Pollen-feeding *Heliconius* butterflies host distinctive adult-stage microbiomes

Tobin J. Hammer[1*], Jacob C. Dickerson[2,3], W. Owen McMillan[2], Noah Fierer[3,4]

1. Department of Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA
2. Smithsonian Tropical Research Institute, Apartado 0843-03092, Panama City, Panama
3. Department of Ecology & Evolutionary Biology, University of Colorado at Boulder, Boulder, CO 80309, USA
4. Cooperative Institute for Research in Environmental Sciences, University of Colorado at Boulder, Boulder, CO 80309, USA

*Corresponding author: tjhammer@gmail.com

Keywords: symbiosis, gut bacteria, Lepidoptera, microbiota, parasites

Abstract

Evolutionary transitions in animal diets often coincide with shifts in the microbiome, but the degree to which diet-microbiome feedbacks vary across host taxa and development is unresolved. We examined these potential feedbacks from the perspective of Lepidoptera (butterflies and moths), a diverse clade in which little is known about adult-stage microbial associations. With the exception of *Heliconius* butterflies, most lepidopteran adults are short-lived and either feed on simple substrates, like nectar, or do not feed at all. *Heliconius* consume pollen as adults, which provides amino acids and allows the butterflies to have an extended lifespan. Using 16S rRNA gene sequencing of 214 field-collected individuals, we found that adult passion-vine butterfly microbiomes exhibited a strong signal of host phylogeny, with a clear distinction between *Heliconius* and non-pollen-feeding relatives. This pattern was largely driven by differing relative abundances of bacterial phylotypes shared among host taxa, as opposed to the presence or absence of host-specific phylotypes. Using shotgun metagenomic sequencing, we also discovered trypanosomatids and microsporidia to be prevalent in butterfly guts, suggesting potential interactions with co-localized gut bacteria. Overall, we show that a major transition in adult-stage lepidopteran diet and life history coincides with a shift in microbiomes, and our work provides a foundation for future tests of microbiome function in adult butterflies.

Introduction

A major goal of contemporary microbial ecology is to identify the factors that influence the structure and function of host-associated microbiomes. In animals, diet has emerged as a consistently strong predictor of microbiome structure [1–3]. In some cases, microbes are also known to directly facilitate evolutionary transitions in host diet through their contributions to digestion, nutrient synthesis, or detoxification [4–6]. However, there are many animal taxa for which comparative microbiome data are either unavailable or only available for a specific phase of their development. This makes it difficult to generalize diet-microbiome interactions, as both microbiomes and diets often change radically during host development [7, 8].

31 Lepidoptera (butterflies and moths) are a diverse and ecologically important group in
32 which microbiome roles are likely to be highly specific to certain host taxa and life stages. For
33 lepidopteran larvae (caterpillars), even major shifts in larval diet do not coincide with
34 functionally relevant shifts in the microbiome [9–11]. Lepidopteran larvae typically harbor a
35 very low number of gut microbes, and most of these microbes are transient with the exception of
36 pathogen infections [9]. However, the microbiomes of insects that undergo complete
37 metamorphosis, such as Lepidoptera, often exhibit profound shifts in size, composition and
38 function across host life stages [7]. A recent survey of various butterfly species found that adult-
39 stage gut bacterial communities were generally abundant (median ~10^8 16S rRNA gene copies
40 per gut) and distinct from diet-associated bacterial communities [12]. It is possible that
41 evolutionary transitions in the diet of adult lepidopterans could cause, or be caused by,
42 concomitant transitions in adult-stage microbiomes.
43 Arguably, the single most dramatic and consequential transition in adult diet among the
44 150,000+ described species of Lepidoptera [13] occurred in the ancestors of neotropical
45 *Heliconius* butterflies (Nymphalidae: Heliconiini). While adults of other lepidopterans either do
46 not feed at all, or feed on comparatively nitrogen-poor and/or inconsistently available substrates
47 [14], *Heliconius* evolved the ability to consume pollen, an abundant resource rich in essential
48 amino acids [15]. Directly or indirectly, pollen-feeding led to a suite of changes in *Heliconius*
49 population structure, behavior, chemical ecology, mimicry, coevolutionary interactions with
50 plants, and life history traits [15–18]. As an example of the latter, *Heliconius* are exceptional
51 among butterflies in their adult longevity, with a potential lifespan of several months [18–20].
52 *Heliconius* have been an influential model system in evolutionary biology for over 150
53 years [17]. As with adult butterflies in general, however, very little is known about *Heliconius*
54 microbiomes. Data for a handful of *Heliconius* individuals and species are available [12, 21, 22],
55 but a comparative context with non-pollen-feeding members of the passion-vine butterfly tribe
56 Heliconiini has been missing. Nectar- and fruit-feeding butterfly species were recently found to
57 host distinct gut microbiomes [12], supporting the possibility that other kinds of shifts in adult
58 butterfly diets could be linked to shifts in microbiomes.
59 Does the ability to feed on pollen coincide with the presence of novel symbionts, changes
60 in community composition, or other microbiome differences? To answer this question, we
61 collected 214 wild adult butterflies representing 23 species and subspecies across the Heliconiini

62 and characterized their microbiomes. We used amplicon sequencing targeting a portion of the
63 16S rRNA gene and assessed the distribution of exact sequence variants (ESVs) across host
64 individuals, species, genera, and the heliconiine phylogeny. We evaluated the reliability of our
65 amplicon sequencing approach by obtaining shotgun metagenomic data from a subset of
66 samples. We also used the metagenomes to assemble near-full-length 16S rRNA genes and
67 evaluate bacterial strain-level patterns that may not be evident from analyses of amplicon
68 sequence data. Finally, we uncovered prevalent gut-associated trypanosomatids and
69 microsporidia. These putatively parasitic microeukaryotes have been little-studied in butterflies,
70 despite their potential interactions with gut bacteria and relevance to host fitness.
71
72 **Materials and Methods**
73
74 *Field collections*
75
76 The wild adult butterflies used for whole-body microbiome sequencing were collected
77 from seven locations in Panama and Ecuador in May-August 2014 (more detail is provided in the
78 supplemental file "Collection_localities.txt"). Butterflies were euthanized with ethyl acetate and
79 stored in DMSO after removal of wings, following [23]. We also stored two DMSO-only blanks
80 to use as negative controls. In June 2016, we collected additional adult butterflies for gut and
81 head/thorax sequencing from Gamboa and Pipeline Road, Panama. For these specimens, we
82 dissected the gut (hindgut, midgut, and the distal ~1/2 of the foregut) using sterilized tools prior
83 to storage in DMSO. The whole head and thorax (including the proximal foregut) were stored
84 separately. Species or subspecies were identified based on morphology. Butterflies were
85 collected under permit # SC/A-7-11 from Panama's Autoridad Nacional del Ambiente and #
86 005-13 IC-FAU-DNB/MA from Ecuador's Ministerio del Ambiente.
87
88 *Sample processing, PCR and sequencing*
89
90 We removed whole bodies and head/thorax samples from DMSO and, after
91 homogenization, used approximately 50 mg subsamples of homogenate for DNA extractions
92 with the MoBio PowerSoil kit following the manufacturer's instructions. We added entire guts

93 directly to DNA extraction tubes, in which they were homogenized during the first bead-beating
94 step of the protocol. Two DMSO blanks and 30 DNA extraction blanks were also processed in
95 tandem with the butterfly samples and sequenced.
96 PCR amplifications (515F/806R primers, V4 region) and 2 X 150 bp Illumina MiSeq
97 sequencing of 16S rRNA genes followed standard Earth Microbiome Project protocols
98 (dx.doi.org/10.17504/protocols.io.nuudeww). For 29 butterfly individuals of nine species, gut
99 and head/thorax samples were also amplified using primers that target the ITS gene region of
100 fungi [24]. Amplification success with these fungal-specific primers (as estimated from gel
101 electrophoresis) was generally very low, suggesting a lack of abundant fungal DNA that was
102 later corroborated with the shotgun metagenomic data (see below). DNA extracts from a subset
103 of 15 amplicon-sequenced gut samples were used for shotgun metagenomic sequencing
104 following the approach described previously [25] with an input DNA concentration of 0.75
105 ng/ul, KAPA HiFi HotStart ReadyMix and bead cleanup with Ampure XP beads at a 0.9x ratio.
106
107 *Amplicon data processing*
108
109 Amplicons from the 2014 whole-body samples and 2016 gut and head/thorax samples
110 were sequenced on separate runs, demultiplexed using idemp (https://github.com/yhwu/idemp),
111 and combined for futher processing. Cutadapt [26] was used to remove primer sequences. We
112 then used the DADA2 pipeline [27] to quality-filter (max EE value = 1) and trim (150 bp
113 forward, 140 bp reverse) reads, infer exact sequence variants (ESVs), merge paired-end reads,
114 and remove chimeras. We classified ESVs using the RDP Naive Bayesian Classifier algorithm
115 [28] against the SILVA training set v. 132 [29].
116 Further data processing and analyses were conducted in R v. 3.6.0 [30]. We used
117 decontam [31] for prevalence-based identification of putative contaminant ESVs based on 34
118 negative controls (DMSO and DNA extraction blanks and PCR no-template controls). The
119 median percentage of contaminant sequences across butterfly samples was 0.09%, but two
120 samples with >10% contaminants were removed from further analysis. ESVs with < 100 total
121 sequences across all samples (out of a combined total of 5.6 million sequences) were removed,
122 as were ESVs classified as mitochondria or chloroplast, or bacteria lacking sub-domain

