



Patterns of Microbiome Variation Among Infrapopulations of Permanent Bloodsucking Parasites

Jorge Doña^{1,2*}, Stephany Virrueta Herrera¹, Tommi Nyman³, Mervi Kunnasranta^{4,5} and Kevin P. Johnson^{1*}

¹ Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, IL, United States, ² Departamento de Biología Animal, Universidad de Granada, Granada, Spain, ³ Department of Ecosystems in the Barents Region, Norwegian Institute of Bioeconomy Research, Svanvik, Norway, ⁴ Department of Environmental and Biological Sciences, University of Eastern Finland, Joensuu, Finland, ⁵ Natural Resources Institute Finland, Joensuu, Finland

OPEN ACCESS

Edited by:

Hirokazu Toju,
Kyoto University, Japan

Reviewed by:

Martin Stoffel,
University of Edinburgh,
United Kingdom
Hideyuki Doi,
University of Hyogo, Japan

*Correspondence:

Jorge Doña
jorged@illinois.edu
Kevin P. Johnson
kpjohnso@illinois.edu

Specialty section:

This article was submitted to
Microbial Symbioses,
a section of the journal
Frontiers in Microbiology

Received: 16 December 2020

Accepted: 26 March 2021

Published: 16 April 2021

Citation:

Doña J, Virrueta Herrera S, Nyman T, Kunnasranta M and Johnson KP (2021) Patterns of Microbiome Variation Among Infrapopulations of Permanent Bloodsucking Parasites. *Front. Microbiol.* 12:642543. doi: 10.3389/fmicb.2021.642543

While interspecific variation in microbiome composition can often be readily explained by factors such as host species identity, there is still limited knowledge of how microbiomes vary at scales lower than the species level (e.g., between individuals or populations). Here, we evaluated variation in microbiome composition of individual parasites among infrapopulations (i.e., populations of parasites of the same species living on a single host individual). To address this question, we used genome-resolved and shotgun metagenomic data of 17 infrapopulations (balanced design) of the permanent, bloodsucking seal louse *Echinophthirius horridus* sampled from individual Saimaa ringed seals *Pusa hispida saimensis*. Both genome-resolved and read-based metagenomic classification approaches consistently show that parasite infrapopulation identity is a significant factor that explains both qualitative and quantitative patterns of microbiome variation at the intraspecific level. This study contributes to the general understanding of the factors driving patterns of intraspecific variation in microbiome composition, especially of bloodsucking parasites, and has implications for understanding how well-known processes occurring at higher taxonomic levels, such as phylosymbiosis, might arise in these systems.

Keywords: genome-resolved metagenomics, host-symbiont, intraspecific variation, lice, microbiota, shotgun metagenomics, symbiont

INTRODUCTION

Patterns of inter- and intraspecific variation in microbiome composition of animals have received much attention because the microbiome may influence many biological processes that have considerable effects on the host (Clemente et al., 2012; Le Chatelier et al., 2013; Rothschild et al., 2018; Rudman et al., 2019; Velazquez et al., 2019). For instance, particular microbiome compositions have been found to drive genomic adaptation (Rudman et al., 2019) or to confer protection against pathogens (Velazquez et al., 2019).

In general, both stochastic (e.g., dispersal, or ecological drift) and deterministic (e.g., host immunological regulation, or microbe–microbe interactions) processes operate across multiple spatial scales to shape the composition of animal microbiomes (Adair and Douglas, 2017;

Kohl, 2020). In particular, among the many determinants shaping microbiome composition, host species identity has been repeatedly identified as a key factor determining the composition of animal microbiomes (Brooks et al., 2016; Mazel et al., 2018; Nishida and Ochman, 2018; Knowles et al., 2019; Lutz et al., 2019; Lim and Bordenstein, 2020; Song et al., 2020). In other words, microbiomes of individuals of the same species tend to be more similar than to those of another species. This pattern is generally the result of filtering microbial taxa by the host (e.g., through host diet, habitat, or immune system, Adair and Douglas, 2017) or result from host–microbe coevolution (Lim and Bordenstein, 2020). When this process exhibits phylogenetic signal, the pattern is known as phylosymbiosis (i.e., microbial community relationships that recapitulate the phylogeny of their host, Brucker and Bordenstein, 2013; Brooks et al., 2016; Lim and Bordenstein, 2020). Nonetheless, several aspects of the variation of animal microbiomes are yet to be better understood (Lim and Bordenstein, 2020). In particular, for non-human animals, there is still much to learn about how microbiomes vary at scales below the species level, such as between populations (Blekhman et al., 2015; Kohl et al., 2018; Rothschild et al., 2018; Campbell et al., 2020; Fountain-Jones et al., 2020) or ecotypes (Agany et al., 2020).

An area of focus on understanding intraspecific variation in microbiome composition has been bloodsucking parasites. In these parasites, previous studies have consistently found a major role of the host species in shaping microbiome composition in the parasites (Osei-Poku et al., 2012; Zhang et al., 2014; Swee and Kwan, 2017; Zolnik et al., 2018; Díaz-Sánchez et al., 2019; Landesman et al., 2019; Lee et al., 2019; Muturi et al., 2019). However, in ticks (*Ixodes scapularis*), host individual identity of the blood meal was even more important than host species identity in explaining microbiome composition (Landesman et al., 2019). These results suggested that individual host identity of the blood meal might be an important factor that shapes parasite microbiomes at the intraspecific level (Landesman et al., 2019). In theory, microbiomes of individual bloodsucking parasites could vary due to: (1) the individual parasite immune system that may impose selection on different bacterial taxa (Blekhman et al., 2015; Suzuki et al., 2019); (2) differences in the source of the blood meal that may transfer or disperse particular bacterial taxa, or modulate bacteria by creating specific conditions during digestion (Rothschild et al., 2018); (3) microbe–microbe interactions (Hassani et al., 2018); and (4) stochastic processes (e.g., ecological drift) (Lankau et al., 2012). However, for most species, and for bloodsucking parasites in particular, the nature of intraspecific variation in microbiomes and the relative importance of factors shaping this variation remain understudied.

Sucking lice (Phthiraptera: Anoplura) are permanent blood-feeding ectoparasites that live in the fur or hairs of mammals. Anopluran lice have been shown to host intracellular bacterial endosymbionts that are likely to help to complement deficiencies in their diet, and these symbionts tend to be located on specialized structures known as mycetomes (Buchner, 1965; Boyd and Reed, 2012; Sasaki-Fukatsu et al., 2006; Perotti et al., 2007, 2008). Previous studies have found that members of Anoplura host a

single endosymbiont, but belonging to different bacterial genera depending on louse species, including *Riesia* (Sasaki-Fukatsu et al., 2006; Kirkness et al., 2010; Boyd et al., 2014), *Sodalis* (Boyd et al., 2016), and *Legionella* (Říhová et al., 2017). We also know from these studies that, as in other arthropod endosymbionts, louse endosymbionts tend to have reduced genomes (Kirkness et al., 2010; Boyd et al., 2017). On the other hand, processes such as replacement (i.e., the substitution of one endosymbiotic species by another) and independent acquisitions of different endosymbionts can occur across evolutionary time scales (i.e., millions of years) (Sasaki-Fukatsu et al., 2006; Allen et al., 2007, 2016; Hypša and Koižek, 2007; Fukatsu et al., 2009). Thus, while there is some background knowledge on louse endosymbionts, several aspects are yet to be understood. For instance, how microbiomes vary across organs, systems, or individuals of lice from the same species is mostly unknown, with a single study to date providing bacterial community data for different individual lice of the same species (Říhová et al., 2019).

The sucking lice of pinnipeds (seals, sea lions, and walrus) are of particular interest because of their need to adapt to the aquatic lifestyle of their hosts (Durden and Musser, 1994; Leonardi et al., 2013). There is evidence that the sucking lice of seals and sea lions have codiversified with their hosts (Kim, 1971, 1975, 1985; Leonardi et al., 2019). Indeed, the sucking lice of pinnipeds represent an interesting system in which to study the variation in microbiome composition and the drivers of this variation at an intraspecific level because: (1) these lice have well defined, isolated populations (infrapopulations) on individual seal hosts, due to an expected low rate of horizontal dispersal among host individuals, which is only possible during the seals' haul-out periods on land or ice (Kim, 1985; Leonardi et al., 2013, 2019); and (2) these lice feed only upon the blood of their host (Snodgrass, 1944; Kim, 1985), so that it can be assumed that individuals from the same infrapopulation feed upon “exactly” the same resource (i.e., the blood of the individual seal on which they occur). In addition, previous studies conducted on seal microbiomes have found that while factors such as species identity, age, sex, and diet play a role in shaping seal microbial communities, seals show evidence of a core microbiome with which they have co-evolved (Nelson et al., 2013; Acquarone et al., 2020; Kim et al., 2020; Stoffel et al., 2020).

Here, we used genome-resolved approaches (the construction of draft microbial genomes from short-read shotgun sequencing data; Bowers et al., 2017; Uritskiy et al., 2018) and metagenomic classification tools (taxonomic classification of individual sequencing reads; Menzel et al., 2016) to infer patterns of microbiome variation among individuals of the sucking seal louse *Echinophthirius horridus* (von Olfers, 1816) inhabiting individual Saimaa ringed seals *Pusa hispida saimensis* (Nordquist, 1899). These two approaches have different limitations and strengths. For example, the genome-resolved approach allows the assembly of multiple highly complete bacterial genomes, but only for organisms with enough coverage to be assembled and binned. On the other hand, metagenomic classification of reads may offer a more comprehensive picture of community composition because of higher database completeness or less strict thresholds to analyze data. However, read classification

is limited by the fact that it is based only on the fraction of reads that map to reference databases (Quince et al., 2017). Our sampling design, involving analysis of two individual lice from each of 17 seals, allowed us to evaluate the degree to which variation in microbiome composition among individual lice is explained by the infrapopulation (the identity of the seal host).

MATERIALS AND METHODS

Sampling, DNA Extraction, and Sequencing

Thirty-four individual lice were sampled from 17 individual Saimaa ringed seals (*Pusa hispida saimensis*), which is an endemic endangered landlocked subspecies of the ringed seal living in freshwater Lake Saimaa in Finland (e.g., Nyman et al., 2014). Individual lice were collected from seals found dead or from seals that were live-captured for telemetry studies (e.g., Niemi et al., 2019), and placed in 2-ml screw-cap tubes with 99.5% ethanol. Lice from a single seal individual were put in the same tube. Prior to DNA extraction, each louse individual was rinsed with 95% ethanol and placed alone in a new sterile vial; then, the remaining ethanol was evaporated at room temperature.

