1 # Seasonally increasing parasite load is associated

2 # with microbiome dysbiosis in wild bumblebees

3

4 Mark G. Young, Josefine Just, Ye Jin Lee, Thomas McMahon, James Gonzalez, Suegene Noh,

5 and David R. Angelini *

6

7 Colby College, Department of Biology, 5700 Mayflower Hill, Waterville, ME 04901, USA

8

9 * Corresponding author email: david.r.angelini@gmail.com

10 **ORCID**

11   0009-0001-4111-9329 Mark G. Young

12   0000-0002-5342-518X   Josefine Just

13   0000-0002-9100-5282 Ye Jin Lee

14   0000-0002-7242-2718 Suegene Noh

15   0000-0002-2776-2158 David R. Angelini

16 **Keywords**

17 *Bombus*, bumble bee, microbiome, *Crithidia*, dysbiosis

18 **Author contributions**

19 - MY, JJ and DRA designed the experiments.

20 - MY conducted the main analysis, creation of draft figures, and writing of the main text.

21 - SN contributed to data analysis and manuscript editing

22 - DRA secured funding, produced some figure components and was involved in

23   manuscript editing for publication.

24 - MY, JJ, YJL, TM, JG collected specimens and were involved in preliminary data

25   analyses.

Abstract

The microbiome is increasingly recognized for its complex relationship with host fitness. Bumblebees are host to a characteristic gut microbiome community that is derived and reinforced through social contact between individuals. The bumblebee microbiome is species-poor, and primarily composed from a small number of core taxa that are associated with the greater tribe of corbiculate bees. Experimental findings support a role for the core bumblebee microbiome in resistance to severe infections by a common trypanosomal parasite, *Crithidia bombi*. However, most studies have been small in scale, often considering just one or two bumblebee species, or making use of commercially-reared bees. To better understand the microbiome diversity of wild populations, we have deeply sampled field populations of ten sympatric species found throughout central and down east Maine in a three-year microbiome field survey. We have used 16S amplicon sequencing to produce microbiome community profiles, and qPCR to screen samples for infections by *Crithidia bombi*. The breadth of our dataset has enabled us to test for seasonal and interspecific trends in the microbiome community. Controlling for these external sources of variation, we have identified microbial factors associated with infection and parasite load that support the role of the core microbiome in resistance to severe infection.

Introduction

44

45   Bumblebees are host to a characteristic gut microbiome community. As eusocial

46   corbiculate bees (alongside honey bees and stingless bees), bumblebees are primarily colonized

47   by the small set of corbiculate core genera, *Snodgrassella, Gilliamella, Bifidobacterium,* and

48   *Lactobacillus* (Firm-4 and Firm-5) [1–3]. Evolution of eusocial behavior was likely intertwined

49   with the establishment of the corbiculate core microbiome, as the core taxa do not colonize

50   related non-eusocial bees [3]. Within colonies, social behavior is the main driver of the

51   homogenization of microbiome communities. Newly eclosed bees must be raised alongside other

52   adults, or their feces, to establish the characteristic microbiome [4,5]. This is in contrast to some

53   other insects, which possess specialized bacteriocyte cells and receive their bacterial

54   endosymbionts through direct vertical inheritance. Between colonies, foraging provides a route

55   for the transmission of microbes [6]. Pollinators deposit microbes on flowers [7], and gut-adapted

56   bacteria are able to persist on flowers [8]. Conspecific microbial transmission is likely common as

57   bumblebees are able to discriminate between flowers, showing preference for specific floral

58   rewards [9–11].

59   The bumblebee microbiome is distinct from that of other eusocial bees despite the

60   potential for microbiota transmission between sympatric bee genera. Besides the corbiculate core

61   microbiota, bumblebees have symbioses with a unique set of taxa, including *Lactobacillus* Firm-

62   3, *Bombiscardovia*, and *Candidatus Schmidhempelia*[12]. Compositionally, bumblebee, honey bee,

63   and stingless bee microbiomes are unique and dissimilar in proportion to the phylogenetic

64   divergence between their hosts [12]. Reflecting host specificity, the strain-level phylogenies of two

65   bee endosymbionts, *Gilliamella* and *Snodgrassella*, correlate more strongly with the phylogenies

66   of their hosts than with their hosts' geographic ranges [12,13]. Although strains of *Snodgrassella*

67   specific to *Apis* and *Bombus* hosts are able to colonize non-native hosts under laboratory

68   conditions, they are unsuccessful at achieving high abundances and are rapidly out-competed by

69   native strains, even at numerical advantages of 10:1[14].

70   Parasitism is a focal point for bumblebee conservation and agricultural bee husbandry.

71   Throughout the 20th century, bumblebee species throughout Europe and North America have

72   contracted in range, a sign of continuing decline in colony survival [15–19]. Parasites do not pose a

73   novel threat to bumblebees, and wild bees successfully mount immune responses against a

74   diverse array of protozoa, nematodes, fungi, mites, prokaryotes, and viruses [20,21]. However,

75   synergy with other stressors exacerbates parasite virulence, causing outsized impacts on colony

76   survival. This "context-dependent virulence" [22] is exemplified by a trypanosome endemic to

77   Europe and North America, *Crithidia bombi*. Where *Crithidia* is endemic colony infection rates

78   can reach 80 to 100% by the end of summer [23–26]. The parasite spreads within nests through

79   coprophagy,[27,28] and between colonies through foraging on the same flowers [29]. Infections only

80   become severe when hosts are exposed to concurrent stressors. A major source of stress for wild

81   queen bees is hibernation, and infection by *C. bombi* reduces the likelihood of successful

82   hibernation by 40%[22]. Context-dependent virulence can be replicated experimentally. When

83   raised under favorable conditions, *Crithdia*-infected bees show no excess mortality in

84   comparison to non-infected controls. When starved, excess mortality increases to 50% [30].

85   Anthropogenic stressors, such as pesticide exposure or habitat loss, likely also drive context-

86   dependent virulence, contributing to modern declines of bumblebee populations.

87   Increasingly, it is accepted that the gut microbiome influences colony survival through its

88   role in immunity. Organisms in the core bumblebee microbiome form a biofilm along the

89   epithelium of the hindgut[31] and have genomes enriched for gut colonization factors [14,5,32,33].

90   Together, they may form a physical barrier against infections, as well as influence host immune

91   responses. In the context of *Crithidia*, experimental results support a mechanistic role in

92   resistance to severe infection. Workers raised under sterile conditions suffer far greater parasite

93   loads than workers with normal microbial communities [4]. Microbiome transplants between

94   *Crithidia*-resistant and susceptible hosts have been shown to alter transcription of immune genes,

95   with susceptible hosts adopting resistant-like gene expression after transplant [34]. In *B. impatiens*,

96   high relative abundances of the core taxa *Apibacter, Lactobacillus* Firm-5, and *Gilliamella* have

97   been shown to inhibit severe infection[35].

98   Observational studies indicate that these results may translate to wild populations [35,36].

99   Diversity of non-core taxa has been associated with increased infection load [36], and the relative

100   abundances of core taxa have been shown to have negative correlations with infection [4,36].

101   However, the studies that we have considered have been limited in scale, and poorly equipped to

102   answer questions of inter-species and inter-site heterogeneity of wild populations. They have

103   also disagreed on the significance of relationships between specific core taxa and infection

104   severity, one even finding no association between microbiome and infection at all [37]. To address

105   this issue, we have conducted a three-year field survey, deeply sampling populations across

106   coastal Maine, and exploring how microbiota change with the landscape. The scale and breadth

107   of our dataset has enabled us to examine trends of microbiome composition with *Crithidia*

108   infection, as well as isolate inter-season, inter-site, and inter-species variation. We found

109   consistent seasonal changes in bumblebee microbiome composition, and identified variation in

110   the relative abundances of core taxa among host species. Controlling for external sources of

111   variation, parasite load had an inverse relationship with the relative abundances of a few core

112   taxa, and infection was characterized by an increase of what appeared to be bacterial

113   opportunists, derived from local environments.

Results

114

Wild bumblebee species share a core microbiome community

115

116   We used 16S ribosomal RNA amplicon sequencing to profile the gut microbiome

117   communities of 638 bumblebees collected during the course of a three year field survey of

118   Maine's wild pollinators. Our sample set included workers (n=505), queens (n=45), and males

119   (n=88) of ten sympatric bumblebee species from 64 ecologically diverse sites (**Figure 1A**). We

120   did not visit all sites evenly; some were longitudinally sampled during each year of the field

121   study, while others were visited opportunistically, or just once (**Figure 1B**). We targeted the v6-

122   v8 region of 16S. Samples had a mean depth of 51,927 paired-end reads after quality control.