123 identification. These ESVs combined typically made up a low proportion of reads from the
124 libraries (median of 2.6% across all samples).
125 As the resulting 16S rRNA gene amplicon datasets were highly variable in read depth
126 across samples, we rarefied all libraries to 5,000 sequences, filtering out 13 samples with lower
127 sequence depth. We also relabeled *Pantoea*, *Erwinia*, *Kluyvera*, *Citrobacter*, *Klebsiella*, and
128 *Cronobacter* taxonomic assignments to *Enterobacter*. This step was taken as genera within
129 Enterobacteriaceae are often polyphyletic and are difficult to resolve from short 16S rRNA gene
130 regions (e.g., [32]), and we wanted to avoid spurious separation of ESVs among genera and
131 resulting idiosyncracies in genus distributions across butterflies.
132
133 *Amplicon data analysis*
134
135 Beta diversity statistics and plots are based on Bray-Curtis dissimilarities and/or
136 UniFrac distances (weighted and unweighted). To obtain a bacterial phylogeny for the latter, we
137 used the fragment-insertion method [33] to place our ESV sequences into the Greengenes
138 reference tree [34]. The butterfly phylogeny is from [35] (TreeBASE #Tr77496). [Note that
139 although this phylogeny nests *aoede* within *Heliconius*, the phylogenetic position of this species
140 is held by some to be uncertain (J. Mallet, pers. comm.)]. Four of the butterfly species in our
141 sample set contained specimens from two distinct subspecies (e.g, *Heliconius sara sara* and
142 *Heliconius sara magdalena*). To include these in the species-level host phylogeny, we inserted
143 subspecies tips halfway along the terminal branches to their sister subspecies. Hereafter we refer
144 to these subspecies as "species" for simplicity.
145 To test for host-phylogenetic signal in microbiomes, we used Mantel tests with 9999
146 permutations to calculate the correlation between microbial community dissimilarities/distances
147 and host phylogenetic distances [36]. Intraspecific variation in microbiomes was handled by
148 averaging the pairwise dissimilarities/distances between all individuals of one species and all
149 individuals of another species. We used the phytools package [37] to visualize concordance
150 between topologies of the host phylogeny and a dendrogram of bacterial community
151 dissimilarities. Nodes were rotated with the "cophylo" function in phytools to maximize tip
152 matching between the two trees.

153 Differences in overall community composition between host genera were tested with
154 PERMANOVA as implemented in the vegan package [38]. Using the "betadisper" function we
155 corroborated that significant test results were due to host genus-level differences in location and
156 not dispersion [39]. We used a nonparametric statistical test (Wilcoxon rank-sum) to identify
157 bacterial genera that differed in relative abundance between host taxa or between sample types
158 (gut versus head/thorax) and applied a false discovery rate correction to the resulting p values.
159
160 *Metagenome data processing and analysis*
161
162 For 15 gut samples, we obtained shotgun metagenomic data to complement the
163 bacterial 16S rRNA gene amplicon dataset. We quality-filtered these reads with sickle [40] and
164 trimmed adapters with cutadapt [26]. We then used Bowtie 2 [41] to filter out reads matching a
165 given sample's corresponding host species' genome, obtained from Lepbase [42]. The two
166 *Dryadula phaetusa* metagenomes were mapped to a genome of the sister species *Dryas iulia* as
167 no *Dryadula* genome was available. Since there was a high proportion of host-derived reads, we
168 focused here on describing microbial diversity using ribosomal RNA gene reads present in the
169 metagenomes. With the host-filtered reads, we used phyloFlash [43] to find and classify
170 eukaryotic and bacterial SSU rRNA reads. Bacterial community composition was compared
171 between the amplicon and shotgun metagenomic datasets using a Mantel test. We also used
172 phyloFlash to assemble 16S rRNA genes from the 150 bp shotgun reads. These longer sequences
173 allowed us to estimate the phylogeny of *Orbus*, the dominant bacterium in these 15 samples.
174 *Orbus* sequences were aligned with MUSCLE [44], curated with Gblocks [45], and used for
175 maximum likelihood reconstruction with the phylogeny.fr implementation [46] of PhyML [47].
176
177 *Data availability*
178
179 Amplicon data, metadata, and R code are available from figshare
180 (figshare.com/projects/Heliconius_butterfly_microbiomes/70520). Metagenomes are available
181 from MG-RAST (project no. MGP89563).
182
183 **Results**

184 Adult heliconiine butterfly microbiomes have a relatively low diversity, with a median
185 of 26 bacterial ESVs per individual. A median of 11 ESVs accounted for 95% of the reads.
186 These communities are also fairly consistent among conspecific butterflies. For example, within
187 our most deeply sampled population (*H. erato petiverana* in Gamboa, Panama; N = 23), a
188 median of 84% of the 16S rRNA gene reads obtained from a given individual's microbiome was
189 assigned to 10 dominant bacterial genera (Fig. S1). Some of these genera are present in roughly
190 similar relative abundances across individuals and even across our two sampling years.
191 Microbiomes from whole, homogenized butterfly bodies are mainly composed of gut-
192 associated taxa. Isolated guts are similar to conspecific whole-body microbiomes in their
193 bacterial community profiles (Fig. S1), with many bacterial genera occurring at similar relative
194 abundances in guts and head/thorax tissue (Table S1). Some bacteria did differ between whole-
195 body and gut samples, such as *Spiroplasma*, which is a known hemolymph (blood) associate [48]
196 and which tends to be enriched in whole bodies (Fig. S1). Likewise, *Acinetobacter* was ~10-fold
197 more relatively abundant in head/thorax samples than in guts (Table S1). *Orbus*, *Enterobacter*,
198 *Asaia*, and some other dominant bacterial genera are clearly enriched in gut tissue (Table S1).
199

Host taxa	Phylum	Class	Order	Family	Genus	Mean prop. gut +/- SEM	Mean prop. head/thorax +/- SEM	p value
Heliconius	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	*Apibacter*	0.036 +/- 0.009	0.002 +/- 0.001	< 0.001
					Chishuiella	0.007 +/- 0.003	0.095 +/- 0.027	0.001
					Chryseobacterium	0.011 +/- 0.004	0.096 +/- 0.024	0.001
	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	0.031 +/- 0.008	0.021 +/- 0.006	0.117
				Streptococcaceae	*Lactococcus*	0.02 +/- 0.006	0.033 +/- 0.009	0.793
	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Asaia*	0.024 +/- 0.007	0.004 +/- 0.002	0.001
					Commensalibacter	0.12 +/- 0.02	0.007 +/- 0.003	< 0.001
					Swaminathania	0.013 +/- 0.01	0.016 +/- 0.014	0.745
		Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	*Variovorax*	0.01 +/- 0.004	0.021 +/- 0.006	0.017
			Enterobacteriales	Enterobacteriaceae	*Enterobacter*	0.233 +/- 0.034	0.031 +/- 0.007	< 0.001
					Serratia	0.037 +/- 0.01	0.027 +/- 0.017	0.013
			Orbales	Orbaceae	*Orbus*	0.163 +/- 0.024	0.021 +/- 0.01	< 0.001
			Pseudomonadales	Moraxellaceae	*Acinetobacter*	0.023 +/- 0.007	0.293 +/- 0.041	< 0.001
				Pseudomonadaceae	*Pseudomonas*	0.086 +/- 0.017	0.024 +/- 0.01	0.291
	Tenericutes	Mollicutes	Entomoplasmatales	Spiroplasmataceae	*Spiroplasma*	0.012 +/- 0.007	0.003 +/- 0.002	0.927
Other Heliconiini	Bacteroidetes	Bacteroidia	Flavobacteriales	Weeksellaceae	*Apibacter*	0.139 +/- 0.045	0.008 +/- 0.004	0.083
					Chishuiella	0.009 +/- 0.003	0.138 +/- 0.029	0.016
					Chryseobacterium	0 +/- 0	0.015 +/- 0.012	0.808
	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Enterococcus*	0.056 +/- 0.027	0.031 +/- 0.011	0.843
				Streptococcaceae	*Lactococcus*	0.036 +/- 0.026	0.012 +/- 0.011	0.098
	Proteobacteria	Alphaproteobacteria	Acetobacterales	Acetobacteraceae	*Asaia*	0.153 +/- 0.039	0.004 +/- 0.001	0.01
					Commensalibacter	0.031 +/- 0.015	0.002 +/- 0.002	0.106
					Swaminathania	0.007 +/- 0.004	0 +/- 0	0.226
		Gammaproteobacteria	Betaproteobacteriales	Burkholderiaceae	*Variovorax*	0.005 +/- 0.002	0.016 +/- 0.004	0.292
			Enterobacteriales	Enterobacteriaceae	*Enterobacter*	0.156 +/- 0.042	0.014 +/- 0.006	0.01
					Serratia	0.02 +/- 0.012	0.001 +/- 0	0.075
			Orbales	Orbaceae	*Orbus*	0.139 +/- 0.043	0.007 +/- 0.004	0.042
			Pseudomonadales	Moraxellaceae	*Acinetobacter*	0.046 +/- 0.018	0.45 +/- 0.054	< 0.001
				Pseudomonadaceae	*Pseudomonas*	0.02 +/- 0.017	0.011 +/- 0.004	0.393
	Tenericutes	Mollicutes	Entomoplasmatales	Spiroplasmataceae	*Spiroplasma*	0.068 +/- 0.038	0.052 +/- 0.052	0.843

200
201 Table S1. Relative abundances and within-body distribution of the top 15 bacterial genera (ranked by mean abundance)
202 across butterflies collected from Gamboa, Panama in 2016. These individuals were dissected to compare microbiomes
203 between isolated gut tissue and the combined head and thorax. Abundances are shown for *Heliconius* (top) and species
204 belonging to other heliconiine genera (bottom) separately. P values are from a nonparametric statistical test of
205 proportions in guts versus head/thorax samples, after FDR correction (see Methods).