Whole lice were ground up individually, and genomic DNA was extracted using the Qiagen QIAamp DNA Micro Kit (Qiagen, Valencia, CA, United States). The standard protocol was modified so that specimens were incubated in ATL buffer and proteinase K at 55 (insert degree) C for 48 h instead of the recommended 1–3 h, as well as by substituting buffer AE with buffer EB (elution buffer). This was done to ensure maximal yield (greater than 5 ng) of DNA from each louse. Each DNA extract was quantified with a Qubit 2.0 Fluorometer (Invitrogen, Carlsbad, CA, United States) following the manufacturer's recommended protocols.

Shotgun genomic libraries were prepared from the extracts with Hyper Library construction kits (Kapa Biosystems, Wilmington, MA, United States). The libraries were quantitated by quantitative PCR (qPCR) on a Bio-Rad CFX Connect Real-Time System (Bio-Rad Laboratories, Inc., CA, United States) and 150 bp pair-end sequenced on either one of two lanes (**Supplementary Table 1**) of an Illumina NovaSeq 6000 sequencer (Albany, New York, United States). FASTQ files from sequence data were generated and demultiplexed with bcl2fastq v.2.20. All library preparations, sequencing, and FASTQ file generation were carried out at the Roy J. Carver Biotechnology Center (University of Illinois, Urbana, IL, United States). Raw reads were subsequently deposited to the NCBI GenBank SRA database (**Supplementary Table 1**).

Metagenomic Analyses

For the genome-resolved metagenomic analyses, we used the metaWRAP v1.1.5 pipeline (Uritskiy et al., 2018) along with all the recommended databases (i.e., Checkm_DB, NCBI_nt, and NCBI_tax). We used the metaWRAP Read_qc module with default parameters to quality trim the reads and to decontaminate each sample from host reads. For decontamination, we ran a *de novo* genome assembly of an individual louse of the same species, not included in this study, and with a high

amount of sequencing data (“Echor52”) in Abyss v2.0.1 (Jackman et al., 2017). Finally, we filtered out all non-bacterial reads from the contig file using Blobtools v1.0.1 (Laetsch and Blaxter, 2017) and used this file to decontaminate all the other samples with the metaWRAP Read_qc module. See **Supplementary Table 1** and Data Availability section for more details on the data preprocessing results. Next, we co-assembled reads from all the samples with the metaWRAP Assembly module (–use metaspades option) (Nurk et al., 2017). For this assembly, and because of memory limitations, we used BBNorm¹ before assembly to reduce the coverage of the concatenated FASTQ file to a maximum of 100× and to discard reads with coverage under 3×. We binned reads with the metaWRAP Binning module (–maxbin2 –concoct –metabat2 options) (Alneberg et al., 2014; Wu et al., 2016; Kang et al., 2019) and then consolidated the resulting bins into a final bin set with both metaWRAP's Bin_refinement module (–c 50 –x 10 options) and the Reassemble_bins module. We quantified the bins resulting from the Bin_refinement module with Salmon (Patro et al., 2017) using the Quant_bins module with default parameters. Finally, we classified bins using the Classify_bins module. This module uses Taxator-tk, which gives highly accurate but conservative classifications (Dröge et al., 2015). Accordingly, we also uploaded our final metagenome-assembled genomes (MAGs) to MiGA for a complementary analysis to determine the most likely taxonomic classification and novelty rank of the bin (Rodriguez-R et al., 2018). We used the NCBI Genome database (Prokaryotes; February 26, 2020 version) for this analysis.

For the metagenomic classification of reads, we used the metagenomic classifier Kaiju (Menzel et al., 2016) with Reference database: nr (Bacteria and Archaea; Database date: 2017-05-16). We used the default parameters for these analyses (SEG low complexity filter: yes; Run mode: greedy; Minimum match length: 11; Minimum match score: 75; Allowed mismatches: 5). We then converted Kaiju's output files into a summary table at the genus and species level and filtered out taxa with low abundances (<0.1% of the total reads). We also removed poorly identified taxa because they would artificially increase the similarity between our samples. Specifically, the following taxa were excluded: “NA,” “Viruses,” “archaeon,” “uncultured bacterium,” “uncultured Gammaproteobacteria bacterium” (**Supplementary Tables 2, 3**).

Lastly, we used Decontam v1.2.1 to filter out bacterial taxa exhibiting known statistical properties of contaminants (Davis et al., 2018). We used the frequency method (*isContaminant* function) which is based on the inverse relationship between the relative abundance of contaminants and sample DNA concentration, and also has been found suitable for samples dominated by host DNA (Willner et al., 2012; Lusk, 2014; Salter et al., 2014; Jervis-Bardy et al., 2015; Hooper et al., 2019; McArdle and Kaforou, 2020). As input for Decontam analyses, we used the aforementioned total DNA concentration values. Then, as recommended, we explored the distribution of scores assigned by Decontam to assign the threshold according to bimodality between very low and high scores (Davis et al., 2018). For the

¹<https://sourceforge.net/projects/bbmap/>

MAGs matrix, no bimodality was found, and thus we used the 0.1 default value (**Supplementary Figure 1A**). None of the MAGs were classified as contaminants, according to Decontam. For Kaiju matrices, a 0.3 threshold value was selected for the species-level matrix (**Supplementary Figure 1B**) and 0.31 for the genus-level matrix (**Supplementary Figure 1C**). Decontam filtered out a single species (*Clostridia* bacterium k32) from the species matrix and two genera (*Cupriavidus* and *Massilia*) from the genus matrix.

Statistical Analyses

For the genome-resolved metagenomic analyses, we used the normalized MAGs compositional matrices resulting from Salmon. Specifically, these MAG counts are standardized to the individual sample size (MAG copies per million reads) and thus allow between-sample comparisons. For the Kaiju analyses, we used the `rarefy_even_depth` function of `phyloseq` (without replacement as in the hypergeometric model) to rarefy samples to the smallest number of classified sequences per individual observed (85,513 and 71,267 reads in genus and species matrices, respectively) (Weiss et al., 2017).

To visualize similarities of microbiome composition among louse individuals from the same or different individual seal hosts, we constructed non-metric multidimensional scaling (NMDS) ordinations based on Bray–Curtis and Jaccard (binary = T) dissimilarities using the `phyloseq` v1.26-1 R package (McMurdie and Holmes, 2013). Also, to remove subjective bias when interpreting the results of our main NMDS ordinations, we also analyzed whether Bray–Curtis dissimilarity distances between samples from the same infrapopulation were lower than those of comparisons with samples from other infrapopulations. We statistically tested these results using two-sided Sign tests, with which we evaluated whether the difference between within- versus among-infrapopulation medians was significantly different from 0.

To assess the influence of individual host identity on the microbiome composition of louse individuals, we conducted a permutational multivariate analysis of variance (PERMANOVA) (Anderson and Walsh, 2013; Anderson, 2014). PERMANOVA analyses were done using the `adonis2` function in `vegan` v2.5–4 (Oksanen et al., 2019), based on Bray–Curtis and Jaccard distance matrices with 100 iterations. In PERMANOVA analyses, for the individual host identity factor, our within-group sample size ($n = 2$) was smaller than both the total number of groups ($n = 17$) and the total sample size ($n = 34$). From the perspective of statistical power for testing effects related to host individuals, the relatively high number of hosts, in essence, offsets the low number of lice per host. Nevertheless, to account for a potential deviation in F-statistics and R^2 values (Kelly et al., 2015), we wrote an R simulation that randomly subsampled the infrapopulations from which the louse came (5 infrapopulations per iteration). We ran 10 iterations and ran a PERMANOVA analysis for each iteration. Note that, for a few iterations, subsampled samples were too similar and PERMANOVA could not be done. In addition, we ran PERMANOVA analyses to explore additional factors (louse sex: male, female; sequencing lane: 1, 2; and host status: dead, alive) that may explain variance in microbiome

composition. Furthermore, we included significant factors as the first factors of the host identity PERMANOVA models (i.e., to obtain the variance explained by host identity after accounting for the variance explained by that factor). We also restricted the groups in which permutations could be done to only those with the same value of that vector using the `strata` argument (e.g., for a sample collected from a dead host, and for the host-status factor, permutations could only be done among other dead hosts). Lastly, we ran a Mantel test using the `mantel` function in `vegan` (method = spearman, permutations = 9999) to explore if host locality (i.e., the coordinates in which each host was sampled) correlated with the microbiome composition of louse individuals. For this analysis, we ran 10 iterations of an R simulation in which we randomly subsampled one louse sample for each individual host and ran a Mantel test for each iteration. The following packages were used to produce the plots: `ggplot2` v3.1.0.9 (Wickham, 2016), `grid` v3.5.3 (R Core Team, 2019), `gridExtra` v2.3 (Auguie, 2016), `ggrepel` v0.8.0 (Slowikowski et al., 2019), `ggpubr` v0.2.5 (Kassambara, 2018), and `ggsci` v2.9 (Xiao, 2018).

RESULTS

From the genome-resolved metagenomics pipeline, 13 high-quality bacterial metagenome-assembled genomes (MAGs) were obtained (**Table 1** and **Figure 1**). According to MiGA analyses, 10 of them (77%) likely belong to a species not represented in the NCBI Genome database.

Kaiju analyses recovered a higher diversity of microorganisms than did the genome-resolved approach (**Figure 2** and **Supplementary Figure 2**). These differences are likely because of the quality-filtering parameters used in the genome-resolved metagenomics pipeline (i.e., these taxa may have been discarded because the completeness values of their bins were lower than 50% and/or their contamination values were higher than 10%). Nevertheless, bacterial taxa found in the genome-resolved metagenomic approach were generally found also in Kaiju and with similar relative abundances (**Figure 2**), and a similar pattern was found also at the species level (**Supplementary Figure 2**).

Non-metric multidimensional scaling ordinations and PERMANOVA analyses show a major role of infrapopulation identity in explaining microbiome composition for both quantitative and presence–absence data. In the genome-resolved metagenomic dissimilarity matrices, the Bray–Curtis-based NMDS ordination evidenced a strong pattern of clustering by infrapopulation identity (**Figure 3A**). This pattern was not noticeable in the Jaccard-based NMDS ordination because dissimilarity was too low among samples (**Supplementary Figure 3A**). PERMANOVA analyses indicated that most (>84% in all cases) of the variance was explained by infrapopulation identity (PERMANOVA: Bray–Curtis, $R^2 = 0.857$, $F = 6.419$, $P = 0.001$, **Figure 3A**; Jaccard, $R^2 = 0.842$, $F = 5.671$, $P = 0.001$; **Supplementary Figure 3A**). The analyses of differences in pairwise distances showed highly consistent results, as pairwise distances between samples from the same infrapopulations were lower than the median of among-infrapopulation pairwise comparisons in 16 out of 17 infrapopulations (94%; Sign test:

TABLE 1 | Statistics of the MAGs assembled.

