

RESEARCH ARTICLE

Deep Sequencing and Ecological Characterization of Gut Microbial Communities of Diverse Bumble Bee Species

Haw Chuan Lim^{1*}, Chia-Ching Chu², Manfredo J. Seufferheld^{1*}, Sydney A. Cameron^{1*}

1 Department of Entomology, University of Illinois, Urbana, IL, 61801, United States of America, **2** Department of Crop Sciences, University of Illinois, Urbana, IL, 61801, United States of America

✉ Current address: Smithsonian Institution, National Museum of Natural History, Washington, DC, 20560, United States of America.

* limhc@si.edu (HCL); mseufferheld@gmail.com (MJS); scameron@life.illinois.edu (SAC)



OPEN ACCESS

Citation: Lim HC, Chu C-C, Seufferheld MJ, Cameron SA (2015) Deep Sequencing and Ecological Characterization of Gut Microbial Communities of Diverse Bumble Bee Species. PLoS ONE 10(3): e0118566. doi:10.1371/journal.pone.0118566

Academic Editor: Fabio S. Nascimento, Universidade de São Paulo, Faculdade de Filosofia Ciências e Letras de Ribeirão Preto, BRAZIL

Received: November 6, 2014

Accepted: January 17, 2015

Published: March 13, 2015

Copyright: © 2015 Lim et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All sequencing data is in Genbank under accession numbers KP412082-KP412155.

Funding: The work was supported by a U.S. Department of Agriculture, National Institute of Food and Agriculture grant (#2010-65104-05992) to SAC and by a U.S. Department of Agriculture, National Institute of Food and Agriculture grant (#2009-35505-06012) to MJS. Both grants supported the expenses for pyrosequencing, DNA fingerprinting, and materials/reagents used in this study. The

Abstract

Gut bacterial communities of bumble bees are correlated with defense against pathogens. Further understanding this host-microbe association is vitally important as bumble bees are currently experiencing global population declines, potentially due in part to emergent diseases. In this study, we used pyrosequencing and community fingerprinting (ARISA) to characterize the gut microbial communities of nine bumble species from across the *Bombus* phylogeny. Overall, we delimited 74 bacterial taxa (operational taxonomic units or OTUs) belonging to Betaproteobacteria, Gammaproteobacteria, Bacilli, Actinobacteria, Flavobacteria and Alphaproteobacteria. Each bacterial community was taxonomically simple, containing an average of 1.9 common (relative abundance per sample > 5%) bacterial OTUs. The most abundant and prevalent (occurring in 92% of the samples) bacterial OTU, based on 16S rRNA sequences, closely matched that of the previously described Betaproteobacteria species *Snodgrassella alvi*. Bacteria that were first described in bee-related external environments dominated a number of gut bacterial communities, suggesting that they are not strictly dependent on the internal gut environment. The ARISA data showed a correlation between bacterial community structures and the geographic locations where the bees were sampled, suggesting that at least a subset of the bacterial species may be transmitted environmentally. Using light and fluorescent microscopy, we demonstrated that the gut bacteria form a biofilm on the internal epithelial surface of the ileum, corroborating results obtained from *Apis mellifera*.

Introduction

Recent studies have shown that bacterial symbionts of animals play significant roles in shaping the physiology, ecology and evolution of their hosts [1,2]. In insects, these symbionts allow specialization on particular diets, such as termites on wood and aphids on phloem, even to the extent that some host and bacterial genomes have become complementary and compensatory

fundors had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

with respect to genes that underlie biosynthesis pathways [3,4]. Moreover, resident gut bacteria may protect hosts against invasion by microbial pathogens through production of antimicrobial compounds, stimulation of host-immune systems or competition for space and nutrients [5–7]. Such colonization resistance conferred by resident microbiota has been documented in a wide variety of insects, including flies [8], caterpillars [9], termites [10], and locusts [11] and bumble bees [12].

Recent studies of gut microbiota in bees belonging to the corbiculate clade (honey bees [Apini], bumble bees [Bombini], stingless bees [Meliponini] and orchid bees [Euglossini]) have advanced understanding of the taxonomic composition and structure of microbial communities [13–15], elucidated important health effects and roles in defense against host enemies [12,16–19] and revealed ecological factors, such as host species, and environmental chemical substances that affect bacterial communities [20,21]. Understanding the contributions of intestinal microbiota to host fitness of bees is important and timely as many species of these globally important pollinators are currently under threat from disease [22–25], agrochemicals [26,27], habitat degradation [28,29] or a combination of these factors [30,31]. To date, most studies of bumble bees (*Bombus*), which are much more speciose than honey bees (*Apis*), have focused on specific bacterial lineages [32] of only a few host species [15], often relying on 16S rRNA sequencing of a small number of molecular clones, limiting the scope of our knowledge of the microbial diversity and structure of their gut microbiota [33]. To further our understanding of the composition and structure of bumble bee gut bacterial communities, we conducted pyrosequencing and community fingerprinting of gut bacterial microbiota of nine bumble bee species from across the known *Bombus* phylogeny—to date, the gut microbiota of most of our study species have not been characterized at all. We applied naïve Bayesian classification and phylogenetic analysis to the pyrosequences. To examine the effects of bumble bee species and sampling locality on bacterial community profile, we conducted community fingerprinting on a subset of taxa, using larger sampling of each species from more localities. Finally, we used light and fluorescent microscopy to visualize the distribution of bacteria in the lumen of the gut of *Bombus impatiens* to compare with the pattern reported for honey bees [34].

Materials and Methods

Sampling of bumble bees

To conduct gut bacterial 16S rRNA amplicon sequencing and Automated Ribosomal Intergenic Spacer Analysis (ARISA), we assembled a total of 57 individuals of nine species (*B. atratus*, *B. bohemicus*, *B. convexus*, *B. fraternus*, *B. impatiens*, *B. pensylvanicus*, *B. terrestris*, *B. terricola* and *B. ussurensis*) from seven subgenera. These include species that are known to be undergoing widespread or regional population declines in the US (*B. [Thoracobombus] pensylvanicus*, *B. [Bombus] terricola* and *B. [Cullumanobombus] fraternus*) [35,36], a US species with healthy populations (*B. [Pyrobombus] impatiens*) [35], a European species used for commercial pollination (*B. [Bombus] terrestris*), a socially parasitic species (*B. [Psithyrus] bohemicus*) and two early diverging species (*B. [Mendacibombus] convexus* and *B. [Diversobombus] ussurensis*) [37] (Table 1). All individuals were workers collected in the field with aerial nets. For 16S rRNA amplicon sequencing, we examined all nine species to understand the taxonomic diversity of bacteria found in taxa selected from different clades across the *Bombus* phylogeny. For this portion of the study, we sequenced 1–2 pools of whole-gut DNA for each species. Each pool of DNA was the result of equimolar mixing of DNA from two to three conspecific bees that were collected from the same locality at the same time (Table 1). Hereafter, each such pool will be referred to as a pyrosequencing sample. In total, 29 individuals from 13 distinct locations were used for sequencing, resulting in 13 pyrosequencing samples. The sequencing IDs

Table 1. Bumble bees used in amplicon sequencing and ARISA, and their geographic origins.

ARISA ID	Pyrosequencing sample ID	<i>Bombus</i> species	State/Province	Country	Latitude	Longitude	Field ID (if available)
	ATRATUS	<i>B. atratus</i>	Boyaca	Columbia	5.64	-72.9	
	ATRATUS	<i>B. atratus</i>	Boyaca	Columbia	5.64	-72.9	
	BOHEMICUS	<i>B. bohemicus</i>	Uppland	Sweden	60.48	17.62	
	BOHEMICUS	<i>B. bohemicus</i>	Uppland	Sweden	60.48	17.62	
	BOHEMICUS	<i>B. bohemicus</i>	Uppland	Sweden	60.48	17.62	
	CONVEXUS	<i>B. convexus</i>	Sichuan	China	32.79	102.54	
	CONVEXUS	<i>B. convexus</i>	Sichuan	China	32.79	102.54	
	CONVEXUS	<i>B. convexus</i>	Sichuan	China	32.79	102.54	
B.FR7	FRATERNUS_FL	<i>B. fraternus</i>	Florida	USA	30.13	-83.35	FL11.0011
B.FR9	FRATERNUS_FL	<i>B. fraternus</i>	Florida	USA	30.13	-83.35	FL11.0009
B.FR10		<i>B. fraternus</i>	Florida	USA	30.13	-83.35	FL11.0010
B.FR2		<i>B. fraternus</i>	Georgia	USA	31.35	-82.02	GA11.0031
B.FR3		<i>B. fraternus</i>	Georgia	USA	31.35	-82.02	GA11.0028
B.FR4		<i>B. fraternus</i>	Georgia	USA	31.35	-82.02	GA11.0032
B.FR6		<i>B. fraternus</i>	North Carolina	USA	34.67	-77.03	NC11.0032
B.FR8		<i>B. fraternus</i>	North Carolina	USA	34.67	-77.03	NC11.0036
B.IM10	IMPATIENS_GA	<i>B. impatiens</i>	Georgia	USA	31.85	-82.36	GA11.0015
B.IM12	IMPATIENS_GA	<i>B. impatiens</i>	Georgia	USA	31.85	-82.36	GA11.0014
B.IM3	IMPATIENS_MI	<i>B. impatiens</i>	Michigan	USA	44.98	-84.28	MI11.0028
	IMPATIENS_MI	<i>B. impatiens</i>	Michigan	USA	44.98	-84.28	MI11.0031
B.IM1		<i>B. impatiens</i>	Michigan	USA	44.98	-84.28	MI11.0030
B.IM2		<i>B. impatiens</i>	Michigan	USA	44.98	-84.28	MI11.0029
B.IM5		<i>B. impatiens</i>	Michigan	USA	44.98	-84.28	MI11.0032
B.IM6		<i>B. impatiens</i>	West Virginia	USA	37.77	-80.23	WV11.0007
B.IM8		<i>B. impatiens</i>	West Virginia	USA	37.77	-80.23	WV11.0004
B.IM9		<i>B. impatiens</i>	West Virginia	USA	37.77	-80.23	WV11.0006
B.PE4	PENNSYLVANICUS_FL	<i>B. pensylvanicus</i>	Florida	USA	29.89	-82.67	FL11.0001
	PENNSYLVANICUS_FL	<i>B. pensylvanicus</i>	Florida	USA	29.89	-82.67	FL11.0003
B.PE1	PENNSYLVANICUS_NC	<i>B. pensylvanicus</i>	North Carolina	USA	34.48	-78.63	NC11.0042
B.PE5	PENNSYLVANICUS_NC	<i>B. pensylvanicus</i>	North Carolina	USA	34.48	-78.63	NC11.0040
B.PE2		<i>B. pensylvanicus</i>	Florida	USA	29.89	-82.67	FL11.0002
B.PE3		<i>B. pensylvanicus</i>	South Carolina	USA	34.23	-79.15	SC11.0005
B.PE6		<i>B. pensylvanicus</i>	South Carolina	USA	34.23	-79.15	SC11.0004
B.PE9		<i>B. pensylvanicus</i>	South Carolina	USA	34.23	-79.15	SC11.0003
B.TRS.7	TERRESTRIS_ARG	<i>B. terrestris</i>	Río Negro	Argentina	-43.06	-71.53	BARI2012.003
B.TRS.8	TERRESTRIS_ARG	<i>B. terrestris</i>	Río Negro	Argentina	-43.06	-71.53	BARI2012.004
B.TRS.1	TERRESTRIS_FR	<i>B. terrestris</i>	Pyrenees-Orientalis	France	42.55	2.97	Fr10.350
B.TRS.2	TERRESTRIS_FR	<i>B. terrestris</i>	Pyrenees-Orientalis	France	42.55	2.97	Fr10.351
B.TRS.3		<i>B. terrestris</i>	Pyrenees-Orientalis	France	42.55	2.97	Fr10.347
B.TRS.4		<i>B. terrestris</i>	Pyrenees-Orientalis	France	42.55	2.97	Fr10.352
B.TRS.5		<i>B. terrestris</i>	Río Negro	Argentina	-43.06	-71.53	BARI2012.001
B.TRS.6		<i>B. terrestris</i>	Río Negro	Argentina	-43.06	-71.53	BARI2012.002
B.TE10	TERRICOLA_MI	<i>B. terricola</i>	Michigan	USA	46.16	-88.77	MI11.0014
B.TE14	TERRICOLA_MI	<i>B. terricola</i>	Michigan	USA	46.16	-88.77	MI11.0001
B.TE12	TERRICOLA_WI	<i>B. terricola</i>	Wisconsin	USA	45.73	-89.54	WI11.0002
B.TE2	TERRICOLA_WI	<i>B. terricola</i>	Wisconsin	USA	45.73	-89.54	WI11.0009

