Structure and function of the bacterial and fungal gut flora of Neotropical butterflies

Alison Ravenscraft*[1,2], Michelle Berry[1], Tobin Hammer[3], Kabir Peay[1], Carol Boggs[4]

[1] Department of Biology, Stanford University, Stanford, CA 94305 USA

[2] Current address: Center for Insect Science, University of Arizona, Tucson, AZ 85721 US

[3] Ecology and Evolutionary Biology, University of Colorado Boulder, Boulder, CO 80309 USA

[4] Department of Biological Sciences, University of South Carolina, Columbia, SC 29208 USA

*Author for correspondence: aravenscraft@gmail.com; tel 314-791-0738

Author Contributions: AR, MB, KP and CB designed the experiments. AR and MB performed the field work. AR conducted the lab work for the Illumina sequencing. AR, MB, and TH analyzed the data. AR wrote the manuscript and all authors provided additional writing and editorial feedback.

Abstract

The relationship between animals and their gut flora is simultaneously one of the most common and most complex symbioses on Earth. Despite its ubiquity, our understanding of this invisible but often critical relationship is still in its infancy. We employed adult Neotropical butterflies as a study system to ask three questions: First, how does gut microbial community composition vary across host individuals, species and dietary guilds? Second, how do gut flora compare to food microbial communities? Finally, are gut flora functionally adapted to the chemical makeup of host foods? To answer these questions we captured nearly 300 Costa Rican butterflies representing over 50 species, six families and two feeding guilds: frugivores and nectivores. We characterized the bacteria and fungi in guts, wild fruits and wild nectars via amplicon sequencing and assessed the catabolic abilities of the gut flora via culture-based assays.

Gut communities were distinct from food communities, suggesting that the gut environment acts as a strong filter on potential colonists. Nevertheless, gut flora varied widely among individuals and species. On average, a pair of butterflies shared 21% of their bacterial species and 6% of their fungi. Host species explained 25-30% of total variation in microbial communities while host diet explained 4%. However, diet was still relevant at the individual microbe level—half of the most abundant microbial species differed in abundance between frugivores and nectivores. Diet was also related to the functional profile of gut flora: compared to frugivores, nectivores' gut flora exhibited increased catabolism of sugars and sugar alcohols and decreased catabolism of amino acids, carboxylic acids and dicarboxylic acids. Since fermented juice contains more amino acids and less sugar than nectar, it appears that host diet filters the gut flora by favoring microbes that digest compounds abundant in foods.

By quantifying the degree to which gut communities vary among host individuals,

species and dietary guilds and evaluating how gut microbial composition and catabolic potential are related to host diet, this study deepens our understanding of the structure and function of one of the most complex and ubiquitous symbioses in the animal kingdom.

Key words: microbiota, Lepidoptera, diet, feeding guild, symbiosis, catabolism

Introduction

Microbes have been detected in the gut of almost every animal studied to date. This ubiquity is underpinned by the myriad functions these microbes serve: Gut microbes can assist animals with the uptake, synthesis and recycling of nutrients, breakdown of toxic or recalcitrant chemicals, and resistance to pathogens (Dillon and Dillon 2004). Despite their prevalence and importance, our knowledge of how and why these symbiotic communities change across host individuals, species, functional ecological groups, and geographical locations is still in its infancy.

At the most basic level, variation in the community composition of the gut flora can result either from exposure of the host to different pools of potential microbial colonists (including microbes transmitted vertically from parents to offspring), or from selective filtering of microbes by the physical and chemical conditions of the gut. Differences in diet or gut physiology between host dietary guilds, host species, and individual host organisms can affect both colonization by, and survival of, microbes in the gut.

Across disparate animal groups, host dietary guild has frequently attracted interest as a potential determinant of gut community composition. Different foods contain different microbial flora, and it is often assumed that consumption of foods can directly alter the composition of the gut community by introducing microbial colonists to the gut. This role appears to be most important during initial colonization. Once established the gut community is normally resilient, and most food-borne microbes pass through without becoming residents (e.g. Robinson et al 2010b; McNulty et al 2011). In addition to exposing hosts to alternative pools of microbes, diet can alter the composition of both horizontally acquired and vertically transmitted gut microbes by determining nutrient availability in, or affecting the chemical conditions of, the gut habitat (Robinson et al 2010a). Ingestion of a particular nutrient will promote the growth of gut

67 microbes that digest that nutrient, leading to feedbacks between diet and microbial community
68 function. For example, anaerobic fungi use lignocellulose as a food source; consumption of high
69 fiber foods by ruminant species leads to greater abundance of these fungi than in non-ruminants
70 (Gordon and Phillips 1998; Solomon et al 2016). These two mechanisms—food-borne microbial
71 colonists and the chemical composition of the diet—are generally expected to result in similar
72 gut flora among hosts that eat similar foods. Indeed, in many systems, gut community
73 composition has been shown to be more similar between species that eat similar diets, even when
74 these species belong to evolutionarily divergent groups (Colman et al 2012; Delsuc et al 2014).
75 For example, herbivorous, omnivorous, and carnivorous mammals and fish host gut communities
76 that are more similar within than between these three diet categories (Ley et al 2008a; Muegge et
77 al 2011; Sullam et al 2012). In insects, species belonging to the wood-feeding and detritivorous
78 guilds both host convergent gut flora (Colman et al 2012).

79 But host dietary guild is only one of many factors that may affect gut microbial
80 community composition. Differences in colonization can also result from living in different
81 habitats (e.g. Xiang et al 2006; Belda et al 2011), which results in sampling of divergent
82 microbial pools by hosts. Additionally, the chemical conditions of the gut vary among species,
83 independent of diet, and may act as a filter that selects for a specific gut community (Rawls et al
84 2006). For example, the guts of small animals, such as insects, range from fully aerobic to
85 anaerobic and therefore favor different sets of microbes (Johnson and Barbehenn 2000).
86 Similarly, gut pH varies among species and likely determines which microbes can establish in
87 the gut (Beasley et al 2015).The host immune system can also regulate potential gut colonists
88 (McFall-Ngai 2007; Salzman et al 2009).

Gut flora vary among host species not only in the composition of the core community, but also in the degree of variation among individuals. Such intraspecies variation can result from inter-individual differences in all of the factors mentioned above, including individual hosts' diets, microbial source pools, gut chemistries, and immune systems. Species with the least variable communities are often those that depend on their gut flora to survive on particularly poor or recalcitrant diets, such as termites or herbivorous ants; these species also often have elaborately structured guts that both facilitate digestion and provide microenvironments that support a stable gut community (Engel and Moran 2013). In contrast, omnivorous hosts with simply-structured, tube-like guts, such as *Drosophila*, tend to have gut flora that vary more among individuals (Engel and Moran 2013). In general, gut membership is expected to be more stable when the gut flora serve a more crucial function for their host.

Insects offer a tractable system to understand the causes and consequences of variation in gut microbial communities due to their variation in dietary guild, their abundance and species diversity, and their often relatively simple gut flora. Further, insects are major primary consumers, pollinators, and disease vectors in terrestrial ecosystems. Their gut flora may therefore have a large, yet hidden, impact on ecosystem function (e.g. Nardi et al 2002). Indeed, the ability of many insects to subsist on nutritionally unbalanced or recalcitrant foods often stems directly from contributions of their gut flora (e.g. Warnecke et al 2007). Despite the importance of, and increasing interest in, the insect gut flora in general, the adult lepidopteran microbiome has been largely ignored. Although a few culture-based studies confirmed the presence of bacteria in the adult gut decades ago (Steinhaus 1941; Kingsley 1972), culture-independent profiling is only beginning. Such data exist for only a single species, *Heliconius erato*

(Nymphalidae), and indicate that about 12 bacterial species dominate the adult gut (Hammer et al 2014).