123   Consistent with prior studies, a small number of taxa accounted for nearly all of the

124   bumblebee microbiome. Of 2388 OTU99s detected within our sample set, just 82 were observed

125   in multiple samples at an average relative abundance greater than 1% (**Figure 1C**;

126   **Supplementary Table 1**). Samples were overwhelmingly colonized by OTUs classified as

127   Gammaproteobacteria (median 94% relative abundance), and a small number of genera

128   previously annotated as being part of the bumblebee core microbiome (**Table 1**), three of which

129   (*Snodgrassella, Candidatus Schmidhempelia,* and *Gilliamella*) accounted for 65% average

130   composition (**Figure 1D**). Samples with lower relative abundance of Gammaproteobacteria were

131   characterized by higher Faith's phylogenetic diversity (**Figure 1E**; spearman $r = -0.39$).

132   We tested for stratification of microbiome structure by host species, host caste, collection

133   site, and collection month with marginal permanova. All factors except for host caste explained

134   significant variation in the weighted UniFrac dissimilarity between samples, after multiple

135   hypothesis correction (p<0.05; **Table 2**). However, the size of the effects were small, and

136   samples did not show obvious clustering on principal coordinate plots (**Supplementary Figure**

137   **1**). To test for monotonicity, we examined correlations between β-diversity and the phylogenetic

138   dissimilarity (sum of branch lengths) between host species, the distance between sites, and

139   differences in collection dates (day-of-year) with Mantel tests (**Table 3**). The relationship

140   between β-diversity and days between collection was significant (spearman r=0.10, p < 0.001,

141   **Figure 1F**), indicating bumblebee gut microbiota changed consistently across each summer.

The phylogenetic divergence of *Bifidobacterium, Saccharibacter,* and *Gilliamella* correlate with those of their hosts

144   Many of the microbiome taxa identified in the field survey dataset were composed of

145   multiple unique OTUs, including each of the 28 taxa that were fully classified to the genus level

146   and were at least moderately prevalent (>10% of samples) (**Supplementary Figure 2**).

147   Potentially OTUs of the same genus corresponded to species or strain-level diversity. We

148   quantified the relative abundance-weighted phylogenetic dissimilarity between the genus-

149   specific fractions of samples with weighted UniFrac. For the core genera *Bifidobacterium* and

150   *Gilliamella,* there was a Spearman correlation between dissimilarity in OTU composition

151   between samples and host phylogenetic divergence, even after controlling for inter-site variation

152   (Mantel test stratified by collection site, p < 0.05, **Supplementary Table 2**). The association was

153   also significant for OTUs classified as belonging to the genus *Saccharibacter*. It is possible that

154   some or all of the OTUs classified as *Saccharibacter* belonged to the closely related core genus

155   *Parasaccharibacter*, which was not included in our classifier. These correlations indicate

156   phylogenetic diversification of endosymbionts alongside their bumblebee hosts, and that for

157   these three genera, host specificity had a greater effect on microbiome composition than local

158   transmission.

Crithidia infections are common throughout Maine and increase in prevalence over the course of the summer

161   Hosts from all three summers of the field survey were infected with the parasite *Crithidia*

162   *bombi*. The distribution of estimated infection loads was bimodal, with a left peak below our

163   predetermined detection threshold (1000 copies/ng), potentially corresponding to low-level,

164   latent infections **(Supplementary Figure 3)**. We consider samples passing our originally defined

165   threshold as positive infections in our analysis.

166   The overall *C. bombi* infection positivity rate was 45%, but varied significantly between

167   local populations. To compare between locations, we used five sites that had been deeply

168   sampled (≥15 samples/summer) over multiple summers of the field survey: two geographically

169   isolated offshore islands (Vinalhaven and Allen Island), two islands reachable by bridge (Great

170   Wass and Swans Island), and one mainland site (Colby College). The single summer positivity

171   rates were significantly different between sites (χ^2 test, $p < 2.7\times10^{-12}$) and greatest for the island

172   sites (**Figure 1B**). Over 50% of samples collected from Allen Island, Great Wass Island, and

173   Swans Island tested positive, while only 15.6% from Colby College were positive for *Crithidia*.

174   Infection positivity rate also consistently increased over the course of each summer. We

175   used a generalized linear mixed effects model to measure the dependence of infection status on

176   host collection date and caste, with random effects to capture inter-site and inter-species

177   variation **(Supplementary Table 3)**. There was no difference in infection positivity rate between

178   castes, but the odds of infection increased by 2.6% ($p < 0.001$) for each day of the summer.

179   Visually, the increase appeared to continue until collections ended in early September (**Figure**

180   **2A**), at odds with midsummer peaks in infection rate reported in Popp et al. (2012) [38]. Consistent

181   with the inter-site comparison of infection rates above, the site random effect explained a large

182   and significant proportion of total variation (likelihood ratio test; $p < 1\times10^{-12}$). To capture inter-

183   species differences, we used independent and identically distributed random intercepts (i.i.d.), as

184   well as a correlation structure proportional to the phylogenetic dissimilarity between hosts. Only

185   the i.i.d. encoding explained significant variation (LRT; $p < 0.0008$), indicating that the

186   differences in infection rate between species were driven by something other than their

187   phylogenetic relationships. The variation explained by species was relatively small, roughly 14%

188   of that explained by differences in site (**Supplementary Table 3**).

Crithidia infections were associated with changes to the relative abundances of a small number of taxa

191   The relative abundances of a small number of microbial taxa were different between

192   *Crithidia* infected and non-infected (healthy) bees, suggestive of infection-associated dysbiosis.

193   To test for differential abundance, we used linear mixed effects models, as described in Maaslin

194   [39], with fixed and random effects to control for the effects of host species, collection date, and

7

195   collection site. We tested whether *Crithidia* infection was associated with the 38 genus-level and

196   9 class-level taxonomic bins found across 10% or more samples, and an 'other' bin of less

197   prevalent or unassigned taxa.The 'other' bins contained highly diverse assemblages of low

198   abundance taxa. The genus level bin contained 620 unique taxa, representative of 2.2% of

199   average relative abundance. The class level bin contained 74 groupings, for 0.7% of average

200   relative abundance. The genera *Pseudomonas* and *Hafnia-Obesumbacterium,* as well as the

201   class-level 'other' bin had significant positive associations with infection (**Supplementary**

202   **Table 4**). Though at a weaker significance (p < 0.10), the class Gammaproteobacteria was

203   positively associated with infection, while the genera *Apibacter* and *Gilliamella* and class

204   Bacteroidia were negatively associated with infection. Qualitatively, the health-associated taxa

205   appeared to be consistent and prevalent members of the bumblebee microbiome, while besides

206   Gammaproteobacteria, infection-associated taxa were only found at high relative abundance in

207   small clusters of primarily *Crithidia*-positive samples **(Figure 2B)**. Infected samples also had

208   greater Faith's phylogenetic diversity (**Figure 2C**), indicating the presence of additional non-

209   core taxa.

210   Many taxa changed in relative abundance over the course of the summer, consistent with

211   the compositional changes noted above, but associations with collection date did not discriminate

212   between infection- and health-associated microbes. At the genus level, the infection-associated

213   *Hafnia-Obesumbacterium* and grouping of 'other' taxa both increased in relative abundance with

214   collection date, mirroring the date dependent rise in *Crithidia* infections (p< 0.05). However,

215   relative abundances of the health-associated *Apibacter* and *Gilliamella* did as well. For the full

216   microbiome, the effect of collection date was relatively large, but sporadic. Twenty-five of 38

217   genus level taxa, representative of 42% of the average relative abundance, were significantly

218   associated with date (positive or negative), while only three relatively minor classes were,

219   Bacteroidia, Bacilli, and Cyanobacteria (**Supplementary Table 4**, **Supplementary Figure 4**).

220   The random effect structures from the differential abundance tests helped to determine

221   the contributions of host specificity and local transmission to the relative abundances of specific

222   taxa. The two taxa with the largest ratio of variance explained by host vs. site were unknown

223   genera of the families Bifidobacterium and Orbaceae **(Figure 2D)**, both of which are well

224   adapted to the tribe of corbiculate bees [3]. Relative abundances of other core taxa, including

225   *Gilliamella*, *Snodgrassella*, and *Candidatus Schmidhempelia*, were also significantly dependent

226   on host phylogeny (likelihood ratio test; p< 0.05), reflecting adaptation to the bumblebee gut. In

227   contrast, neither *Pseudomonas* nor *Hafnia-Obesumbacterium* showed any association with host

228   phylogeny (p = 1), while instead being largely dependent on collection site (p < 1×10^{-6},

229   **Supplementary Table 4, Figure 2D**), suggesting they were environmentally derived

230   opportunists, rather than parasites specifically adapted to bumblebee hosts.

Severe infection is associated with increasingly dysbiotic microbiome composition

233   Among *Crithidia*-positive samples, there was great heterogeneity in infection severity.

234   Positive samples exhibited a 230-fold range in infection load (**Supplementary Figure 2**). We

235   used linear mixed effects models to assess the relationship between infection severity and the

236   relative abundances of bacterial taxa. We fit models of the same form as described above, with

237   the binary indicator for infection exchanged for log-transformed infection load. We used only the

238   subset of samples passing the positivity threshold.