(Continued)

Table 1. (Continued)

ARISA ID	Pyrosequencing sample ID	<i>Bombus</i> species	State/Province	Country	Latitude	Longitude	Field ID (if available)
B.TE13		<i>B. terricola</i>	Michigan	USA	46.16	-88.77	MI11.0002
B.TE5		<i>B. terricola</i>	Michigan	USA	46.16	-88.77	MI11.0018
B.TE3		<i>B. terricola</i>	Minnesota	USA	47.89	-92.77	MN11.0043
B.TE4		<i>B. terricola</i>	Minnesota	USA	47.89	-92.77	MN11.0027
B.TE6		<i>B. terricola</i>	Minnesota	USA	47.89	-92.77	MN11.0026
B.TE7		<i>B. terricola</i>	Minnesota	USA	47.89	-92.77	MN11.0038
B.TE9		<i>B. terricola</i>	Minnesota	USA	47.89	-92.77	MN11.0040
B.TE1		<i>B. terricola</i>	Wisconsin	USA	45.73	-89.54	WI11.0008
	USSURENSIS	<i>B. ussurensis</i>	North Chungcheong	South Korea	38.1	127.07	
	USSURENSIS	<i>B. ussurensis</i>	North Chungcheong	South Korea	38.1	127.07	
	USSURENSIS	<i>B. ussurensis</i>	North Chungcheong	South Korea	38.1	127.07	

For amplicon sequencing, bumble bees whose DNAs were pooled prior to Roche 454 library preparation were given the same pyrosequencing sample ID.

doi:10.1371/journal.pone.0118566.t001

for the samples are capitalized and enclosed within quotes in the following sections. To complement the amplicon sequencing, we used ARISA—a bacterial community fingerprinting technique relying on length variation in the rRNA internal transcribed spacer—to examine the effects of host species and geography on dissimilarity in bacterial communities. We applied ARISA to gut bacterial communities from four North American species (*B. fraternus*, *B. impatiens*, *B. pensylvanicus*, *B. terricola*) and one from Europe (*B. terrestris*), samples of which were taken from 2–3 different geographic locations. A total of 44 individuals were used for ARISA. Individuals were treated separately in this portion of the study (i.e., DNA from multiple individuals was not pooled) and each bee was considered an independent sample (Table 1).

Ethics statement

Our collections included no endangered invertebrate species. No national permits were required to collect the bumble bee samples from public lands in the United States, Europe, Argentina, and other locations. For the samples collected from China, permits were obtained via Dr. Xue-xin Chen, Professor of Entomology, Institute of Insect Science, Zhejiang University, Zijingang Campus, Hangzhou 310058, China.

Gut and DNA extraction

Prior to abdomen excision, each bee was kept either in 95% (vol/vol) ethanol at 4°C or at –80°C. To remove the entire gut from each bee, we surface-sterilized the bee in 70% alcohol for 30 s. This was followed by a 5-min washing in 0.1% of Triton-X-100 and two rinses in distilled water. Gut tissue was then excised in Ringer’s solution under a stereomicroscope. Total DNA from the excised gut was extracted using the FastDNA SPIN kit for soil (MP Biomedicals).

Pyrosequencing library preparation

Based on Wang and Qian [38], we designed and optimized primers that target the V4–V6 variable regions of bacterial 16S rRNA. These primers (E515f: GTGCCAGCAGCCGCGGTA and E1063r: CTCACGRCACGAGCTGACG) span approximately 567 bp of the rRNA and produce amplicon sizes that are suitable for Roche 454 GS FLX+ sequencing. Further, these variable rRNA regions have been shown to be useful for taxonomic assignments and phylogenetic placement [39,40]. To produce sequencing libraries, we conducted two rounds of PCR using

pooled DNA as templates (for details, see [supporting information](#)). For the first round of PCR, we used basic primers (as listed above) with universal tails attached to the 5' ends (for E515f, GTTGTAACGACGGCCAGT, and for E1063r, CACAGGAAACAGCTATGACC). Upon completion of the first round of PCR, we took 5–10 ng of the product and conducted a second PCR to attach Roche 454 Lib-L sequencing primers and barcodes for identification. After gel extraction (eventual amplicon size after round-two PCR ~677 bp) and quality assessment, the libraries were submitted to the University of Illinois W. M. Keck Center for Comparative and Functional Genomics for 1/4th PTP sequencing on a Roche 454 GS FLX+ machine.

Pyrosequencing data processing

After preliminary signal processing using Roche software v. 2.6, we used the mothur v 1.2 [41] suite of modules to conduct demultiplexing, quality filtering, and data processing and analysis (see [supporting information](#) for details). To align and subsequently classify the 16S rRNA sequences generated, we used a custom, taxonomically annotated 16S rRNA training set composed of: 1) a SILVA reference alignment (14,956 sequences, downloaded from <http://www.mothur.org> in September 2012) and 2) an alignment of 16S rRNA sequences ($n = 276$) derived from honey bee (*Apis mellifera*) [42]. After chimera screening and weeding, we carried out taxonomic classification of the remaining sequences using the Ribosomal Database Project's (RDP) naïve Bayesian classifier [40]. To assess the confidence of each taxonomic designation, we conducted bootstrap confidence estimation (1,000 iterations) following Wang et al. [40]. Sequences with taxonomic designations matching Cyanobacteria or chloroplast (likely contaminant sequences derived from pollens) and those represented by only one read (singletons) were removed. The remaining sequences were used to conduct average neighbor clustering [41] to assign sequences to clusters (distance threshold = 3%). From each OTU, we randomly selected one representative read and used it to conduct taxonomic classification to the family level using RDP's naïve Bayesian Classifier described above. In our custom training set, we followed a previous study [42], regarding bacterial phylotypes or clades named in honey bees (e.g., alpha-2.1), as representing family groups. We used the customized training set to provide a general picture of the classification of delimited OTUs to the family level. For OTUs with the highest relative abundance in each pyrosequencing sample, we took sequences of dominant OTUs and conducted BLASTn searches against the GenBank nr database.

16S rRNA rarefaction, phylogenetic and community analyses

To assess the completeness of bacterial OTU sampling, we carried out intra- and inter-sample rarefaction analysis (details in supporting information). To estimate a phylogeny of the 74 bacterial OTUs, we aligned the representative sequence for each OTU and aligned it against 276 near full-length 16S rRNA sequences from a previous study [42] using a template-based alignment method implemented in CRWAlign [43]. Two sequences (CP000459 and CP001562) from their original set could not be aligned using this method. Next, we used RNAalifold [44] to predict consensus secondary structure based on the aligned sequences (RIBOSUM scoring, other settings default) and incorporated this information into MrBayes 3.2 [45] for phylogeny estimation. Specifically, we identified RNA stem-pairs as those receiving covariance scores greater than or equal to 90, and coded them as such under MrBayes' doublet model (nucmodel = doublet); the remaining nucleotide bases were placed in another data partition with the following evolutionary parameters: nucmodel = 4by4, nst = 6 and rates = invgamma. We ran MrBayes for five million generations (sufficient to reach stationarity, split frequencies < 0.05), two parallel runs, three heated chains and one cold chain per run, sampling the chains every 1,000 generations, and calculated the 50% majority-rule consensus tree. Liu et al. [46] showed

that short reads, when added to a phylogeny based on full length sequences, can be placed accurately within the phylogeny.

We used the Unique Fraction (UniFrac) [47] metric to assess inter-sample microbial differences. This method employs information contained within the microbial phylogeny (i.e., the Bayesian consensus tree) to determine the fraction of evolution (branch length) that is unique to each microbial community under investigation. We performed the “all environments together” and the “each environment individually” UniFrac tests using the program’s web server [48]. The “all environments together” test determined if sequences differed significantly across all 13 pyrosequencing samples while the “each environment individually” test determined if each community was significantly different from the remaining ones. In each case, statistical significance was established by comparing the observed UniFrac distance against distances obtained when community membership of sequences was randomized (number of permutations = 100). We further calculated normalized weighted inter-sample UniFrac distances and used these distances to conduct non-metric multidimensional scaling (NMDS) and hierarchical clustering in *mothur* [41]. To assess robustness of the cluster analysis, we conducted jackknife resampling, keeping 972 sequences (= 75% of the total number of sequences in the sample with the fewest reads [“BOHEMICUS”]) for each repetition (100 total repetitions).

ARISA protocol and data analysis

We conducted ARISA following Fisher and Triplett [49] with slight modifications. We carried out PCR amplifications on 30 ng of DNA using primer pair ITSf and ITSr from a previous study [50]. Primer ITSf was 5'-labeled with the phosphoramidite dye 6-FAM. PCR conditions were set at 94°C-2 min, followed by 30 cycles of 94°C-45 s, 55°C-45 s, 72°C-1 min, and 72°C-10 min. Amplified samples were subjected to capillary electrophoresis-based fragment analysis with an ABI 3730xl genetic analyzer (Applied Biosystems Inc.). We used the Genemapper software (Applied Biosystems Inc.) to output fluorescence intensity data of all fragments that fell between 0.1 and 1 kb. We conducted noise filtering (peak height, standard deviation multiplier = 2) and fragment binning (clustering threshold = 1.2 bp) with the T-REX software [51]. Hellinger transformation [52] was applied to the dataset and transformed data were used to calculate among-sample Bray-Curtis dissimilarity measures using the “vegan” package in R [53]. Multivariate analyses of the dissimilarity measures were conducted using PRIMER6 and PERMANOVA+ (PRIMER-E Ltd.) [54]. Non-metric multidimensional scaling was used to project the relationships among samples into lower dimensional space. To study factors influencing gut microbiotas of bumble bees, we conducted permutational multivariate analysis of variance (PERMANOVA) using an unbalanced nested model that included host species (fixed effect), and sampling location (random effect nested in species) as factors [54]. The PERMANOVA test was conducted using type III sums of squares; permutations (9,999 iterations) of residuals were conducted under a reduced model.

Light and fluorescent microscopy of bumble bee gut tissues

To visualize bacteria located in portions of the ileum of bumble bees using light and fluorescent microscopy, we dissected randomly selected workers from a commercial *B. impatiens* colony obtained from Koppert Biological Systems Inc. The relatively abundant *B. impatiens* was chosen as a representative species of the genus. A previous study of honey bees showed that the ileum (together with the rectum) was a key part of the intestinal tract that harbored bacteria (bacteria were largely absent from the crop and the ventriculus/midgut) [34]. Details regarding gut sample preparation are described in the supporting information. Both light and fluorescent images were taken with a Nikon E600 microscope with universal objective lenses.

DNA sequence accession numbers

Representative sequences of the OTUs are available in GenBank under the following accession numbers: KP412082–KP412155.