In contrast to adults, the gut flora of several larval lepidopterans—particularly those of economic importance—have been characterized (Broderick et al 2004; Robinson et al 2010b; Pinto-Tomás and Sittenfeld 2011). However, the adult gut flora differ from those of larvae (Hammer et al 2014). This difference may derive from the fact that during pupation, the contents of the guts are voided, antimicrobial peptides are secreted into the gut lumen, and gut itself is replaced, presumably eliminating a large portion of the larval gut flora (Russell and Dunn 1996; Hakim et al 2010; Johnston and Rolff 2015). Indeed, in a culture-based study, Kingsley (1972) found that the density of colony-forming units derived from the guts of monarch butterflies (*Danaus plexippus*) decreased 1000-fold between pupae and freshly emerged adults. Existing data on larval lepidopterans therefore do not shed light on the gut flora of adult butterflies. Butterflies are biologically and economically important as herbivores and pollinators, and in areas outside of gut community ecology they are well-studied model organisms in the fields of population and nutritional ecology (Boggs et al 2003). Characterization of the causes and consequences of variation in the adult lepidopteran gut flora will therefore pave the way for connections between microbiome research and classical questions in ecology and evolutionary biology.

Here we use adult butterflies as a novel study system to characterize patterns of variation in the gut flora at the level of host individuals, species, and feeding guilds. We focus on how host nutritional ecology affects the community composition and functional capacity of the gut flora. The Neotropical community of butterflies allows for a high degree of replication both within and between species while controlling for geographic origin. Neotropical butterflies also exhibit

134 well-described ecological variation that likely structures the gut community, particularly in
135 feeding behavior: adults of some species feed on nectar, while others feed on the juice of rotting
136 fruits. Adults are easily captured in the wild, permitting examination of the natural gut flora and
137 avoiding the microbial community shifts frequently observed in captive animals (e.g. Chandler et
138 al 2011). These characteristics make Neotropical butterflies a useful system to elucidation
139 patterns in the composition and function of the gut microbial community.

140 We asked: (1) What bacteria and fungi are present in the guts of adult butterflies? (2)
141 How does gut microbial community composition vary among individuals, species, and dietary
142 guilds? (3) How does gut microbial composition compare to that of butterfly foods? (4) How
143 does community composition translate into functional abilities—specifically, catabolism of
144 various carbohydrates and amino acids?

145 Based on previous studies of other animal taxa (Ley et al 2008a; Colman et al 2012), we
146 expected butterfly gut communities to cluster according to both host diet and host species, such
147 that microbial beta diversity would be lowest between conspecific hosts, intermediate between
148 hosts that belong to different species but the same feeding guild, and highest between
149 heterospecific hosts that belong to different feeding guilds. We hypothesized that we would
150 observe signals of environmental acquisition of the gut flora. Specifically, we predicted that most
151 of the gut community would be a subset of the microbes present in the host's food, and that
152 frugivores, which feed on more "contaminated" foods (e.g. rotting fruit), would have more dense
153 gut communities than nectivores. Finally, we hypothesized that the digestive abilities of the gut
154 community would differ between feeding guilds, and that gut community catabolism should be
155 highest for nutrients abundant in the host's diet.
156

Methods

Study site

Butterflies were collected at La Selva Biological Research Station (10° 26'N, 83° 59' W). La Selva is located in the lowlands on the Caribbean versant of the Cordillera Central in Costa Rica. The station owns 1,600 hectares, 55% of which is primary tropical wet forest; the remainder comprises secondary forest and disturbed habitats (e.g. abandoned pastures and plantations) in various states of regrowth. This habitat diversity allowed us to collect a large diversity of butterfly species while controlling for the geographic origin of our samples. La Selva also provided the facilities necessary for both the culture-dependent and -independent components of our project (e.g. a wet lab and laminar flow hood). From January to March 2013, we captured butterflies and sampled butterfly foods along the trail system in primary forest, secondary forest and recovering pastures. We returned to La Selva in March 2014 to collect additional food samples (fruits and nectars).

Study organisms

Neotropical lepidopterans belong to two main feeding guilds: nectar feeders and fruit feeders. The Papilionidae, Pieridae, Lycaenidae, Riodinidae, Hesperiidae, and some Nymphalidae feed primarily on flower nectar, while several subfamilies of the Nymphalidae—Satyrinae, Morphinae, Charaxinae, and some members of the Nymphalinae—feed primarily on rotting fruits or other non-floral liquids (DeVries 1987, DeVries 1988). Our sampling included representatives of all of the nectivorous families—though most individuals derived from the Pieridae and Nymphalidae—as well as representatives from all the frugivorous nymphalid

179 subfamilies described above (Supplementary Table 1). Three *Heliconius* (Nymphalidae:
180 Heliconiinae) species were sampled; this genus supplements its nectivorous diet with pollen.
181
182 *Sample collection and processing*
183 We used aerial nets to catch both nectivorous and frugivorous species. To capture frugivores we
184 also used Van Someren-Rydon traps (Daily and Ehrlich 1995) baited with a fermented mixture
185 of fruit (primarily mango), molasses and rum. Traps were checked at least once every two days.
186 Upon capture, butterflies were placed in glassine envelopes and transported to the lab.
187 We identified butterflies to species using *The Butterflies of Costa Rica and Their Natural History*
188 (Devries 1987). Butterflies were fed a 5 to 10 microliter droplet of a filter-sterilized solution of
189 sugar and nigrosin dye (Sigma 198285) to assist with visualization of their guts during
190 dissection. Butterflies were euthanized with ethyl acetate, washed with 70% ethanol, and
191 dissected in a laminar flow hood to minimize contamination with environmental bacteria and
192 fungi. Guts were removed and homogenized in a small amount of inoculating fluid (described
193 below) in a sterile 2 mL tube. A portion of the homogenate was cultured to assess gut community
194 catabolic profile (described below). The remainder of the gut homogenate was suspended in cetyl
195 trimethyl ammonium bromide (CTAB) and refrigerated at 4°C until transportation to Stanford
196 University for DNA extraction. CTAB is an effective preservative for storage of insect samples
197 prior to microbiota analysis (Hammer et al 2015).
198 We also collected samples of the butterflies' foods to characterize potential source pools
199 of microbial gut communities. Trap baits were sampled in 2013 and 2014 by suspending
200 approximately 0.25 to 0.5 mL of rotten fruit slurry in CTAB. In 2014, nectar samples were also
201 collected. We used capillary tubes to extract nectar from flowers in the field, obtaining between

0.5 and 10 µL per flower. Nectars were also preserved in CTAB. We only collected nectar from species at which we directly observed at least one instance of butterfly feeding. In 2013, we sampled fruits of *Dipteryx oleifera* (Fabaceae) by scraping flesh from the surface of the fruit into CTAB. In 2014, we rinsed the fruit with 5 mL of sterile water and preserved 0.25 to 0.5 mL of the rinsate, since butterflies only feed on surface juices and do not consume the pulp. *Dipteryx oleifera* is the second-most abundant leguminous tree in primary forest at La Selva (McDade et al 1994) and was the only wild fruit we observed butterflies feeding upon during both visits to the site.

Research was carried out under permit numbers 202-2012-SINAC and R-016-2014-OT-CONAGEBIO.

Community catabolic profiling

A portion of each gut homogenate was seeded into 10 mL of inoculating fluid (IF-A, Biolog Catalog No. 72401) and cultured in microbial catabolic phenotype plates (Gen III microplates, Biolog Inc., Hayward, CA). The volume of gut homogenate used depended on the size of the butterfly: we used 25 microliters for most animals, but took less volume from the largest species (whose gut homogenate was more concentrated, since the guts themselves were large), and more volume from the smallest species (whose gut homogenate was less concentrated, since the guts were small). This was meant to correct for the grossest differences in inoculation density; further corrections are described below.

The catabolic phenotyping plates allowed simultaneous assessment of catabolism of 71 carbon and nitrogen sources (e.g. glucose, fructose, ammonia, uric acid; Table S5) via a colorimetric assay. Plates were incubated at ambient temperature. Absorbance at 590 nm was

measured twice per day using a 96-well plate reader (Chromate-4300, Awareness Technology Inc., Palm City, FL) until color remained stable (5 days on average).