MAG name	Completeness (%)	Contamination (%)	N50 (bp)	Size (bp)	Taxator tk ID	MiGA ID	RDP ID	Taxonomic novelty
bin.1	100	1.07	57370	1869975	Flavobacteriaceae	Flavobacteriaceae*	NA	Species****
bin.4	99.26	0.24	81315	2500734	Flavobacteriaceae	Chryseobacterium*	Chryseobacterium (100.0%)	Species****
bin.2	98.51	0.42	36844	3101576	Deinococcus	Deinococcus grandis*	Deinococcus (100.0%)	Subspecies****
bin.7	97.75	0	16123	2650064	Moraxellaceae	Psychrobacter sp. PRwf-1*	NA	Subspecies****
bin.3	97.41	1.33	32961	4014303	Neisseriales	Pseudogulbenkiania*	NA	Species****
bin.11	95.65	0.92	69243	2786419	Moraxellaceae	Psychrobacter*	NA	Species****
bin.10	95.12	0	13409	2459723	Deinococcaceae	Deinococcus*	NA	Species****
bin.12	93.14	0.85	24793	2851493	Deinococcaceae	Deinococcus*	NA	Species****
bin.6	88.74	1.45	7283	1988194	Micrococcales	Arthrobacter*	NA	Species****
bin.13	77.11	0.64	3045	2627969	Deinococcaceae	Deinococcus*	NA	Species****
bin.5	74.13	0.61	24837	1635952	Moraxellaceae	unclassified Moraxellaceae*	Alkanindiges (99%)	Species****
bin.8	67.76	0	10934	2837743	Deinococcaceae	Deinococcus*	NA	Species****
bin.9	61.13	0.30	2210	2110411	Janthinobacterium	Janthinobacterium sp. SNU WT3***	NA	Subspecies****

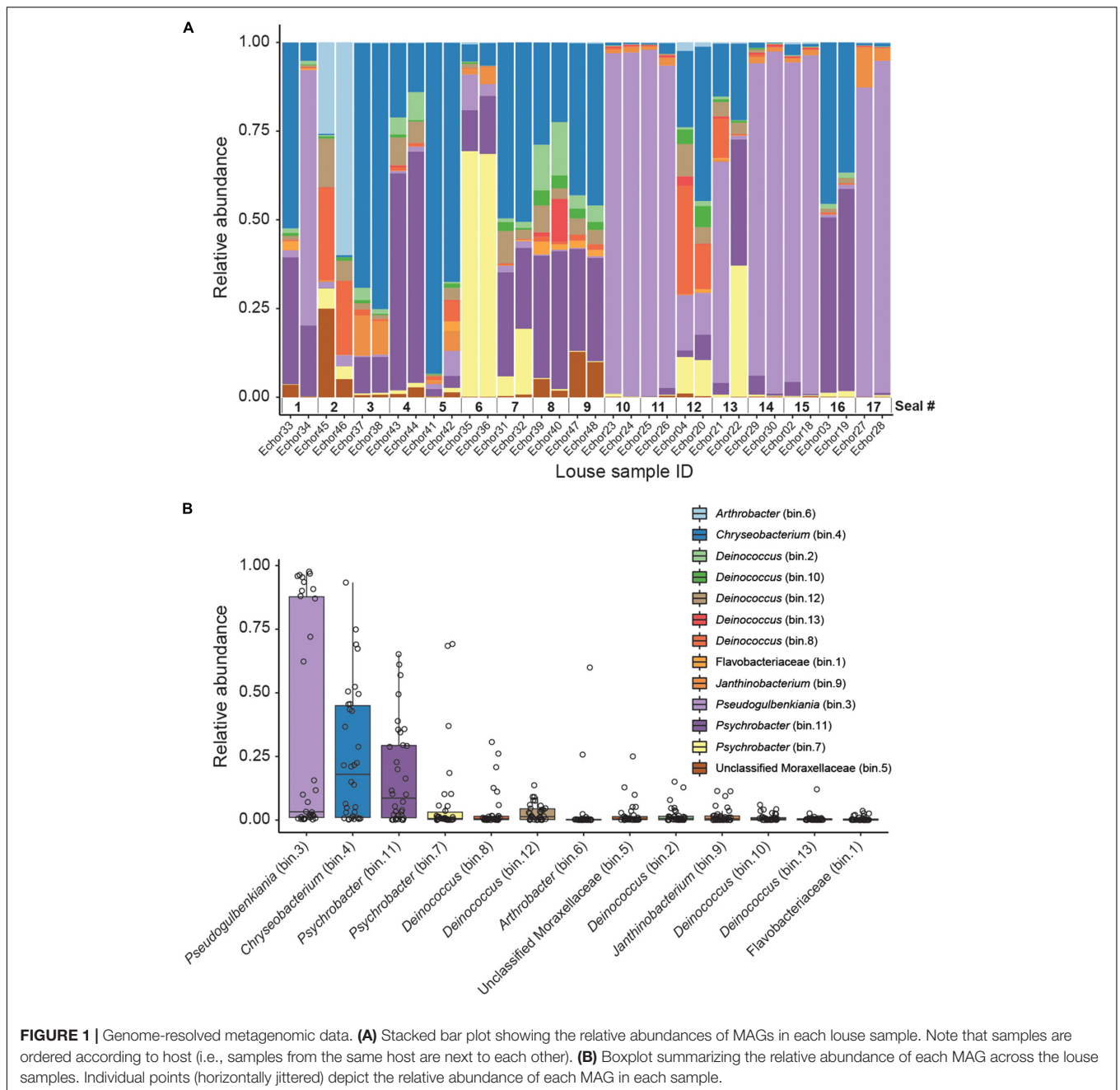
MAG name indicates the name given to that bin for this study (e.g., in **Figure 1**). The highest taxonomic rank with p -value ≤ 0.5 is shown in MiGA ID. RDP ID is the result of the identification analysis using rRNA genes (16S) implemented in MiGA; % indicates confidence in identification. Taxonomic novelty is a MiGA analysis that indicates the taxonomic rank at which the MAG represents a novel taxon with respect to the NCBI Genome database; highest taxonomic rank with p -value ≤ 0.01 are shown. Significance at p -value below: *0.5, ***0.05, ****0.01.

$P < 0.001$; **Supplementary Figures 4, 5**). Results from the simulations were in line with the results of the regular model, and thus support that our results were not biased by the sampling design [PERMANOVA: Bray–Curtis, R^2 (min = 0.65, max = 0.98, mean = 0.78); P (min = 0.001, max = 0.019, $n < 0.05 = 10/10$); Jaccard, R^2 (min = 0.66, max = 1, mean = 0.86), P (min = 0.001, max = 0.106, $n < 0.05 = 5/7$)]. From all the additional factors examined, only host status (i.e., dead, alive) explained a significant amount of variance (**Table 2**). Including host status in PERMANOVA analyses did not alter the results on the major influence of host identity in explaining microbiome composition (PERMANOVA: Host identity, Bray–Curtis, $R^2 = 0.57$, $F = 4.58$, $P = 0.001$; Jaccard, $R^2 = 0.71$, $F = 5.09$, $P = 0.002$).

Similarly, in Kaiju matrices collapsed at the species level, NMDS ordinations showed a pattern of clustering by infrapopulation identity (**Figure 3B** and **Supplementary Figure 3B**). In the same vein, most (>80% in all cases) of the variance was also explained by infrapopulation identity (PERMANOVA: Bray–Curtis, $R^2 = 0.804$, $F = 4.346$, $P = 0.001$, **Figure 3B**; Jaccard, $R^2 = 0.803$, $F = 4.319$, $P = 0.001$; **Supplementary Figure 3B**). The analyses of differences in pairwise distances again supported the results: pairwise distances between samples from the same infrapopulations were lower than the median of among infrapopulation pairwise comparisons in 15 out of 17 infrapopulations (88%; Sign test: $P < 0.01$; **Supplementary Figures 5, 6**). Furthermore, results from simulations were similar [PERMANOVA: Bray–Curtis, R^2 (min = 0.62, max = 0.88, mean = 0.75); P (min = 0.003, max = 0.058, $n < 0.05 = 9/10$); Jaccard, R^2 (min = 0.63, max = 0.95, mean = 0.76), P (min = 0.002, max = 0.09, $n < 0.05 = 9/10$)]. Of all the other factors examined, only host

status explained a significant amount of variance (**Table 2**). PERMANOVA analysis accounting for host status did not alter the relevance of host identity in explaining a significant amount of variance (PERMANOVA: Bray–Curtis, $R^2 = 0.52$, $F = 1.78$, $P = 0.007$; Jaccard, $R^2 = 0.59$, $F = 3.37$, $P = 0.001$).

Furthermore, results were consistent when using matrices collapsed at the genus level. Samples appeared clustered by infrapopulation identity in NMDS ordinations (**Supplementary Figures 7A,B**) and > 77% of variance was explained in all cases by this factor (PERMANOVA: Bray–Curtis, $R^2 = 0.865$, $F = 6.804$, $P = 0.001$, **Supplementary Figure 7A**; Jaccard, $R^2 = 0.774$, $F = 3.634$, $P = 0.001$; **Supplementary Figure 7B**). Once again, results from simulations were similar [PERMANOVA: Bray–Curtis, R^2 (min = 0.68, max = 0.96, mean = 0.8); P (min = 0.002, max = 0.073, $n < 0.05 = 9/10$); Jaccard, R^2 (min = 0.54, max = 0.86, mean = 0.73), P (min = 0.003, max = 0.061, $n < 0.05 = 9/10$)]. Additionally, of all the others factors examined, only host status explained a significant amount of variance [PERMANOVA: Bray–Curtis, Host status: $R^2 = 0.3$, $F = 14$, $P = 0.001$, Louse sex: $R^2 = 0.05$, $F = 0.51$, $P = 0.851$, Sequencing lane: $R^2 = 0.01$, $F = 0.39$, $P = 0.753$; Jaccard, Host-status: $R^2 = 0.18$, $F = 7.19$, $P = 0.002$, Louse sex: $R^2 = 0.07$, $F = 0.75$, $P = 0.53$, Sequencing lane: $R^2 = 0.01$, $F = 0.40$, $P = 0.75$; Mantel test, locality, Bray–Curtis: ρ (min = 0.09, max = 0.09, mean = 0.09), P (min = 0.720, max = 0.734, $n < 0.05 = 0/10$); Jaccard: ρ (min = 0.02, max = 0.02, mean = 0.02), P (min = 0.404, max = 0.425, $n < 0.05 = 0/10$)]. Likewise, PERMANOVA analysis accounting for host status did not alter the results on the relevance of host identity (PERMANOVA: Bray–Curtis, $R^2 = 0.56$, $F = 4.73$, $P = 0.001$; Jaccard, $R^2 = 0.59$, $F = 2.96$, $P = 0.001$).