Results

Bacterial phylogeny and taxonomic classification

After data quality control and filtering, we obtained a total of 106,196 pyrosequencing reads of bacterial 16S rRNA, with an average read length of 445.0 bp. Seventy four bacterial OTUs were delimited, and the number of reads included under each OTU ranged from two to 57,931 (S1 Table). Representative sequences of OTUs received high-confidence taxonomic assignments from the naïve Bayesian classifier. The average bootstrap values at the levels of Class, Order and Family were 92.4%, 88.6% and 83.5%, respectively (S1 Table). These values do not include samples that were not classified. The proportion of representative sequences lacking classification was highest at the level of Family (35.1%), followed by Order (2.7%) and Class (1.4%), reflecting our generally poor understanding of bacterial taxonomy at lower taxonomic levels (S1 Table and Table 2). One representative sequence was assigned with very low confidence to an unclassified bacterium clone (ML635J-21) that originated from a meromictic soda lake [55].

Concordance between the taxonomy assigned to sequences and their placement in the bacterial phylogeny was strong, as sequences possessing the same taxonomic classifications generally clustered together in the tree (Fig. 1, S1 Fig. and S2 Fig.). This was especially true for the full-length sequences from a previous work [42], where there were no classification-phylogeny mismatches at any taxonomic level, except for one case (at the Family level) in which a sequence assigned to Anaplasmataceae was phylogenetically nested within Rickettsiaceae, a sister Family within the same Order, Rickettsiales (S2 Fig.). Sequences assigned to Family-level bacterial phylotypes commonly found in honey bees were found in expected positions in the phylogeny. For example, sequences assigned to alpha-2.2 were found within Rhodospirillales, and those assigned to beta were found within Neisseriales (S1 Fig.). Of the 74 bacterial OTUs we delimited, representative sequences from only six (6.1%) showed classification-phylogeny mismatches, indicating that the short sequences made it more difficult to conduct classification and/or phylogeny placement (Fig. 1). These six sequences received relatively low bootstrap confidence scores from the naïve Bayesian classifier (at the Class level, average = 74.5%). Sequences belonging to OTUs with mismatches constitute a small portion of the total pool (0.24%) and thus have negligible effects on downstream analysis (S1 Table).

Sampling adequacy and community diversity

For each pyrosequencing sample, we obtained 1,296 to 12,408 high-quality reads representing seven to 19 bacterial OTUs (Table 3). Sample-based rarefaction analysis showed that OTU richness curves largely plateaued. On the other hand, inter-sample rarefaction indicated that more bacterial OTUs are likely to be found with additional pyrosequencing samples (S3 Fig.). Three pyrosequencing samples, “IMPATIENS_MI”, “PENSYLVANICUS_FL” and “TERRESTRIS_FR”, had the highest estimated bacterial richness (10.4–11.9) when sequencing depth was normalized to that of “BOHEMICUS” ($n = 1,296$), the sample with the smallest number of reads (Table 3). While each pyrosequencing sample contained several OTUs (range: 7–19), many of the OTUs were rare (Fig. 2, S4 Fig.). In any particular sample, the maximum number of OTUs with relative abundance above 1% was five (average = 2.54 ± 1.20 [standard deviation, SE]) while the average number of OTUs with relative abundance above 5% was 1.92 ± 0.95 (S1 Table).

Table 2. Results of taxonomic classification (Class to Family) conducted on representative sequences of 74 bacterial OTUs.

Class	Order	Family	Number of OTUs with this Classification	Relative abundance (%) averaged across 13 pyrosequencing samples
Actinobacteria	Bifidobacteriales	unclassified	11	2.746
Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	1	3.800
		alpha-2.2	1	0.002
Bacilli	Lactobacillales	firm-5	1	0.123
		Lactobacillaceae	2	0.007
		Leuconostocaceae	3	3.933
		unclassified	13	14.424
		unclassified	unclassified	1
Betaproteobacteria	Burkholderiales	Burkholderiaceae	1	0.064
	Neisseriales	beta	19	52.594
Flavobacteria	Flavobacteriales	Flavobacteriaceae	3	8.805
Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	9	2.315
		gamma-1	5	10.314
	Oceanospirillales	Oceanospirillaceae	3	0.869
ML635J-21	unclassified	unclassified	1	0.002

Families listed in bold correspond to phylotypes typically found in western honey bees (*Apis mellifera*).

doi:10.1371/journal.pone.0118566.t002

Gut bacterial taxa

Sequencing of all 13 pyrosequencing samples revealed bacteria from only six classes: Actinobacteria, Alphaproteobacteria, Bacilli, Betaproteobacteria, Flavobacteria and Gammaproteobacteria (Table 2). Overall, Betaproteobacteria was the most dominant class, accounting for $52.7 \pm 42.2\%$ of the bacterial sequences when abundance was averaged across 13 pyrosequencing samples. This was followed by Bacilli ($18.5 \pm 36.0\%$) and Gammaproteobacteria ($13.5 \pm 24.9\%$) (Fig. 3, S1 Table). The number of OTUs found in each class varied widely; it ranged from 20 each in Bacilli and Betaproteobacteria, to two in Alphaproteobacteria. Bacterial Orders containing OTUs with high abundance included Neisseriales (Betaproteobacteria), Lactobacillales (Bacilli) and Enterobacteriales (Gammaproteobacteria). At the opposite end of the spectrum, bacterial orders with OTUs with low overall abundance included Burkholderiales (Betaproteobacteria), Oceanospirillales (Gammaproteobacteria) and Bifidobacteriales (Actinobacteria) (Table 2).

Overall, only six bacterial OTUs from five different classes were found to be dominant across all 13 samples (i.e., OTUs possessing the highest relative read abundance in each pyrosequencing sample, range: 47.3%–99.3%). Based on BLASTn results, five of these OTUs have sequences that were highly similar (98.7%–100%) to named bacteria known either from the gut of bees, flowers or the immediate environment of bees (e.g., bee bread) (Table 4). The most prevalent of these dominant OTUs matched *Snodgrassella alvi*, found in 12 pyrosequencing samples. Further, it was the most dominant OTU in eight sequencing samples (Fig. 2, S4 Fig.). The other five dominant OTUs matched sequences of *Saccharibacter floricola*, *Fructobacillus tropaeoli*, *Lactobacillus kunkeei*, *Gilliamella apicola*, and an unnamed bacterium belonging to Family Flavobacteriaceae (Table 4). Other than the dominant OTUs, we uncovered some OTUs that were less dominant (rarely accounting for more than 5% of the reads in a sample), such as one classified as phylotype firm-5 (HUSOE0101BP7S6, detected in seven bee species) and a Bifidobacteriales OTU (HUSOE0101AJ02K, detected in five *Bombus* species) (S1 Table).

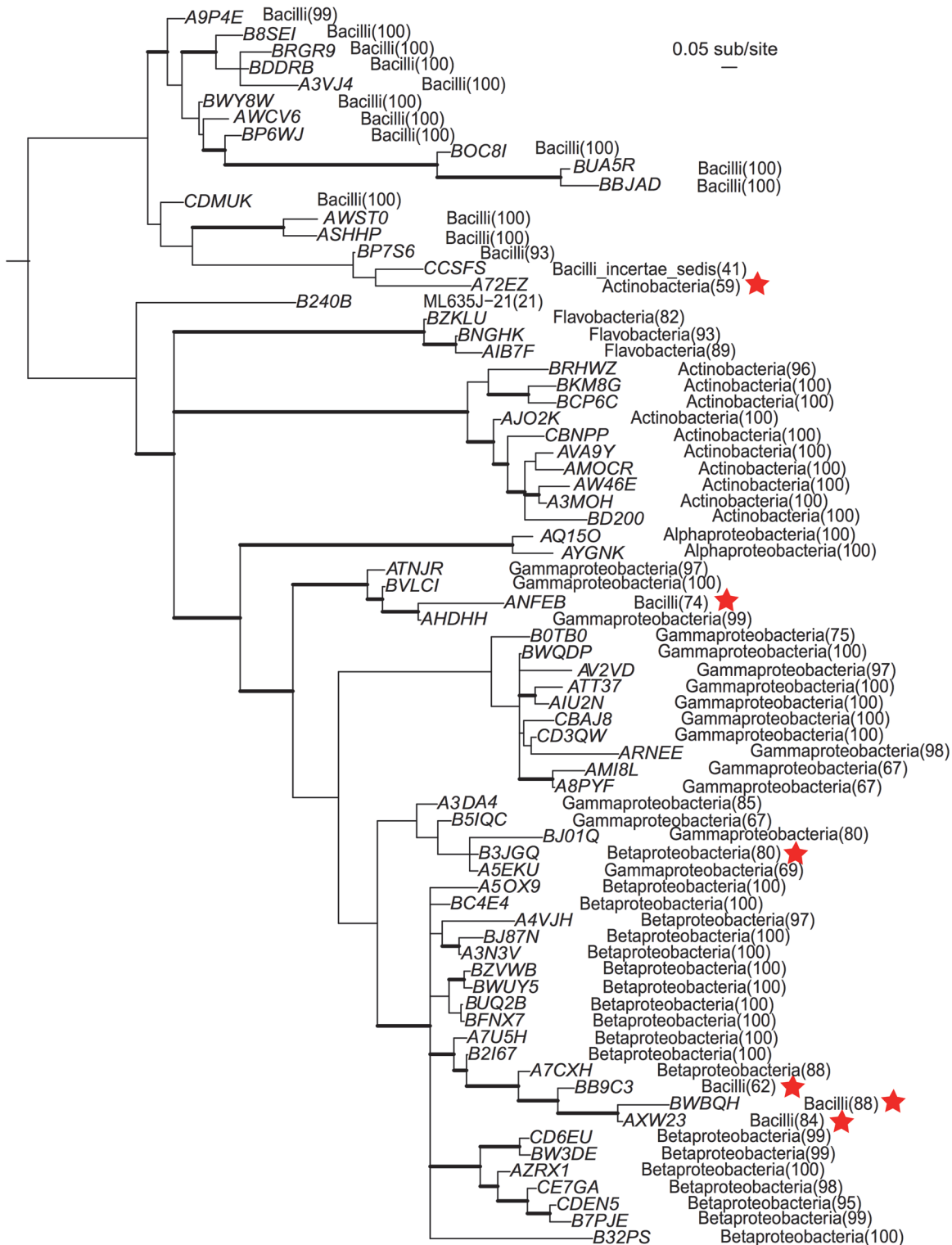


Fig 1. Consensus Bayesian inference tree showing phylogenetic relationships of representative sequences of 74 bacterial OTUs, and their classifications at the level of Class (classification bootstrap confidence in parenthesis). The IDs of the OTUs are abbreviated by only keeping their unique, last five characters (e.g., HUSOE0101A9P4E to A9P4E). The tree was inferred by combining these sequences with 274 16S rRNA sequences from [42], followed by dropping the latter sequences from the tree using the drop.tip function found in the R package “ape”. Thickened branches indicate posterior probabilities of 90% or higher; stars signify sequences that have classification-phylogeny incompatibility. See S1 Fig. for classifications of the OTUs at various taxonomic levels. See S2 Fig. for the full trees that include sequences from [42].

doi:10.1371/journal.pone.0118566.g001

Table 3. Sequencing depth and number of OTUs detected for each pyrosequencing sample.