To reduce bias due to differences in inoculum density between the samples, the data were standardized as in Garland et al (2001). Briefly, based on visual inspection of the distribution of plate averages, we chose a target average absorbance value per plate of 190. For each plate, we selected data from the timepoint at which the average absorbance was closest to this target. These measurements were standardized by subtracting the value of the negative control and dividing by the plate's average absorbance.

Microbial community characterization

Samples were shipped to Stanford University, where they were homogenized via bead beating, extracted with chloroform, and cleaned using the DNeasy Blood and Tissue Kit (Qiagen, Germantown, MD). Our DNA extraction protocol is described in detail in Peay et al (2007).

Bacterial DNA was amplified with primer set 515f (5'-GTGCCAGCMGCCGCGGTAA-3') and 806r (5'- GGACTACHVGGGTWTCTAAT-3'), which amplifies the V4 hypervariable region of the 16S rRNA with few taxonomic biases (Bergmann et al 2011). The bacterial amplicons were indexed with barcoded forward and reverse primers to (Caporaso et al 2012; Kozich et al 2013). The PCR reaction contained 0.2 µM forward primer, 0.2 µM reverse primer, 0.2 mM dNTP, 0.65 U OneTaq HotStart (New England Biolabs) and 1X Thermopol buffer (New England Biolabs) in a volume of 25 µL. The thermocycler program began with denaturation at 94 C for 3 minutes followed by 35 cycles of denaturation at 94 C for 45 seconds, annealing at 50 C for 60 seconds, and extension at 68 C for 90 seconds, with a final extension of 68 C for 10 minutes (adapted from the Earth Microbiome Project protocol; Gilbert et al 2014).

Fungal DNA was amplified with primers ITS1f (5'-CTTGGTCATTTAGAGGAAGTAA-3') and ITS2 (5'-GCTGCGTTCTTCATCGATGC-3') (White et al 1990; Gardes and Bruns 1993). The reverse primer was barcoded to index the samples (Smith and Peay 2014). The PCR recipe was identical to the bacterial PCR recipe above. Thermocycler conditions were denaturation at 95 C for 1 minute followed by 35 cycles of 94 C for 30 seconds, 52 C for 60 seconds, and 68 C for 60 seconds, with a final extension of 68 C for 5 minutes. Each bacterial and fungal amplification was run in triplicate to minimize the effects of stochastic amplification.

Bacterial and fungal PCR products were cleaned with AMPure XP magnetic beads (Agencourt A63881) and the final concentration of each sample was quantified fluorometrically with a 96-well plate reader (Qubit dsDNA HS kit, Thermo Fisher Q32854). An equal mass of DNA from each sample was added to the bacterial or fungal library, respectively. Bacterial and fungal libraries were sequenced in separate runs on an Illumina MiSeq platform at the Stanford Functional Genomics Facility (Stanford, CA), with 2 x 250 chemistry for bacteria and 2 x 300 chemistry for fungi.

Illumina sequence data processing and cleaning

We used the program cutadapt (Martin 2011) to remove priming sites and poor quality bases at the 5' and 3' ends of the sequences. Sequences were merged and clustered at a 97% similarity cutoff with UPARSE (Edgar 2013). Both de-novo and reference-based chimera checking were performed in UPARSE; bacterial reads were compared to the RDP Gold database and fungal reads to the UNITE database. Bacterial taxonomy was initially assigned using the RDP classifier (Wang et al 2007) with Greengenes (McDonald et al 2012) as the training set. We used the RDP

271 classifier with the Warcup (Deshpande et al 2016) training set to assign fungal taxonomy. We
272 checked and revised these assignments by aligning our representative sets of bacterial and fungal
273 sequences against the NCBI nucleotide collection using BLAST.

274 OTUs that were identified as lepidopteran 18S, archaeans, mitochondria, or chloroplasts
275 were excluded. (These accounted for less than 0.1%, less than 0.1%, 0.6%, and 4.8% of the raw
276 total reads, respectively, and 98% of the chloroplast reads derived from samples of butterflies'
277 foods.) Additionally, DNA extraction kits and other laboratory reagents are known to contain
278 microbial DNA that can contaminate microbiome analyses (Salter et al 2014). We found fourteen
279 bacterial OTUs that were present at higher abundance in the negative controls than the butterfly
280 gut samples (Supplementary Table 3); these were classified as contaminants and removed from
281 the dataset prior to statistical analyses. The fungal dataset (especially nectivore samples)
282 contained many ectomycorrhizal lineages that were unlikely to be members of the butterfly gut
283 community. This suggested that the butterfly gut samples (particularly those of nectivores) had
284 low fungal biomass, resulting in detection of ambient lab and environmental contaminants. We
285 removed lineages that were likely contaminants from the analysis: we omitted 135 OTUs
286 assigned to the genus *Mortierella*, which are decay fungi that inhabit soil. Wood decomposers
287 and ectomycorhizal, arbuscular or ericoid mycorrhizal fungi were also removed from the dataset,
288 as were lineages that were present at higher abundance in the negative controls than in the
289 samples (Supplementary Table 4). We retained lineages that were known phylloplane
290 inhabitants, plant pathogens, and fruit decomposers, as well as most ascomycete and
291 basidiomycete yeasts.

292 Frugivorous butterflies were captured either with traps baited with fermented fruit or with
293 aerial nets. For those captured in baited traps, a portion of the gut flora likely derived from

294 microbes in the trap baits. To identify these microbes, we tested for differential abundance of
295 bacterial and fungal OTUs between trapped frugivores and netted frugivores using the *mt*
296 function in the R package "phyloseq" (McMurdie and Holmes 2013), which includes an FDR
297 correction for multiple testing. No bacterial OTUs differed in abundance between netted and
298 trapped animals, but three fungal OTUs were significantly more abundant in the trapped
299 butterflies. These OTUs were all strains of the yeast *Kazachstania exigua*, which was also by far
300 the most abundant fungus detected in the traps. These three OTUs accounted for 43.7% of all
301 fungal reads from the fermented baits and only 0.3% of reads from wild fruits. We concluded
302 that the OTUs had been introduced into the trapped butterflies' guts via the fermented baits and
303 removed them from the data derived from butterfly samples.
304 To control for differences among samples in sequencing depth, we rarefied the
305 sequencing data using phyloseq. Bacteria were rarefied to 1000 and fungi to 200 sequences per
306 sample. These cutoffs allowed us to retain as many samples in the dataset as possible while still
307 profiling the dominant microbes. All sequencing results reported are based on rarefied data,
308 unless otherwise noted.
309
310 *Quantification of total bacterial abundance*
311 Bacterial DNA was quantified with qPCR using SYBR green fluorescent chemistry (iCycler IQ,
312 Bio-Rad, Hercules, CA). Since the 515f/806r primer set sometimes amplifies butterfly 18S in
313 addition to bacterial 16S rRNA, we designed a PNA clamp to block 18S amplification (Lundberg
314 et al 2013). The clamp sequence was GCCCGCTTTGAGCACTCT and it was synthesized by
315 PNA Bio (Thousand Oaks, CA). The reaction volume was 20 µL and the recipe was 7.5 µM
316 PNA clamp, 0.2 µM 515f primer, 0.2 µM 806r primer, and 1X PerfeCTa SYBR Green FastMix

317 for iQ (Quanta Biosciences, Gaithersburg, MD). Thermocycler settings were an initial 10-minute
318 denaturation at 95 C followed by 45 cycles of denaturation at 95 C for 15 sec, PNA clamp
319 annealing at 76 C for 10 sec, primer annealing at 50 C for 30 sec, and extension at 68 C for 30
320 sec. To check that amplified fragments were the expected length, we performed a melt curve
321 ramping from 55 C to 95 C in 0.5 C increments at 10 second intervals. PCR products were also
322 visualized using gel electrophoresis.

323 To calculate the starting number of 16S rRNA copies, each sample's threshold cycle was
324 compared to an internal standard curve ranging from an initial 10 to 10^7 copies per µL of *E. coli*
325 16S rRNA amplicons. (Generation of these standards is described below.) Each sample and
326 standard was run in triplicate and the results were averaged. The correlation coefficients of the
327 standard curves for all qPCR plates ranged from 0.973 to 0.988.