DISCUSSION

Two different metagenomic approaches support a major role of infrapopulation identity (ringed seal host individual) in explaining microbiome variation among individuals of the seal louse. In addition, highly similar results were found for approaches using either presence–absence or quantitative matrices, suggesting that not only is bacterial composition, but also bacterial abundance explained by infrapopulation identity. Our analyses were done on whole louse individuals and, thus, we cannot confidently differentiate between bacterial taxa inhabiting the lice (e.g., *Wolbachia* or *Hodgkinia*) from transient taxa

present in the host blood meal (e.g., *Chlamydia*). Nevertheless, in line with current evidence on the determinants of microbiome composition of bloodsucking parasites, the louse blood meal from individual seals is the most likely candidate in explaining the patterns of microbiome variation across the focal louse infrapopulations. Indeed, many of the taxa found in our analyses have already been found in other bloodsucking parasites, thus supporting the influence of blood in shaping the composition of parasite microbiomes studied here (Jiménez-Cortés et al., 2018).

However, other factors in addition to blood may have contributed to the similarity of microbiomes between individual lice from the same seal host individual. Some similarity may

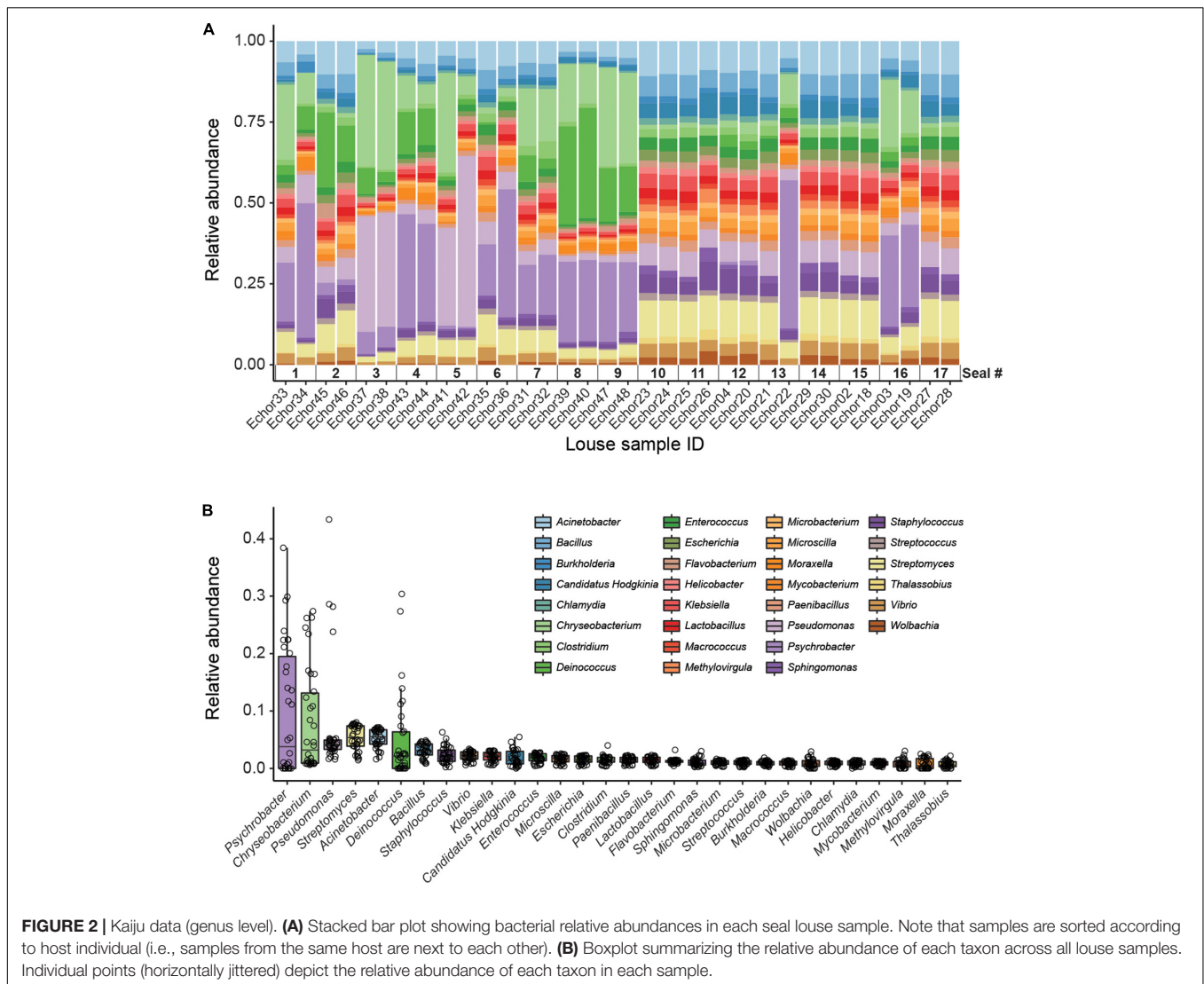


FIGURE 2 | Kaiju data (genus level). **(A)** Stacked bar plot showing bacterial relative abundances in each seal louse sample. Note that samples are sorted according to host individual (i.e., samples from the same host are next to each other). **(B)** Boxplot summarizing the relative abundance of each taxon across all louse samples. Individual points (horizontally jittered) depict the relative abundance of each taxon in each sample.

have arisen from shared environmental bacteria, i.e., those on the surface of the louse from a shared environment (skin and fur of the host), or contamination between louse individuals in screw-cap tubes, and not filtered by our decontamination procedures. While some potential contamination sources are nearly impossible to avoid, possible contamination between louse individuals in screw-cap tubes could have been avoided in this study should the louse from the same individuals have been placed in separate screw-cap tubes. We believe it seems unlikely, especially for some bacterial taxa (e.g., gut bacteria adhered to the gut epithelial cells; Narasimhan and Fikrig, 2015), that once in ethanol, these bacteria could have gone out of the louse individuals and reached the other louse interior. Nevertheless, our ethanol rinses, procedures to extract DNA (i.e., crashing whole louse individuals), and the bioinformatic decontamination filtering ensure this process does not mainly drive our results. There may also be insect-specific bacterial taxa, independent from the host blood, that are shared horizontally between individual lice from the same infrapopulation. Finally, louse

infrapopulations are known to typically be highly inbred, with a high level of relatedness between individuals (Koop et al., 2014; DiBlasi et al., 2018; Virrueta Herrera et al., in prep.). It may be that genetic factors of the lice interact with the microbiome to produce a specific composition (Blekhman et al., 2015; Dobson et al., 2015; Suzuki et al., 2019).

Our results are congruent with previous findings on the influence of host blood on microbiomes of bloodsucking parasites. Specifically, several studies have found a major role of the specific host species from which a blood meal is taken in shaping microbiomes of other bloodsucking organisms, such as ticks (*Ixodes scapularis*, *Ixodes pacificus*) and mosquitoes (*Aedes aegypti*) (Swei and Kwan, 2017; Landesman et al., 2019; Muturi et al., 2019). Furthermore, Landesman et al. (2019) showed that microbiomes of deer tick (*Ixodes scapularis*) nymphs were largely explained by the individual hosts of the tick, a result similar to the one obtained here. Interestingly, in that study, the percentage of variation explained was considerably lower (45%) than that found here (> 77%). It may be that differences in parasite ecology,

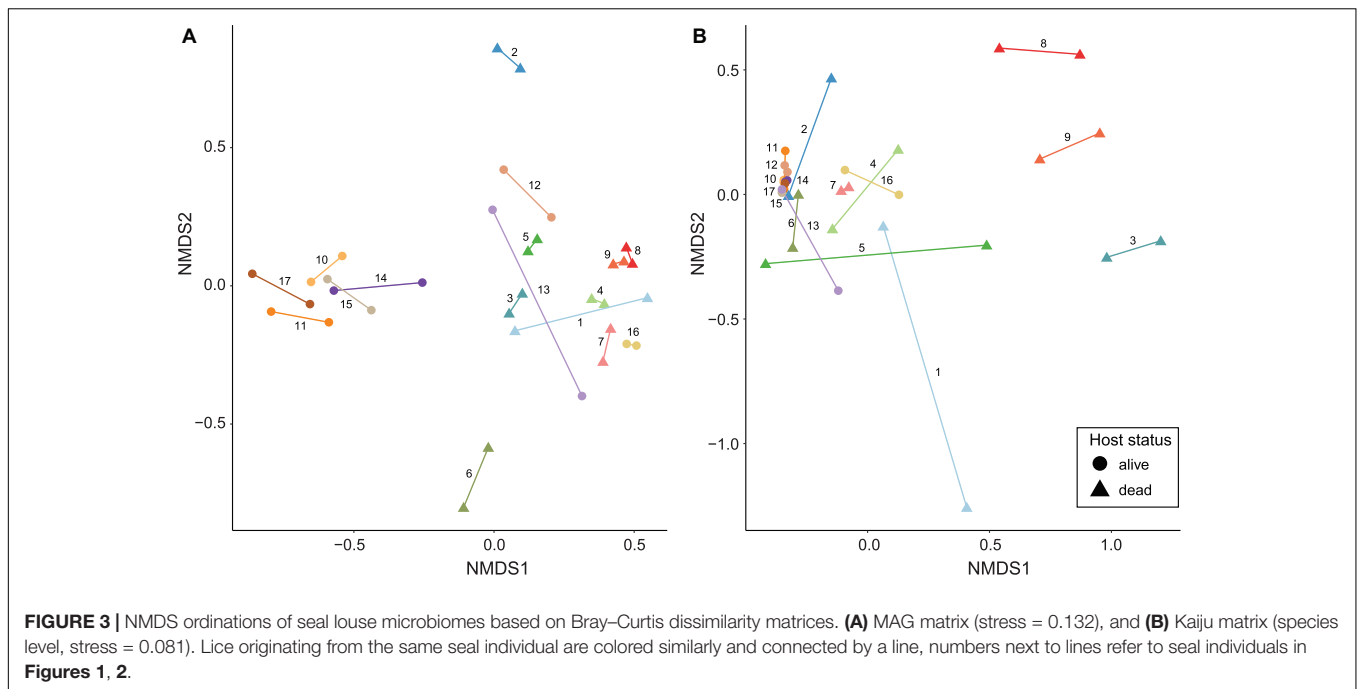


TABLE 2 | PERMANOVA results from the factors in addition to infrapopulation that were evaluated to potentially influence microbiome variation among samples.