Pyrosequencing sample ID	<i>Bombus</i> species	Number of sequences	Number of OTUs	Estimated number of OTU at 1,296 sequences (rarefaction based)
ATRATUS	<i>B. atratus</i>	9434	8	5.49
BOHEMICUS	<i>B. bohemicus</i>	1296	7	7.00
CONVEXUS	<i>B. convexus</i>	6982	8	5.93
FRATERNUS_FL	<i>B. fraternus</i>	11284	12	7.45
IMPATIENS_GA	<i>B. impatiens</i>	11556	9	5.48
IMPATIENS_MI	<i>B. impatiens</i>	4551	16	11.91
PENNSYLVANICUS_NC	<i>B. pensylvanicus</i>	5005	15	8.46
PENNSYLVANICUS_FL	<i>B. pensylvanicus</i>	12084	19	10.40
TERRESTRIS_FR	<i>B. terrestris</i>	5520	16	11.01
TERRESTRIS_ARG	<i>B. terrestris</i>	12408	13	7.47
TERRICOLA_MI	<i>B. terricola</i>	4616	10	7.96
TERRICOLA_WI	<i>B. terricola</i>	9227	11	6.08
USSURENSIS	<i>B. ussurensis</i>	12233	11	5.49

doi:10.1371/journal.pone.0118566.t003

At the level of individual pyrosequencing samples, the group of OTUs comprising phylotype beta (or *S. alvi*) appeared to be particularly diverse. Each sample that contained HUSOE0I01A3N3V (the most abundant of the beta/*S. alvi* OTUs) generally also contained at least two other closely related beta OTUs. Most of these rarer beta OTUs were unlikely to be variants caused by sequencing errors because they have non-trivial abundance (> 5 reads) and were found in more than one sample.

Community distinctiveness based on phylogenetic distances among bacteria

Based on the UniFrac “all environments together” test, there was an overall significant difference among the pyrosequencing samples, a difference that was based on a combination of their community composition and the phylogenetic distances among the OTUs within the communities ($P \leq 0.01$, after correcting for multiple comparisons). This difference was primarily driven by pyrosequencing samples “FRATERNUS_FL”, “USSURENSIS” and “CONVEXUS” (Table 5). The distinctiveness of these microbial communities was supported by NMDS and jackknife hierarchical clustering (Fig. 4 and S5 Fig.). The sample “CONVEXUS” was dominated by three bacterial OTUs, all of which were either absent or present at very low abundance in other pyrosequencing samples. These three OTUs included two closely related taxa belonging to the order of Lactobacillales (relative abundance = 50.3% and 37.5%, respectively) and one belonging to Oceanospirillales (11.2%) (Fig. 2, S4 Fig.). The sample “USSURENSIS” was dominated by a member of Lactobacillales (relative abundance = 99.3%), which occurred in significant abundance in only one other pyrosequencing sample (“TERRESTRIS_ARG”). The third highly distinctive pyrosequencing sample was “FRATERNUS_FL”. It contained two principal OTUs, belonging to Flavobacteriales (76.2%) and Enterobacteriales (17.1%), respectively. Both of these OTUs occurred in significant abundance in only 1–2 other gut bacterial communities (Fig. 2, S4 Fig.). The communities of five pyrosequencing samples (“TERRESTRIS_FR”, “IMPATIENS_GA”, “PENSYLVANICUS_FL”, “PENSYLVANICUS_NC”, and “ATRATUS”) were highly similar; they were all dominated (relative abundance > 94%) by an OTU whose representative sequence (HUSOE0I01A3N3V) closely matched that of *S. alvi* (see Gut bacterial taxa).

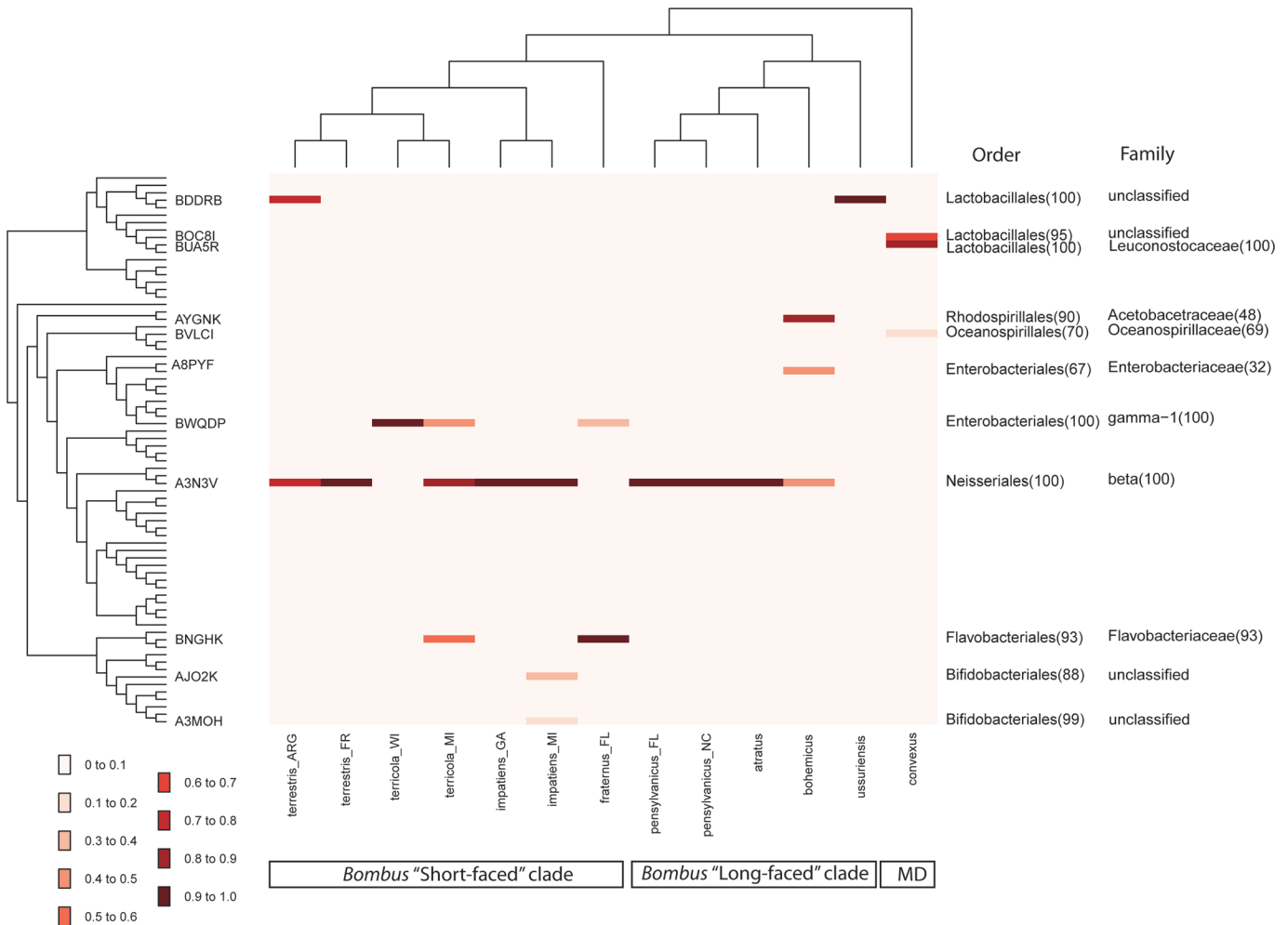


Fig 2. Heatmap showing relative abundance of each bacterial OTU (rows) in each of the pyrosequencing sample (columns). On the left is a cladogram showing phylogenetic relationships of the OTUs (tip labels = representative sequence ID). The IDs of the OTUs are abbreviated by only keeping their unique, last five characters (e.g., HUSOE0101BDDRB to BDDRB). On the right are classifications of the OTUs at two taxonomic levels, if available. For clarity, only labels for OTUs with relative abundance > 10% are shown. Atop the heatmap, a cladogram indicates phylogenetic relationships of the bee species investigated. The bars at the bottom of the graph indicate the phylogenetic clade in which each bee species falls. MD = *Mendacibombus* (subgenus). See S4 Fig. for fully labeled tips.

doi:10.1371/journal.pone.0118566.g002

Effects of species and locality on gut microbiota structure

After noise-filtering of the ARISA data, we observed an average of 125 fragments per sample. A 3-D plot of the NMDS analysis based on Bray-Curtis dissimilarity measures showed that samples tended to cluster by bumble bee species, and by locations within species (Fig. 5). The plot also showed trends resembling that of the 2-D NMDS plot based on the sequencing data (Fig. 4); microbial communities of *B. fraternus* were relatively distinct from those of the other four *Bombus* species analyzed (*B. pensylvanicus*, *B. terricola*, *B. terrestris*, and *B. impatiens*). A two-way unbalanced PERMANOVA including “species” (fixed effect) and “locality” (random effect nested in species) showed that both factors significantly influenced the gut microbiota structures of the bees examined (Species, $P = 0.0001$; Locality, $P = 0.0001$).

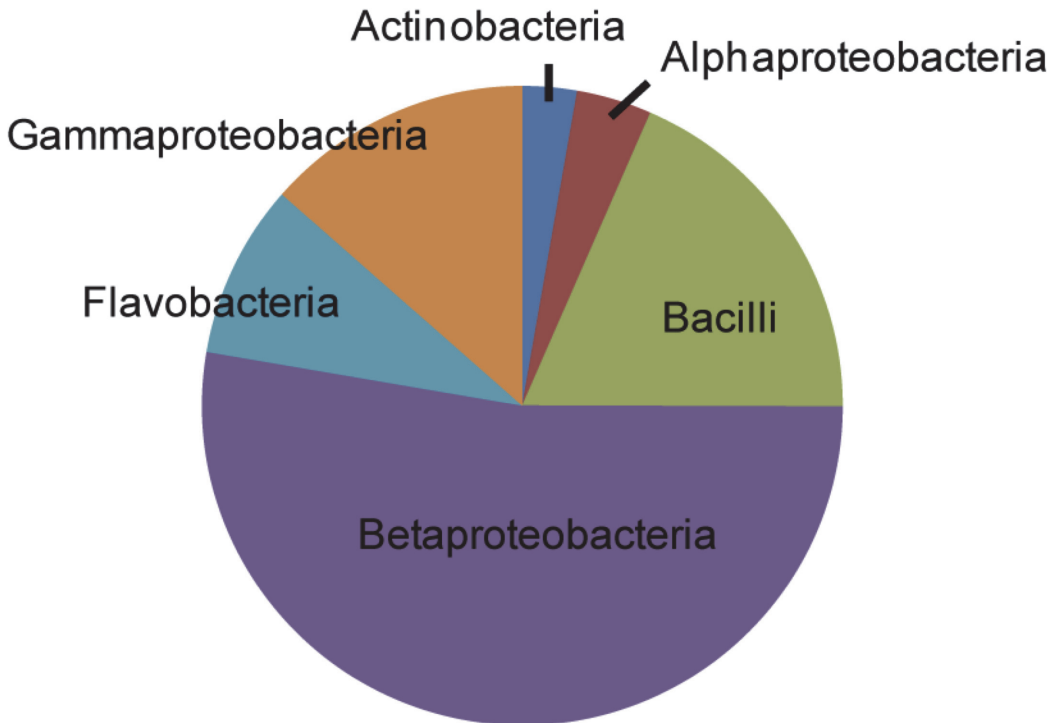


Fig 3. Relative abundance (averaged across 13 pyrosequencing samples) of different bacterial classes. The average number of sequences obtained from each pyrosequencing sample was 15,170.

doi:10.1371/journal.pone.0118566.g003

Microscopy results

In *B. impatiens*, we observed through light microscopy both yeast and bacterial cells in the lumen of the anterior part of the ileum (Fig. 6B). In the middle and posterior sections of the ileum, bacteria alone were found in the lumen (Fig. 6C and 6D). Based on fluorescent microscopy, these bacteria appeared to form a biofilm on the cuticular intima, a lining separating the lumen from the epithelial cells (Fig. 6E, F). Because of the presence of longitudinal folds in the posterior portion of the ileum, the distribution of bacteria in the posterior portion appeared to be less even when compared to the anterior portion.