239   Across the full set of taxa, there was a strong correlation between the fitted values from

240   the two modeling approaches (Pearson's r = 0.80), indicating that changes in relative abundances

241   with increased severity largely mirrored the differences between non-infected and infected bees.

242   Severity was associated with an increase in the relative abundance of *Pseudomonas* (p=1×10^{-8}),

243   and decreases in *Gilliamella* (p=0.0001) and *Apibacter* (p=0.0002). There were also additional

244   core microbiota that decreased in relative abundance with severity, that were not associated with

245   binary infection status: *Snodgrassella* (p = 0.06), *Candidatus Schmidhempelia* (p = 0.02), and an

246   unknown genus of family *Orbaceae* (p = 0.02) **(Figure 3**; **Supplementary Table 5)**. At the class

247   level, the relative abundance of Alphaproteobacteria increased with infection severity, despite

248   not showing a significant relationship with binary positivity, and the relative abundance of

249   Gammaproteobacteria decreased, despite being slightly greater in infected bees on average. The

250   fold changes associated with Gammaproteobacteria were relatively small, likely reflective of its

251   high average relative abundance in the bumblebee gut. For both differential abundance tests, the

252   bin of 'other' taxa had the largest magnitude of association with *Crithidia* infection, reflective of

253   an increase in randomness of community composition accompanying the depletion of core taxa.

254 # Discussion

255 **Evidence for a core bumblebee microbiome**

256   The findings of our Maine bumblebee field survey were consistent with a model of a

257 robust and coevolved bumblebee gut microbiome. Samples from all ten *Bombus* species were

258 primarily colonized by bacterial genera previously annotated as being part of the core bumblebee

259 microbiome, mostly of the class Gammaproteobacteria (**Table 1**; **Figure 1D**). The resolution of

260 our study is limited by our use of short read 16S amplicons, and the taxa represented in the Silva

261 16S database (release 132.99) we used for taxonomic classification. We were unable to

262 differentiate between phylotypes of the genus *Lactobacillus,* nor identify the genus

263 *Parasaccharibacter*, both considered low abundance members of the core microbiome[12]. In spite

264 of these limitations, we still retained a high degree of taxonomic sensitivity. Across the full

265 dataset, we identified 2,388 unique OTU99s, which corresponded to 659 genus-level taxonomic

266 bins. *Snodgrassella* and *Gilliamella* were the most frequently observed microbes, colonizing just

267 over 98% and 96% of samples, respectively (**Table 1**). Similarly high prevalences of these taxa

268 have been observed in other surveys of field caught bees [35,36,40].

269   The field survey dataset additionally provided evidence for microbiome divergence

270 between sympatric bumblebee species. Prior characterizations of the bumblebee microbiome

271 have focused on the genus as a whole, often contrasted against the other tribes of corbiculate

272 bees. We found that members of the 'bumblebee core microbiome' (including *Snodgrassella,*

273 *Gilliamella,* and *Candidatus Schmidhempelia*) varied in relative abundance between sympatric

274 bumblebee species. The degree of microbiome divergence was proportional to the phylogenetic

275 dissimilarity between hosts (**Figure 2D**). At a more granular level, we also found evidence of

276 host specificity of OTUs of three genera, *Bifidobacterium, Saccharibacter,* and *Gilliamella*

277 (**Supplementary Table 2**). Similar phylogenetic associations have been reported in large-scale

278 comparisons of corbiculate bees [12], and confirmed experimentally, with strains isolated from

279 bumblebees and honeybees [31]. To our knowledge this is the first observation of host specificity

280 within bumblebees. Notably, we did not observe OTU diversity within *Snodgrassella*, with one

281 major OTU accounting for >99.9% of all *Snodgrassella* relative abundance (**Supplementary**

282 **Figure 2**). This could reflect a below average mutation rate at the 16S loci chosen for

283 identification (V6-V8) or a technical artifact of the choice to cluster ASVs into OTU99s.

Gut flora change with the seasons

Seasonal variation in bumblebee ecology was reflected in the microbiome. The bumblebee life cycle can be divided into two major phases, colony formation in the summer, and hibernation in the winter. Only the gynes (reproductive females) hibernate, before emerging in the spring to form colonies of their own. Over the winter, core taxa decrease in relative abundance, while the microbiome increases in richness and evenness [13,41]. In the summer phase of the life cycle, the bumblebee microbiome undergoes selective pressure from diet, and is homogenized between colonies through pollination-mediated exchange of microbes. Consistent with local exchange, we found that the collection site of samples was a significant determinant of their microbiome communities (**Table 2**). We also found date-dependent variation in community structure that was consistent across all three years of the field survey, implicating the consistent annual trends in lifecycle and floral rewards as driving factors (**Figure 1F**; **Table 2**). Despite their different ecological roles, we found no significant inter-caste variation in microbiome composition **(Table 2)** when controlling for confounding effects of collection date, suggesting homogeneity between members of the same colonies.

Crithidia infection is associated with dysbiosis of the gut microbiome

Crithdia infections were seasonal and highly prevalent within our dataset. Positivity rate climbed linearly from around 10% in May to near 70% in September, consistent with continual parasite transmission within and between local colonies (**Figure 2A**). Infection positivity rates were also highly variable between sites, possibly as a result of variable colonization by the parasite or population bottlenecks caused by overwintering. In agreement with prior characterizations of context-dependent virulence [22] parasite load was bimodally distributed, with separate peaks possibly corresponding to mild and virulent infections (**Supplemental Figure 2**).

Parasite loads greater than our threshold for positivity (1000 copies/ ng gut) were associated with the abundance of non-core taxa. We observed significant increases in the relative abundance of 'other' taxa (too low-abundance for individual testing), as well as an increase in alpha diversity with infection **(Figure 2)**, recapitulating prior findings of associations between the non-core microbiome and *Crithidia* [35,36,42]. However, we did not find any strong candidates for indicator species for infection. The two taxa with increased relative abundance in infected

11

313   samples, *Pseudomonas* and *Hafnia-Obesumbacterium*, varied significantly between collection

314   sites, and were not consistent indicators of infection **(Figure 2D)**.

315   Changes to the core microbiome were significant in the contrast between weak and strong

316   infections. Core microbes *Gilliamella, Apibacter, Candidatus Schmidhempelia,* and an unknown

317   genus of the family *Orbaceae* were significantly negatively associated with parasite load.

318   (**Figure 3**). These changes were consistent with experimental support for the role of the core

319   microbiome in resistance to virulent infection [4,34]. However, the lack of strong associations with

320   positivity suggests that the core microbiome does little to protect against the spread of non-

321   severe infection.

Conclusion

322

323   Here, we report the results of a three year field survey of the gut microbiome dynamics of

324   wild bumblebees, in the context of infection by an endemic trypanasome, *Crithidia bombi*. We

325   found that the relative abundances of core taxa were inversely related with infection severity, and

326   that non-core microbial taxa were at increased relative abundance in infected samples.

327   Additionally, we found evidence for diversification in the microbiome among sympatric

328   bumblebee species, and patterns of seasonal variation in microbiome community structure that

329   was consistent across all three years. These results are evidence of strong associations between

330   bumblebees and their endosymbionts, and the contribution of the microbiome community to

331   resistance to severe parasitism.

332

Materials and Methods

333

Sample collection

334

335   Foraging bumblebees were collected throughout the state of Maine during summer 2017, 2018,

336   and 2019. Individuals were photographed at the time of capture and stored at -80˚C before

337   dissection to remove the gut. Species and caste were confirmed with reference to Williams et al.

338   (2014) [43]. Maxwell 16 DNA purification kits (Promega, Madison, WI) were used for DNA

339   extraction from gut samples. DNA extractions were performed separately after each field season.

340   Extracted DNA was stored at -20˚C prior to sequencing.

Screening and quantification of *Crithidia bombi*

342   We screened bumble bee gut DNA samples for *Crithidia bombi* infections with quantitative PCR

343   on a CFX96 Touch real-time thermocycler using iTaq Universal SYBR Green Supermix

344   (BioRad, Hercules, CA) and primers for *C. bombi GAPDH* (Cb'gapdh-52F

345   GCGTACCAGATGAAGTTTGATACG; Cb'gapdh-147R AAGCACATCCGGCTTCTTCA).

346   We used 1000 copies/ng DNA as a threshold for positivity. Samples were initially screened in

347   batches, in order to reduce the total number of reactions. We used a batch size of four, and ran

348   batches in duplicate. We then individually ran samples from positive batches, in triplicate.