328 Since a fraction of each sample was removed for culture-based catabolic profiling prior to
329 DNA preservation and extraction, the qPCR estimates were corrected accordingly. For example,
330 if 50% of the gut homogenate from a given sample was removed prior to DNA preservation, that
331 sample's qPCR estimate was multiplied by two to obtain the estimated total 16S rRNA count for
332 the entire gut.

333 The 16S rRNA standard curve was generated as follows: 16S rRNA was amplified from
334 *E. coli* using primers 27f and 1492r, ligated into a plasmid vector, and cloned into chemically
335 competent cells. Colonies were screened for inserts of the expected size by PCR amplification
336 using M13f and M13r primers followed by gel electrophoresis. To further screen for the correct
337 insert, several colonies were bidirectionally sequenced. One colony that passed both screens was
338 selected for the standard curve. This clone was grown to saturation in LB + kanamycin selective
339 media. Plasmids were extracted using the Qiagen Plasmid Mini kit and linearized using the

restriction enzyme SpeI (FastDigest, Thermo Fisher FD1253). DNA was purified using a Qiagen PCR cleanup column and run on a gel to verify complete linearization. DNA concentration was quantified with PicoGreen and copy number per µL was calculated as the molecular weight of plasmid plus insert (g/molecule, calculated as length in bp * 660 Da/bp / $6.02*10^{23}$) divided by the measured DNA concentration (g/µL). Via serial dilution of the raw extract with sterile TE, we generated a standard curve ranging from 10^7 copies/µL to 10 copies/µL.

We were not able to quantify total fungal abundance because our ITS primers amplified fragments of vastly different lengths (ranging from 250 base pairs to over 1000 bp).

Statistical analyses

All models had Gaussian error structure and were fit in R version 3.2.2 (R Core Team 2015). Models without random effects were fit using the *lm* command in the "stats" package. Models with random effects were fit using the *lmer* command in the package "lme4" (Bates et al 2014).

a. Variation in total bacterial load among host species and feeding guilds

A log transformation was applied to the 16S rRNA counts prior to analysis. We used linear and linear mixed effect models (Zuur et al 2009) to assess whether total bacterial load differed among host species or between host feeding guilds. We used a likelihood ratio test to compare models with and without a random effect of host species, including fixed effects of host diet, wing length (as a proxy for the animal's size), and capture date. We then tested for significance of the fixed effects using backwards model selection with likelihood ratio tests, including a random effect of host species in all models.

363 *b. Variation in OTU richness*

364 To test whether observed bacterial and fungal richness differed among host species, we again
365 used a likelihood ratio test to compare models with and without a random effect of host species,
366 including a fixed effect of host diet to control for overall diet-based differences. We tested for
367 correlation between host diet and observed bacterial and fungal richness by comparing models
368 with a fixed effect of diet to models without the fixed effect; both models had a random effect of
369 host species. We omitted data from host species represented by a single individual from these
370 analyses.

371

372 *c. Variation in microbial community composition*

373 *Between feeding guilds:* To test for a relationship between butterfly feeding guild and gut
374 microbial community composition, we used ordination (NMDS) plots to visualize and
375 perMANOVAs to test for dissimilarity in microbial community composition (Anderson 2001).
376 Both operations were performed on the Bray-Curtis dissimilarities between rarefied samples. We
377 performed two perMANOVAs (the *adonis* test in the R package "vegan"; Oksanen et al 2015) to
378 test for dissimilarity in bacterial and fungal communities, respectively. These tests included
379 terms for both host diet and host species, thus simultaneously testing whether gut communities
380 differed between host species and feeding guilds. Host species represented by a single individual
381 were omitted from these analyses.

382 If communities differ in variance, perMANOVA results can be unreliable (Anderson and
383 Walsh 2013). To verify the results of the tests, we therefore tested for differences in dispersion
384 among feeding guilds using the *betadisper* function from the package "vegan." For bacteria, the
385 feeding guilds did not differ in dispersion (anova: df =1, F=0.12, p=0.73) therefore the data

386 conformed to the assumptions of perMANOVA. For fungi, the feeding guilds did differ in spread
387 (anova: *df* =1, F=4.4, p=0.04). However, after removing the five host species with extreme
388 dispersions, the feeding guilds no longer differed in spread (anova: *df* =1, F=1.3, p=0.21). The
389 perMANOVA results with these five species removed were qualitatively identical and
390 numerically similar to those obtained from all host species, therefore we report the test results for
391 all species (omitting those represented by a single individual).

392 In order to understand the effect of host diet on individual OTU abundances, we modeled
393 the relative abundance of a bacterial or fungal OTU (counts out of 1000 or 200, respectively) as a
394 function of the interaction between host diet and OTU identity, with a random effect of host
395 species. Since it was not feasible to model abundances of every OTU, we selected the 20
396 bacterial or 20 fungal OTUs that were present at the highest abundances in the pooled data and
397 were detected in at least 10% of either frugivore or nectivore samples. Prior to the analyses,
398 abundance data were transformed by log(x+1) to homogenize variance.

399 *Among host species:* We used perMANOVAs (described above) to test for dissimilarity
400 in the gut flora among host species while accounting for differences between feeding guilds.
401 Since perMANOVA assumes homogeneity of variance among communities, we again used the
402 *betadisper* function to test for differences in dispersion. Host species' bacterial and fungal
403 communities did differ in dispersion (bacterial anova: *df* =36, F=2.4, p<0.001; fungal anova: *df*
404 =28, F=11.4, p<0.001). For bacteria the difference was driven by five species that had unusually
405 low variance; when these were removed, the remaining 32 species no longer differed in spread
406 (anova: *df* =31, F=0.9, p=0.65). For fungi, the difference was also driven by five extreme
407 species. The remaining 24 species in the fungal dataset no longer differed in variance (anova: *df*
408 =23, F=1.3, p=0.16). For both bacteria and fungi, we re-ran the perMANOVA tests using the

409 subset host species that did not differ in dispersion. Results were again qualitatively the same as
410 those for all host species, so we report the results for all species.

411 *Among individuals:* To quantify beta diversity among individual butterflies' gut flora, we
412 calculated the median percentage of OTUs shared between individuals, where the percentage
413 shared between individuals A and B was calculated as 0.5 * (number of A's OTUs shared by B /
414 total OTUs in A + number of B's OTUs shared by A / total OTUs in B).

415 *Between food microbial communities and gut communities:* We used NMDS plots (on
416 Bray-Curtis distances between rarefied samples) to visualize differences between the butterfly
417 gut flora and the microbial composition of butterfly foods. To test whether gut flora were more
418 similar to the microbial communities of one food than to another food, we calculated the
419 pairwise Bray-Curtis distances between each butterfly gut community and each food community
420 and performed t-tests on the per-butterfly averages of these distances for the three foods (fruits,
421 nectars and baits). To correct the resulting p-values for multiple testing, we used the Benjamini-
422 Hochberg false discovery rate (FDR) method as implemented by the *p.adjust* function in R
423 (Benjamini and Hochberg 1995).

424

425 *d. Variation in microbial community function*

426 The Hellinger transformation was applied to the standardized catabolic data prior to all analyses.

427 *With gut community composition:* To test whether differences in microbial community
428 function were correlated with differences in microbial species composition, Mantel tests were
429 performed between the Bray-Curtis distances between bacterial or fungal microbial communities
430 and the Euclidean distances between catabolic profiles. Since trap bait-derived fungal OTUs
431 (three strains of *Kazachstania exigua*) were introduced into the Biolog plates via frugivores'

guts, these OTUs were included in the calculation of Bray-Curtis distances between fungal communities. However, results were qualitatively equivalent when these three OTUs were omitted from the Bray-Curtis calculations.