Discussion

Using Roche 454 pyrosequencing, we conducted deep sampling to characterize the gut microbiota of bumble bee species from multiple subgenera. Classification of the 74 bacterial OTUs included bacterial families that were consistent with other reports on gut microbes of hymenopteran insects [15,23], suggesting good classification results using our customized training set. Regardless of the bumble bee species, the gut bacterial communities were not taxonomically diverse, with only 1–3 common (relative abundance > 5%) bacterial OTUs within the each microbiota. Among the 74 bacterial OTUs we delimited, the two most widespread and abundant bacteria (*S. alvi* and *G. apicola*) were also identified as core gut bacterial species in previous studies of honey bees and several bumble bee species [15,33]. The detection of these two bacterial species in most of our pyrosequencing samples, including many previously unexamined bumble bee species spanning large phylogenetic distances, suggests they play important functional roles in bees. These two bacterial species also have been shown to exhibit co-

Table 4. Taxonomic assignments of six OTUs based on naïve Bayesian classification of 16S rRNA sequence.

Representative sequence (preceded by HUSOE0101)	Class-Order-Family (bootstrap confidence scores in parenthesis)	Number of samples in which OTU has highest relative abundance	Number of samples in which OTU was detected	Avg. relative abundance (% of OTU across 13 samples)	ID of GenBank match	% identity (length of query in bp)	Name of bacterium	Documented source of bacterium	References describing the species
AYGNK	Alphaproteobacteria (100)- Rhodospirillales (90)- Acetobacteraceae (48)	1	2	3.8	JF794031	98.9(445)	<i>Saccharibacter floricola</i>	Pollen	[57]
BUA5R	Bacilli (100)- Lactobacillales (100)- Leuconostocaceae (100)	1	4	3.9	AB542054	100.0 (446)	<i>Fructobacillus tropaeoli</i>	Flower	[56]
BDDRB	Bacilli (100)- Lactobacillales (100)- unclassified	1	5	11.4	KF600556	100.0 (446)	<i>Lactobacillus kunkeei</i>	Gut of <i>Apis mellifera</i> , nectar, honey and beebread	[58]
A3N3V	Betaproteobacteria (100)- Neisseriales (100)- beta (100)	8	12	52.1	JQ746648	98.7(445)	<i>Snodgrassella alvi</i>	Gut of <i>Bombus vagans</i>	[76]
BNGHK	Flavobacteria (93)- Flavobacteriales (93)- Flavobacteriaceae (93)	1	7	8.8	DQ837639	99.1(443)	unnamed	Gut of <i>Apis mellifera</i>	NA
BWQDP	Gammaproteobacteria (100)- Enterobacteriales (100)- gamma-1 (100)	1	9	10.3	JQ936675	99.8(445)	<i>Gilliamella apicola</i>	Gut of <i>Bombus birmaculatus</i>	[76]

In parentheses are bootstrap confidence of each assignment. Each of these OTUs was the most dominant OTU in at least one pyrosequencing sample. Also included are information related to each OTUs closest matching GenBank sequence reported by a published paper.

doi:10.1371/journal.pone.0118566.t004

Table 5. Significance of UniFrac tests that compared a single pyrosequencing sample against all other samples combined.

Pyrosequencing sample	P-value
ATRATUS	1.00
BOHEMICUS	0.40
CONVEXUS	0.02
FRATERNUS_FL	0.05
IMPATIENS_GA	0.63
IMPATIENS_MI	0.12
PENNSYLVANICUS_FL	0.17
PENNSYLVANICUS_NC	0.38
TERRESTRIS_ARG	0.50
TERRESTRIS_FR	0.67
TERRICOLA_MI	0.73
TERRICOLA_WI	0.61
USSURENSIS	0.02

A significant P-value ($P < 0.05$, in bold) indicates that the bacterial community tested has more unique branch length than expected by chance.

doi:10.1371/journal.pone.0118566.t005

phylogenetic patterns with their hosts [32], suggesting long-term coevolution with their hosts and a tendency for vertical transmission.

Of the remaining four dominant OTUs, three have high (> 98.9%) 16S rRNA sequence similarity to previously named bacterial species: *S. floricola*, *F. trophaeoli* and *L. kunkeei*. These three bacteria are aerotolerant, osmophilic/osmotolerant lactic acid or acetic acid bacteria that, unlike *S. alvi* and *G. apicola*, were discovered and described from environments associated with bees, including flowers, hive, bee bread (processed, packed pollens) and grape juice [56–59]. This suggests that some of the gut bacteria have greater physiological tolerance for living outside the gut and may be amenable to environmental transmission. Our ARISA results showing that not only species but also geographic location of the host affects its gut bacterial community, again supports existing evidence that at least some members of the gut microbial community appear to be environmentally transmitted [60–63]. This may explain the relatively different gut microbial communities found between populations of same *Bombus* species collected in different geographic regions, such as the microbiota of *B. terrestris* from France and Argentina (Fig. 4 and S5). European colonies of *B. terrestris* were shipped to South America for commercial pollination and became naturalized since the 1990s. The species then spread to Argentina, and may have been exposed to environmental resources considerably different to those in Europe [64,65].

In identifying common bacteria found in the gut of bumble bees, we can rely on previously conducted *in vivo* and *in vitro* studies (often focusing on related bacteria in honey bees) to infer the roles they play in contributing to colony and individual health. Outside of the gut, bacteria such as *L. kunkeei* and *S. floricola* may benefit bumble bees by producing weak acids that aid in the fermentation and preservation of stored pollens in the nest [66,67]. The gut is likely to be an important source of these bacteria because worker bumble bees first moisten collected pollen with regurgitated nectar before packing it onto the hind leg pollen-carrying structure (corbicula) [59,68]. In the ileum, bacteria may enhance nutrient absorption by helping to break down food molecules. This is supported by a functional genomics study that found that the gut microbiome of honey bees (which share a number of gut bacterial species with bumble

(A) Minimum stress = 0.05

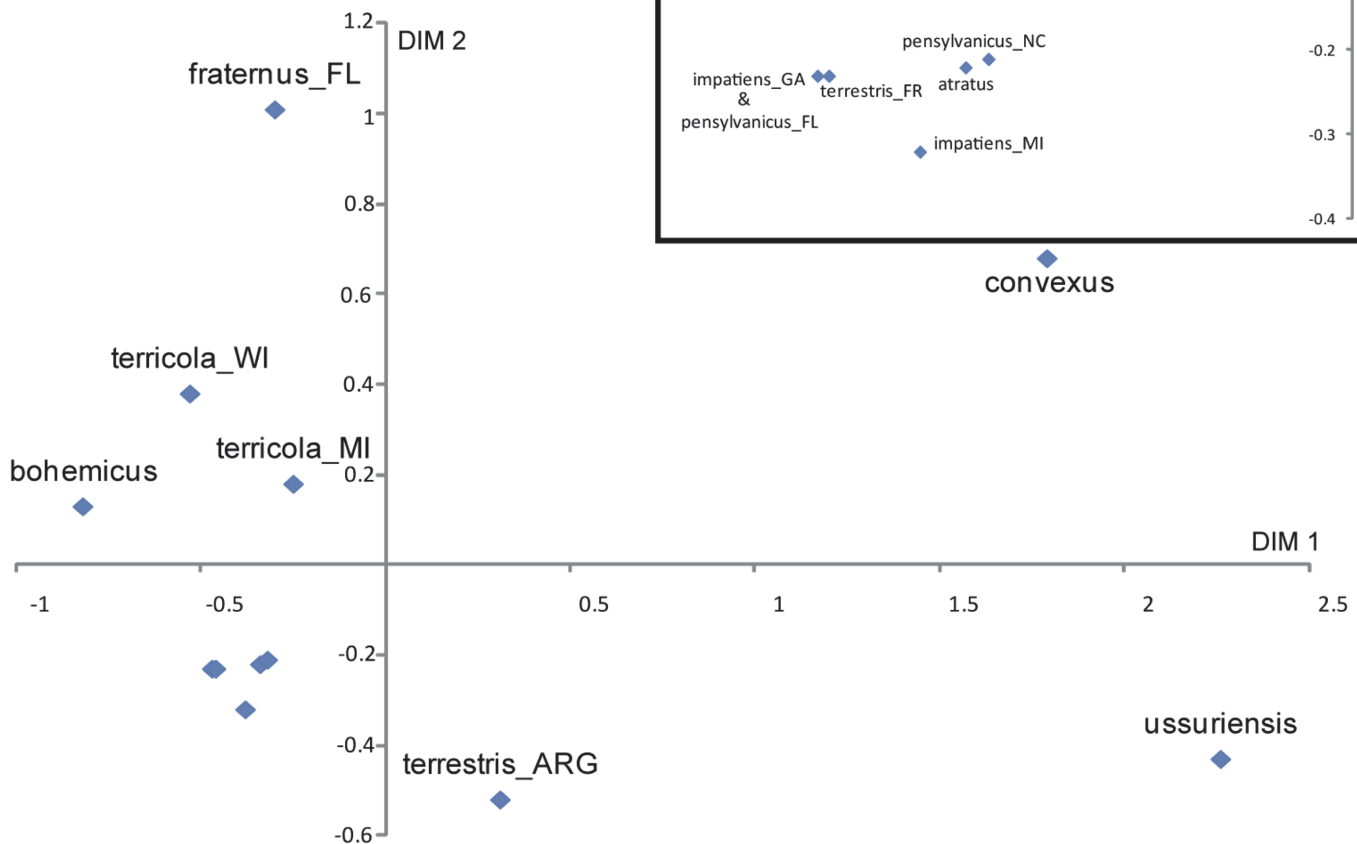


Fig 4. Plot of pyrosequencing samples in the 2-D nMDS space based on inter-sample UniFrac distances (A). Minimum stress after 999 iterations is 0.05. The cluster of samples in the lower-left quadrant (ones without labels) is expanded in the top right corner (B).

doi:10.1371/journal.pone.0118566.g004

bees) was enriched for genes associated with carbohydrate breakdown (including those involved in pectin breakdown) and transport [17]. In addition to nutrition, evidence is mounting that gut bacteria benefit their bee hosts by contributing to the latter’s defenses against non-native bacteria and protistan enemies. For example, Vasquez et al. [19] found that lactic acid bacteria (*Lactobacillus* spp. and *Bifidobacterium* spp.) isolated from honey bees have inhibition properties on microorganisms such as yeasts and other bacteria (including *Melissococcus pluto-nius*, the causative agent of European foulbrood), with *L. kunkeei* (dominant in “USSURENSIS” and “TERRESTRIS_ARG”) showing the strongest effects. Similarly, studies have shown that bacteria cultures from honey bee crops or larvae inhibited the growth of bacterial pathogens of bees [69,70]. As one of the few studies that focused on bumble bee protistan parasites, Koch and Schmid-Hempel [12] showed that bees first treated with antibiotics then subsequently fed with feces from nest mates have lower parasite load (Trypanosomatidae: *Crithidia bombi*) compared to bees not given viable bacterial inoculation. While the study was groundbreaking, the mechanisms by which gut bacteria exert their protective functions against enteric protists are not entirely clear. Similar to honeybees, we found that bacteria in the ileum of the bumble bee species *B. impatiens* forms biofilms on the inside surface of the gut epithelium (Fig. 6). This