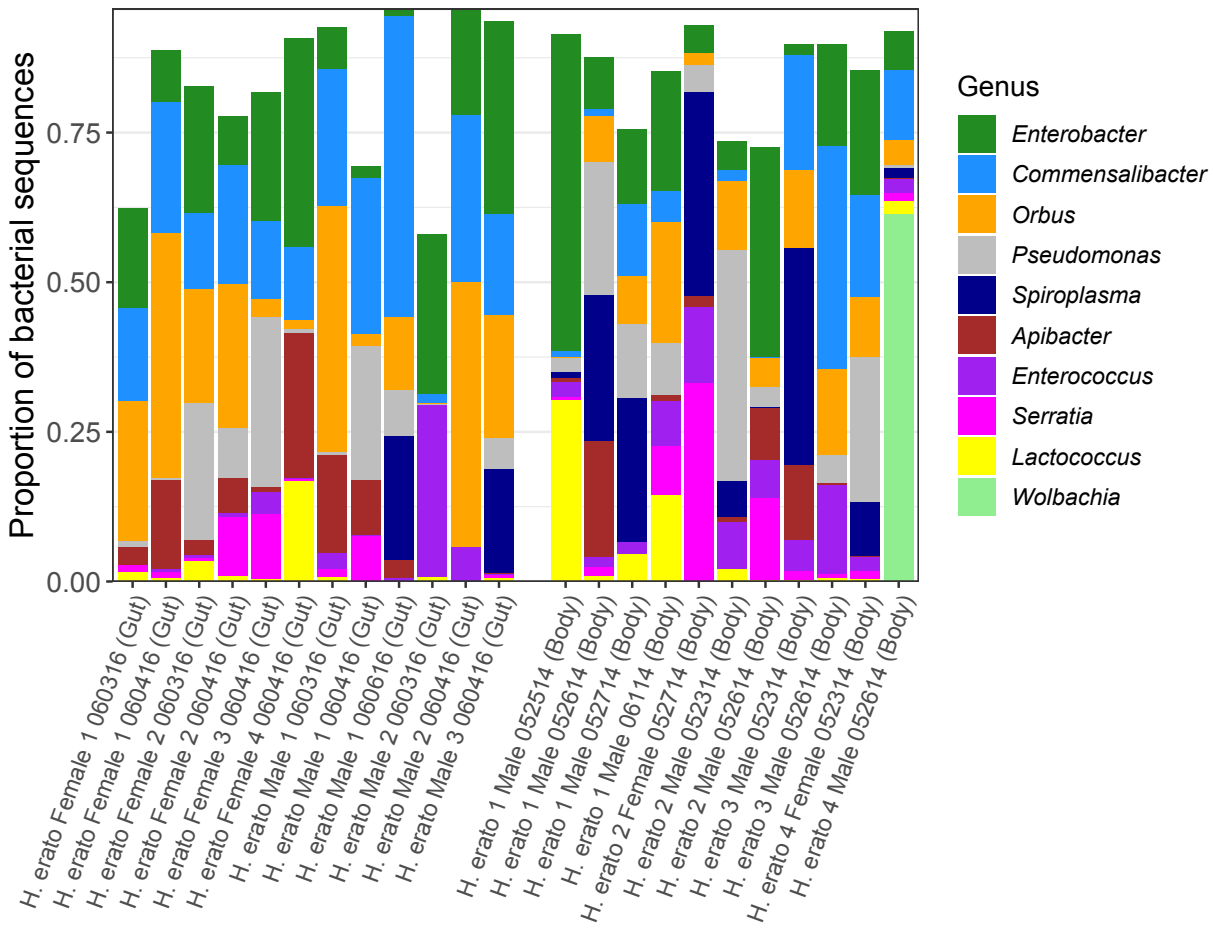
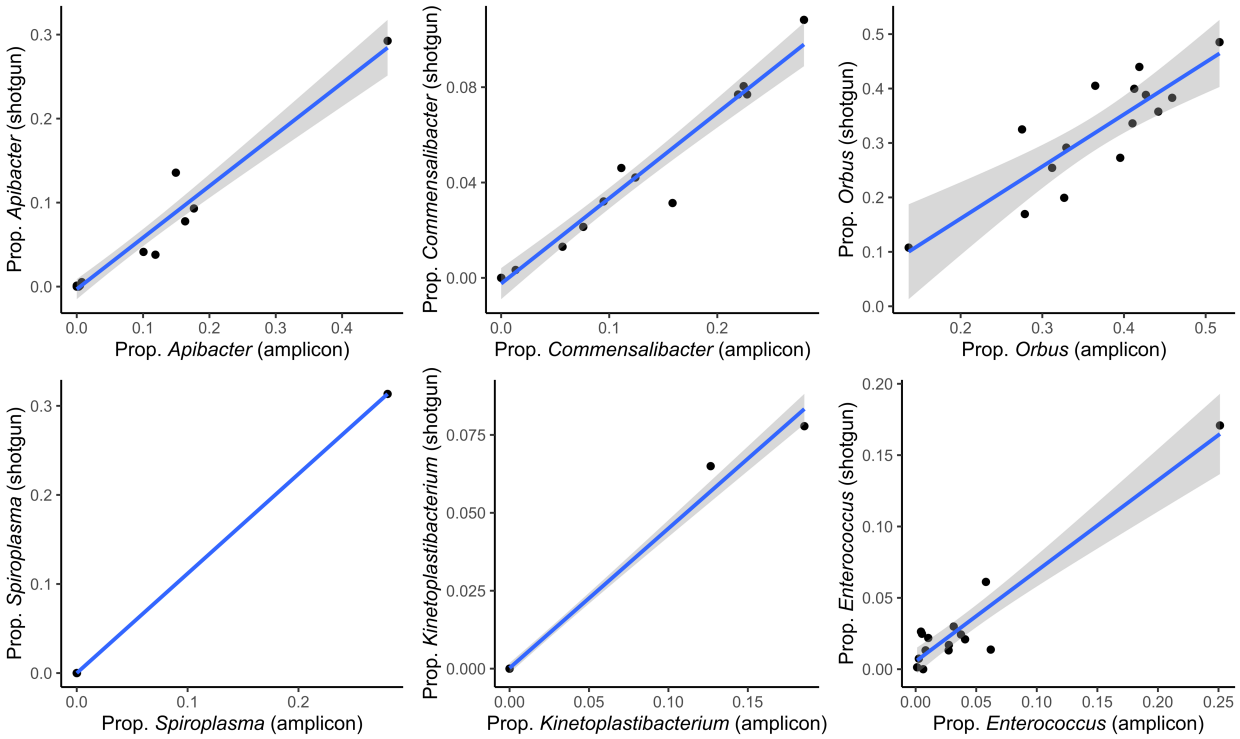


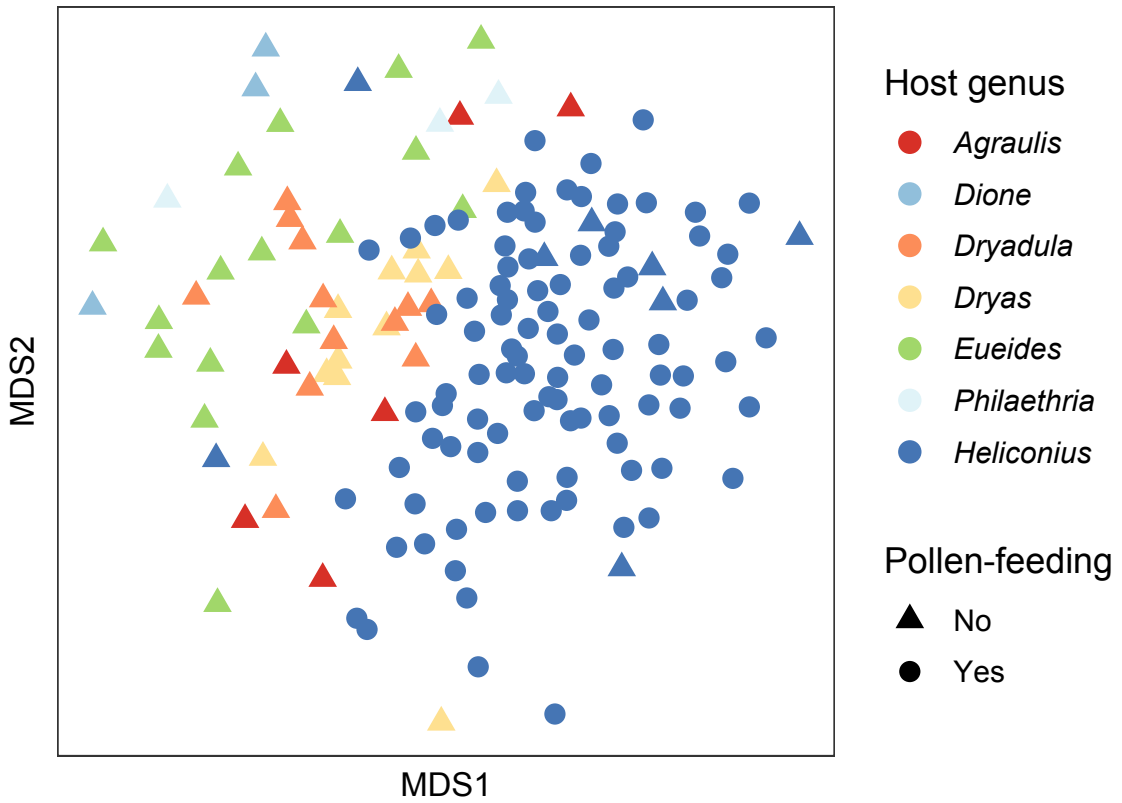
Figure S1. Whole-body bacterial communities are consistent among individuals and predominantly represent gut-associated taxa. Shown are the relative abundances of the top 10 bacterial genera (ranked by mean abundance) across *H. erato petiverana* collected from Gamboa, Panama. Remaining white space represents sequences belonging to other genera; these made up a median 16% of sequence libraries across individuals. Samples on the left are isolated guts from individuals collected in 2016, while samples on the right are whole-body homogenates from individuals collected in 2014. Note that *Enterobacter* here includes sequences originally assigned as *Klebsiella* and some other closely related Enterobacteriaceae genera (see Methods).