Microbiome sequencing

350   DNA samples were shipped on ice packs to the Centre for Comparative Genomics and

351   Evolutionary Bioinformatics at Dalhousie University (Halifax, Canada). Samples were used to

352   prepare paired-end 2 × 300-bp MiSeq libraries, with PCR amplification targeting variable

353   regions 6-8 of the bacterial 16S ribosomal RNA gene [44]. Of 732 bumble bees collected, 638

354   samples passed standard benchmarks for library quality (**Supplementary Table 6**).

OTU Clustering

356   We processed the demultiplexed amplicon sequencing data with QIIME2 (v. 2019.4)[45]. We

357   trimmed the first 34 bp from each read and truncated forward and reverse mates to 280 and 240

358   bp, respectively, before forming ASVs and removing chimeras with DADA2 (v. 1.1.0)[46]. We

359   then clustered ASVs into OTU99s with VSEARCH (v. 2.21.1)[47], to reduce what was seen as

360   unnecessary dataset complexity. The ASV table produced by DADA2 contained 3,752 total

361   ASVs, including 2,766 that were only found in a single sample (74%). On average, individual

362   ASVs were only found across 4.77 of 638 samples. Clustering to 99% identity with VSEARCH

363   reduced the dimensionality to 2,388 features. After clustering, a similar proportion of features

364   were found in single samples, (1,823 ASVs, 76%), and the average prevalence of individual

365   features was slightly greater (5.69 samples/feature). Because inter-year and inter-site trends in

366   microbiome diversity were the primary focus of our analysis, we were comfortable with the

367   potential loss of some specificity provided by the unclustered ASV sequences. Since we

368   performed the secondary clustering, we refer to features as OTUs throughout the text.

Taxonomic classification

370   We classified OTU representative sequences with a Naive Bayes classifier trained on the SILVA

371   138 SSU NR database [48]. We used taxonomic labels with minimum posterior probabilities of 0.7

372   and classified OTUs to the species level when possible. The Silva database does not contain

373   representative sequences for the different *Lactobacillus* phylotypes, nor the genus

374   *Parasaccaribacter*, which are core members of the *Bombus* microbiome.

Host filtering

376   Four OTUs were assigned to the order Hymenoptera by the taxonomic classifier. We confirmed

377   the classification with BLASTn alignments to the NCBI NR database. At least one of the four

378   OTUs was present in 453 of 638 samples (71%). The mean relative abundance of the

379   Hymenoptera OTUs was 0.16%, and varied widely between samples (**Supplementary Table 1**).

380   We removed these OTU from the feature table before diversity metric calculation and

381   differential abundance testing.

Diversity calculation

383   We used the phylogeny-aware weighted UniFrac distance for β-diversity analyses. To build a *de-*

384   *novo* phylogeny, we created a multiple sequence alignment of all OTUs with MAFFT (v.

385   7.505)[49] and masked positions with conservation less than 40%. We then created an unrooted

386   tree with FastTree2 [50] and midpoint-rooted this tree. UniFrac distance calculations were

387   implemented in QIIME2. For α-diversity, we calculated Faith's phylogenetic diversity, also

388   using the rooted tree and a feature table down-sampled to a depth of 4,000 OTUs, which

389   included 631 (98.9%) of 638 samples.

Host metadata

391   We measured the dependence of microbiome composition on host species, collection site, and

392   collection date. To quantify the phylogenetic dissimilarity between bumblebee species, we used

393   the sum of branch lengths from the bumblebee phylogeny reported in Cameron and Hines (2007)

394   [51]. To quantify the difference in collection site between samples, we used the distance in

14

395   kilometers between individual bumblebee collection sites, and also categorically encoded

396   collection sites. Categorical sites had maximum radii of 5 kilometers, which has been reported to

397   be the maximum foraging range of worker bees[20]. We used day-of-year encoding for collection

398   dates, as well as categorical offsets for each year of the field survey.

399   **Statistical analysis**

400   All statistical analyses were performed in R v. 4.1.3 [52]. We used linear models to assess variation

401   in *Crithidia* infection rate, and to test for differential abundance of bacterial taxa. For both

402   applications, we fit models with the *pglmm* function from the R package *phyr* [53]. We used *pglmm*

403   for its support of generalized linear mixed effects modeling with phylogenetic random effects.

404   For modeling variation in *Crithidia* infection rates, we fit the model:

405

406     $Crithdia \sim days\_Since\_May1 + caste + (1|species\_) + (1|site\_year)$, with a binomial link function. The variables

407   were defined as:

408   *Crithdia:* Infection status (positive/negative) of individual samples

409   *days_Since_May1*: The year-independent collection date of samples, measured as the the

410   days between sample collection and May 1st of the same year.

411   *caste:* sample caste (male/worker/queen)

412   *(1|species__)*: The phylogenetic effect, consisting of independent and identically

413   distributed intercepts for bumblebee species, as well as a covariance structure

414   proportional to the phylogenetic dissimilarity between species (sum of branch lengths).

415   *(1|site_year)*: Independent and identically distributed random intercepts for collection

416   sites. Collection sites were encoded separately for each year (ex: AllenIsland_2017,

417   AllenIsland_2018, AllenIsland_2019), in order to capture year-specific variation.

418

419   Prior to log-transformation of relative abundances, we additively smoothed zero values with a

420   pseudocount equal to half the smallest non-zero relative abundance. We used Benjamini-

421   Hochberg FDR-corrected p-values to assess significance of associations. We corrected for false

422   discovery rate separately for the genus and class-level models. For all modeling applications, we

423   used likelihood ratio tests (implemented in *phyr*) to quantify the significance of the variation

424   explained by random effects. In the discussion of differential abundance testing results, by

15

425 'species random effect', we refer to the covariance structure proportional to phylogenetic

426 dissimilarities between host species. All figures were created with Python 3.8.16 (*seaborn*

427 v.0.12.2*, matplotlib* v.3.6.0).

Data availability

429 Code and data involved in this analysis are available online at

430 https://github.com/aphanotus/bombus.landscape

Acknowledgements

Works Cited

437 1. Koch, H. & Schmid-Hempel, P. Bacterial Communities in Central European Bumblebees: Low

438 Diversity and High Specificity. *Microb. Ecol.* **62**, 121–133 (2011).

439 2. Martinson, V. G. *et al.* A simple and distinctive microbiota associated with honey bees and bumble

440 bees. *Mol. Ecol.* **20**, 619–628 (2011).

441 3. Kwong, W. K. *et al.* Dynamic microbiome evolution in social bees. *Science advances* **3**, (2017).

442 4. Koch, H. & Schmid-Hempel, P. Socially transmitted gut microbiota protect bumble bees against an

443 intestinal parasite. *Proc. Natl. Acad. Sci. U. S. A.* **108**, (2011).

444 5. Martinson, V. G., Magoc, T., Koch, H., Salzberg, S. L. & Moran, N. A. Genomic Features of a

445 Bumble Bee Symbiont Reflect Its Host Environment. *Appl. Environ. Microbiol.* **80**, 3793 (2014).

446 6. Keller, A. *et al.* (More than) Hitchhikers through the network: the shared microbiome of bees and

447 flowers. *Current opinion in insect science* **44**, (2021).

448 7. Zemenick, A. T., Vannette, R. L. & Rosenheim, J. A. Linked networks reveal dual roles of insect

449 dispersal and species sorting for bacterial communities in flowers. *Oikos* **130**, 697–707 (2021).

450 8. McFrederick, Q. S. *et al.* Flowers and Wild Megachilid Bees Share Microbes. *Microb. Ecol.* **73**,

451 188–200 (2017).

452   9.   Müller, H. The Effect of the Change of Colour in the Flowers of 'Pulmonaria offcinalis' upon its
453   Fertilisers. *Nature* **28**, 81–81 (1883).

454   10. Chittka, L., Gumbert, A. & Kunze, J. Foraging dynamics of bumble bees: correlates of movements
455   within and between plant species. *Behav. Ecol.* **8**, 239–249 (1997).

456   11. Cresswell, J. E. & Robertson, A. W. Discrimination by Pollen-Collecting Bumblebees among
457   Differentially Rewarding Flowers of an Alpine Wildflower, Campanula rotundifolia
458   (Campanulaceae). *Oikos* **69**, 304–308 (1994).

459   12. Kwong, W. K. *et al.* Dynamic microbiome evolution in social bees. *Science advances* **3**, (2017).

460   13. Koch, H., Abrol, D. P., Li, J. & Schmid-Hempel, P. Diversity and evolutionary patterns of bacterial
461   gut associates of corbiculate bees. *Mol. Ecol.* **22**, 2028–2044 (2013).

462   14. Kwong, W. K., Engel, P., Koch, H. & Moran, N. A. Genomics and host specialization of honey bee
463   and bumble bee gut symbionts. *Proc. Natl. Acad. Sci. U. S. A.* **111**, (2014).

464   15. Kosior, A. *et al.* The decline of the bumble bees and cuckoo bees (Hymenoptera: Apidae: Bombini)
465   of Western and Central Europe. *Oryx* **41**, 79–88 (2007).