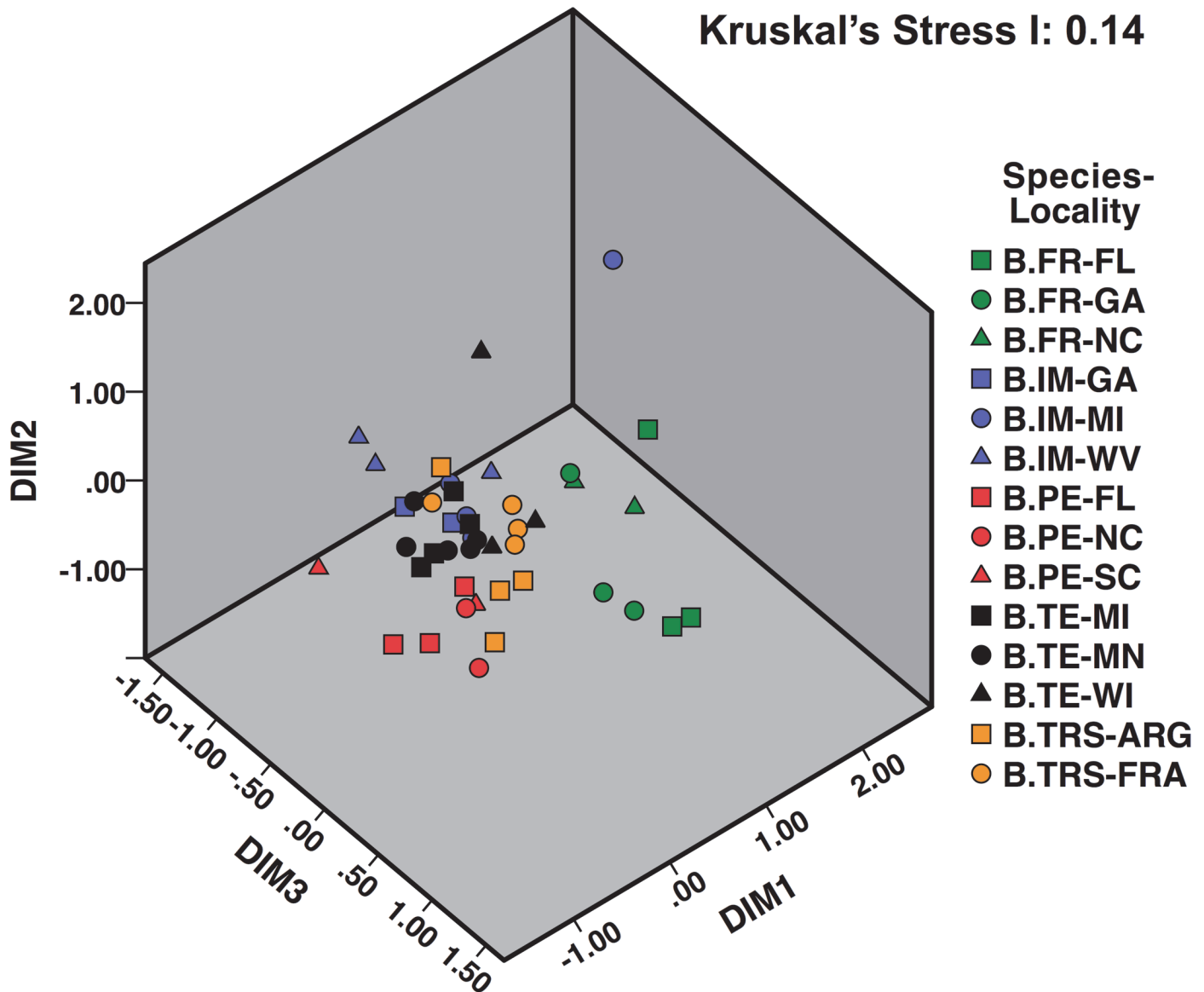


Fig 5. Three-dimensional (DIM1–3) nMDS ordination of ARISA profiles of different bumble bee gut DNA samples. Corresponding bumble bee species and populations (collection sites) are illustrated by the colors and shapes of the markers on the right. Abbreviations: B.FR = *B. fraternus*; B.IM = *B. impatiens*; B.PE = *B. pensylvanicus*; B.TE = *B. terricola*; B.TRS = *B. terrestris*. Abbreviations following species names represent different collection sites, sorted by states in the US (FL, GA, MI, MN, NC, SC, WI, and WV), or across countries (ARG and FRA). ARG: Argentina, FRA: France.

doi:10.1371/journal.pone.0118566.g005

suggests that bumble bee gut microbiotas may have specific functions in their hosts, such as providing nutritional benefits or defense against intestinal pathogens. Future research in these areas will have important implications for the conservation of corbiculate bees worldwide, as numerous species and populations are purportedly threatened by invasive parasites or deficient nutrition [71–73].

In addition to the dominant bacterial OTUs, we also uncovered numerous rare ones. It is possible that these could be the result of sequencing errors despite our stringent computational pipeline, removal of singletons and the clustering of sequences into OTUs. The inflation of microbial diversity as a result of sequencing errors is a recognized methodological artifact of next-generation sequencing [74]. Nonetheless, several of the rare OTU appear authentic and likely

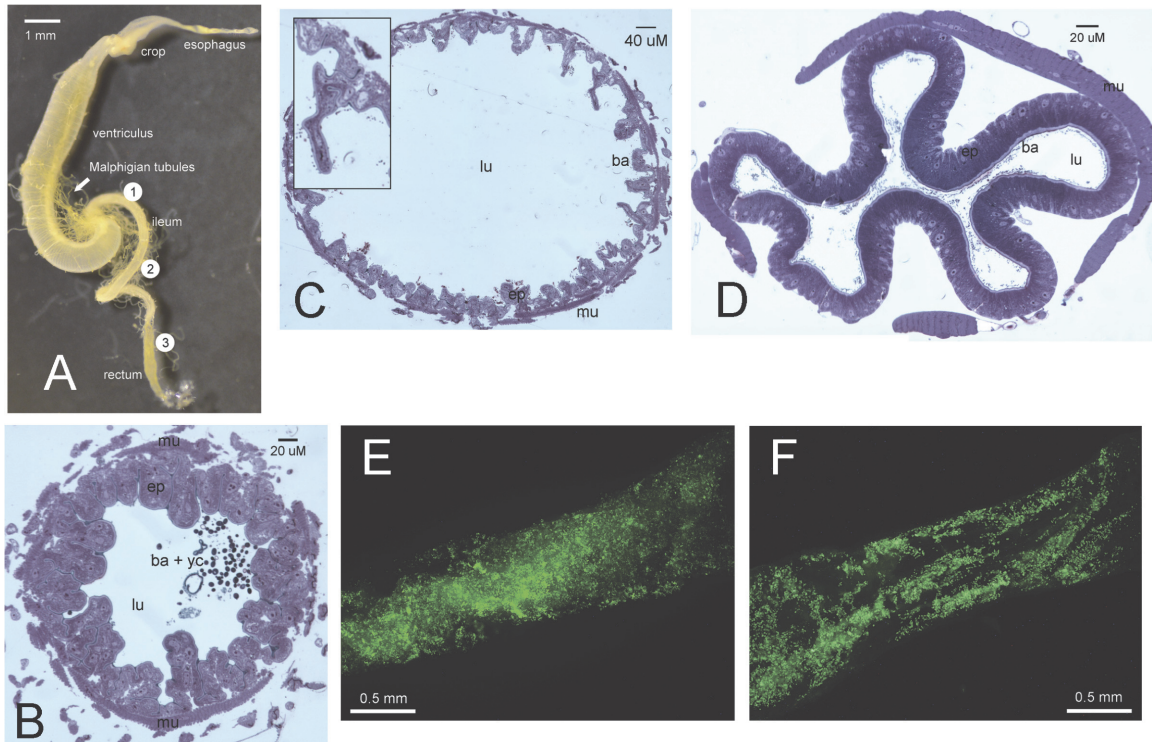


Fig 6. Images of *B. impatiens* gut (A) from light (B, C and D) and fluorescent microscopy (E and F). (A) shows the whole alimentary canal and general locations where sectioning and imaging were made (multiple individuals were used to create the dissected gut and microscopy images). Section made at (1), (2) and (3) were used to produce images (B), (C) and (D), respectively. Fluorescent images (E) and (F) were produced by visualizing a portion of the ileum located between (1) and (2), and between (2) and (3), respectively. (B) shows both yeast and bacterial cells in the lumen, whereas (C) and (D) show only bacterial cells in the lumen. Villi and longitudinal folds composed of epithelial cells can be observed in (C) and (D), respectively. Insert in (C) shows a villus and bacterial cells (black dots) in greater detail. In (E) and (F), each fluorescing green spot represents one bacterial cell. Abbreviations: ba = bacteria; ep = epithelial cells; lu = lumen; mu = transverse muscle; yc = yeast cell.

doi:10.1371/journal.pone.0118566.g006

have important functions because of their ubiquity or moderate abundance in a small number of samples. Rare OTUs classified under beta/*S. alvi* are particularly interesting because a number of them always coexist with the dominant beta/*S. alvi* OTU when the latter is present. A similar situation was reported in honey bees in which a number of closely related *S. alvi* genomes coexisted at the level of colony, if not at the level of individual bees [17]. Given that gene content of bacteria can differ significantly even when their 16S rRNA sequences are very similar [75], these minor beta/*S. alvi* variants may play complementary functional roles to the dominant bacterial strains.

Supporting Information

S1 Fig. Phylograms (Bayesian consensus trees) based on 16S rRNA alignment. The trees show classifications (A—class, B—order, C—family) of representative sequences of 74 bacterial OTUs delimited in this study. Numbers in parentheses next to names indicate bootstrap confidence scores of the classifications.

(TIF)

S2 Fig. Consensus Bayesian inference phylogenetic tree based on a bacterial 16S rRNA alignment. The tree comprises: 1) 274 full or almost full-length sequences from a previous work (7) and 2) 74 sequences representing 74 bacterial OTUs (this study). The 274 sequences (NR) are part of a training set that also contains sequences from the SILVA database. Tip labels

for the trees denote: (A) sequence ID and (B-D) taxonomic information at the class, order and family levels. Taxonomic information is either provided (NR sequences) or generated using a naïve Bayesian classifier (this study, all sequence ID's start with "HUSOE"). Taxonomic information may be missing for some NR sequences at lower taxonomic levels. Sequences from this study may be "unclassified" because of missing taxonomic information in the training set. In (A), node labels represent Bayesian posterior probabilities of nodes. For clarity, labels for some nodes closer to the tips are not shown. Sequences with a star (sequences from this study, $n = 6$) or circle (NR sequences, $n = 1$) have mismatches between their placement on the phylogenetic tree and their taxonomic classification at one or more levels. For sequences generated in this study, numbers in parentheses represent bootstrap confidence scores for taxonomic classification (B-D). Tree is midpoint rooted.

(PDF)

S3 Fig. Rarefaction curves showing saturation of sampling at the level of reads and pyrosequencing samples. (A) Individual rarefaction curves of 13 pyrosequencing samples showing estimated number of OTUs detected for each additional 100 sequences. (B) Rarefaction curve showing the number of distinct OTUs discovered as more sequencing samples were added. Error bars depict 95% confidence intervals.

(PDF)

S4 Fig. Heatmap showing relative abundance of each bacterial OTU (rows) in each of the sequencing samples (columns). On the left is a cladogram showing phylogenetic relationships of the OTUs (tip labels = representative sequence ID). On the right are classifications of the OTUs at two taxonomic levels (left = order, right = family), if available. Atop the heatmap, a cladogram indicates phylogenetic relationships of the bee species investigated.

(PDF)

S5 Fig. Hierarchical clustering of bumble bee gut microbial community profiles. The clustering is based on the profiles' normalized weighted inter-sample unique fraction (UniFrac) distances derived from pyrosequencing data. Jackknife resampling was carried out (values next to each branch) to assess the robustness of the cluster analysis.