We used metagenomes (N = 15) to examine the accuracy of amplicon sequencing for describing bacterial community composition in our broader sample set. Interindividual microbiome variation was highly correlated between the amplicon-based dataset and the metagenomic dataset (Mantel $r = 0.63$, $p < 0.001$). Furthermore, for six of the most abundant bacterial genera, relative abundances in amplicon libraries were highly predictive of relative

220 abundances in metagenomes (Fig. S2). These results support the use of amplicon sequencing for
221 estimating bacterial community composition in adult heliconiine butterflies.
222



223
224 Figure S2. Bacterial genus-level relative abundances are highly correlated between the amplicon sequence libraries
225 and metagenomes. Shown are relative abundances of six of the most abundant genera across the 15 butterfly gut
226 samples for which we obtained both amplicon and shotgun sequence data.
227
228 Compared to other passion-vine butterflies, adult *Heliconius* harbor clearly distinct
229 whole-body bacterial communities (Fig. 1). Microbiomes clustered by host genera to varying
230 degrees depending on the distance metric used. The effect was strongest with taxonomic or
231 phylogenetic metrics that incorporate information on the relative abundances of exact sequence
232 variants (ESVs) (Bray-Curtis: $R^2 = 0.13$, $p = 0.001$; weighted UniFrac: $R^2 = 0.18$, $p = 0.001$) as
233 opposed to purely presence/absence-based metrics (Jaccard: $R^2 = 0.11$, $p = 0.001$; unweighted
234 UniFrac: $R^2 = 0.10$, $p = 0.001$). Six individuals of the non-pollen-feeding species *Heliconius*
235 *aoede* (formerly *Neruda aeode*; see Methods) had bacterial communities more similar to those
236 found in pollen-feeding *Heliconius* individuals, while two individuals had communities more
237 similar to those found in other butterfly genera (Fig. 1).



238

239 Figure 1. Adult *Heliconius* butterflies host distinct bacterial communities compared with related non-pollen-feeding
240 genera. Shown is an ordination of microbiome variation (Bray-Curtis dissimilarities) among all whole-body samples
241 (N = 104 *Heliconius*, 52 other Heliconiini). With the exception of *H. aoede* (dark blue triangles), pollen-feeding is
242 exclusive to, and ubiquitous within *Heliconius*. However, the phylogenetic placement of *H. aoede* is not fully settled
243 (see Methods).

244

245 As suggested by the host genus-level clustering, we found that variation in butterfly
246 microbiome composition was associated with host relatedness (phylogenetic distance). More
247 closely related butterfly lineages harbored more similar microbiomes (Bray-Curtis: Mantel r =
248 0.40, p = 0.01), and the topologies of the butterfly phylogeny and the dendrogram of microbiome
249 similarity were moderately congruent (Fig. 2). This pattern was still statistically significant when
250 microbiome variation was measured using weighted UniFrac distances (Mantel r = 0.31, p =
251 0.03), but not when using unweighted UniFrac (Mantel r = 0.03, p = 0.41) or Jaccard distances
252 (Mantel r = 0.19, p = 0.08), which do not incorporate relative abundance information.

253

254

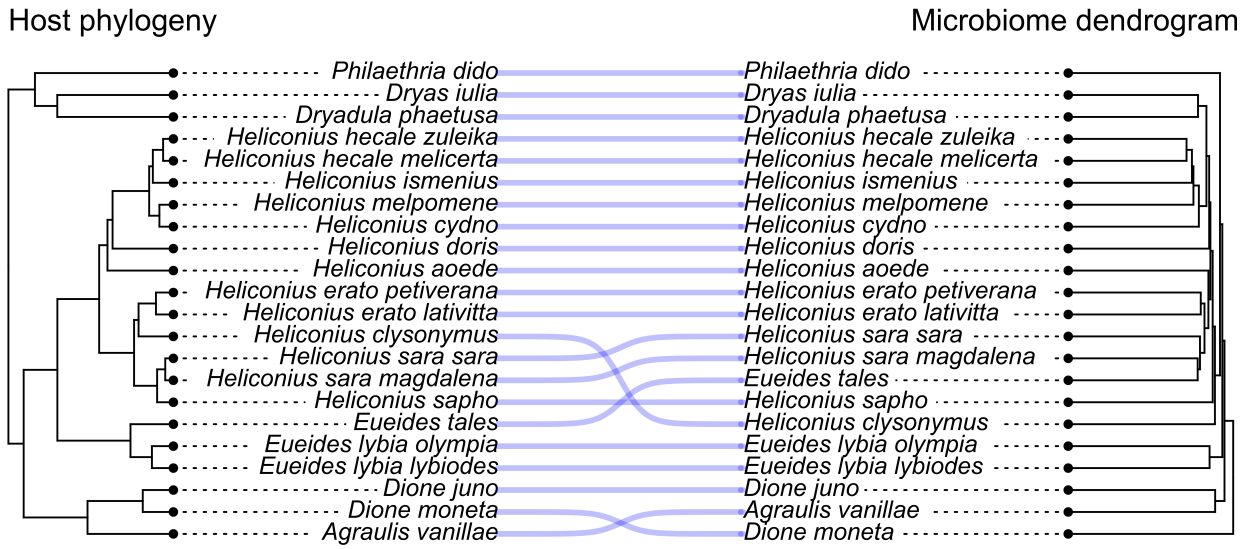


Figure 2. Heliconiine butterfly microbiomes exhibit a strong signal of host relatedness. Shown is the concordance between the host phylogeny (from [35]) and a dendrogram representation of microbiome variation among species (Bray-Curtis dissimilarities). Here, all nodes have been rotated to maximize tip matching. Note, however, that sets of parallel lines connecting tips do not always correspond to matching clades.

We then determined which bacterial taxa contribute to the observed taxonomic (Fig. 1) and phylogenetic (Fig. 2) structuring in overall community composition. The dominant bacterial genera *Enterobacter*, *Orbus*, and *Pseudomonas* (or *Serpens* [49]) were proportionally more abundant in *Heliconius* versus the non-pollen-feeding butterfly genera, while *Asaia* and *Apibacter* showed the opposite pattern (Fig. 3). None of these bacterial genera, however, were exclusively restricted to one host feeding guild, genus, or species, and their relative abundances were occasionally highly variable (Fig. 3). We also analyzed the data at the ESV-level, including all individual ESVs that were reasonably abundant within at least one host species (≥ 5% within-species average). We did not find evidence for prevalent host species- or genus-restricted ESVs (Fig. 4). Most ESVs are present, albeit with varying relative abundances, across host genus and species boundaries. However, ESV_22 (*Chryseobacterium*) is notable as being abundant in nearly all pollen-feeding species and rare or absent from non-pollen-feeding host species, including *H. aoede*.

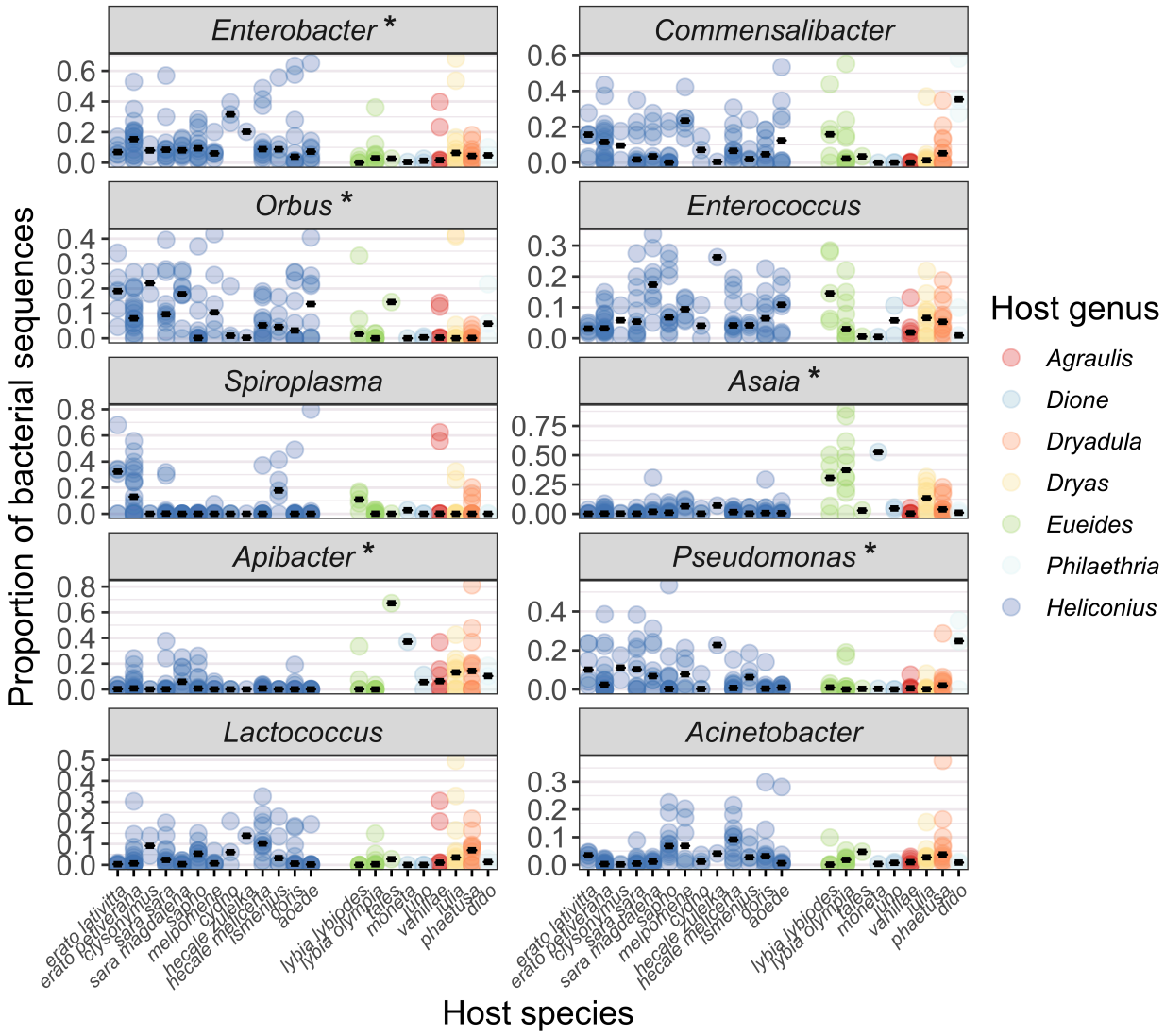


Figure 3. Dominant bacterial genera are largely shared among *Heliconius* and other butterflies, although some are differentially abundant. Shown are the relative abundances of the top 10 bacterial genera, ranked by mean abundance, in whole-body microbiomes. Dots indicate replicate individuals, and black bars indicate median proportions within a host species. Starred bacterial genera differed significantly in relative abundance between *Heliconius* and non-*Heliconius* butterflies ($p < 0.05$ after FDR correction). The arrangement of host species on the x axis corresponds to the phylogeny shown in Fig. 4. Note that *Enterobacter* here includes sequences originally assigned as *Klebsiella* and some other closely related Enterobacteriaceae genera (see Methods).