Data type	MAGs						Kallu					
	Bray–Curtis			Jaccard			Bray–Curtis			Jaccard		
Dissimilarity distance	R ² /ρ	F	P	R ² /ρ	F	P	R ² /ρ	F	P	R ² /ρ	F	P
Host status	0.28	12.72	0.001	0.13	4.93	0.002	0.22	9.03	0.001	0.21	8.73	0.001
Louse sex	0.08	0.9	0.554	0.03	0.28	0.867	0.08	0.81	0.564	0.08	0.88	0.497
Sequencing lane	0.01	0.38	0.878	0	0.01	1	0.01	0.35	0.859	0.01	0.4	0.825
Locality	–0.09;	–	0.875;	–0.29;	–	0.97;	0.04;	–	0.564;	–0.03;	–	0.534;
	–0.09;		0.887; 0	–0.29;		0.978; 0	0.04;		0.579; 0	–0.03;		0.549; 0
	–0.09			–0.29			0.04			–0.03		

Note that the influence of locality was tested using multiple Mantel tests, and therefore ρ (min, max, and mean) and P (min, max, n < 0.05) simulation values are given (see section “Materials and Methods” for further details).

such as the whether the parasite is a permanent or a recurrent feeder (as are both the case in sucking lice) may modulate the extent to which host blood shapes parasite microbiomes. The differences in the proportion of variance explained by infrapopulation identity between the two studies could also be due to differences in experimental design, such as the number of sampled infrapopulations (3 in ticks, and 17 in the seal lice here) and whether the sample design is balanced (i.e., the same number of individual parasites sampled per infrapopulation).

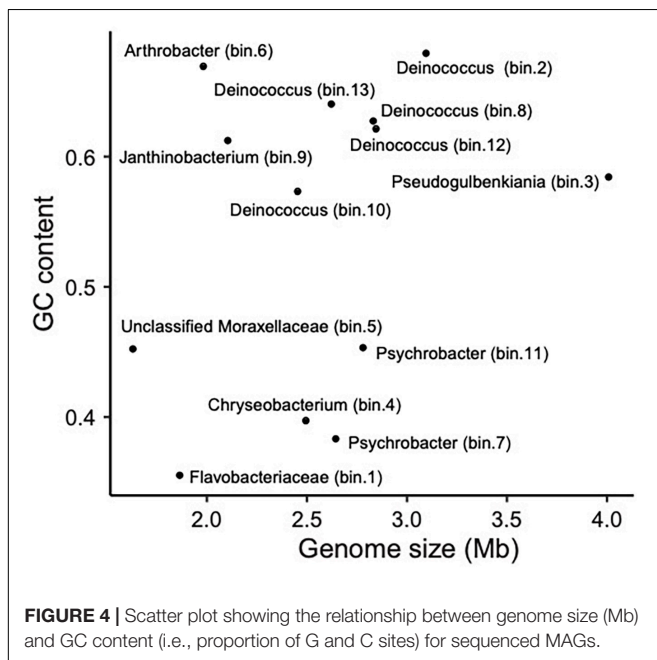
The knowledge that intraspecific variation in the blood of seals may lead to similarity of the microbiomes of lice feeding on the same host individual can potentially provide new insights into the influence of host blood on parasites. At least two not necessarily mutually exclusive processes may explain the influence of a host individual’s blood on louse microbiomes. First, the blood from a particular host individual may contain a specific composition of bacterial loads that enter the louse during feeding. Indeed, anopluran lice might have a high likelihood of being colonized by bacteria from host blood because they do not possess a peritrophic membrane, an extracellular layer in the midgut that is

composed of chitin, proteoglycans, and proteins, which in most other insects surrounds the ingested food bolus and separates the gut content, including bacteria, from the epithelium (Terra, 2001; Waniek, 2009). Indeed, the idea that a lack of a peritrophic membrane facilitates colonization of blood-feeding parasites by bacteria present in the host blood has potentially also been supported by work on mouse fleas (*Rhadinopsylla dahurica*), which also lack this membrane (Li et al., 2018). In this case, there was evidence of homogenization (i.e., similar bacterial communities) between the host blood and the parasite (whole flea individuals). The lack of a peritrophic membrane is often associated with permanent parasites, such as blood-feeding lice, for which the continual availability of food means that there is less selection for efficient digestion. Therefore, the presence versus absence of a peritrophic membrane may explain the differences between lice and ticks (of which the latter possess a peritrophic membrane) with regards to the influence of host blood on the composition of the parasite microbiome.

A second possibility that could explain why host blood may influence louse microbiome composition is that the conditions

during blood digestion may alter bacterial taxa that were already present in the louse. The specifics of blood digestion may have an individual host-specific signature. Specifically, catabolism of blood meal leads to the generation of reactive oxygen species that are known to alter the midgut bacterial composition and diversity of bloodsucking parasites (Souza et al., 1997; Wang et al., 2011; Muturi et al., 2019). Also, the blood of different host species is known to differ in composition (e.g., total protein, hemoglobin, and hematocrit content), and these differences may lead to a differential proliferation of microbial taxa during digestion by parasites (Souza et al., 1997; Wang et al., 2011; Muturi et al., 2019). It may be the case that differences in blood composition among individuals even within the same host species may shape the bacterial composition of lice in a manner that is specific to host individuals.

Bloodsucking organisms, and anopluran lice in particular, are well known to rely on mutualistic endosymbionts to complement deficiencies in their diet (Perotti et al., 2008; Boyd and Reed, 2012; Boyd et al., 2017; Jiménez-Cortés et al., 2018). Notwithstanding that several of the bacterial taxa we found may not be stable inhabitants of lice, we found evidence for the presence of several louse-specific bacterial taxa (i.e., taxa that are highly unlikely to come from environmental contamination). These include the obligate intracellular arthropod bacteria *Wolbachia* (Werren, 1997) and *Hodgkinia* (for which only endosymbionts of cicadas are known; McCutcheon et al., 2009). Accordingly, we explored our MAGs for genome characteristics typical of endosymbionts. In particular, because endosymbiont genomes typically are small and have an AT bias, we explored the relative position of the observed MAGs in a “Genome size ~ GC content” correlation plot (Wernegreen, 2015; Figure 4). Bin 1 appears to be the best candidate to be a mutualistic endosymbiont, according to its relative position in the correlation plot. This MAG was 100%



complete (according to CheckM; Parks et al., 2015), detected in most samples (prevalence = 71%; similar endosymbiont prevalences have been found in Allen et al., 2016 and Říhová et al., 2019), and classified with confidence as Flavobacteriaceae. MiGA analyses suggest it may even belong to *Chryseobacterium* (p -value 0.585). Endosymbionts belonging to *Chryseobacterium* are known in other arthropods (e.g., termites, mosquitoes, cockroaches, and ticks; Eutick et al., 1978; Dugas et al., 2001; Campbell et al., 2004; Montasser, 2005; Burešová et al., 2006). Additionally, we conducted a preliminary investigation of the metabolic capabilities of this bacterium by investigating the completeness of metabolic pathways using GhostKOALA (Kanehisa et al., 2016) and KEGG-Decoder (Graham et al., 2018). This MAG has complete routes for synthesis of vitamin B (riboflavin), an essential amino acid (lysine), and several non-essential amino acids (e.g., serine; see **Supplementary Table 4**), as well as many fully or partially missing routes that may be redundant or potentially shared (or synthesized along) with the louse (**Supplementary Table 4**).

Overall, these results are congruent with what has been found for endosymbionts of bloodsucking parasites (Moriyama et al., 2015; Boyd et al., 2016; Santos-Garcia et al., 2017; Duron et al., 2018). Nevertheless, it is worth mentioning that the relative abundance of this putative endosymbiont MAG was the lowest among all the MAGs studied here. However, in contrast to what it is known for other bloodsucking parasites (e.g., ticks; Narasimhan and Fikrig, 2015), little is known in lice about the abundance and prevalence patterns of the microbiome of different organs and systems. In addition, apart from this study, whole-genome metagenomic data from different individuals are not available. Further research on the individual-level microbiomes of lice is needed to understand the relative abundances of bacteria in lice.

Another anopluran pinniped louse (*Proechinophthirus fluctus*) has been found to have a *Sodalis* endosymbiont (Boyd et al., 2016), but we found no evidence of *Sodalis* in *Echinophthirus horridus*. Other species of Anoplura have yet other endosymbionts (Boyd et al., 2017; Říhová et al., 2017, 2019), suggesting that endosymbiont replacement is an ongoing and relatively common process within the order. Population-scale studies are needed to better understand endosymbiont dynamics in populations of lice. Also, phylogenomic studies aiming at elucidating the phylogenetic placement of the potentially mutualistic *Chryseobacterium* found here and studies using fluorescence *in situ* hybridization (FISH) to ascertain its location in louse individuals are needed to get deeper insight into the interaction of this bacterium with *E. horridus*.

Lastly, the methodology used in this study (i.e., a dual metagenomic approach that combines genome-resolved metagenomics with metagenomic classification tools and state-of-the-art bioinformatic decontamination procedures) opens the door to further studies of the microbiomes of both parasites and free-living organisms for which WGS data are available. Here, this approach allowed us to characterize the variation of microbiomes among individuals of the same parasite species, and to identify factors underlying the observed variation. We were also able to identify potential endosymbionts and to

recover high-quality genomic data from them. Interestingly, in the current field of genomics, excessive accumulation of data and the resultant ever-increasing demand for data-storage capacity are worrying trends (Stephens et al., 2015). Thus, the possibility of using the same genome-level sequence datasets to address multiple different research questions (e.g., data generated for population-genomic analyses and cophylogenomics later leveraged to investigate introgression dynamics; Doña et al., 2020) and in different contexts (e.g., host-derived population-genomic data to infer bacterial composition; Hooper et al., 2019, this study) allows for more efficient use of existing genomic resources.

DATA AVAILABILITY STATEMENT

Raw sequence reads for all samples are available at SRA (see **Supplementary Table 1**). Metagenomic assemblies and FastQC reports (before and after preprocessing) are available at Figshare (doi: 10.6084/m9.figshare.12366575). Metagenomic assemblies are also available from NCBI Genome (SAMN18543074–SAMN18543086).