With host species and diet: First, we tested for correlation between gut community catabolic profile and both host diet and host species by performing a perMANOVA on the Euclidean distances between the gut catabolic profiles. (Singleton host species were omitted from this analysis.) To investigate differences between frugivores and nectivores in more detail, we next modeled substrate catabolism as a function of the interaction between host diet and substrate identity, including a fixed effect for the log of the number of 16S rRNA copies in the host's gut to control for differences in bacterial density, and a random effect for the identity of the host butterfly to control for overall differences in catabolic activity of the gut community among individuals. Finally, to test whether catabolism of certain classes of substrate (for example, amino acids or sugars) was consistently up- or down-regulated with host diet, we modeled substrate catabolism as a function of the interaction between host diet and substrate class, with a fixed effect for log(16S rRNA counts) and a random intercepts of host identity as before, plus an additional random intercept and random diet effect (slope term) for each substrate. Thirteen substrates belonged to classes with three or fewer substrates and were therefore omitted from the class-level analysis. For both models, least-square means and the contrasts between them were calculated using the *lsm* command in the "lsmeans" package (Lenth 2016) to assess the whether a substrate or substrate class was catabolized differently between frugivores and nectivores. The resulting p-values for each substrate or substrate type, respectively, were FDR corrected to control for multiple testing.

Results

Below we first describe the overall composition of the adult butterfly gut flora, then address patterns of variation in microbial community composition at the level of host feeding guild, host species, and host individual. Gut communities are then compared to the microbes present in butterfly foods. Functional differences between gut microbial communities are described last.

1. Bacteria and fungi present in the butterfly gut

We sequenced the bacterial flora of 306 butterflies, of which 290 were retained after rarefaction, and the fungal gut communities of 247 butterflies, 161 of which were retained after rarefaction. (See Table 1 for a full sample size breakdown and Supplementary Table 1 for counts per butterfly host species). After sequence processing, quality filtering, and removal of contaminant OTUs, 7.2 million bacterial sequences and 1.5 million fungal sequences were obtained from butterfly guts. After rarefying, a total of 958 bacterial OTUs and 880 fungal OTUs were observed across all butterflies. (Note that numbers of bacterial and fungal OTUs cannot be compared because bacteria and fungi were rarefied to different read depths.) Individual butterflies hosted a mean of 30 (interquartile range:18-37) bacterial OTUs (number observed at a 1000 read per sample cutoff) and 21 (IQR: 12-28) fungal OTUs (number observed at a 200 read per sample cutoff).

The most prevalent bacterial phyla and classes across all butterflies were Proteobacteria:Gammaproteobacteria (40% of all butterfly-derived bacterial sequences), Proteobacteria:Alphaproteobacteria (28%), Firmicutes:Bacilli (12%), Bacteroidetes:Flavobacteriia (7%), and Tenericutes:Mollicutes (4%). The dominant fungal class was Saccharomycetes (62% of all butterfly-derived fungal sequences) in the phylum

478 Ascomycota. All of these Saccharomycetes belonged to the order Saccharomycetales, which
479 defines the "budding yeasts" or "true yeasts." Other common fungal clades were
480 Basidiomycota:Tremellomycetes (8%), Zygomycota:Mucoromycotina (6%),
481 Ascomycota:Dothideomycetes (6%) and Basidiomycota:Microbotryomycetes (5%).
482 The 20 most abundant bacterial and fungal OTUs across all butterflies (excluding those
483 detected in less than 10% of frugivores or nectivores) are listed in Tables 2 and 3. The
484 predominant bacteria included known associates of butterflies (*Orbus* spp.; Kim et al 2013;
485 Hammer et al 2014) and other insects with sugar-rich diets (e.g. *Erwinia* sp., *Asaia* sp.,
486 *Commensalibacter intestini;* Crotti et al 2010), putative insect parasites and pathogens (e.g.
487 *Spiroplasma* sp., *Serratia* spp., *Wolbachia* sp.), common gut inhabitants (e.g.
488 *Klebsiella*/*Enterobacter, Vagococcus/Enterococcus*), and common fermenting bacteria (e.g.
489 *Lactococcus* sp., *Acetobacter* sp.). The most abundant fungi were largely ascomycotous yeasts
490 (especially members of the genera *Hanseniaspora*, *Pichia*, *Kazachstania* and *Candida*), many of
491 which are known to associate with insects and/or plants— particularly nectars, decaying fruits,
492 and beetle guts (Suh et al 2005; Kurtzman et al 2011). One OTU was most closely identified as a
493 *Rhizopus* species; this genus is commonly associated with decaying vegetable matter (Kirk et al
494 2008). Likely plant pathogens (e.g. *Clavariopsis* sp. and a Letiomycetes species) and plant-
495 associated basidiomycotous yeasts (OTUs that aligned most closely to the genera *Bensingtonia*)
496 were also common.
497 Several of the most abundant OTUs (those with the greatest number of counts across all
498 butterflies) were also the most prevalent (detected in the largest number of butterflies). The most
499 frequently detected bacterial OTUs were a *Vagococcus/Enterococcus* species found in 83% of
500 the butterflies, a *Klebsiella/Enterobacter* species found in 82%, and an *Orbus* species found in

73%. The most prevalent fungal OTUs were *Hanseniaspora uvarum* found in 45%, *Hanseniaspora opuntiae* found in 42%, and *Hanseniaspora guilliermondii* found in 40%.

1a. Total bacterial load in the butterfly gut

We quantified the total abundance of bacteria in the guts of 261 butterflies. The total number of 16S rRNA copies in a butterfly's gut ranged from 5×10^5 to 1×10^{11}, with a median of 7.5×10^8 (interquartile range: $1.2\times10^8 - 2.7\times10^9$). Larger butterfly individuals hosted greater numbers of bacteria (Supplementary Figure 1). Bacterial load differed among host species (Figure 1; random host species intercept term: $df = 1$, $\chi^2 = 17.5$, $p < 0.001$) but did not differ between feeding guilds after accounting for host species and size (fixed diet term: $df = 1$, $\chi^2 = 1.8$, $p = 0.18$).

2. Variation in gut microbial community composition between host dietary guilds, species, and individuals.

2a. Variation between the feeding guilds

The gut flora of frugivores and nectivores did not differ from each other in observed OTU richness (Figure 2; bacteria fixed feeding guild term: $df = 1$, $\chi^2 = 0.98$, p = 0.32; fungi fixed feeding guild term: $df = 1$, $\chi^2 = 0.09$, p = 0.76). However, gut microbial community composition did differ between the feeding guilds (bacteria: permanova, F= 13.2, R^2= 0.040, p=0.001, Figure 3a; fungi: permanova, F= 7.4, R^2= 0.038, p=0.001, Figure 3b). Host feeding guild explained 4% of the variation in both bacterial and fungal community compositions.

Despite substantial inter-individual and inter-specific variation in microbial community composition (described below), average relative abundances of 10 of the 20 most abundant bacteria (Figure 4) and fungi (Figure 5) systematically differed between the feeding guilds. Of

524 the bacteria, *Swaminanthia/Asaia* sp., *Bartonella* sp. and a strain of *Commensalibacter intestini*
525 were at higher relative abundance in nectivores than frugivores. Another strain of
526 *Commensalibacter intestini*, *Wolbachia* sp., a Porphyromonadaceae species, *Gilliamella/Orbus*
527 sp., *Orbus* sp. and *Acetobacter* sp. were at higher relative abundance in frugivores than
528 nectivores. Of the fungi, an unidentified fungus (OTU 7), a Letiomycetes species and a
529 Pleosporales species were more abundant in nectivores, while an unidentified ascomycete (OTU
530 20), two strains of *Kazachstania exigua*, *Hanseniaspora guilliermondii*, *H. uvarum*, *H. opuntiae*,
531 and *Pichia fermentans* associated more strongly with frugivores.
532
533 *2b. Variation among host species*
534 Observed OTU richness differed among butterfly species (Supplementary Figure 2; bacteria
535 random host species intercept term: $df = 1$, $\chi^2 = 43.3$, $p < 0.001$; fungi random host species
536 intercept term: $df = 1$, $\chi^2 = 52.3$, $p < 0.001$). Gut microbial community composition also differed
537 among host species (bacteria: permanova host species term, $df = 35$, $F = 2.3$, $R^2 = 0.244$, $p < 0.001$;
538 fungi: permanova host species term, $F = 2.3$, $R^2 = 0.319$, $p < 0.001$). Host species explained 24%
539 and 32% of the variation in bacterial and fungal community composition, respectively. Gut flora
540 were more similar within than between species (Figure 6).
541
542 *2c. Variation among individuals*
543 The vast majority of variation in community composition was expressed among individuals:
544 After accounting for species- and guild-level differences, residual variation in community
545 composition was 71.7% for bacteria and 64.4% for fungi (permanova). On average, a pair of
546 butterflies shared 23% (IQR 13%-31%) of their bacterial OTUs and 10% (IQR 0%-16%) of their