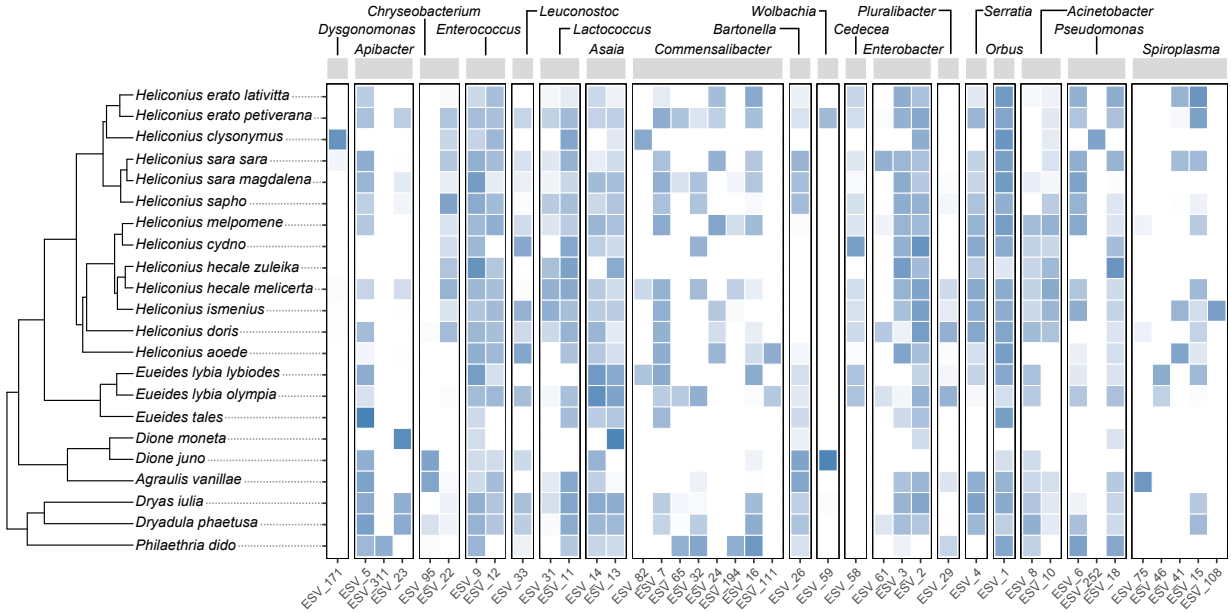


Figure 4. Most bacterial exact sequence variants (ESVs) are not host species- or host genus-specific. Shown are all ESVs in the dataset that had ≥ 5% mean relative abundance across conspecific individuals for one or more host species. For each ESV, the species-level mean relative abundance in whole-body samples, after log transformation, is indicated by the color of the cells (white = not detected in that species). Note that bacterial genera (labeled at the top) contained varying numbers of ESVs that met the aforementioned prevalence threshold.

Amplicon-derived ESVs are limited in their ability to resolve bacterial strains as they represent only a short fragment of the 16S rRNA gene (here, ~250 bp). To test for potential host specificity at a finer level of resolution, we obtained near-full-length 16S rRNA gene sequences from the bacterium *Orbus*, which is highly prevalent across gut and whole-body samples (Fig. 3, Fig. S1) and almost exclusively composed of a single ESV (Fig. 4). These sequences were reconstructed from the 12 metagenomes in which *Orbus* was sufficiently abundant. A phylogeny based on these sequences and other Orbaceae shows that butterfly-associated *Orbus* form a single, well-supported clade (Fig. 5). Host-phylogenetic or geographic structure was not evident within this clade. In fact, many of the Panamanian heliconiine butterflies harbored *Orbus* that have identical or nearly-identical 16S rRNA gene sequences to an *Orbus* strain isolated from an East Asian butterfly, *Sasakia charonda* [50].

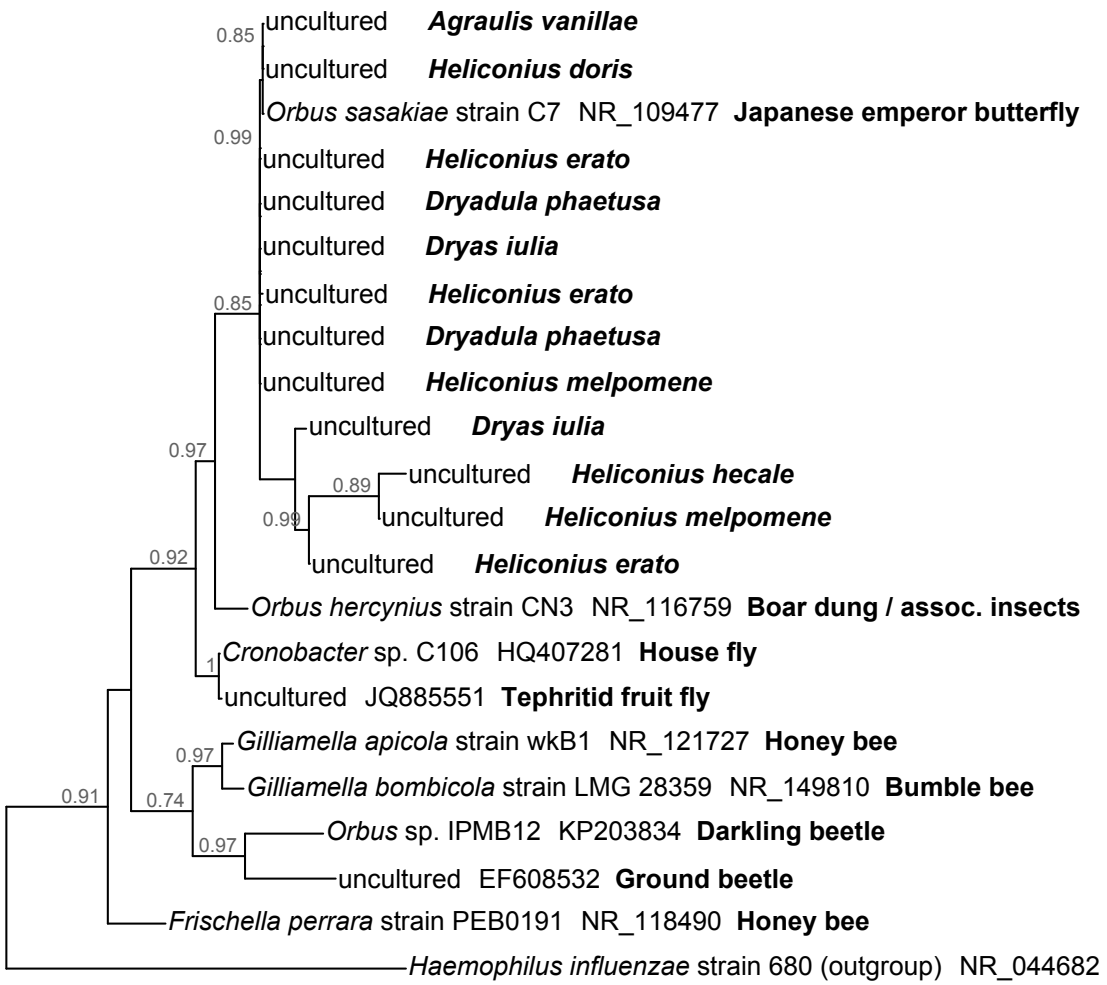
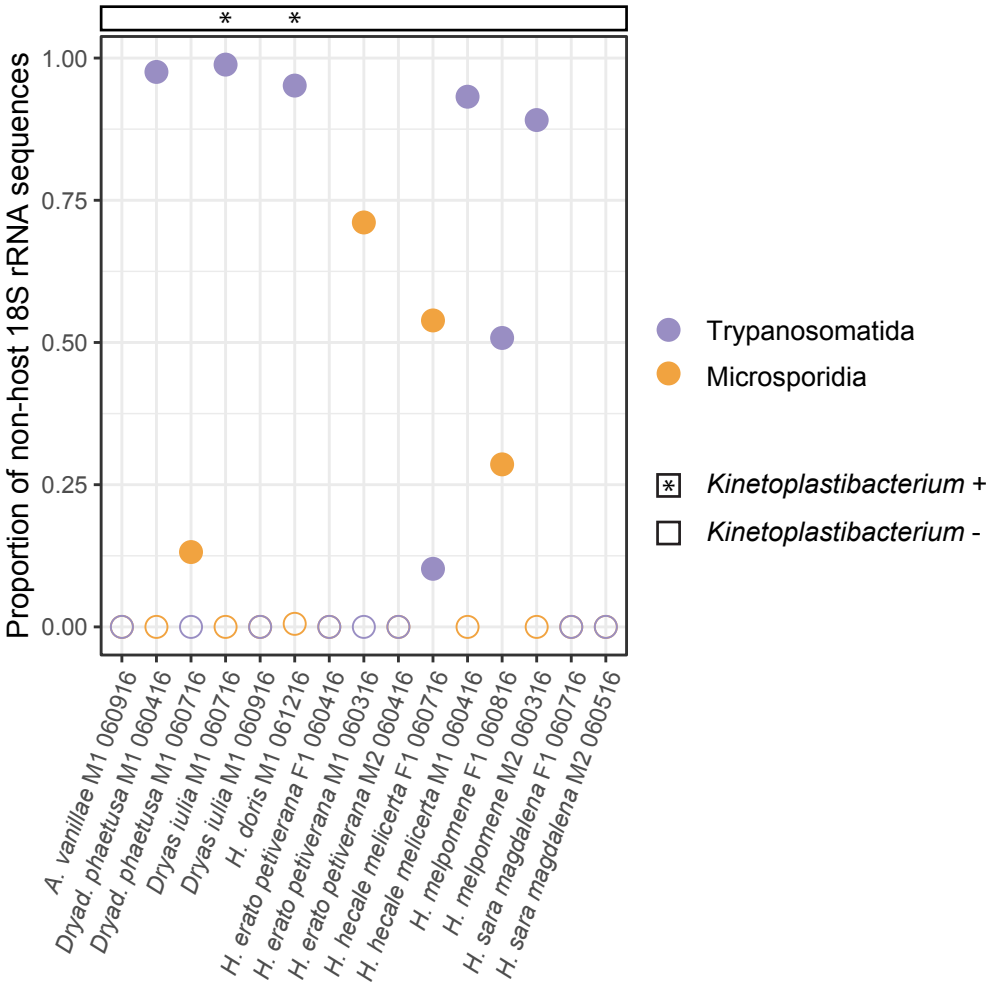