466   16. Colla, S. R. & Packer, L. Evidence for decline in eastern North American bumblebees
467   (Hymenoptera: Apidae), with special focus on Bombus affinis Cresson. *Biodivers. Conserv.* **17**,
468   1379–1391 (2008).

469   17. Goulson, D., Lye, G. C. & Darvill, B. Decline and conservation of bumble bees. *Annu. Rev.*
470   *Entomol.* **53**, 191–208 (2008).

471   18. Williams, P. H. & Osborne, J. L. Bumblebee vulnerability and conservation world-wide. *Apidologie*
472   **40**, 367–387 (2009).

473   19. Cameron, S. A. *et al.* Patterns of widespread decline in North American bumble bees. *Proc. Natl.*
474   *Acad. Sci. U. S. A.* **108**, 662–667 (2011).

475   20. Goulson, D. *Bumblebees: Behaviour, Ecology, and Conservation*. (OUP Oxford, 2010).

476   21. Pascall, D. J., Tinsley, M. C., Clark, B. L., Obbard, D. J. & Wilfert, L. Virus Prevalence and Genetic
477   Diversity Across a Wild Bumblebee Community. *Front. Microbiol.* **12**, 650747 (2021).

478   22. Brown, M. J. F., Schmid-Hempel, R. & Schmid-Hempel, P. Strong context-dependent virulence in a
479   host–parasite system: reconciling genetic evidence with theory. *J. Anim. Ecol.* **72**, 994–1002 (2003).

480   23. Shykoff, J. A. & Schmid-Hempel, P. Incidence and effects of four parasites in natural populations of
481   bumble bees in Switzerland. *Apidologie* **22**, 117–125 (1991).

482   24. Imhoof, B. & Schmid-Hempel, P. Colony success of the bumble bee, Bombus terrestris, in relation
483   to infections by two protozoan parasites, Crithidia bombi and Nosema bombi. *Insectes Soc.* **46**, 233–
484   238 (1999).

485   25. Averill, A. L., Couto, A. V., Andersen, J. C. & Elkinton, J. S. Parasite Prevalence May Drive the

17

Biotic Impoverishment of New England (USA) Bumble Bee Communities. *Insects* **12**, (2021).

26. Gillespie, S. Factors affecting parasite prevalence among wild bumblebees. *Ecol. Entomol.* **35**, 737–747 (2010).

27. Schmid-Hempel, P. & Schmid-Hempel, R. Transmission of a Pathogen in Bombus terrestris, with a Note on Division of Labour in Social Insects. *Behav. Ecol. Sociobiol.* **33**, 319–327 (1993).

28. Otterstatter, M. C. & Thomson, J. D. Contact networks and transmission of an intestinal pathogen in bumble bee (Bombus impatiens) colonies. *Oecologia* **154**, 411–421 (2007).

29. Durrer, S. & Schmid-Hempel, P. Shared use of flowers leads to horizontal pathogen transmission. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **258**, 299–302 (1997).

30. Brown, M. J. F., Loosli, R. & Schmid-Hempel, P. Condition-dependent expression of virulence in a trypanosome infecting bumblebees. *Oikos* **91**, 421–427 (2000).

31. Martinson, V. G., Moy, J. & Moran, N. A. Establishment of characteristic gut bacteria during development of the honeybee worker. *Appl. Environ. Microbiol.* **78**, 2830–2840 (2012).

32. Ellegaard, K. M. *et al.* Extensive intra-phylotype diversity in lactobacilli and bifidobacteria from the honeybee gut. *BMC Genomics* **16**, 284 (2015).

33. Kwong, W. K., Steele, M. I. & Moran, N. A. Genome Sequences of Apibacter spp., Gut Symbionts of Asian Honey Bees. *Genome Biol. Evol.* **10**, 1174–1179 (2018).

34. Näpflin, K. & Schmid-Hempel, P. Immune response and gut microbial community structure in bumblebees after microbiota transplants. *Proc. Biol. Sci.* **283**, (2016).

35. Mockler, B. K., Kwong, W. K., Moran, N. A. & Koch, H. Microbiome Structure Influences Infection by the Parasite Crithidia bombi in Bumble Bees. *Appl. Environ. Microbiol.* **84**, (2018).

36. Cariveau, D. P., Elijah Powell, J., Koch, H., Winfree, R. & Moran, N. A. Variation in gut microbial communities and its association with pathogen infection in wild bumble bees (Bombus). *ISME J.* **8**, 2369–2379 (2014).

37. Straw, E. A., Mesnage, R., Brown, M. J. F. & Antoniou, M. N. No impacts of glyphosate or Crithidia bombi, or their combination, on the bumblebee microbiome. *Sci. Rep.* **13**, 8949 (2023).

38. Popp, M., Erler, S. & Lattorff, H. M. G. Seasonal variability of prevalence and occurrence of multiple infections shape the population structure of Crithidia bombi, an intestinal parasite of bumblebees (Bombus spp.). *Microbiologyopen* **1**, 362 (2012).

39. Mallick, H. *et al.* Multivariable association discovery in population-scale meta-omics studies. *PLoS Comput. Biol.* **17**, e1009442 (2021).

40. Powell, E., Ratnayeke, N. & Moran, N. A. Strain diversity and host specificity in a specialized gut symbiont of honeybees and bumblebees. *Mol. Ecol.* **25**, 4461–4471 (2016).

41. Bosmans, L. *et al.* Hibernation Leads to Altered Gut Communities in Bumblebee Queens (Bombus

520   terrestris). *Insects* **9**, (2018).

521   42. Koch, H., Cisarovsky, G. & Schmid-Hempel, P. Ecological effects on gut bacterial communities in

522   wild bumblebee colonies. *J. Anim. Ecol.* **81**, 1202–1210 (2012).

523   43. Williams, P. H., Thorp, R. W., Richardson, L. L. & Colla, S. R. *Bumble Bees of North America: An*

524   *Identification Guide*. (Princeton University Press, 2014).

525   44. Comeau, A. M., Li, W. K. W., Tremblay, J.-É., Carmack, E. C. & Lovejoy, C. Arctic Ocean

526   Microbial Community Structure before and after the 2007 Record Sea Ice Minimum. *PLoS One* **6**,

527   e27492 (2011).

528   45. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using

529   QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

530   46. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat.*

531   *Methods* **13**, 581–583 (2016).

532   47. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool

533   for metagenomics. *PeerJ* **4**, e2584 (2016).

534   48. Quast, C. *et al.* The SILVA ribosomal RNA gene database project: improved data processing and

535   web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2012).

536   49. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7:

537   improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

538   50. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for

539   large alignments. *PLoS One* **5**, e9490 (2010).

540   51. Cameron, S. A., Hines, H. M. & Williams, P. H. A comprehensive phylogeny of the bumble bees

541   (Bombus). *Biol. J. Linn. Soc. Lond.* **91**, 161–188 (2007).

542   52. R Core Team. R: A Language and Environment for Statistical Computing. Preprint at

543   https://www.R-project.org/ (2022).

544   53. Li, D., Dinnage, R., Nell, L. A., Helmus, M. R. & Ives, A. R. phyr: An r package for phylogenetic

545   species-distribution modelling in ecological communities. *Methods Ecol. Evol.* **11**, 1455–1463

546   (2020).