ETHICS STATEMENT

Telemetry studies have been approved by the local environmental authority Centre for Economic Development, Transport and the Environment (permit numbers: ESAELY/433/07.01/2012 and ESA-2008-L-519-254) and the Animal Experiment Board in Finland (permit numbers: ESAVI/8269/04.10.07/2013 and ESAVI-2010-08380/Ym-23).

REFERENCES

- Acquarone, M., Salgado-Flores, A., and Sundset, M. A. (2020). The Bacterial Microbiome in the Small Intestine of Hooded Seals (*Cystophora cristata*). *Microorganisms* 8:1664. doi: 10.3390/microorganisms8111664
- Adair, K. L., and Douglas, A. E. (2017). Making a microbiome: the many determinants of host-associated microbial community composition. *Curr. Opin. Microbiol.* 35, 23–29. doi: 10.1016/j.mib.2016.11.002
- Agany, D. D. M., Potts, R., Hernandez, J. L. G., Gnimpieba, E. Z., and Pietri, J. E. (2020). Microbiome differences between human head and body lice ecotypes revealed by 16S rRNA Gene Amplicon Sequencing. *J. Parasitol.* 106, 14–24. doi: 10.1645/19-132
- Allen, J. M., Burleigh, J. G., Light, J. E., and Reed, D. L. (2016). Effects of 16S rDNA sampling on estimates of the number of endosymbiont lineages in sucking lice. *PeerJ* 4:e2187.
- Allen, J. M., Reed, D. L., Perotti, M. A., and Braig, H. R. (2007). Evolutionary Relationships of “*Candidatus* RIESIA spp.” Endosymbiotic Enterobacteriaceae Living within Hematophagous Primate Lice. *Appl. Environ. Microbiol.* 73, 1659–1664. doi: 10.1128/aem.01877-06
- Alneberg, J., Bjarnason, B. S., de Bruijn, I., Schirmer, M., Quick, J., Ijaz, U. Z., et al. (2014). Binning metagenomic contigs by coverage and composition. *Nat. Methods* 11, 1144–1146. doi: 10.1038/nmeth.3103
- Anderson, M. J. (2014). *Permutational Multivariate Analysis of Variance (PERMANOVA)*. *Wiley StatsRef: Statistics Reference Online*. (Hoboken, NJ: John Wiley & Sons, Ltd), 1–15. doi: 10.1002/9781118445112.stat07841
- Anderson, M. J., and Walsh, D. C. I. (2013). PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous dispersions: What null hypothesis are you testing? *Ecol. Monogr.* 83, 557–574. doi: 10.1890/12-2010.1

AUTHOR CONTRIBUTIONS

JD, SV, and KJ conceived the study. TN and MK obtained samples. SV and KJ collected the data. JD analyzed the data and wrote the manuscript. TN, MK, and KJ obtained financial support for the project. All authors contributed to editing the manuscript.

FUNDING

This study was supported by the United States National Science Foundation (DEB-1239788, DEB-1342604, and DEB-1926919 to KJ). Sequencing costs were supported by grants from the Oskar Öflund Foundation, the Betty Väänänen Foundation, Societas Pro Fauna et Flora Fennica, and the Nestori Foundation.

ACKNOWLEDGMENTS

We thank all researchers and students who collected the lice analyzed in this study. Especially, we would like to thank researchers Miina Auttila, Vincent Biard, Meeri Koivuniemi, Lauri Liukkonen, Marja Niemi, Sari Oksanen, Mia Valtonen, and Eeva Ylinen for keeping an eye out for seal lice during their own studies on the Saimaa ringed seal.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2021.642543/full#supplementary-material>

- Auguie, B. (2016). *gridExtra: Miscellaneous Functions for “Grid” Graphics. R package version 2.3*. Available online at: <https://CRAN.R-project.org/package=gridExtra>
- Blekhman, R., Goodrich, J. K., Huang, K., Sun, Q., Bukowski, R., Bell, J. T., et al. (2015). Host genetic variation impacts microbiome composition across human body sites. *Genome Biol.* 16:191. doi: 10.1186/s13059-015-0759-1
- Bowers, R. M., Kyrpides, N. C., Stepanauskas, R., Harmon-Smith, M., Doud, D., Reddy, T. B. K., et al. (2017). Minimum information about a single amplified genome (MISAG) and a metagenome-assembled genome (MIMAG) of bacteria and archaea. *Nat. Biotechnol.* 35, 725–731. doi: 10.1038/nbt.3893
- Boyd, B. M., Allen, J. M., de Crécy-Lagard, V., and Reed, D. L. (2014). Genome Sequence of *Candidatus* RIESIA pediculischaeffii, Endosymbiont of Chimpanzee Lice, and Genomic Comparison of Recently Acquired Endosymbionts from Human and Chimpanzee Lice. *G3* 4:g3.114.012567. doi: 10.1534/g3.114.012567
- Boyd, B. M., Allen, J. M., Koga, R., Fukatsu, T., Sweet, A. D., Johnson, K. P., et al. (2016). Two Bacterial Genera, *Sodalis* and *Rickettsia*, Associated with the Seal Louse *Proechinophthirus fluctus* (Phthiraptera: Anoplura). *Appl. Environ. Microbiol.* 82, 3185–3197. doi: 10.1128/aem.00282-16
- Boyd, B. M., Allen, J. M., Nguyen, N.-P., Vachaspati, P., Quicksall, Z. S., Warnow, T., et al. (2017). Primates, lice and bacteria: speciation and genome evolution in the symbionts of hominid lice. *Mol. Biol. Evol.* 34, 1743–1757. doi: 10.1093/molbev/msx117
- Boyd, B. M., and Reed, D. L. (2012). Taxonomy of lice and their endosymbiotic bacteria in the post-genomic era. *Clin. Microbiol. Infect.* 18, 324–331. doi: 10.1111/j.1469-0691.2012.03782.x
- Brooks, A. W., Kohl, K. D., Brucker, R. M., van Opstal, E. J., and Bordenstein, S. R. (2016). Phyllosymbiosis: relationships and functional effects of microbial