547 fungal OTUs. No OTU was present in all butterfly individuals. Each bacterial OTU was found in
548 a mean of 9 individuals (IQR: 1-6) and each fungal OTU was found in a mean of 4 individuals
549 (IQR: 1-3).
550
551 *3. Comparison of the gut flora to food microbial communities*
552 We sequenced the bacterial flora of 87 food samples, 83 of which were retained after
553 rarefaction, and the fungal flora of 75 foods, 66 of which were retained after rarefaction (Table 1;
554 Supplementary Table 2). After sequence processing and quality filtering we obtained 7.7 million
555 bacterial and 3.2 fungal sequences from potential food sources (fruits, nectars, and trap baits).
556 After rarefying, a total of 1205 bacterial OTUs and 529 fungal OTUs were observed across all
557 the food samples. (Again note that numbers of bacterial and fungal OTUs cannot be compared
558 due to differences in rarefaction depth.)
559 The microbial communities in wild fruit juice, the fermented trap baits, and wild nectars
560 were distinct from each other and from butterfly gut communities (Figure 7; Figure 8). In fact,
561 butterfly gut communities were highly dissimilar to the microbial communities on their food
562 sources. For example, frugivores and wild fruits shared 7% of their bacterial OTUs on average
563 (IQR 3%-9%), and the median Bray-Curtis dissimilarity between their bacterial communities
564 was 0.91 (Figure 8a). Despite these pronounced differences, frugivores' gut bacteria more
565 closely resembled the microbial composition of the trap baits than nectivores' gut bacteria did,
566 and similarly, nectivores' gut communities were more similar to nectar bacteria than frugivores'
567 gut flora were (Figure 8a). Frugivores' and nectivores' bacterial communities, however, were
568 equivalently distinct from those inhabiting the surfaces of wild fruits (Figure 8a).
569 Fungal communities in frugivore guts more closely resembled those of wild fruits and

570 trap baits than did those of nectivore guts, as expected (Figure 8b). Surprisingly, frugivores' gut
571 fungal communities were more similar, on average, to fungi in nectar than nectivores' gut
572 communities were (Figure 8b), though the effect size was extremely small.

573 While gut communities were distinct from food microbial flora in terms of species
574 composition and relative abundances, the percentage of bacterial gut reads assigned to OTUs that
575 were also present in the host's diet was generally high. On average, about 85% percent of gut
576 bacterial reads derived from OTUs present in their host's food. Fungal OTUs present in food,
577 however, accounted for a comparatively small 24% to 40% of butterfly gut fungal reads.

578

579 *4. Functional abilities: gut microbial community catabolism*

580 We measured the catabolic profiles of 248 gut communities via culture-based assays. Gut
581 communities digested an average of 51 substrates (IQR: 41-63) out of 71 total. As two
582 communities diverged in bacterial species composition, their catabolic profiles also became more
583 different; however, the effect size was very small (Mantel test: $r = 0.067$, $p = 0.013$). Catabolism
584 was not correlated with fungal community composition (Mantel test: $p = 0.50$).

585 Across the entire dataset, host diet explained just 3% of the total variation in community
586 catabolic profile (permanova on Euclidean distances between catabolic profiles; df =1, F= 7.6,
587 R^2=0.029, p=0.001), while host species identity explained 18% of the variation (df= 33, F= 1.4,
588 R^2=0.180, p=0.001). After accounting for diet- and species-level differences, residual variation
589 among individuals in catabolic profile was 79.1%.

590 Lack of strong correlation between catabolic profile and host diet could be due to the fact
591 that not all catabolic functions are relevant for bacteria in butterfly guts. To investigate diet-
592 based differences in catabolic profile in more detail, we modeled catabolism as a function of host

diet, substrate identity, and substrate class, controlling for butterfly identity and number of 16S rRNA copies in the butterfly. We also tested an alternate random effects structure, nesting butterfly individuals within butterfly species, but this did not improve model fit (anova: $df = 1$, $\chi^2 = 0.258$ p =0.61). Catabolism of several individual substrates differed markedly between the feeding guilds (feeding guild by substrate interaction term: $df = 70$, $\chi^2 = 428.9$ p <<0.001). Frugivorous gut flora catabolized 15 substrates more actively than nectivorous gut communities did, with the most marked difference being in catabolism of D-serine, saccharic acid, mucic acid, L-serine and lactic acid (Supplementary Table 5). Nectivorous gut flora digested 10 substrates more actively, with the most extreme differences in catabolism of mannitol, sucrose, glucose, fructose, and maltose (Supplementary Table 5). Entire classes of nutrients were also catabolized differently between the guilds (feeding guild by substrate class interaction term: $df = 6$, $\chi^2 = 61.7$, p <<0.001). Specifically, nectivores' gut flora catabolized sugars and sugar alcohols more actively than frugivorous gut flora, while frugivores' gut flora were more successful at digesting amino acids, carboxylic acids, and dicarboxylic acids (Figure 9; Supplementary Table 6).

Total number of 16S rRNA copies was not significantly correlated with catabolism (p= 0.09 for the model of substrate identity and p=0.07 for the model of substrate class), suggesting that our standardization procedure adequately controlled for any initial differences in cell inoculation densities.

Discussion

We evaluated how microbial community structure and catabolic function varied among host feeding guilds, species and individuals within a community of over 50 species of Neotropical butterflies. On average, 30 bacterial OTUs and 21 fungal OTUs were detected per

gut. Although the bacterial and fungal communities in adult butterfly guts varied substantially among individual hosts, they did differentiate among host species and between host feeding guilds. Gut community composition varied less between host feeding guilds than among host species, suggesting that other non-dietary aspects of host biology play a large role in structuring the butterfly gut flora. Gut communities were distinct in composition from the microbial flora found in butterfly foods, indicating that the adult butterfly gut environment strongly filters potential colonists. Despite high variability in the gut microbial community, host dietary guild was nevertheless associated with consistent changes in both the relative abundances of dominant microbes and the catabolic capacities of the gut flora. Frugivorous gut communities were better at digesting amino acids and carboxylic acids, while nectivorous gut flora outperformed frugivorous communities in catabolism of sugars and sugar alcohols. These catabolic patterns are congruent with the relative nutritional makeup of butterfly foods (Ravenscraft and Boggs 2016), which suggests that gut flora specialize in digestion of compounds abundant in the host's diet.

1. Microbes present in the guts of adult butterflies

1a. Bacteria

The composition of the butterfly's bacterial flora is broadly similar to that of other insects: at the phylum-level, Proteobacteria comprised 68% and Firmicutes 12% of all reads while in comparison, a meta-analysis of the insect gut flora reported that 57% were Proteobacteria and 22% were Firmicutes (Colman et al 2012). At finer taxonomic resolution (see Table 2), the most abundant bacterial OTUs were mostly known gut colonists, particularly those associated with sugar-rich diets, and OTUs related to insect parasites and pathogens. (Individual OTUs are discussed in more detail below.) Total bacterial load varied among species, but there

639 were no consistent differences between the feeding guilds after controlling for differences in size
640 among individual hosts. Larger butterfly individuals hosted greater numbers of bacteria. This
641 conforms previous work which demonstrated that total microbial load scales with host size (Kieft
642 and Simmons 2015).