(PDF)

S1 File. Supplemental Methods.

(DOCX)

S1 Table. Taxonomic classification of representative sequences of 74 OTUs obtained from 16S rRNA amplicon sequencing of 13 pyrosequencing samples. Also provided are number of sequences that belong to each OTU. Numbers in parenthesis next to each taxonomic classification represents bootstrap confidence.

(PDF)

Acknowledgments

We thank J. Lozier, I. Stewart, C. Morales and R. Thorp for assistance in collecting samples. We also thank I. Newton for providing newly generated honey bee 16S rRNA sequence and a pre-publication classification training set. We are indebted to J. Nardi for conducting the microscopy work.

Author Contributions

Conceived and designed the experiments: HCL MJS SAC. Performed the experiments: HCL CCC. Analyzed the data: HCL CCC. Contributed reagents/materials/analysis tools: HCL CCC. Wrote the paper: HCL CCC SAC.

References

1. Fraune S, Bosch TCG (2010) Why bacteria matter in animal development and evolution. *BioEssays* 32: 571–580. doi: [10.1002/bies.200900192](https://doi.org/10.1002/bies.200900192) PMID: [20544735](https://pubmed.ncbi.nlm.nih.gov/20544735/)
2. Round JL, Mazmanian SK (2009) The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 9: 313–323. doi: [10.1038/nri2515](https://doi.org/10.1038/nri2515) PMID: [19343057](https://pubmed.ncbi.nlm.nih.gov/19343057/)
3. Gil R, Sabater-Munoz B, Latorre A, Silva FJ, Moya A (2002) Extreme genome reduction in *Buchnera* spp.: Toward the minimal genome needed for symbiotic life. *Proc Natl Acad Sci USA* 99: 4454–4458. PMID: [11904373](https://pubmed.ncbi.nlm.nih.gov/11904373/)
4. Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, et al. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450: 560–565. PMID: [18033299](https://pubmed.ncbi.nlm.nih.gov/18033299/)
5. Rolfe RD (1997) Colonization resistance. In: Mackie RI, White B.A. and Isaacson R.E., editor. *Gastrointestinal microbiology*. London: Chapman and Hall. pp. 501–536.
6. Reid G, Howard J, Gan BS (2001) Can bacterial interference prevent infection? *Trends Microbiol* 9: 424–428. PMID: [11553454](https://pubmed.ncbi.nlm.nih.gov/11553454/)
7. King JG, Hillyer JF (2012) Infection-induced interaction between the mosquito circulatory and immune systems. *PLoS Pathog* 8: e1003058. doi: [10.1371/journal.ppat.1003058](https://doi.org/10.1371/journal.ppat.1003058) PMID: [23209421](https://pubmed.ncbi.nlm.nih.gov/23209421/)
8. Greenberg B, Kowalski JA, Klowden MJ (1970) Factors affecting the transmission of *Salmonella* by flies: Natural resistance to colonization and bacterial interference. *Infect Immun* 2: 800–809. PMID: [16557919](https://pubmed.ncbi.nlm.nih.gov/16557919/)
9. Jarosz J (1979) Gut flora of *Galleria mellonella* suppressing ingested bacteria. *J Invertebr Pathol* 34: 192–198. PMID: [119813](https://pubmed.ncbi.nlm.nih.gov/119813/)
10. Veivers PC, O'Brien RW, Slaytor M (1982) Role of bacteria in maintaining the redox potential in the hindgut of termites and preventing entry of foreign bacteria. *J Insect Physiol* 28: 947–951.
11. Dillon RJ, Charnley AK (1986) Inhibition of *Metarhizium anisopliae* by the gut bacterial flora of the desert locust, *Schistocerca gregaria*: Evidence for an antifungal toxin. *J Invertebr Pathol* 47: 350–360.
12. Koch H, Schmid-Hempel P (2011) Socially transmitted gut microbiota protect bumble bees against an intestinal parasite. *Proc Natl Acad Sci USA* 108: 19288–19292. doi: [10.1073/pnas.1110474108](https://doi.org/10.1073/pnas.1110474108) PMID: [22084077](https://pubmed.ncbi.nlm.nih.gov/22084077/)
13. Jeyapakash A, Hoy MA, Allsopp MH (2003) Bacterial diversity in worker adults of *Apis mellifera capensis* and *Apis mellifera scutellata* (Insecta: Hymenoptera) assessed using 16S rRNA sequences. *J Invertebr Pathol* 84: 96–103. PMID: [14615218](https://pubmed.ncbi.nlm.nih.gov/14615218/)
14. Mohr KI, Tebbe CC (2006) Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. *Environ Microbiol* 8: 258–272. PMID: [16423014](https://pubmed.ncbi.nlm.nih.gov/16423014/)
15. Martinson VG, Danforth BN, Minckley RL, Rueppell O, Tingek S, et al. (2011) A simple and distinctive microbiota associated with honey bees and bumble bees. *Mol Ecol* 20: 619–628. doi: [10.1111/j.1365-294X.2010.04959.x](https://doi.org/10.1111/j.1365-294X.2010.04959.x) PMID: [21175905](https://pubmed.ncbi.nlm.nih.gov/21175905/)
16. Anderson KE, Sheehan TH, Eckholm BJ, Mott BM, DeGrandi-Hoffman G (2011) An emerging paradigm of colony health: Microbial balance of the honey bee and hive (*Apis mellifera*). *Insectes Soc* 58: 431–444.
17. Engel P, Martinson VG, Moran NA (2012) Functional diversity within the simple gut microbiota of the honey bee. *Proc Natl Acad Sci USA* 109: 11002–11007. doi: [10.1073/pnas.1202970109](https://doi.org/10.1073/pnas.1202970109) PMID: [22711827](https://pubmed.ncbi.nlm.nih.gov/22711827/)
18. Koch H, Schmid-Hempel P (2012) Gut microbiota instead of host genotype drive the specificity in the interaction of a natural host-parasite system. *Ecol Lett* 15: 1095–1103. doi: [10.1111/j.1461-0248.2012.01831.x](https://doi.org/10.1111/j.1461-0248.2012.01831.x) PMID: [22765311](https://pubmed.ncbi.nlm.nih.gov/22765311/)
19. Vasquez A, Forsgren E, Fries I, Paxton RJ, Flaberg E, et al. (2012) Symbionts as major modulators of insect health: Lactic acid bacteria and honeybees. *PLoS ONE* 7: e33188. doi: [10.1371/journal.pone.0033188](https://doi.org/10.1371/journal.pone.0033188) PMID: [22427985](https://pubmed.ncbi.nlm.nih.gov/22427985/)

20. Koch H, Cisarovsky G, Schmid-Hempel P (2012) Ecological effects on gut bacterial communities in wild bumblebee colonies. *J Anim Ecol* 81: 1202–1210. doi: [10.1111/j.1365-2656.2012.02004.x](https://doi.org/10.1111/j.1365-2656.2012.02004.x) PMID: [22708631](https://pubmed.ncbi.nlm.nih.gov/22708631/)
21. Babendreier D, Joller D, Romeis J, Bigler F, Widmer F (2007) Bacterial community structures in honeybee intestines and their response to two insecticidal proteins. *FEMS Microbiol Ecol* 59: 600–610. PMID: [17381517](https://pubmed.ncbi.nlm.nih.gov/17381517/)
22. Klee J, Besana AM, Genersch E, Gisder S, Nanetti A, et al. (2007) Widespread dispersal of the microsporidian *Nosema ceranae*, an emergent pathogen of the western honey bee, *Apis mellifera*. *J Invertebr Pathol* 96: 1–10. PMID: [17428493](https://pubmed.ncbi.nlm.nih.gov/17428493/)
23. Cox-Foster DL, Conlan S, Holmes EC, Palacios G, Evans JD, et al. (2007) A metagenomic survey of microbes in honey bee colony collapse disorder. *Science* 318: 283–287. PMID: [17823314](https://pubmed.ncbi.nlm.nih.gov/17823314/)
24. Sammataro D, Gerson U, Needham G (2000) Parasitic mites of honey bees: Life history, implications, and impact. *Annu Rev Entomol* 45: 519–548. PMID: [10761588](https://pubmed.ncbi.nlm.nih.gov/10761588/)
25. Genersch E (2010) American foulbrood in honeybees and its causative agent, *Paenibacillus larvae*. *J Invertebr Pathol* 103, Supplement: S10–S19. doi: [10.1016/j.jip.2009.06.015](https://doi.org/10.1016/j.jip.2009.06.015) PMID: [19909971](https://pubmed.ncbi.nlm.nih.gov/19909971/)
26. Whitehorn PR, O'Connor S, Wackers FL, Goulson D (2012) Neonicotinoid pesticide reduces bumble bee colony growth and queen production. *Science* 336: 351–352. doi: [10.1126/science.1215025](https://doi.org/10.1126/science.1215025) PMID: [22461500](https://pubmed.ncbi.nlm.nih.gov/22461500/)
27. Henry M, Béguin M, Requier F, Rollin O, Odoux J-F, et al. (2012) A common pesticide decreases foraging success and survival in honey bees. *Science* 336: 348–350. doi: [10.1126/science.1215039](https://doi.org/10.1126/science.1215039) PMID: [22461498](https://pubmed.ncbi.nlm.nih.gov/22461498/)
28. Brown MJF, Paxton RJ (2009) The conservation of bees: A global perspective. *Apidologie* 40: 410–416.
29. Potts SG, Biesmeijer JC, Kremen C, Neumann P, Schweiger O, et al. Global pollinator declines: trends, impacts and drivers. *Trends Ecol Evol* 25: 345–353. doi: [10.1016/j.tree.2010.01.007](https://doi.org/10.1016/j.tree.2010.01.007) PMID: [20188434](https://pubmed.ncbi.nlm.nih.gov/20188434/)
30. Alaux C, Brunet J-L, Dussaubat C, Mondet F, Tchamitchan S, et al. (2010) Interactions between *Nosema* microspores and a neonicotinoid weaken honeybees (*Apis mellifera*). *Environ Microbiol* 12: 774–782. doi: [10.1111/j.1462-2920.2009.02123.x](https://doi.org/10.1111/j.1462-2920.2009.02123.x) PMID: [20050872](https://pubmed.ncbi.nlm.nih.gov/20050872/)
31. Bromenshenk JJ, Henderson CB, Wick CH, Stanford MF, Zulich AW, et al. (2010) Iridovirus and microsporidian linked to honey bee colony decline. *PLoS ONE* 5: e13181. doi: [10.1371/journal.pone.0013181](https://doi.org/10.1371/journal.pone.0013181) PMID: [20949138](https://pubmed.ncbi.nlm.nih.gov/20949138/)
32. Koch H, Abrol DP, Li J, Schmid-Hempel P (2013) Diversity and evolutionary patterns of bacterial gut associates of corbiculate bees. *Mol Ecol* 22: 2028–2044. doi: [10.1111/mec.12209](https://doi.org/10.1111/mec.12209) PMID: [23347062](https://pubmed.ncbi.nlm.nih.gov/23347062/)
33. Koch H, Schmid-Hempel P (2011) Bacterial communities in Central European Bumblebees: Low diversity and high specificity. *Microb Ecol* 62: 121–133. doi: [10.1007/s00248-011-9854-3](https://doi.org/10.1007/s00248-011-9854-3) PMID: [21556885](https://pubmed.ncbi.nlm.nih.gov/21556885/)
34. Martinson VG, Moy J, Moran NA (2012) Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl Environ Microbiol* 78: 2830–2840. doi: [10.1128/AEM.07810-11](https://doi.org/10.1128/AEM.07810-11) PMID: [22307297](https://pubmed.ncbi.nlm.nih.gov/22307297/)
35. Cameron SA, Lozier JD, Strange JP, Koch JB, Cordes N, et al. (2011) Patterns of widespread decline in North American bumble bees. *Proc Natl Acad Sci USA* 108: 662–667. doi: [10.1073/pnas.1014743108](https://doi.org/10.1073/pnas.1014743108) PMID: [21199943](https://pubmed.ncbi.nlm.nih.gov/21199943/)
36. Grixti JC, Wong LT, Cameron SA, Favret C (2009) Decline of bumble bees (*Bombus*) in the North American Midwest. *Biol Conserv* 142: 75–84.
37. Cameron SA, Hines HM, Williams PH (2007) A comprehensive phylogeny of the bumble bees (*Bombus*). *Biol J Linn Soc* 91: 161–188.
38. Wang Y, Qian P-Y (2009) Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies. *PLoS ONE* 4: e7401. doi: [10.1371/journal.pone.0007401](https://doi.org/10.1371/journal.pone.0007401) PMID: [19816594](https://pubmed.ncbi.nlm.nih.gov/19816594/)
39. Liu ZZ, DeSantis TZ, Andersen GL, Knight R (2008) Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers. *Nucleic Acids Res* 36.
40. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 73: 5261–5267. PMID: [17586664](https://pubmed.ncbi.nlm.nih.gov/17586664/)
41. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, et al. (2009) Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 75: 7537–7541. doi: [10.1128/AEM.01541-09](https://doi.org/10.1128/AEM.01541-09) PMID: [19801464](https://pubmed.ncbi.nlm.nih.gov/19801464/)