- communities across host evolutionary history. *PLoS Biol.* 14:e2000225. doi: 10.1371/journal.pbio.2000225
- Brucker, R. M., and Bordenstein, S. R. (2013). The hologenomic basis of speciation: gut bacteria cause hybrid lethality in the genus *Nasonia*. *Science* 341, 667–669. doi: 10.1126/science.1240659
- Buchner, P. (1965). *Endosymbiosis of Animals with Plant Microorganisms*. (New York, NY: Interscience Publishers), 909.
- Burešová, V., Franta, Z., and Kopáček, P. (2006). A comparison of *Chryseobacterium indologenes* pathogenicity to the soft tick *Ornithodoros moubata* and hard tick *Ixodes ricinus*. *J. Invertebr. Pathol.* 93, 96–104. doi: 10.1016/j.jip.2006.05.006
- Campbell, C. L., Mummey, D. L., Schmidtman, E. T., and Wilson, W. C. (2004). Culture-independent analysis of midgut microbiota in the Arbovirus vector *Culicoides sonorensis* (Diptera: Ceratopogonidae). *J. Med. Entomol.* 41, 340–348. doi: 10.1603/0022-2585-41.3.340
- Campbell, T. P., Sun, X., Patel, V. H., Sanz, C., Morgan, D., and Dantas, G. (2020). The microbiome and resistome of chimpanzees, gorillas, and humans across host lifestyle and geography. *ISME J.* 14, 1584–1599. doi: 10.1038/s41396-020-0634-2
- Clemente, J. C., Ursell, L. K., Parfrey, L. W., and Knight, R. (2012). The impact of the gut microbiota on human health: an integrative view. *Cell* 148, 1258–1270. doi: 10.1016/j.cell.2012.01.035
- Davis, N. M., Proctor, D. M., Holmes, S. P., Relman, D. A., and Callahan, B. J. (2018). Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. *Microbiome* 6:226. doi: 10.1186/s40168-018-0605-2
- Díaz-Sánchez, S., Estrada-Peña, A., Cabezas-Cruz, A., and de la Fuente, J. (2019). Evolutionary insights into the tick hologenome. *Trends Parasitol.* 35, 725–737. doi: 10.1016/j.pt.2019.06.014
- DiBlasi, E., Johnson, K. P., Stringham, S. A., Hansen, A. N., Beach, A. B., Clayton, D. H., et al. (2018). Phoretic dispersal influences parasite population genetic structure. *Mol. Ecol.* 27, 2770–2779. doi: 10.1111/mec.14719
- Dobson, A. J., Chaston, J. M., Newell, P. D., Donahue, L., Hermann, S. L., Sannino, D. R., et al. (2015). Host genetic determinants of microbiota-dependent nutrition revealed by genome-wide analysis of *Drosophila melanogaster*. *Nat. Commun.* 6:6312. doi: 10.1038/ncomms7312
- Doña, J., Sweet, A. D., and Johnson, K. P. (2020). Comparing rates of introgression in parasitic feather lice with differing dispersal capabilities. *Commun. Biol.* 3:610. doi: 10.1038/s42003-020-01345-x
- Dröge, J., Gregor, I., and McHardy, A. C. (2015). Taxator-tk: precise taxonomic assignment of metagenomes by fast approximation of evolutionary neighborhoods. *Bioinformatics* 31, 817–824. doi: 10.1093/bioinformatics/btv745
- Dugas, J. E., Zurek, L., Paster, B. J., Keddie, B. A., and Leadbetter, E. R. (2001). Isolation and characterization of a *Chryseobacterium* strain from the gut of the American cockroach, *Periplaneta americana*. *Arch. Microbiol.* 175, 259–262. doi: 10.1007/s002030000243
- Durden, L. A., and Musser, G. G. (1994). The mammalian hosts of the sucking lice (Anoplura) of the world: a host-parasite list. *J. Vector Ecol.* 19, 130–168.
- Duron, O., Morel, O., Noël, V., Buysse, M., Binetruy, F., Lancelot, R., et al. (2018). Tick-bacteria mutualism depends on b vitamin synthesis pathways. *Curr. Biol.* 28, 1896–1902.e5. doi: 10.1016/j.cub.2018.04.038
- Eutick, M. L., O'Brien, R. W., and Slaytor, M. (1978). Bacteria from the gut of Australian termites. *Appl. Environ. Microbiol.* 35, 823–828.
- Fountain-Jones, N. M., Clark, N. J., Kinsley, A. C., Carstensen, M., Forester, J., Johnson, T. J., et al. (2020). Microbial associations and spatial proximity predict North American moose (*Alces alces*) gastrointestinal community composition. *J. Anim. Ecol.* 89, 817–828. doi: 10.1111/1365-2656.13154
- Fukatsu, T., Hosokawa, T., Koga, R., Nikoh, N., Kato, T., Hayama, S., et al. (2009). Intestinal endocellular symbiotic bacterium of the macaque louse *Pedicinus obtusus*: distinct endosymbiont origins in anthropoid primate lice and the old world monkey louse. *Appl. Environ. Microbiol.* 75, 3796–3799. doi: 10.1128/aem.00226-09
- Graham, E. D., Heidelberg, J. F., and Tully, B. J. (2018). Potential for primary productivity in a globally-distributed bacterial phototroph. *ISME J.* 12, 1861–1866. doi: 10.1038/s41396-018-0091-3
- Hassani, M. A., Durán, P., and Hacquard, S. (2018). Microbial interactions within the plant holobiont. *Microbiome* 6:58. doi: 10.1186/s40168-018-0445-0
- Hooper, R., Brealey, J. C., Valk, T., Alberdi, A., Durban, J. W., Fearnbach, H., et al. (2019). Host-derived population genomics data provides insights into bacterial and diatom composition of the killer whale skin. *Mol. Ecol.* 28, 484–502. doi: 10.1111/mec.14860
- Hypša, V., and Kořízek, J. (2007). Molecular evidence for polyphyletic origin of the primary symbionts of sucking lice (Phthiraptera, Anoplura). *Microb. Ecol.* 54, 242–251. doi: 10.1007/s00248-006-9194-x
- Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., et al. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* 27, 768–777. doi: 10.1101/gr.214346.116
- Jervis-Bardy, J., Leong, L. E. X., Marri, S., Smith, R. J., Choo, J. M., Smith-Vaughan, H. C., et al. (2015). Deriving accurate microbiota profiles from human samples with low bacterial content through post-sequencing processing of Illumina MiSeq data. *Microbiome* 3:19. doi: 10.1186/s40168-015-0083-8
- Jiménez-Cortés, J. G., García-Contreras, R., Bucio-Torres, M. I., Cabrera-Bravo, M., Córdoba-Aguilar, A., Benelli, G., et al. (2018). Bacterial symbionts in human blood-feeding arthropods: patterns, general mechanisms and effects of global ecological changes. *Acta Trop.* 186, 69–101. doi: 10.1016/j.actatropica.2018.07.005
- Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006
- Kang, D. D., Li, F., Kirton, E., Thomas, A., Egan, R., An, H., et al. (2019). MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. doi: 10.7717/peerj.7359
- Kassambara, A. (2018). *Package 'ggpubr': ggplot2-based Publication Ready Plots. Version 0.2*. Available online at: <https://rpkgs.datanovia.com/ggpubr>
- Kelly, B. J., Gross, R., Bittinger, K., Sherrill-Mix, S., Lewis, J. D., Collman, R. G., et al. (2015). Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. *Bioinformatics* 31, 2461–2468. doi: 10.1093/bioinformatics/btv183
- Kim, K. C. (1971). The Sucking Lice (Anoplura: Echinophthiriidae) of the Northern Fur Seal; Descriptions and Morphological Adaptation. *Ann. Entomol. Soc. Am.* 64, 280–292. doi: 10.1093/aesa/64.1.280
- Kim, K. C. (1975). Specific antiquity of the sucking lice and evolution of otariid seals. *Rapp. Proc.* 169, 544–549.
- Kim, K. C. (1985). *Coevolution of Parasitic Arthropods and Mammals*. New York, NY: Wiley.
- Kim, M., Cho, H., and Lee, W. Y. (2020). Distinct gut microbiotas between southern elephant seals and Weddell seals of Antarctica. *J. Microbiol.* 58, 1018–1026. doi: 10.1007/s12275-020-0524-3
- Kirkness, E. F., Haas, B. J., Sun, W., Braig, H. R., Perotti, M. A., Clark, J. M., et al. (2010). Genome sequences of the human body louse and its primary endosymbiont provide insights into the permanent parasitic lifestyle. *Proc. Natl. Acad. Sci. U.S.A.* 107, 12168–12173. doi: 10.1073/pnas.1003379107
- Knowles, S. C. L., Eccles, R. M., and Baltrūnaitė, L. (2019). Species identity dominates over environment in shaping the microbiota of small mammals. *Ecol. Lett.* 22, 826–837. doi: 10.1111/ele.13240
- Kohl, K. D. (2020). Ecological and evolutionary mechanisms underlying patterns of phyllosymbiosis in host-associated microbial communities. *Philos. Trans. R. Soc. B* 375:20190251. doi: 10.1098/rstb.2019.0251
- Kohl, K. D., Varner, J., Wilkening, J. L., and Dearing, M. D. (2018). Gut microbial communities of American pikas (*Ochotona princeps*): evidence for phyllosymbiosis and adaptations to novel diets. *J. Anim. Ecol.* 87, 323–330. doi: 10.1111/1365-2656.12692
- Koop, J. A., DeMatteo, K. E., Parker, P. G., and Whiteman, N. K. (2014). Birds are islands for parasites. *Biol. Lett.* 10:20140255. doi: 10.1098/rsbl.2014.0255
- Laetsch, D. R., and Blaxter, M. L. (2017). BlobTools: interrogation of genome assemblies. *F1000Res.* 6:1287. doi: 10.12688/f1000research.12232.1
- Landesman, W. J., Mulder, K., Allan, B. F., Bashor, L. A., Keesing, F., LoGiudice, K., et al. (2019). Potential effects of blood meal host on bacterial community composition in *Ixodes scapularis* nymphs. *Ticks Tick Borne Dis.* 10, 523–527. doi: 10.1016/j.ttbdis.2019.01.002
- Lankau, E. W., Hong, P., and Mackie, R. I. (2012). Ecological drift and local exposures drive enteric bacterial community differences within species of Galápagos iguanas. *Mol. Ecol.* 21, 1779–1788. doi: 10.1111/j.1365-294x.2012.05502.x