643 The most abundant gut residents in our data were similar to those found in adult
644 *Heliconius erato* butterflies in Panama (Hammer et al 2014). The genera *Orbus*, *Enterobacter*,
645 *Asaia, Enterococcus*, *Lactococcus* and *Commensalibacter* were abundant in both studies, and
646 *Orbus* has also been isolated from the gut of the butterfly *Sasakia charonda* in South Korea (Kim
647 et al 2013). This suggests that adult butterflies do form consistent associations with some gut
648 microbes. What drives these associations is uncertain. Butterflies might maintain specific
649 relationships with some microbes via mechanisms such as vertical transmission. Gut flora can be
650 transmitted from mother to offspring via internal migration of gut microbes into the eggs, as
651 observed in the moth *Galleria mellonella*, or via the smearing of gut microbes on egg surfaces,
652 which has been hypothesized to occur in tobacco hornworm (*Manduca sexta*) (Brinkmann et al
653 2008; Freitak et al 2014). The relative frequency of such inter-generational transmission is
654 unclear, however, because the majority of the larval gut flora are purged during metamorphosis
655 (Kingsley 1972; Hammer et al 2014; but see Johnston and Rolff 2015). Other members of the gut
656 flora may be common in both butterflies and other insects because they are generally adapted to
657 the insect gut environment (e.g. Chouaia et al 2012) and opportunistically colonize many species.
658 Many of the most abundant genera we detected may fall into this category.
659
660 *1b. Gut fungi were primarily yeasts and common plant associates.*
661 The most abundant fungal OTUs were common plant and insect associates. Two thirds of

all fungal sequences belonged to the order Saccharomycetales, the "budding yeasts" or "true yeasts." These single-celled fungi are often associated with sugar-rich environments and decaying vegetable matter. Most of the remaining OTUs with reliable taxonomic assignments were likely plant associates or pathogens. Studies characterizing the gut flora often focus on bacteria and overlook fungi, but available evidence suggests that some insect guts host a rich diversity of fungi, especially yeasts (Suh et al 2005). Some of these fungi may play an active role in the gut, while others may simply use insects as a means of dispersal (Blackwell and Jones 1997; Starmer and Lachance 2011).

2. Variation in gut microbial community composition between host dietary guilds, among species, and among individuals

2a. A smaller degree of variation was attributable to differences between feeding guilds.

Frugivores and nectivores did not differ in total bacterial load nor in bacterial or fungal species richness, but did differ in the composition of their microbial flora. The degree of variation explained by dietary guild was relatively small (about 4% for bacteria and fungi), however it was comparable to that observed in many studies of microbial communities and is likely biologically relevant: relative abundances of half of the most common OTUs differed systematically between the feeding guilds. In general, the fungi more abundant in nectivores were closely related to plant pathogens, while the fungi more abundant in frugivores were yeasts associated with fruit decay, suggesting that this variation likely derived from exposure to different microbial source pools. In contrast, there was no pattern in the types of bacteria that differed in abundance between frugivores and nectivores.

686 *2b. A moderate degree of variation was attributable to differences among host species.*

687 About 24% and 32% percent of the variation in bacterial and fungal OTU compositions,
688 respectively, was attributed to differences among host species. Total bacterial load and bacterial
689 and fungal OTU richness also differed among butterfly species. Interspecies variation in the
690 butterfly gut flora could result from behavioral differences that expose butterflies to different
691 microbial pools. Such species traits could include habitat preferences. For example, species that
692 favor the forest understory could be exposed to more soil and wood decay fungi than those that
693 prefer the canopy. Additionally, many, but not all species drink from mud, dung, or carrion, in a
694 behavior known as "puddling" (Adler and Pearson 1982; Sculley and Boggs 1996), and variation
695 in these auxiliary feeding substrates could introduce different microbes to the gut. Differences
696 could also arise from variation in gut chemistry: for example, the guts of larger butterflies might
697 be more anoxic, or there could be interspecies differences in gut pH.

698 At least some of these factors could exhibit evolutionary signal in butterflies, resulting in
699 correlation between the butterfly phylogeny and gut community composition. Alternatively or
700 additionaly, vertical transmission of a fraction of the gut community from parent to offspring
701 could result in parallel divergence of microbiota and hosts. The potential for such evolutionary
702 patterns to contribute to interspecies differences in gut community composition is beyond the
703 scope of the current study, but will be addressed in a future paper (Ravenscraft et al, in prep).

704

705 *2c. Variation in microbial community composition among individual butterflies is high, but not*
706 *unusual.*

707 There was a great deal of variation among individual butterflies in microbial community
708 composition: the most prevalent bacterial and fungal OTUs were detected in only 83% and of

709 45% butterflies, respectively, and half of all OTUs were found in only one individual. Although
710 this degree of variation may seem extreme, it is comparable to that observed in other animals:
711 human colon microbiota were 34% similar among individuals in one study (Green et al 2006),
712 and in another, no OTUs were universally shared and 79% were found in only a single individual
713 (Tap et al 2009). A study of mammalian gut flora found that, on average, 56% of OTUs in an
714 individual mammal were unique to that individual (Ley et al 2008a). It has been proposed that
715 the high degree of variation observed in gut communities is the result of functional redundancy
716 at the OTU level: many microbes may require similar gut environmental conditions and/or serve
717 similar functional roles, leading to communities that differ in OTU membership but not in
718 overall function (Ley et al 2006; Peay et al 2016). However, it is also possible that variation in
719 membership *does* lead to significant variation in function (Peay et al 2016) and that the butterfly
720 gut flora do not play a consistent role across individuals.
721
722 *2d. The effect sizes of variables related to gut community composition are rarely reported or*
723 *compared.*
724 We had expected that a butterfly's feeding guild would explain a larger amount of
725 variation in gut community membership than host species identity, but we found the opposite.
726 The relative importance of dietary guild as a determinant of gut community composition has
727 been obscured in the literature by a trend of reporting only the significance (p-value) and not the
728 strength (either the effect size or R-squared) of these relationships. Diet has been shown to
729 significantly affect the community composition of the gut flora in many systems, including
730 mammals (Ley et al 2008a), fish (Sullam et al 2012), and insects (Colman et al 2012), but the
731 strength of this correlation was only reported for the last case. In a broad survey of the insect gut

732 flora, feeding guild (e.g. herbivore, omnivore, detritivore, pollenivore) explained 23% of
733 variation in the gut flora. The comparatively small amount of variation which diet explains in the
734 butterfly gut flora could be due in part to the broadly similar nature of butterfly foods (both being
735 sugary liquids).
736
737 3. Comparison of gut flora to food microbial communities.
738 *3a. Butterfly gut communities were highly dissimilar to fruit and nectar communities, but*
739 *generally shared more OTUs with the diet of their host than the diet of the other feeding guild.*
740 Host diet may shape the gut flora by serving as a source pool of microbes. Since it is
741 likely that adult butterflies acquire a substantial fraction of their gut flora from the environment
742 after emergence from the pupa (Hammer et al 2014), the microbial composition of adult butterfly
743 foods could be especially relevant. We asked whether the gut flora of Neotropical butterflies
744 might derive from their foods, and tested whether differences between frugivore and nectivore
745 gut communities were driven, in part, by differences in the microbes present in fruits and nectars.
746 Butterfly gut communities were more dissimilar to food microbial communities than we had
747 anticipated, with median percentage OTUs shared between a gut community and the microbes
748 residing in its host's food ranging from 5 to 16% for bacteria and 0 to 14% for fungi. However,
749 gut flora were generally more similar to the community on their host's food than to microbes on
750 the food of the other feeding guild. This suggests that a limited number of food-derived bacteria
751 and fungi either pass through or colonize the gut. Indeed, comparison of frugivores caught with
752 aerial nets to those caught via baited traps indicated that the yeast *Kazachstania exigua* was
753 transferred from the fermented baits into trapped butterflies' guts.