Figure 5. Maximum likelihood phylogenetic reconstruction of the bacterial family Orbaceae (Gammaproteobacteria: Pasteurellales). Some host taxonomic structure is apparent at the order level (i.e. Lepidoptera, beetles, bees, flies) but not within butterflies. 16S rRNA gene sequences from *Agraulis*, *Dryadula*, *Dryas*, and *Heliconius* were assembled from short metagenomic reads. Other 16S rRNA gene sequences are from GenBank. *Haemophilus influenzae* (Pasteurellaceae) was used as the outgroup.

We also used the shotgun metagenomes to search for microbial eukaryotes in butterfly gut samples, and found that microsporidia and trypanosomatids were prevalent (27% and 47% of the 15 individual samples analyzed, respectively) (Fig. S3). When trypanosomatids were detected in a given sample, we also sometimes detected 16S rRNA gene reads in the metagenomic data classified as *Kinetoplastibacterium*, an obligate bacterial endosymbiont of certain trypanosomatids [51]. *Kinetoplastibacterium* was not detected without its corresponding trypanosomatid host (Fig. S3). This observation led us to reexamine the larger, amplicon-based

325 bacterial dataset. We found that 9% of *Heliconius* individuals and 4% of individuals belonging to
326 other heliconiine genera were infected with a trypanosomatid, as inferred by the presence of
327 *Kinetoplastibacterium*. These proportions were not significantly different (Fisher's Exact Test, p
328 = 0.28). They are also likely underestimates, given the aforementioned prevalence of
329 trypanosomatids among metagenomes (47%) and the potential for false negatives—in five
330 metagenomes, trypanosomatids were detected without *Kinetoplastibacterium* (Fig. S3). These
331 results point to a need for more targeted analyses of butterfly-associated parasite diversity.



332

333 Figure S3. Microsporidia and Trypanosomatidae (related to *Crithidia*, *Leishmania*, and *Trypanosoma*) are prevalent
334 in the 15 butterfly individuals with sequenced gut metagenomes. Points show the proportion of each microeukaryotic
335 taxon out of all non-butterfly 18S rRNA sequences identified by phyloFlash. Circles are filled if the proportion of
336 microsporidia or trypanosomatids exceeded 0.01. Asterisks indicate samples in which *Kinetoplastibacterium*, a
337 bacterial endosymbiont of trypanosomatids, was detected. Other microeukaryotes not shown here comprise a variety
338 of very low-abundance taxa (i.e. ≤ 30 total reads per metagenome) classified mainly as coccidia, ascomycete fungi,
339 acanthamoeba, and algae.

340 **Discussion**

341

342 The evolution of pollen-feeding in *Heliconius* butterflies underlies a number of unique
343 traits and likely facilitated the rapid diversification of this genus. Although *Heliconius* have been
344 a major study system for decades, their microbiomes were unexplored beyond a few species [12,
345 21, 22]. In particular, it was unknown whether they evolved distinctive microbiomes
346 concomitant with this major shift in diet and life history. This question is particularly interesting
347 in butterflies, in which even major diet shifts in the larval stage (caterpillars) do not depend on
348 microbiomes [9–11]. *Heliconius* is arguably the single lineage of Lepidoptera most reliant on
349 adult-stage nutrition, and thus might be likely to engage in microbial symbioses as adults.
350 Using extensive field collections of 21 species and subspecies of heliconiine
351 butterflies, we found that adult *Heliconius* have relatively low-diversity bacterial communities
352 that are compositionally distinct from those found in their non-pollen-feeding relatives (Fig. 1,
353 Fig. 3). Most of the dominant bacterial taxa were present in isolated gut tissue, but others were
354 enriched elsewhere in the body, such as the hemolymph or structures in the head and thorax (Fig.
355 S1, Table S1). Beyond the clear separation of *Heliconius* from other heliconiine genera, there
356 was also strong a phylogenetic signal in adult heliconiine butterfly microbiomes (Fig. 2). The
357 strength of this signal was comparable to that observed in mammalian gut microbiomes [3] and
358 other host groups [36].
359 Host-phylogenetic structure in microbiomes can arise from host-microbe co-
360 diversification and/or contemporary host filtering of environmental microbes [4, 36]. Given that
361 the dominant bacterial taxa appear to be largely shared among host lineages (discussed below),
362 we suggest that host filtering is likely to be particularly important in heliconiine butterflies. For
363 example, the shift to pollen-feeding in adult *Heliconius* represents a major change in nutrient
364 inputs, potentially altering which resident bacteria are able to proliferate in the gut. Pollen itself
365 also harbors microbes [52], some of which may colonize the gut. Community assembly in the
366 adult stage could also be influenced, in part, by dynamics occurring in earlier life stages [7].
367 Some bacterial taxa present in mature, adult *H. erato* are already present in freshly emerged
368 (unfed) adults, pupae and larvae, and are potentially carried over from the larval host plant
369 through metamorphosis [22]. Moreover, Heliconiini larvae show marked variation in the species

370 identity and the age (young vs. old foliage) of host plants they consume [17], traits reported to
371 structure leaf microbiomes [53–55].
372 Certain bacterial taxa were enriched in *Heliconius* butterflies and their abundances may
373 be associated with with the evolution of pollen-feeding. For example, *Orbus* was the most
374 dominant bacterial genus among *Heliconius* and was more than twice as proportionally abundant
375 in *Heliconius* as compared with other heliconiine butterflies. The biology of the Orbaceae is
376 largely unknown, excepting the specialized bee gut symbionts *Gilliamella* and *Frischella* [56,
377 57]. In honey bees, which also consume pollen, some *Gilliamella* strains can degrade and
378 ferment structural compounds in pollen walls [58, 59]. Adding pollen to honey bee diets also
379 increases gut colonization by *Gilliamella* and *Frischella* [60]. However, a key difference is that
380 pollen digestion in *Heliconius*, unlike bees, occurs extra-orally. Pollen grains collected on the
381 proboscis are steeped in exuded saliva, and amino acids and other compounds released from the
382 grains are then ingested [61]. Host-encoded enzymes that likely play a role in pollen digestion
383 have been identified [62, 63]. Whether microbes contribute to this process is still unclear, but we
384 identified candidate bacterial taxa that are overrepresented in the head and thorax (Table S1),
385 structures that contain the proboscis and salivary gland. The exact localization and role of these
386 microbes is worth more detailed investigation.
387 Based on the amplicon-based data, we found that microbiome differences across
388 heliconiine butterflies were largely driven by shifts in the relative abundance of shared bacterial
389 genera and ESVs, as opposed to the presence or absence of host-specific bacterial taxa. These
390 bacteria might be widely distributed among larval host plants or adult nectar sources, and
391 independently acquired by different butterflies. Sympatric butterfly species might also exchange
392 bacteria, given their frequent co-occurrence at nectar sources. Flowers can act as hubs for the
393 horizontal transmission of microbes among pollinators [64, 65].
394 The amplicon data were based on a short region of the 16S rRNA gene and might obscure
395 finer-scale diversity and bacterial distribution patterns among hosts. We examined this
396 possibility using near-full-length 16S rRNA gene sequences from the bacterium *Orbus*. As
397 inferred from the amplicon data, this genus was represented by a single ESV that was nearly
398 ubiquitous across all butterflies (Fig. 4). We found additional diversity among the longer
399 sequences, but within butterflies, different *Orbus* phylotypes did not segregate by butterfly
400 phylogeny or geography (Fig. 5). At a broader scale, the Orbaceae do appear to exhibit some