547   # Figure Legends

548   **Main Text**
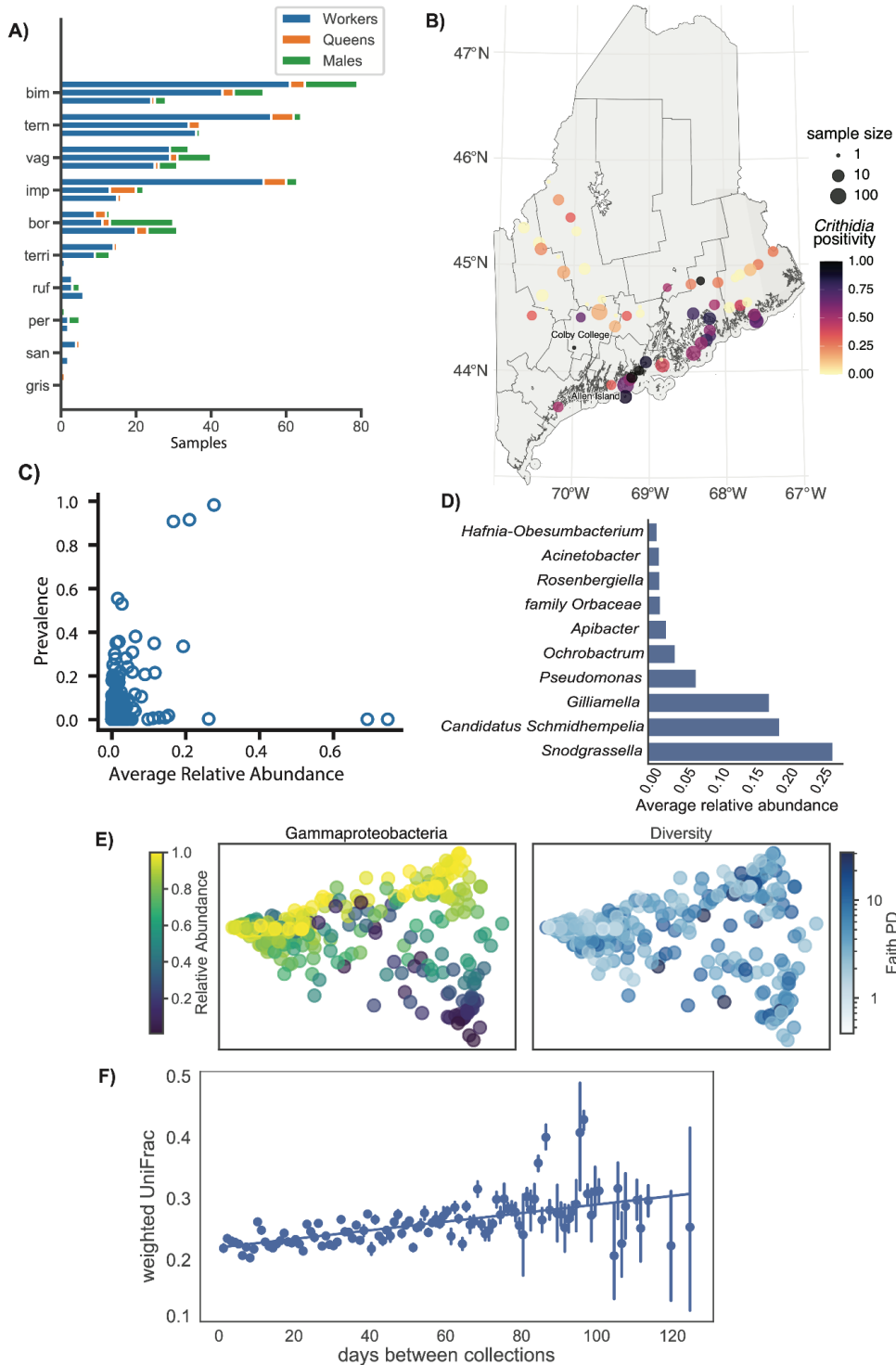
550 **Figure 1: Wild bumblebees share a genus specific microbiome community.** A) Wild species
551 and castes are unequally represented in the field survey dataset. The prevalence of species is
552 represented by horizontal bars, broken down by caste (color) and year (vertical stacks). B) Map
553 of Maine with markers for collection sites, sized by number of samples collected per site, and
554 colored by *Crithidia* infection prevalence. C) Prevalence across samples versus average relative
555 abundance distribution of OTUs. The plotted value for average relative abundance is the average
556 non-zero relative abundance. D) Average relative abundance of most frequently identified
557 microbial genera across the full dataset, including samples where they were not detected. E)
558 PCoA plot of weighted UniFrac diversity, shaded by Gammaproteobacteria relative abundance
559 (left) and Faith's phylogenetic diversity (right). Gammaproteobacteria relative abundance and
560 Faith's PD are inversely correlated (Spearman's r = -0.39). F) The difference in microbiome
561 composition between samples was dependent on the difference in their dates of collection
562 (Spearman's r = 0.10, Mantel test, p < 0.001). Difference in collection date was calculated from
563 day-of-year dates, in order to test for trends consistent across the years of the field survey.
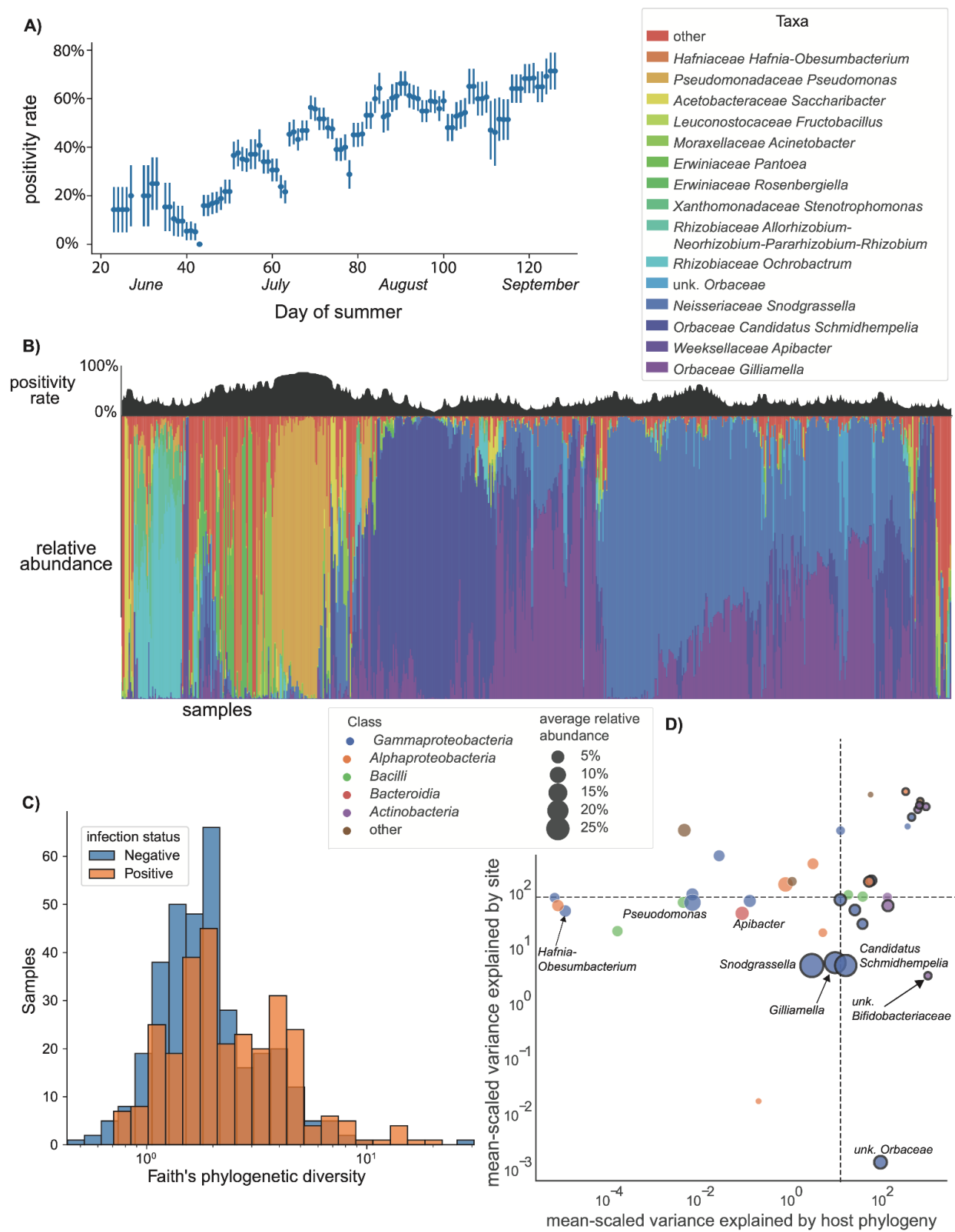
Figure 2: Bumblebee microbiome community is associated with host ecology and *Crithidia* infection status. A) *Crithdia* infection rate increases over the course of the summer. Positivity

567 rates were estimated as the proportion of positive samples collected within five days before or

568 after a given day. Rate estimates include samples from all three years of the field survey.

569 Standard deviations of estimates are shown as vertical bars. B) *Crithidia* infections are associated

570 with microbiome dysbiosis. The relative abundances of the 15 most abundant microbial taxa are

571 shown for all 638 samples. Samples are sorted along the horizontal axis by agglomerative

572 clustering on weighted unifrac distance. *Crithidia* positivity rate was estimated using a sliding

573 window with a radius of 20 samples, and is shown above the taxonomy plot. C) *Crithidia*

574 positive samples had greater Faith's phylogenetic diversity. Faith's phylogenetic diversity, a

575 phylogeny-aware alpha diversity metric, was calculated for microbiome communities with

576 qiime2. Horizontal axis is log scaled. D) Linear mixed effects models were used to assess the

577 association between the relative abundances of individual taxa with *Crithidia* infection,

578 collection date, collection site, and host species. Points represent the variance in relative

579 abundance explained by the phylogenetic and site random effects. Mean-scaled variance is

580 plotted. Medians are indicated by dashed lines. Taxa for which significant variation is explained

581 by host phylogeny ($\alpha = 0.05$) are given black outlines.