- Le Chatelier, E., Nielsen, T., Qin, J., Prifti, E., Hildebrand, F., Falony, G., et al. (2013). Richness of human gut microbiome correlates with metabolic markers. *Nature* 500, 541–546. doi: 10.1038/nature12506
- Lee, S., Kim, J. Y., Yi, M., Lee, I.-Y., Fyumagwa, R., and Yong, T.-S. (2019). Comparative microbiomes of ticks collected from a black rhino and its surrounding environment. *Int. J. Parasitol. Parasites Wildl.* 9, 239–243. doi: 10.1016/j.ijppaw.2019.05.008
- Leonardi, M. S., Crespo, E. A., Raga, J. A., and Aznar, F. J. (2013). Lousy mums: patterns of vertical transmission of an amphibious louse. *Parasitol. Res.* 112, 3315–3323. doi: 10.1007/s00436-013-3511-3
- Leonardi, M. S., Herrera, S. V., Sweet, A., Negrete, J., and Johnson, K. P. (2019). Phylogenomic analysis of seal lice reveals codivergence with their hosts. *Syst. Entomol.* 44, 699–708. doi: 10.1111/syen.12350
- Li, H., Li, T., and Qu, J. (2018). Stochastic processes govern bacterial communities from the blood of pikas and from their arthropod vectors. *FEMS Microbiol. Ecol.* 94:fy082. doi: 10.1093/femsec/fiy082
- Lim, S. J., and Bordenstein, S. R. (2020). An introduction to phylosymbiosis. *Proc. R. Soc. B* 287:20192900. doi: 10.1098/rspb.2019.2900
- Lusk, R. W. (2014). Diverse and widespread contamination evident in the unmapped depths of high throughput sequencing data. *PLoS One* 9:e110808. doi: 10.1371/journal.pone.0110808
- Lutz, H. L., Jackson, E. W., Webala, P. W., Babyesiza, W. S., Peterhans, J. C. K., Demos, T. C., et al. (2019). Ecology and host identity outweigh evolutionary history in shaping the bat microbiome. *mSystems* 4:e00511-19. doi: 10.1128/mSystems.00511-19
- Mazel, F., Davis, K. M., Loudon, A., Kwong, W. K., Groussin, M., and Parfrey, L. W. (2018). Is host filtering the main driver of phylosymbiosis across the tree of life? *mSystems* 3:e00097-18. doi: 10.1128/mSystems.00097-18
- McArdle, A. J., and Kaforou, M. (2020). Sensitivity of shotgun metagenomics to host DNA: abundance estimates depend on bioinformatic tools and contamination is the main issue. *Access Microbiol.* 2:acmi000104. doi: 10.1099/acmi.0.000104
- McCutcheon, J. P., McDonald, B. R., and Moran, N. A. (2009). Origin of an alternative genetic code in the extremely small and GC-Rich genome of a bacterial symbiont. *PLoS Genet.* 5:e1000565. doi: 10.1371/journal.pgen.1000565
- McMurdie, P. J., and Holmes, S. (2013). phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One* 8:e61217. doi: 10.1371/journal.pone.0061217
- Menzel, P., Ng, K. L., and Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nat. Commun.* 7:11257. doi: 10.1038/ncomms11257
- Montasser, A. A. (2005). Gram-negative bacteria from the camel tick *Hyalomma dromedarii* (Ixodidae) and the chicken tick *Argas persicus* (Argasidae) and their antibiotic sensitivities. *J. Egypt. Soc. Parasitol.* 35, 95–106.
- Moriyama, M., Nikoh, N., Hosokawa, T., and Fukatsu, T. (2015). Riboflavin provisioning underlies *Wolbachia*'s fitness contribution to its insect host. *mBio* 6:e01732-15. doi: 10.1128/mbio.01732-15
- Muturi, E. J., Dunlap, C., Ramirez, J. L., Rooney, A. P., and Kim, C.-H. (2019). Host blood-meal source has a strong impact on gut microbiota of *Aedes aegypti*. *FEMS Microbiol. Ecol.* 95:fy213. doi: 10.1093/femsec/fiy213
- Narasimhan, S., and Fikrig, E. (2015). Tick microbiome: the force within. *Trends Parasitol.* 31, 315–323. doi: 10.1016/j.pt.2015.03.010
- Nelson, T. M., Rogers, T. L., Carlini, A. R., and Brown, M. V. (2013). Diet and phylogeny shape the gut microbiota of Antarctic seals: a comparison of wild and captive animals. *Environ. Microbiol.* 15, 1132–1145. doi: 10.1111/1462-2920.12022
- Niemi, M., Liukkonen, L., Koivuniemi, M., Auttila, M., Rautio, A., and Kunasranta, M. (2019). Winter behavior of Saimaa ringed seals: non-overlapping core areas as indicators of avoidance in breeding females. *PLoS One* 14:e0210266. doi: 10.1371/journal.pone.0210266
- Nishida, A. H., and Ochman, H. (2018). Rates of gut microbiome divergence in mammals. *Mol. Ecol.* 27, 1884–1897. doi: 10.1111/mec.14473
- Nordquist, O. (1899). Beitrag zur kenntniss der isolierten formen der ringelrobbe (*Phoca foetida* Fabr.). *Acta Societ. Pro Faun. Flora Fenn.* 15, 1–43.
- Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P. A. (2017). metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834. doi: 10.1101/gr.213959.116
- Nyman, T., Valtonen, M., Aspi, J., Ruokonen, M., Kunasranta, M., and Palo, J. U. (2014). Demographic histories and genetic diversities of Fennoscandian marine and landlocked ringed seal subspecies. *Ecol. Evol.* 4, 3420–3434. doi: 10.1002/ece3.1193
- Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., O'hara, R. B., Simpson, G. L., et al. (2019). *Vegan: Community Ecology Package*. R package version 2.5-4. Available online at: <http://CRAN.R-project.org/package=vegan>
- Osei-Poku, J., Mbogo, C. M., Palmer, W. J., and Jiggins, F. M. (2012). Deep sequencing reveals extensive variation in the gut microbiota of wild mosquitoes from Kenya. *Mol. Ecol.* 21, 5138–5150. doi: 10.1111/j.1365-294x.2012.05759.x
- Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P., and Tyson, G. W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* 25, 1043–1055. doi: 10.1101/gr.186072.114
- Patro, R., Duggal, G., Love, M. I., Irizarry, R. A., and Kingsford, C. (2017). Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* 14, 417–419. doi: 10.1038/nmeth.4197
- Perotti, M. A., Allen, J. M., Reed, D. L., and Braig, H. R. (2007). Host-symbiont interactions of the primary endosymbiont of human head and body lice. *FASEB J.* 21, 1058–1066. doi: 10.1096/fj.06-6808.com
- Perotti, M. A., Kirkness, E. F., Reed, D. L., and Braig, H. R. (2008). "Endosymbionts of lice," in *Insect Symbiosis*, Vol. 3, ed. T. M. A. K. Bourtzis (Boca Raton, FL: CRC Press), 223–238.
- Quince, C., Walker, A. W., Simpson, J. T., Loman, N. J., and Segata, N. (2017). Shotgun metagenomics, from sampling to analysis. *Nat. Biotechnol.* 35, 833–844. doi: 10.1038/nbt.3935
- R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Řihová, J., Batani, G., Rodríguez-Ruano, S. M., Martinů, J., Nováková, E., and Hypša, V. (2019). A new symbiotic lineage related to *Neisseria* and *Snodgrassella* arises from the dynamic and diverse microbiomes in sucking lice. *bioRxiv* [Preprint]. doi: 10.1101/867275
- Řihová, J., Nováková, E., Husník, F., and Hypša, V. (2017). *Legionella* becoming a mutualist: adaptive processes shaping the genome of symbiont in the louse *Polyplax serrata*. *Genome Biol. Evol.* 9:evx217. doi: 10.1093/gbe/evx217
- Rodríguez-R, L. M., Gunturu, S., Harvey, W. T., Rosselló-Mora, R., Tiedje, J. M., Cole, J. R., et al. (2018). The Microbial Genomes Atlas (MiGA) webserver: taxonomic and gene diversity analysis of Archaea and Bacteria at the whole genome level. *Nucleic Acids Res.* 46:gky467. doi: 10.1093/nar/gky467
- Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., et al. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature* 555, 210–215. doi: 10.1038/nature25973
- Rudman, S. M., Greenblum, S., Hughes, R. C., Rajpurohit, S., Kiratli, O., Lowder, D. B., et al. (2019). Microbiome composition shapes rapid genomic adaptation of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* 116, 20025–20032. doi: 10.1073/pnas.1907787116
- Salter, S. J., Cox, M. J., Turek, E. M., Calus, S. T., Cookson, W. O., Moffatt, M. F., et al. (2014). Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. *BMC Biol.* 12:87. doi: 10.1186/s12915-014-0087-z
- Santos-García, D., Silva, F. J., Morin, S., Dettner, K., and Kuechler, S. M. (2017). The all-rounder *Sodalis* - a new bacteriome-associated endosymbiont of the lygaeoid bug *Henestaris halophilus* (Heteroptera: Henestarinae) and a critical examination of its evolution. *Genome Biol. Evol.* 9:evx202. doi: 10.1093/gbe/evx202
- Sasaki-Fukatsu, K., Koga, R., Nikoh, N., Yoshizawa, K., Kasai, S., Mihara, M., et al. (2006). Symbiotic bacteria associated with stomach discs of human lice. *Appl. Environ. Microbiol.* 72, 7349–7352. doi: 10.1128/aem.01429-06
- Slowikowski, K., Schep, A., Hughes, S., Lukauskas, S., Irisson, J. O., and Kamvar, Z. N. (2019). *ggrepel: Automatically Position Non-Overlapping Text Labels with ggplot2* 2018 R package version 0.8.0. Available online at: <https://CRAN.R-project.org/package=ggrepel>
- Snodgrass, R. E. (1944). *The Feeding Apparatus of biting and Sucking Insects Affecting Man and Animals*. Smithsonian Miscellaneous Collections. Washington, DC: Smithsonian Institution.
- Song, S. J., Sanders, J. G., Delsuc, F., Metcalf, J., Amato, K., Taylor, M. W., et al. (2020). Comparative analyses of vertebrate gut microbiomes reveal convergence between birds and bats. *mBio* 11:e02901-19. doi: 10.1128/mbio.02901-19

- Souza, A. V. G., Petretski, J. H., Demasi, M., Bechara, E. J. H., and Oliveira, P. L. (1997). Urate protects a blood-sucking insect against hemin-induced oxidative stress. *Free Radic. Biol. Med.* 22, 209–214. doi: 10.1016/s0891-5849(96)00293-6
- Stephens, Z. D., Lee, S. Y., Faghri, F., Campbell, R. H., Zhai, C., Efron, M. J., et al. (2015). Big data: Astronomical or Genomical? *PLoS Biol.* 13:e1002195. doi: 10.1371/journal.pbio.1002195
- Stoffel, M. A., Acevedo-Whitehouse, K., Morales-Durán, N., Grosser, S., Chakarov, N., Krüger, O., et al. (2020). Early sexual dimorphism in the developing gut microbiome of northern elephant seals. *Mol. Ecol.* 29, 2109–2122. doi: 10.1111/mec.15385
- Suzuki, T. A., Phifer-Rixey, M., Mack, K. L., Sheehan, M. J., Lin, D., Bi, K., et al. (2019). Host genetic determinants of the gut microbiota of wild mice. *Mol. Ecol.* 28, 3197–3207. doi: 10.1111/mec.15139
- Swei, A., and Kwan, J. Y. (2017). Tick microbiome and pathogen acquisition altered by host blood meal. *ISME J.* 11, 813–816. doi: 10.1038/ismej.2016.152
- Terra, W. R. (2001). The origin and functions of the insect peritrophic membrane and peritrophic gel. *Arch. Insect. Biochem.* 47, 47–61. doi: 10.1002/arch.1036
- Uritskiy, G. V., DiRuggiero, J., and Taylor, J. (2018). MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158. doi: 10.1186/s40168-018-0541-1
- Velazquez, E. M., Nguyen, H., Heasley, K. T., Saechao, C. H., Gil, L. M., Rogers, A. W. L., et al. (2019). Endogenous *Enterobacteriaceae* underlie variation in susceptibility to *Salmonella* infection. *Nat. Microbiol.* 4, 1057–1064. doi: 10.1038/s41564-019-0407-8
- von Olfers, I. F. M. (1816). De vegetativis et animatis corporibus in corporibus animatis reperiundis commentarius. *Berol. Dissert. Pt. 1*, 79–97.
- Wang, Y., Gilbreath, T. M., Kukutla, P., Yan, G., and Xu, J. (2011). Dynamic gut microbiome across life history of the malaria mosquito *Anopheles gambiae* in Kenya. *PLoS One* 6:e24767. doi: 10.1371/journal.pone.0024767
- Waniek, P. J. (2009). The digestive system of human lice: current advances and potential applications. *Physiol. Entomol.* 34, 203–210. doi: 10.1111/j.1365-3032.2009.00681.x
- Weiss, S., Xu, Z. Z., Peddada, S., Amir, A., Bittinger, K., Gonzalez, A., et al. (2017). Normalization and microbial differential abundance strategies depend upon data characteristics. *Microbiome* 5:27. doi: 10.1186/s40168-017-0237-y
- Wernegreen, J. J. (2015). Endosymbiont evolution: predictions from theory and surprises from genomes. *Ann. N.Y. Acad. Sci.* 1360, 16–35. doi: 10.1111/nyas.12740
- Werren, J. H. (1997). Biology of Wolbachia. *Annu. Rev. Entomol.* 42, 587–609. doi: 10.1146/annurev.ento.42.1.587
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer.
- Willner, D., Daly, J., Whiley, D., Grimwood, K., Wainwright, C. E., and Hugenholtz, P. (2012). Comparison of DNA extraction methods for microbial community profiling with an application to pediatric Bronchoalveolar lavage samples. *PLoS One* 7:e34605. doi: 10.1371/journal.pone.0034605
- Wu, Y.-W., Simmons, B. A., and Singer, S. W. (2016). MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32, 605–607. doi: 10.1093/bioinformatics/btv638
- Xiao, N. (2018). *ggsci: Scientific Journal and Sci-fi Themed Color Palettes for 'ggplot2'. R package version 2.9*. Available online at: <https://github.com/nanxstats/ggsci>
- Zhang, X.-C., Yang, Z.-N., Lu, B., Ma, X.-F., Zhang, C.-X., and Xu, H.-J. (2014). The composition and transmission of microbiome in hard tick, *Ixodes persulcatus*, during blood meal. *Ticks Tick Borne Dis.* 5, 864–870. doi: 10.1016/j.ttbdis.2014.07.009
- Zolnik, C. P., Falco, R. C., Daniels, T. J., and Kolokotronis, S.-O. (2018). Transient influence of blood meal and natural environment on blacklegged tick bacterial communities. *Ticks Tick Borne Dis.* 9, 563–572. doi: 10.1016/j.ttbdis.2018.01.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Doña, Virrueta Herrera, Nyman, Kunnasranta and Johnson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.