401 host-taxonomic specificity (Fig. 5; also see [49]). Potentially, the Orbaceae have diversified into
402 loosely insect-order-specific lineages which at finer scales (e.g. within butterflies) frequently
403 disperse among host taxa and regions. However, we note that the 16S rRNA gene evolves
404 slowly, making it difficult to differentiate environmental transmission from co-diversification in
405 a young host radiation such as the Heliconiini [35] without additional information [4].
406 Adult butterflies clearly structure the composition of their bacterial communities, but the
407 reciprocal effect of these microbes on adult butterfly biology is completely unknown. A recent
408 study found nectar- and fruit-feeding butterflies to host gut microbes with different abilities to
409 metabolize substrates *in vitro* [12], but whether these differences are expressed *in vivo* and
410 whether they would matter for host nutrition remains unclear. A subsequent experiment on the
411 butterfly *Speyeria mormonia* using antibiotics did not find a strong relationship between gut
412 bacterial loads and measures of adult fitness [66]. However, antibiotic-based studies could fail to
413 test a role of the microbiome in resistance to parasites and pathogens if the latter are also
414 suppressed by antibiotics.
415 Colonization resistance is a common feature of many animal gut microbiomes [4, 67, 68]
416 and could be important in adult butterflies. Here, we discovered a high prevalence of putative
417 eukaryotic parasites (trypanosomatids and microsporidia) across adult heliconiine butterflies.
418 Butterfly-associated microeukaryotes are almost unknown, and there is very little research on
419 any adult-stage pathogens or parasites (with the exception of the neogregarine *Ophryocystis*
420 *elektroscirrha* in monarch butterflies [69]). *Serratia* were also widespread among heliconiine
421 butterflies and are frequently pathogenic in insects [70, 71]. Given their co-localization in
422 butterfly guts, we expect direct interactions and potential antagonism between these putative
423 parasites and other gut microbes. For example, in some pyrrhocorid bugs, tsetse flies and social
424 bees, gut bacteria provide an important layer of defense against trypanosomatid parasites and
425 *Serratia* [72–75].
426
427 **Conclusions**
428
429 This study provides the first in-depth characterization of *Heliconius* and other passion-
430 vine butterfly microbiomes, adding a new dimension to a classic model system in evolutionary
431 biology. Our finding that a major diet shift (pollen-feeding) coincides with a shift in adult-stage

microbiomes advances previous work on Lepidoptera feeding ecology and microbial interactions, most of which has focused on the larval stage. The ecological dynamics of microbiomes appear to be very different between larval and adult life stages, at least for lepidopteran groups that feed as adults. We also identified a strong host phylogenetic signal in passion-vine butterfly microbiomes, and speculate that it arises from host filtering of environmental microbes with conserved dietary or behavioral traits, as opposed to co-diversification.

The unique microbiome in *Heliconius* likely represents a response to the nutritional and microbial inputs from pollen, but may also represent a functional contribution to host pollen digestion. However, despite these microbial community-level differences, individual bacterial taxa were typically shared across host species and diet guilds. We suggest there is also overlap in microbiome function among hosts. In particular, the common occurrence of putative parasites in guts of various host species points to colonization resistance as a candidate function of adult microbiomes in passion-vine butterflies. Manipulative experiments can now be performed, informed by our characterization of natural microbial community structure across butterfly hosts. Such studies would fill a critical gap in our understanding of butterflies, a diverse group of considerable ecological, societal, and scientific importance [76].

Acknowledgments

This material is based on work supported by the Cooperative Institute for Research in Environmental Sciences, the Ecology and Evolutionary Biology Department at the University of Colorado Boulder, a National Science Foundation Graduate Research Fellowship (1144083), and a postdoctoral fellowship from the USDA National Institute of Food and Agriculture (2018-08156), to TJH. This work was further supported by grants from the Simons Foundation (429440) and by funding from the Smithsonian Tropical Research Institute. We also thank Jessica Henley for assistance with sequencing, the Smithsonian Tropical Research Institute staff for administrative support and the Panamanian and Ecuadorian Authorities for permission to collect butterflies.

463 **References**

464

465 1. Muegge BD, Kuczynski J, Knights D, Clemente JC, González A, Fontana L, et al. Diet
466 drives convergence in gut microbiome functions across mammalian phylogeny and within
467 humans. *Science* 2011; **332**: 970–974.
468 2. Delsuc F, Metcalf JL, Wegener Parfrey L, Song SJ, González A, Knight R. Convergence of
469 gut microbiomes in myrmecophagous mammals. *Mol Ecol* 2014; **23**: 1301–1317.
470 3. Groussin M, Mazel F, Sanders JG, Smillie CS, Lavergne S, Thuiller W, et al. Unraveling
471 the processes shaping mammalian gut microbiomes over evolutionary time. *Nature*
472 *Communications* 2017; **8**: 14319.
473 4. Moran NA, Ochman H, Hammer TJ. Evolutionary and Ecological Consequences of Gut
474 Microbial Communities. *Annu Rev Ecol Evol Syst* 2019; **50**.
475 5. Engel P, Moran NA. The gut microbiota of insects - diversity in structure and function.
476 *FEMS Microbiology Reviews* 2013; **37**: 699–735.
477 6. McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et
478 al. Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the*
479 *National Academy of Sciences* 2013; **110**: 3229–3236.
480 7. Hammer TJ, Moran NA. Links between metamorphosis and symbiosis in holometabolous
481 insects. *Phil Trans R Soc B* 2019; **374**: 20190068.
482 8. Koenig JE, Spor A, Scalfone N, Fricker AD, Stombaugh J, Knight R, et al. Succession of
483 microbial consortia in the developing infant gut microbiome. *Proceedings of the National*
484 *Academy of Sciences* 2011; **108**: 4578–4585.
485 9. Hammer TJ, Janzen DH, Hallwachs W, Jaffe SP, Fierer N. Caterpillars lack a resident gut
486 microbiome. *Proceedings of the National Academy of Sciences* 2017; **114**: 9641–9646.
487 10. Szenteczki MA, Pitteloud C, Casacci LP, Kešnerová L, Whitaker MRL, Engel P, et al.
488 Bacterial communities within *Phengaris (Maculinea) alcon* caterpillars are shifted
489 following transition from solitary living to social parasitism of *Myrmica* ant colonies. *Ecol*
490 *Evol* 2019; **9**: 4452–4464.
491 11. Whitaker M, Pierce N, Salzman S, Kaltenpoth M, Sanders J. Microbial communities of
492 lycaenid butterflies do not correlate with larval diet. *Frontiers in Microbiology* 2016; **7**:
493 1920.

494 12. Ravenscraft A, Berry M, Hammer T, Peay K, Boggs C. Structure and function of the
495 bacterial and fungal gut microbiota of Neotropical butterflies. *Ecological Monographs*
496 2019; **89**: e01346.
497 13. Stork NE. How Many Species of Insects and Other Terrestrial Arthropods Are There on
498 Earth? *Annu Rev Entomol* 2018; **63**: 31–45.
499 14. Scoble MJ. The Lepidoptera. 1992. Oxford University Press.
500 15. Gilbert LE. Pollen Feeding and Reproductive Biology of *Heliconius* Butterflies.
501 *Proceedings of the National Academy of Sciences* 1972; **69**: 1403–1407.
502 16. Cardoso MZ, Gilbert LE. Pollen feeding, resource allocation and the evolution of chemical
503 defence in passion vine butterflies. *Journal of Evolutionary Biology* 2013; **26**: 1254–1260.
504 17. Jiggins CD. The ecology and evolution of *Heliconius* butterflies. 2016. Oxford University
505 Press.
506 18. Ehrlich PR, Gilbert LE. Population Structure and Dynamics of the Tropical Butterfly
507 *Heliconius ethilla*. *Biotropica* 1973; **5**: 69–82.
508 19. Dunlap-Pianka H, Boggs CL, Gilbert LE. Ovarian Dynamics in Heliconiine Butterflies:
509 Programmed Senescence versus Eternal Youth. *Science* 1977; **197**: 487–490.
510 20. Cook LM, Thomason EW, Young AM. Population Structure, Dynamics and Dispersal of
511 the Tropical Butterfly *Heliconius charitonius*. *The Journal of Animal Ecology* 1976; **45**:
512 851–863.
513 21. van Schooten B, Godoy-Vitorino F, McMillan WO, Papa R. Conserved microbiota among
514 young *Heliconius* butterfly species. *PeerJ* 2018; **6**: e5502.
515 22. Hammer TJ, McMillan WO, Fierer N. Metamorphosis of a Butterfly-Associated Bacterial
516 Community. *PLoS ONE* 2014; **9**: e86995.
517 23. Hammer TJ, Dickerson JC, Fierer N. Evidence-based recommendations on storing and
518 handling specimens for analyses of insect microbiota. *PeerJ* 2015; **3**: e1190.
519 24. Gardes M, Bruns TD. ITS primers with enhanced specificity for basidiomycetes -
520 application to the identification of mycorrhizae and rusts. *Mol Ecol* 1993; **2**: 113–118.
521 25. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, Kishony R. Inexpensive
522 Multiplexed Library Preparation for Megabase-Sized Genomes. *PLOS ONE* 2015; **10**:
523 e0128036.

524 26. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads.
525 *EMBnet.journal* 2011; **17**: 10–12.
526 27. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. DADA2:
527 High-resolution sample inference from Illumina amplicon data. *Nat Methods* 2016; **13**:
528 581–583.
529 28. Wang Q, Garrity GM, Tiedje JM, Cole JR. Naive Bayesian classifier for rapid assignment
530 of rRNA sequences into the new bacterial taxonomy. *Applied and Environmental*
531 *Microbiology* 2007; **73**: 5261–5267.
532 29. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal
533 RNA gene database project: improved data processing and web-based tools. *Nucleic Acids*
534 *Research* 2012; **41**: D590–D596.
535 30. R Core Team. R: A language and environment for statistical computing. 2016. R
536 Foundation for Statistical Computing, Vienna, Austria.
537 31. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical
538 identification and removal of contaminant sequences in marker-gene and metagenomics
539 data. *Microbiome* 2018; **6**: 226.
540 32. Brady C, Cleenwerck I, Venter S, Coutinho T, De Vos P. Taxonomic evaluation of the
541 genus *Enterobacter* based on multilocus sequence analysis (MLSA). *Systematic and*
542 *Applied Microbiology* 2013; **36**: 309–319.
543 33. Janssen S, McDonald D, Gonzalez A, Navas-Molina JA, Jiang L, Xu ZZ, et al.
544 Phylogenetic Placement of Exact Amplicon Sequences Improves Associations with Clinical
545 Information. *mSystems* 2018; **3**: e00021-18.
546 34. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, et al. An
547 improved Greengenes taxonomy with explicit ranks for ecological and evolutionary
548 analyses of bacteria and archaea. *The ISME Journal* 2012; **6**: 610–618.
549 35. Kozak KM, Wahlberg N, Neild AF, Dasmahapatra KK, Mallet J, Jiggins CD. Multilocus
550 Species Trees Show the Recent Adaptive Radiation of the Mimetic *Heliconius* Butterflies.
551 *Systematic Biology* 2015; **64**: 505–524.
552 36. Mazel F, Davis KM, Loudon A, Kwong WK, Groussin M, Parfrey LW. Is Host Filtering
553 the Main Driver of Phylosymbiosis across the Tree of Life? *mSystems* 2018; **3**: e00097-18.