582

583
584
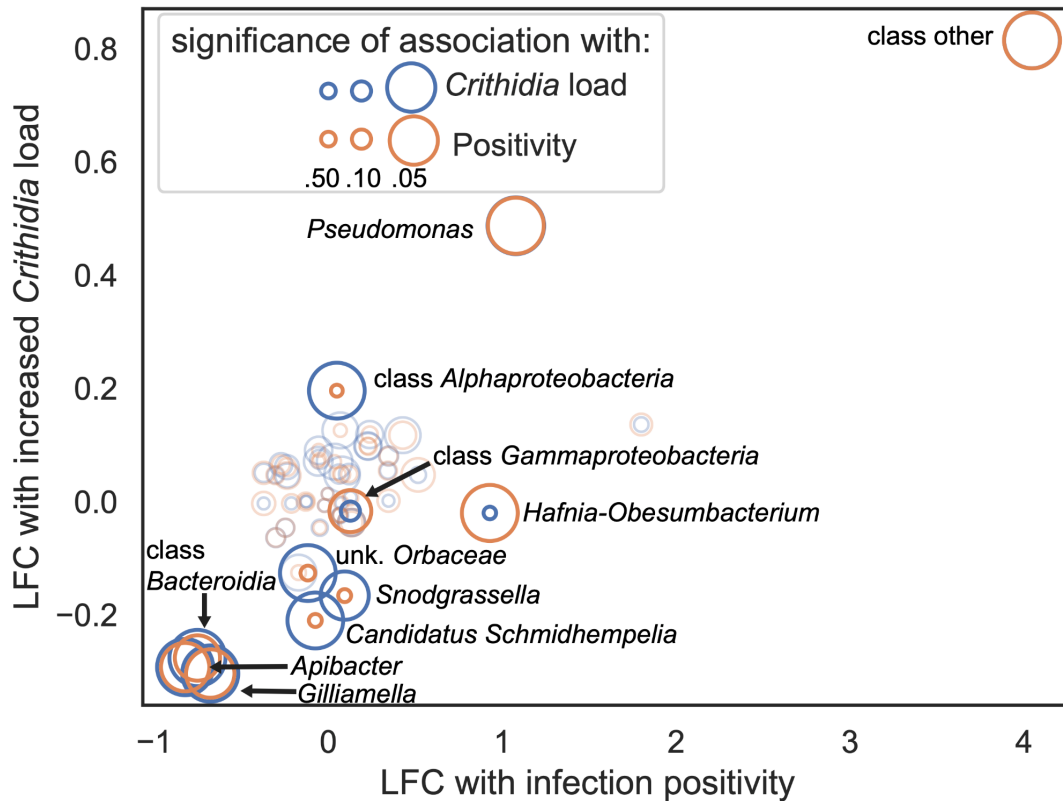
Figure 3: Bacterial taxa associated with infection were also associated with increased infection load. Differential abundance models were used to measure the difference in relative abundance of bacterial taxa between infected and non-infected samples (horizontal axis), as well as to test for changes in relative abundance with increased infection load (vertical axis). Concentric circles represent the FDR-corrected significance of the relationships between taxa relative abundance and infection (orange) and infection load (blue). Larger circles correspond to increased significance. Taxa with weak associations (p<0.1) with either infection load or positivity (or both) are indicated with darker colors. For *Pseudomonas* and 'other' the associations had equal significance, causing the circles to overlap. Differential abundance models were fit for both genus and class-level taxa, fitted values for both are included on the plot.

Supplementary Figures
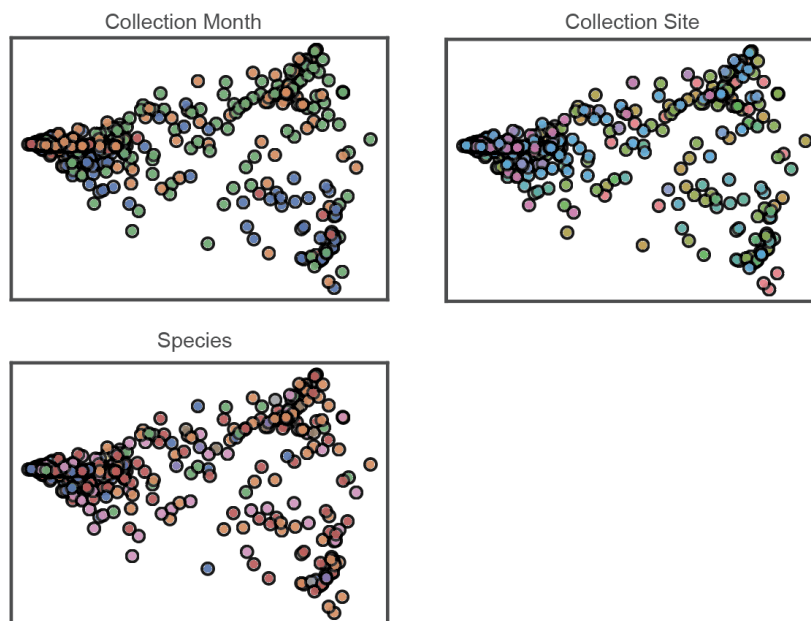
Supplementary Figure 1: Weighted UniFrac PcoA. Collection month, collection site, and species were all significant determinants of bumblebee microbiome composition (permanova, p < 0.05). However, the size of the effects were small, and samples did not visibly cluster. Points on PcoA plots are colored by categorical metadata (month, site, species). Legends are withheld for length, and because samples do not cluster.

603
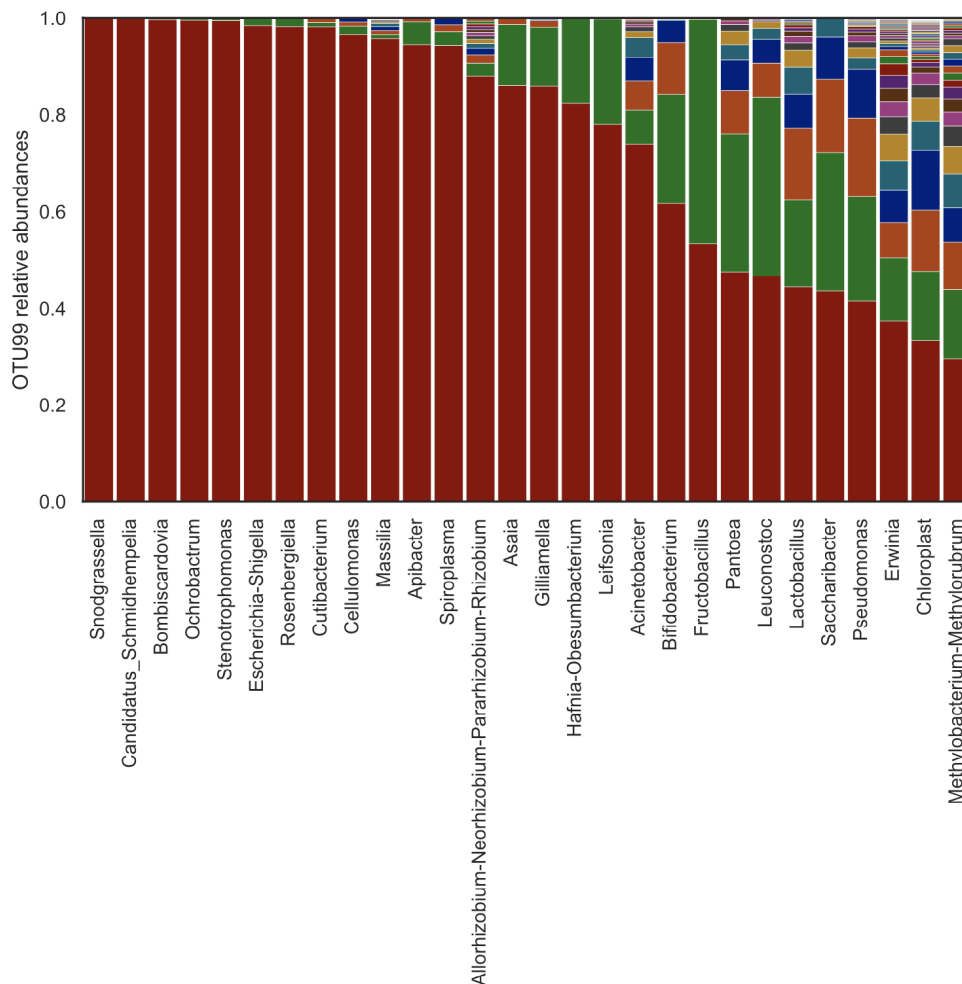
Supplementary Figure 2: Bacterial genera were composed of multiple OTUs. We selected operational taxonomic units by clustering ASVs from DADA2 into OTU99s with Vsearch. There were 28 clusters of OTUs that were able to be fully classified to the genus level and were at least moderately prevalent in our dataset, being found across 10% or more samples. Colored vertical bars are proportional to the average relative abundances of individual OTUs within the 28 genuses. OTUs are sorted vertically by relative abundance.

