12:11464. https://doi.org/10.1038/s41598-022-15681-x

31. Bonnet SI, Binetruy F, Hernández-Jarguín AM, Duron O. 2017. The tick microbiome: why non-pathogenic microorganisms matter in tick biology and pathogen transmission. Front Cell Infect Microbiol 7:236. https://doi.org/10.3389/fcimb.2017.00236

32. Simser JA, Palmer AT, Fingerle V, Wilske B, Kurtti TJ, Munderloh UG. 2002. Rickettsia monacensis sp. nov., a spotted fever group Rickettsia, from ticks (Ixodes ricinus) collected in a European city park. Appl Environ Microbiol 68:4559–4566. https://doi.org/10.1128/AEM.68.9.4559-4566.2002

33. Kurtti TJ, Felsheim RF, Burkhardt NY, Oliver JD, Heu CC, Munderloh UG. 2015. Rickettsia buchneri SP. Nov., a rickettsial endosymbiont of the blacklegged tick Ixodes scapularis. Int J Syst Evol Microbiol 65:965–970. https://doi.org/10.1099/ijs.0.000047

34. Hodosi R, Kazimirova M, Soltys K. 2022. What do we know about the microbiome of Front Cell Infect Microbiol 12:990889. https://doi.org/10.3389/fcimb.2022.990889

35. Estrada-Peña A, Cabezas-Cruz A, Obregón D. 2020. Behind taxonomic variability: the functional redundancy in the tick Microbiome. Microorganisms 8:1829. https://doi.org/10.3390/microorganisms8111829

36. Hand CE, Honek JF. 2005. Biological chemistry of naturally occurring thiols of microbial and marine origin. J Nat Prod 68:293–308. https://doi.org/10.1021/np049685x

37. Hernandez EP, Talactac MR, Fujisaki K, Tanaka T. 2019. The case for oxidative stress habitat molecule involvement in the tick-pathogen interactions -An Omics approach. Dev Comp Immunol 100:103409. https://doi.org/10.1016/j.dci.2019.103409

38. van der Beek SL, Zorzoli A, Çanak E, Chapman RN, Lucas K, Meyer BH, Evangelopoulos D, de Carvalho LPS, Boons G-J, Dorfmueller HC, van Sorge NM. 2019. Streptococcal dTDP-L-Rhamnose biosynthesis enzymes: functional characterization and lead compound identification. Mol Microbiol 111:951–964. https://doi.org/10.1111/mmi.14197

39. Adibekian A, Stallforth P, Hecht M-L, Werz DB, Gagneux P, Seeberger PH. 2011. Comparative bioinformatics analysis of the mammalian and bacterial glycomes. Chem Sci 2:337–344. https://doi.org/10.1039/C0SC00322K

40. Chigwada AD, Mapholi NO, Ogola HJO, Mbizeni S, Masebe TM. 2022. Pathogenic and endosymbiotic bacteria and their associated antibiotic resistance biomarkers in and ticks infesting nguni cattle (Spp.). Pathogens 11:432. https://doi.org/10.3390/pathogens11040432

41. Willis C, Desai D, LaRoche J. 2019. Influence of 16S rRNA variable region on perceived diversity of marine microbial communities of the Northern North atlantic. FEMS Microbiol Lett 366:fnz152. https://doi.org/10.1093/femsle/fnz152

42. Hoffmann A, Fingerle V, Noll M. 2020. Analysis of tick surface decontamination methods. Microorganisms 8:987. https://doi.org/10.3390/microorganisms8070987

43. Rynkiewicz EC, Hemmerich C, Rusch DB, Fuqua C, Clay K. 2015. Concordance of bacterial communities of two tick species and blood of their shared rodent host. Mol Ecol 24:2566–2579. https://doi.org/10.1111/mec.13187

44. Zolnik CP, Prill RJ, Falco RC, Daniels TJ, Kolokotronis S-O. 2016. Microbiome changes through ontogeny of a tick pathogen vector. Mol Ecol 25:4963–4977. https://doi.org/10.1111/mec.13832

45. Hoffman C, Siddiqui NY, Fields I, Gregory WT, Simon HM, Mooney MA, Wolfe AJ, Karstens L, Chia N. 2021. Species-level resolution of female bladder microbiota from 16S rRNA amplicon sequencing. mSystems 6. https://doi.org/10.1128/mSystems.00518-21

46. Aivelo T, Norberg A, Tschirren B. 2019. Bacterial microbiota composition of ticks: the role of environmental variation, tick characteristics and microbial interactions. PeerJ 7:e8217. https://doi.org/10.7717/peerj.8217

47. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. https://doi.org/10.1093/bioinformatics/btu170

48. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS 2nd, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, van der Hooft JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol 37:852–857. https://doi.org/10.1038/s41587-019-0252-6

49. Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, Kightley EP, Thompson LR, Hyde ER, Gonzalez A, Knight R, Gilbert JA. 2017. Deblur rapidly resolves single-nucleotide community sequence patterns. mSystems 2:e00191-16. https://doi.org/10.1128/mSystems.00191-16

50. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. Nucleic Acids Res 41:D590–6. https://doi.org/10.1093/nar/gks1219

51. Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30:772–780. https://doi.org/10.1093/molbev/mst010

52. Price MN, Dehal PS, Arkin AP. 2009. Fasttree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol Biol Evol 26:1641–1650. https://doi.org/10.1093/molbev/msp077

53. Jbisanz. Jbisanz/Qiime2R. Available from: https://github.com/jbisanz/qiime2R. Retrieved 18 2021.

54. McMurdie PJ, Holmes S. 2013. Phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 8:e61217. https://doi.org/10.1371/journal.pone.0061217

55. Ling W, Lu J, Zhao N, Lulla A, Plantinga AM, Fu W, Zhang A, Liu H, Song H, Li Z, Chen J, Randolph TW, Koay WLA, White JR, Launer LJ, Fodor AA, Meyer KA, Wu MC. 2022. Batch effects removal for microbiome data via conditional quantile regression. Nat Commun 13:5418. https://doi.org/10.1038/s41467-022-33071-9

56. Silverman JD, Washburne AD, Mukherjee S, David LA. 2017. A phylogenetic transform enhances analysis of compositional microbiota data. Elife 6:e21887. https://doi.org/10.7554/eLife.21887

57. Tang Z-Z, Chen G, Alekseyenko AV. 2016. PERMANOVA-S: association test for microbial community composition that accommodates confounders and multiple distances. Bioinformatics 32:2618–2625. https://doi.org/10.1093/bioinformatics/btw311

58. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille MGI. 2020. Picrust2 for prediction of metagenome functions. Nat Biotechnol 38:685–688. https://doi.org/10.1038/s41587-020-0548-6

59. Zhou H, He K, Chen J, Zhang X. 2022. Linda: Linear models for differential abundance analysis of microbiome compositional data. Genome Biol 23:95. https://doi.org/10.1186/s13059-022-02655-5