554 37. Revell LJ. phytools: An R package for phylogenetic comparative biology (and other
555 things). *Methods in Ecology and Evolution* 2012; **3**: 217–223.
556 38. Oksanen J, Kindt R, Legendre P. vegan: Community Ecology Package. R package version
557 2.0-7. 2013.
558 39. Anderson MJ, Walsh DCI. PERMANOVA, ANOSIM, and the Mantel test in the face of
559 heterogeneous dispersions: What null hypothesis are you testing? *Ecological Monographs*
560 2013; **83**: 557–574.
561 40. Joshi, NA, Fass, JN. Sickle: A sliding-window, adaptive, quality-based trimming tool for
562 FastQ files. https://github.com/najoshi/sickle.
563 41. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;
564 **9**: 357–359.
565 42. Challis RJ, Kumar S, Dasmahapatra KK, Jiggins CD, Blaxter M. Lepbase: the Lepidopteran
566 genome database. *bioRxiv* 2016; http://dx.doi.org/10.1101/056994.
567 43. Gruber-Vodicka HR, Seah BK, Pruesse E. phyloFlash — Rapid SSU rRNA profiling and
568 targeted assembly from metagenomes. *bioRxiv* 2019; https://doi.org/10.1101/521922.
569 44. Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high
570 throughput. *Nucleic Acids Research* 2004; **32**: 1792–1797.
571 45. Castresana J. Selection of Conserved Blocks from Multiple Alignments for Their Use in
572 Phylogenetic Analysis. *Molecular Biology and Evolution* 2000; **17**: 540–552.
573 46. Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr:
574 robust phylogenetic analysis for the non-specialist. *Nucleic Acids Research* 2008; **36**:
575 W465–W469.
576 47. Guindon S, Delsuc F, Dufayard JF, Gascuel O. Estimating maximum likelihood phylgenies
577 with PhyML. *Methods in Molecular Biology* 2009; **537**: 113–137.
578 48. Blow F, Douglas AE. The hemolymph microbiome of insects. *Journal of Insect Physiology*
579 2019; **115**: 33–39.
580 49. Martinson VG, Carpinteyro-Ponce J, Moran NA, Markow TA. A Distinctive and Host-
581 Restricted Gut Microbiota in Populations of a Cactophilic *Drosophila* Species. *Applied and*
582 *Environmental Microbiology* 2017; **83**: e01551–17.
583 50. Kim JY, Lee J, Shin N-R, Yun J-H, Whon TW, Kim M-S, et al. *Orbus sasakiae* sp. nov., a
584 bacterium isolated from the gut of the butterfly *Sasakia charonda*, and emended description

585 of the genus *Orbus*. *International Journal of Systematic and Evolutionary Microbiology*
586 2012; **63**: 1766–1770.
587 51. De Souza W, Motta MCM. Endosymbiosis in protozoa of the Trypanosomatidae family.
588 *FEMS Microbiology Letters* 1999; **173**: 1–8.
589 52. Ambika Manirajan B, Ratering S, Rusch V, Schwiertz A, Geissler-Plaum R, Cardinale M,
590 et al. Bacterial microbiota associated with flower pollen is influenced by pollination type,
591 and shows a high degree of diversity and species-specificity: The bacterial microbiota of
592 flower pollen. *Environmental Microbiology* 2016; **18**: 5161–5174.
593 53. Redford AJ, Fierer N. Bacterial succession on the leaf surface: a novel system for studying
594 successional dynamics. *Microbial Ecology* 2009; **58**: 189–98.
595 54. Bulgarelli D, Schlaeppi K, Spaepen S, van Themaat EVL, Schulze-Lefert P. Structure and
596 Functions of the Bacterial Microbiota of Plants. *Annual Review of Plant Biology* 2013; **64**:
597 807–838.
598 55. Lindow SE, Brandl MT. Microbiology of the Phyllosphere. *Applied and Environmental*
599 *Microbiology* 2003; **69**: 1875.
600 56. Kwong WK, Moran NA. Cultivation and characterization of the gut symbionts of honey
601 bees and bumble bees: description of *Snodgrassella alvi* gen. nov., sp. nov., a member of
602 the family Neisseriaceae of the Betaproteobacteria, and *Gilliamella apicola* gen. nov., sp.
603 nov., a member of Orbaceae fam. nov., Orbales ord. nov., a sister taxon to the order
604 'Enterobacteriales' of the Gammaproteobacteria. *International Journal of Systematic and*
605 *Evolutionary Microbiology* 2013; **63**: 2008–2018.
606 57. Engel P, Kwong WK, Moran NA. *Frischella perrara* gen. nov., sp. nov., a
607 gammaproteobacterium isolated from the gut of the honey bee, *Apis mellifera*. *International*
608 *Journal of Systematic and Evolutionary Microbiology* 2013; **63**: 3646–3651.
609 58. Engel P, Martinson VG, Moran NA. Functional diversity within the simple gut microbiota
610 of the honey bee. *Proceedings of the National Academy of Sciences* 2012; **109**: 11002–
611 11007.
612 59. Kesnerova L, Mars RAT, Ellegaard KM, Troilo M, Sauer U, Engel P. Disentangling
613 metabolic functions of bacteria in the honey bee gut. *PLoS Biology* 2017; **15**: e2003467.
614 60. Kešnerová L, Emery O, Troilo M, Erkosar B, Engel P. Gut microbiota structure differs
615 between honey bees in winter and summer. *bioRxiv* 2019; https://doi.org/10.1101/703512.

616 61. Eberhard SH, Hikl AL, Boggs CL, Krenn HW. Saliva or Regurgitated Nectar? What
617 *Heliconius* Butterflies (Lepidoptera: Nymphalidae) Use for Pollen Feeding. *Annals of the*
618 *Entomological Society of America* 2009; **102**: 1105–1108.
619 62. Harpel D, Cullen DA, Ott SR, Jiggins CD, Walters JR. Pollen feeding proteomics: Salivary
620 proteins of the passion flower butterfly, *Heliconius melpomene*. *Insect Biochemistry and*
621 *Molecular Biology* 2015; **63**: 7–13.
622 63. Smith G, Macias-Muñoz A, Briscoe AD. Gene Duplication and Gene Expression Changes
623 Play a Role in the Evolution of Candidate Pollen Feeding Genes in *Heliconius* Butterflies.
624 *Genome Biology and Evolution* 2016; **8**: 2581–2596.
625 64. McArt SH, Koch H, Irwin RE, Adler LS. Arranging the bouquet of disease: floral traits and
626 the transmission of plant and animal pathogens. *Ecology Letters* 2014; **17**: 624–636.
627 65. Graystock P, Goulson D, Hughes WOH. Parasites in bloom: flowers aid dispersal and
628 transmission of pollinator parasites within and between bee species. *Proceedings of the*
629 *Royal Society B: Biological Sciences* 2015; **282**: 20151371.
630 66. Ravenscraft A, Kish N, Peay K, Boggs C. No evidence that gut microbiota impose a net
631 cost on their butterfly host. *Molecular Ecology* 2019; **28**: 2100–2117.
632 67. Kamada N, Chen GY, Inohara N, Núñez G. Control of pathogens and pathobionts by the
633 gut microbiota. *Nature Immunology* 2013; **14**: 685–690.
634 68. Ducarmon QR, Zwittink RD, Hornung BVH, van Schaik W, Young VB, Kuijper EJ. Gut
635 Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiol Mol*
636 *Biol Rev* 2019; **83**: e00007-19.
637 69. Altizer SM, Oberhauser KS. Effects of the Protozoan Parasite *Ophryocystis elektroscirrha*
638 on the Fitness of Monarch Butterflies (*Danaus plexippus*). *Journal of Invertebrate*
639 *Pathology* 1999; **74**: 76–88.
640 70. Raymann K, Coon KL, Shaffer Z, Salisbury S, Moran NA. Pathogenicity of *Serratia*
641 *marcescens* Strains in Honey Bees. *mBio* 2018; **9**: e01649-18.
642 71. Grimont F, Grimont PAD. The Genus *Serratia*. In: Dworkin M, Falkow S, Rosenberg E,
643 Schleifer K-H, Stackebrandt E (eds). *The Prokaryotes*. 2006. Springer New York, pp 219–
644 244.

645 72. Onchuru TO, Martinez AJ, Kaltenpoth M. The cotton stainer's gut microbiota suppresses
646 infection of a cotransmitted trypanosomatid parasite. *Molecular Ecology* 2018; **27**: 3408–
647 3419.
648 73. Raymann K, Shaffer Z, Moran NA. Antibiotic exposure perturbs the gut microbiota and
649 elevates mortality in honeybees. *PLoS Biology* 2017; **15**: e2001861.
650 74. Koch H, Schmid-Hempel P. Socially transmitted gut microbiota protect bumble bees
651 against an intestinal parasite. *Proceedings of the National Academy of Sciences* 2011; **108**:
652 19288–19292.
653 75. Weiss BL, Maltz MA, Vigneron A, Wu Y, Walter KS, O'Neill MB, et al. Colonization of
654 the tsetse fly midgut with commensal *Kosakonia cowanii* Zambiae inhibits trypanosome
655 infection establishment. *PLOS Pathogens* 2019; **15**: e1007470.
656 76. Boggs CL, Watt WB, Ehrlich PR. Butterflies: ecology and evolution taking flight. 2003.
657 The University of Chicago Press.