610
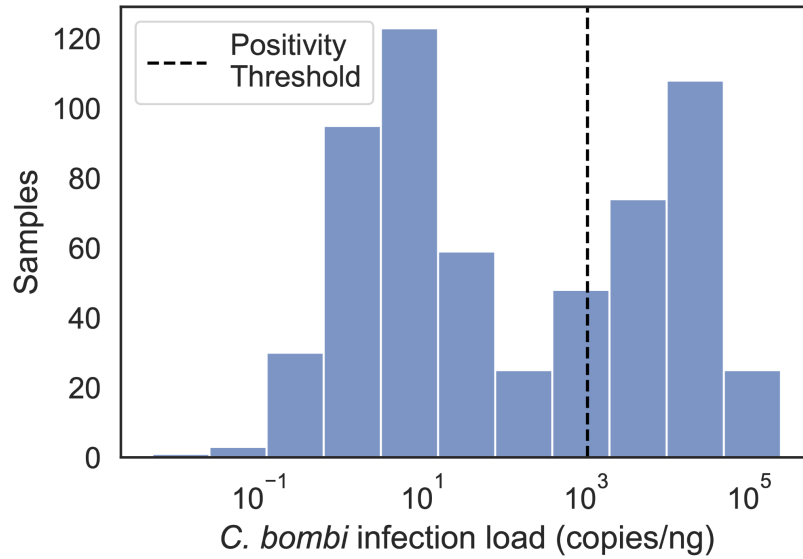
Supplementary Figure 3: Distribution of *C. bombi* infection load. *Crithidia bombi* infection load was measured with qPCR for *C. bombi GAPDH*. Samples were initially screened with a batched protocol (see Supplementary Methods). Only samples from positive batches were screened individually. We used 1000 copies per ng (dashed vertical bar) as the threshold for positivity. The infection loads of samples run individually were bimodally distributed, with a left peak below the threshold.
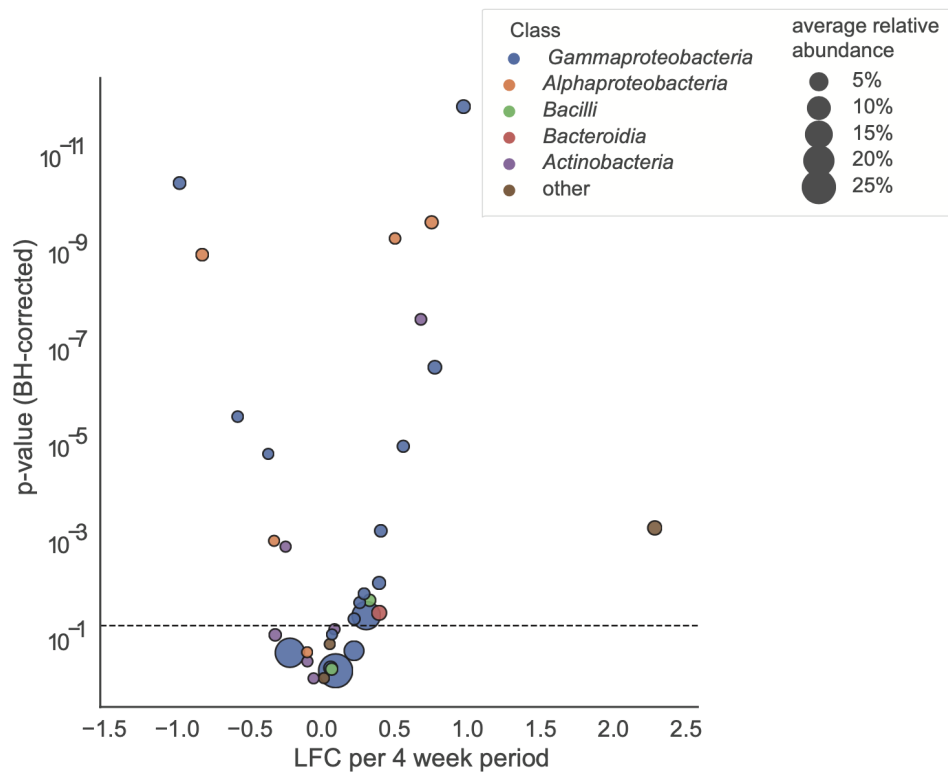
Supplementary Figure 4: Relative abundances of genera were associated with collection date. Points represent specific genera, colored by class and sized by average relative abundance. The chosen significance threshold ($\alpha = 0.05$) is indicated by a dashed horizontal bar. The brown 'other' point corresponds to a grouping of taxa with insufficient prevalence for individual differential abundance testing.

Table 1: Prevalence of core microbes by host species

species	caste	Snodgrassella		Gilliamella		Candidatus Schmidhempelia		Bombiscardovia		Lactobacillus		Apibacter		Bifidobacterium	
	M	25	100%	25	100%	22	88%	15	60%	8	32%	8	32%	10	40%
	Q	8	100%	7	88%	6	75%	6	75%	4	50%	2	25%	3	38%
bimaculatus	W	127	99%	123	96%	119	93%	93	73%	51	40%	42	33%	35	27%
	ALL	160	99%	155	96%	147	91%	114	71%	63	39%	52	32%	48	30%
	M	25	96%	26	100%	22	85%	1	4%	23	88%	11	42%	19	73%
borealis	Q	8	100%	7	88%	6	75%	4	50%	8	100%	5	62%	8	100%
	W	40	100%	39	98%	27	68%	16	40%	37	93%	21	53%	28	70%
	ALL	73	99%	72	97%	55	74%	21	28%	68	92%	37	50%	55	74%
	M	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
grisicolus	Q	1	100%	1	100%	1	100%	1	100%	1	100%	1	100%	1	100%
	W	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
	ALL	1	100%	1	100%	1	100%	1	100%	1	100%	1	100%	1	100%
	M	5	100%	5	100%	5	100%	4	80%	4	80%	5	100%	3	60%
impatiens	Q	13	93%	13	93%	13	93%	6	43%	7	50%	2	14%	2	14%
	W	81	99%	81	99%	81	99%	60	73%	72	88%	61	74%	55	67%
	ALL	99	98%	99	98%	99	98%	70	69%	83	82%	68	67%	60	59%
	M	4	100%	4	100%	4	100%	1	25%	4	100%	2	50%	3	75%
perlexus	Q	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
	W	4	100%	4	100%	4	100%	2	50%	2	50%	0	0%	1	25%
	ALL	8	100%	8	100%	8	100%	3	38%	6	75%	2	25%	4	50%
	M	2	100%	2	100%	1	50%	0	0%	2	100%	1	50%	2	100%
rufocinctus	Q	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
	W	12	100%	12	100%	12	100%	9	75%	7	58%	5	42%	6	50%
	ALL	14	100%	14	100%	13	93%	9	64%	9	64%	6	43%	8	57%
	M	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
sandersoni	Q	1	100%	1	100%	1	100%	0	0%	1	100%	0	0%	0	0%
	W	6	100%	6	100%	5	83%	4	67%	1	17%	2	33%	2	33%
	ALL	7	100%	7	100%	6	86%	4	57%	2	29%	2	29%	2	29%
	M	3	100%	3	100%	3	100%	1	33%	0	0%	1	33%	1	33%
ternarius	Q	9	100%	9	100%	9	100%	6	67%	6	67%	4	44%	3	33%
	W	124	98%	119	94%	121	96%	73	58%	42	33%	58	46%	27	21%
	ALL	136	99%	131	95%	133	96%	80	58%	48	35%	63	46%	31	22%
	M	4	100%	4	100%	4	100%	0	0%	3	75%	0	0%	3	75%
terricola	Q	1	100%	1	100%	1	100%	0	0%	1	100%	1	100%	1	100%
	W	24	100%	24	100%	22	92%	3	12%	14	58%	8	33%	15	62%
	ALL	29	100%	29	100%	27	93%	3	10%	18	62%	9	31%	19	66%
	M	19	100%	18	95%	18	95%	9	47%	6	32%	0	0%	6	32%
vagans	Q	3	100%	3	100%	3	100%	0	0%	1	33%	1	33%	0	0%
	W	78	94%	75	90%	74	89%	41	49%	28	34%	16	19%	15	18%
	ALL	100	95%	96	91%	95	90%	50	48%	35	33%	17	16%	21	20%

Table 2: Weighted UniFrac marginal permanova (adonis2)

factor	df	sum of squares	R^2	F-score	p-value	Bonferroni corrected
species	10	1.863	0.07913	6.6381	0.001	0.004
caste	2	0.1078	0.00458	1.921	0.043	0.172
site	89	4.2785	0.18172	1.7129	0.001	0.004
month	4	0.4305	0.01829	3.8352	0.001	0.004
residual	532	14.9306	0.63416			
total	637	637	23.5439			

Table 3: Weighted UniFrac Mantel Correlation

factor	correlation	p-value	samples
phylogeny	0.00136	0.961	638
geography	-0.00611	0.753	601
day	0.10257	0.001	638