42. Newton I, Roeselers G (2012) The effect of training set on the classification of honey bee gut microbiota using the Naive Bayesian Classifier. *BMC Microbiol* 12: 221. doi: [10.1186/1471-2180-12-221](https://doi.org/10.1186/1471-2180-12-221) PMID: [23013113](https://pubmed.ncbi.nlm.nih.gov/23013113/)
43. Ozer S, Xu W, Gardner DP, Miranker DP, Gutell RR, et al. (2012) An accurate scalable template-based alignment algorithm. *Proceedings of the 2012 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*: IEEE Computer Society.
44. Bernhart SH, Hofacker IL, Will S, Gruber AR, Stadler PF (2008) RNAalifold: Improved consensus structure prediction for RNA alignments. *BMC Bioinformatics* 9: 474. doi: [10.1186/1471-2105-9-474](https://doi.org/10.1186/1471-2105-9-474) PMID: [19014431](https://pubmed.ncbi.nlm.nih.gov/19014431/)
45. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, et al. (2012) MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 61: 539–542. doi: [10.1093/sysbio/sys029](https://doi.org/10.1093/sysbio/sys029) PMID: [22357727](https://pubmed.ncbi.nlm.nih.gov/22357727/)
46. Liu Z, Lozupone C, Hamady M, Bushman FD, Knight R (2007) Short pyrosequencing reads suffice for accurate microbial community analysis. *Nucleic Acids Res* 35: e120. PMID: [17881377](https://pubmed.ncbi.nlm.nih.gov/17881377/)
47. Lozupone C, Knight R (2005) UniFrac: A new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71: 8228–8235. PMID: [16332807](https://pubmed.ncbi.nlm.nih.gov/16332807/)
48. Lozupone C, Hamady M, Knight R (2006) UniFrac—An online tool for comparing microbial community diversity in a phylogenetic context. *BMC Bioinformatics* 7: 371. PMID: [16893466](https://pubmed.ncbi.nlm.nih.gov/16893466/)
49. Fisher MM, Triplett EW (1999) Automated approach for ribosomal intergenic spacer analysis of microbial diversity and its application to freshwater bacterial communities. *Appl Environ Microbiol* 65: 4630–4636. PMID: [10508099](https://pubmed.ncbi.nlm.nih.gov/10508099/)
50. Cardinale M, Brusetti L, Quatrini P, Borin S, Puglia AM, et al. (2004) Comparison of different primer sets for use in automated ribosomal intergenic spacer analysis of complex bacterial communities. *Appl Environ Microbiol* 70: 6147–6156. PMID: [15466561](https://pubmed.ncbi.nlm.nih.gov/15466561/)
51. Culman SW, Bukowski R, Gauch HG, Cadillo-Quiroz H, Buckley DH (2009) T-REX: software for the processing and analysis of T-RFLP data. *BMC Bioinformatics* 10: 171. doi: [10.1186/1471-2105-10-171](https://doi.org/10.1186/1471-2105-10-171) PMID: [19500385](https://pubmed.ncbi.nlm.nih.gov/19500385/)
52. Legendre P, Gallagher ED (2001) Ecologically meaningful transformations for ordination of species data. *Oecologia* 129: 271–280.
53. Oksanen J, Blanchet FG, Kindt R, Legendre P, O'Hara RB, et al. (2011) *Vegan: Community ecology package: R package version 1.17–9*. <http://CRAN.R-project.org/package=vegan>.
54. Anderson MJ, Gorley RN, Clarke KR (2008) *PERMANOVA for PRIMER: Guide to software and statistical methods*. Plymouth, United Kingdom: PRIMER-E
55. Humayoun SB, Bano N, Hollibaugh JT (2003) Depth distribution of microbial diversity in Mono Lake, a meromictic soda lake in California. *Appl Environ Microbiol* 69: 1030–1042. PMID: [12571026](https://pubmed.ncbi.nlm.nih.gov/12571026/)
56. Endo A, Irisawa T, Futagawa-Endo Y, Sonomoto K, Itoh K, et al. (2010) *Fructobacillus tropaeoli* sp. nov., a fructophilic lactic acid bacterium isolated from a flower. *Int J Syst Evol Microbiol* 61: 898–902. doi: [10.1099/ijs.0.023838-0](https://doi.org/10.1099/ijs.0.023838-0) PMID: [20495031](https://pubmed.ncbi.nlm.nih.gov/20495031/)
57. Jojima Y, Mihara Y, Suzuki S, Yokozeki K, Yamanaka S, et al. (2004) *Saccharibacter floricola* gen. nov., sp. nov., a novel osmophilic acetic acid bacterium isolated from pollen. *Int J Syst Evol Microbiol* 54: 2263–2267. PMID: [15545468](https://pubmed.ncbi.nlm.nih.gov/15545468/)
58. Edwards CG, Haag KM, Collins MD, Hutson RA, Huang YC (1998) *Lactobacillus kunkeei* sp. nov.: a spoilage organism associated with grape juice fermentations. *J Appl Microbiol* 84: 698–702. PMID: [9674120](https://pubmed.ncbi.nlm.nih.gov/9674120/)
59. Anderson KE, Sheehan TH, Mott BM, Maes P, Snyder L, et al. (2013) Microbial ecology of the hive and pollination landscape: Bacterial associates from floral nectar, the alimentary tract and stored food of honey bees (*Apis mellifera*). *PLoS ONE* 8: e83125. doi: [10.1371/journal.pone.0083125](https://doi.org/10.1371/journal.pone.0083125) PMID: [24358254](https://pubmed.ncbi.nlm.nih.gov/24358254/)
60. McFrederick QS, Wcislo WT, Taylor DR, Ishak HD, Dowd SE, et al. (2012) Environment or kin: Whence do bees obtain acidophilic bacteria? *Mol Ecol* 21: 1754–1768. doi: [10.1111/j.1365-294X.2012.05496.x](https://doi.org/10.1111/j.1365-294X.2012.05496.x) PMID: [22340254](https://pubmed.ncbi.nlm.nih.gov/22340254/)
61. Aizenberg-Gershtein Y, Izhaki I, Halpern M (2013) Do honeybees shape the bacterial community composition in floral nectar? *PLoS ONE* 8: e67556. doi: [10.1371/journal.pone.0067556](https://doi.org/10.1371/journal.pone.0067556) PMID: [23844027](https://pubmed.ncbi.nlm.nih.gov/23844027/)
62. Fridman S, Izhaki I, Gerchman Y, Halpern M (2011) Bacterial communities in floral nectar. *Environ Microbiol Rep* 4: 97–104. doi: [10.1111/j.1758-2229.2011.00309.x](https://doi.org/10.1111/j.1758-2229.2011.00309.x) PMID: [23757235](https://pubmed.ncbi.nlm.nih.gov/23757235/)
63. Shade A, McManus PS, Handelsman J (2013) Unexpected diversity during community succession in the apple flower microbiome. *mBio* 4.

64. Torretta JP, Medan D, Abrahamovich AH (2006) First record of the invasive bumblebee *Bombus terrestris* (L.) (Hymenoptera, Apidae) in Argentina. *Trans Am Entomol Soc* 132: 285–289.
65. Morales CL, Arbetman MP, Cameron SA, Aizen MA (2013) Rapid ecological replacement of a native bumble bee by invasive species. *Front Ecol Environ* 11: 529–534.
66. Vásquez A, Olofsson TC (2009) The lactic acid bacteria involved in the production of bee pollen and bee bread. *J Apic Res* 48: 189–195.
67. Gilliam M (1997) Identification and roles of non-pathogenic microflora associated with honey bees. *FEMS Microbiol Lett* 155: 1–10.
68. Michener CD, Winston ML, Jander R (1978) Pollen manipulation and related activities and structures in bees of the family Apidae. *Univ Kansas Sci Bull* 51: 575–601.
69. Forsgren E, Olofsson T, Vasquez A, Fries I (2010) Novel lactic acid bacteria inhibiting *Paenibacillus* larvae in honey bee larvae. *Apidologie* 41: 99–108.
70. Evans J, Armstrong T-N (2006) Antagonistic interactions between honey bee bacterial symbionts and implications for disease. *BMC Ecol* 6: 4. PMID: [16551367](#)
71. Arbetman M, Meeus I, Morales C, Aizen M, Smaghe G (2013) Alien parasite hitchhikes to Patagonia on invasive bumblebee. *Biol Invasions* 15: 489–494.
72. Maharramov J, Meeus I, Maebe K, Arbetman M, Morales C, et al. (2013) Genetic variability of the neogregarine *Apicystis bombi*, an etiological agent of an emergent bumblebee disease. *PLoS ONE* 8: e81475. doi: [10.1371/journal.pone.0081475](#) PMID: [24324696](#)
73. Mao W, Schuler MA, Berenbaum MR (2013) Honey constituents up-regulate detoxification and immunity genes in the western honey bee *Apis mellifera*. *Proc Natl Acad Sci USA* 110: 8842–8846. doi: [10.1073/pnas.1303884110](#) PMID: [23630255](#)
74. Tedersoo L, Nilsson RH, Abarenkov K, Jairus T, Sadam A, et al. (2010) 454 Pyrosequencing and Sanger sequencing of tropical mycorrhizal fungi provide similar results but reveal substantial methodological biases. *New Phytol* 188: 291–301. doi: [10.1111/j.1469-8137.2010.03373.x](#) PMID: [20636324](#)
75. Lukjancenko O, Wassenaar T, Ussery D (2010) Comparison of 61 sequenced *Escherichia coli* genomes. *Microb Ecol* 60: 708–720. doi: [10.1007/s00248-010-9717-3](#) PMID: [20623278](#)
76. Kwong WK, Moran NA (2013) Cultivation and characterization of the gut symbionts of honey bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family *Neisseriaceae* of the *Betaproteobacteria*, and *Gilliamella apicola* gen. nov., sp. nov., a member of *Orbaceae* fam. nov., *Orbales* ord. nov., a sister taxon to the order *Enterobacteriales* of the *Gammaproteobacteria*. *Int J Syst Evol Microbiol* 63: 2008–2018. doi: [10.1099/ijs.0.044875-0](#) PMID: [23041637](#)