754 The high degree of differentiation between gut and food communities was surprising
755 because the guts of many nonsocial insects do seem to be colonized by environmental bacteria
756 (Ley et al 2008b; Boissière et al 2012). Indeed, prior studies of lepidopterans have found that
757 lab-reared larvae have a depauperate gut community compared to wild individuals, suggesting
758 that the larval gut *is* populated by microbes acquired from the environment (Xiang et al 2006;
759 Pinto-Tomás and Sittenfeld 2011; Belda et al 2011). Although the adult butterfly gut probably is
760 colonized at least in part by environmental microbes—as suggested by the high percentage of
761 bacterial reads derived from OTUs also found in butterfly foods—the vast majority of ingested
762 microbial species may fail to establish residence. The fact that the majority of the OTUs in
763 butterfly foods were not detected in adult guts emphasizes the strength of the filter the gut exerts
764 on potential microbial colonists, even in non-social hosts such as butterflies. Gut chemistry likely
765 filters potential colonists: most of the microbes in the gut (excluding transients) will belong to
766 the subset that can tolerate the gut's pH and oxygenation conditions. Competition or priority
767 effects from initial gut residents (those carried over from the larval host or established from early
768 adult meals) could also prevent food-derived microbes from successful colonization. Both areas
769 are fruitful avenues for future research aimed at understanding the community assembly of
770 animal gut flora.
771
772 *3b. Known habitat requirements of dominant gut microbes suggest that the adult butterfly gut is*
773 *acidic and potentially aerobic.*
774 The environmental requirements of the gut flora can help us infer chemical conditions in
775 the adult butterfly gut, which can be difficult to measure using traditional means. Acetic acid
776 bacteria (Acetobacteraceae) are obligately aerobic (Komagata et al 2014). The presence of many

777 members of the Acetobacteraceae in the adult gut suggests that it is not entirely anoxic; oxygen
778 may reach the outer layer of the gut via diffusion from gas in the tracheoles. Acetobacteraceae
779 also prefer acidic conditions, as do the yeasts *Hanseniaspora*, *Candida* and *Kazachstania* (Rosa
780 and Peter 2006; Komagata et al 2014). We could find no information about the pH of the adult
781 butterfly's gut, but in adults of the moth *Manduca sexta* it is between 5 and 6 (T. Hammer,
782 unpublished data). The larval gut is often alkaline, ranging from close to neutral pH to over 12
783 (Gross et al 2008), but the prevalence of acetic acid bacteria and acid-tolerant yeasts in our data
784 suggests that the adult gut is in fact acidic across most, if not all, butterfly species.
785
786 4. Functional potential of the butterfly gut flora.
787 *4a. The catabolic potential of the gut flora varies with host diet.*
788 The catabolic capacity of the gut flora was related to host feeding ecology. The gut flora
789 of frugivores exhibited increased catabolism of amino acids, carboxylic acids and dicarboxylic
790 acids compared to nectivores' gut flora, and decreased catabolism of sugars and sugar alcohols.
791 These differences in function likely result from differences in the chemical composition of the
792 diet. Nectars are 90% sugar by dry weight (Luttge 1977). Fruit generally contains more nitrogen
793 than nectar: at our field site, the juice of rotting *Dipteryx oleifera*, the dominant food available to
794 frugivorous butterflies during portions of the year, contains 33 times more essential amino acids
795 and 19 times more non-essential amino acids per unit sugar than the flower nectars fed upon by
796 butterflies (Ravenscraft and Boggs 2016). Frugivorous butterflies therefore provide a
797 comparatively nitrogen-rich gut environment that appears to support microbes capable of
798 catabolizing amino acids, while nectivores provide a more sugar-rich environment that favors
799 microbes specialized on sugar catabolism.

A 96-well plate is a very different environment from a butterfly gut, therefore we cannot conclude that the functions we observed necessarily operate in vivo. However, our results indicate that host species and host diet are related not only to gut community composition, but also the functional capacities of the gut flora, and they suggest that the chemical makeup of the diet selects for microbes that are specialized to digest that diet.

5. Butterfly gut flora include potential mutualists, commensals and pathogens.

A full understanding of the microbial gut community ultimately requires specific knowledge of each of its members. From what is known about the microbial genera present in the butterfly gut, we can begin to speculate whether they may be beneficial, commensal, or detrimental to butterfly hosts.

5a. Bacteria

The genera *Enterobacter*, *Enterococcus*, and *Lactococcus* are commonly present in the digestive tracts of a wide range of animals, from mammals to insects, and likely participate in mutualistic or commensal relationships with their hosts (e.g. Robinson et al 2010; Engel and Moran 2013; Delsuc et al 2014). *Acetobacter, Asaia*, and *Pantoea* are frequent residents of insect guts (Crotti et al 2010; Engel and Moran 2013) and have been shown to directly benefit some hosts. For example, *Acetobacter pomorum* and *Asaia* promote dipteran larval development, probably through nutrient supplementation (Shin et al 2011; Mitraka et al 2013). *Pantoea agglomerans* helps the swarming locust, *Schistocerca gregaria*, synthesize aggregation pheromone (Dillon et al 2002). What these species may do for butterflies is currently unknown.

822 We also detected several genera that might be detrimental to butterflies. *Serratia* species
823 are often insect pathogens (Grimont and Grimont 2006). The genus *Bartonella* contains
824 opportunistic animal pathogens that can infect insects and mammals. Many are vectored by
825 insects— primarily flies (Minnick and Anderson 2000)— but the genus has also been detected in
826 the guts of other non-biting insects including honey bees, ants, and carrion beetles (Jeyaprakash
827 et al 2003; Stoll et al 2007; Kaltenpoth and Steiger 2014). *Spiroplasma* and *Wolbachia* are well-
828 known as reproductive parasites of insects; some kill males to promote their own spread via the
829 female line. *Spiroplasma* are known to colonize the gut lumen. *Wolbachia* are generally
830 associated with the reproductive organs, but have also been found in insect guts (Frost et al 2014;
831 Berasategui et al 2016). We dissected out the guts of the butterflies in our study, but we were not
832 always able to completely remove attached fat, Malphigian tubules, and reproductive tissue.
833 Although we are confident that most of the OTUs reported here were gut-derived, the origin of
834 the *Wolbachia* in our study is uncertain.
835 Although several of the most abundant bacterial genera we detected are often assumed to
836 be harmful to their hosts and could well be detrimental to butterflies, evidence suggests these
837 genera can sometimes be beneficial. *Spiroplasma* and *Wolbachia* can both protect their hosts
838 from pathogens or parasites (e.g. Hedges et al 2008; Jaenike et al 2010) and *Wolbachia* has been
839 shown to engage in nutritional symbioses, providing B vitamins to bedbugs (Hosokawa et al
840 2010). The high prevalence and vertical transmission of *Bartonella* in several fly species suggest
841 that this genus could also be beneficial to some insect hosts (Halos et al 2004). Overall,
842 butterflies host both potential mutualists and pathogens, but the balance seems shifted towards
843 more beneficial than detrimental bacteria.

844 The data that do exist regarding the community composition of the adult butterfly gut
845 flora are largely congruent with our findings. *Enterobacter*, *Enterococcus*, *Lactococcus,*
846 *Commensalibacter* and *Orbus* were dominant members of the adult butterfly gut flora in both our
847 data and in a study of the butterfly *Heliconius erato* in Panama (Hammer et al 2014).
848 *Commensalibacter* has also been detected in *Drosophila* guts, where it may protect the flies from
849 infection by pathogens (Roh et al 2008; Ryu et al 2008). It may play a beneficial role in
850 butterflies, as well: in a related study, we found that abundance of *Commensalibacter* was
851 positively correlated with increased adult lifespan in the temperate butterfly *Speyeria mormonia*
852 (Ravenscraft et al, in review). Little is known about the biology of *Orbus*, but the genus or its
853 close relatives have been detected in the guts of flies (Chandler et al 2011), bees (Kwong and
854 Moran 2013), and geographically disparate butterflies including *Speyeria mormonia* in Colorado
855 and *Sasakia charonda* in South Korea (Kim et al 2013). The apparently widespread associations
856 between butterflies and bacteria in the genera *Commensalibacter* and *Orbus* suggest that these
857 microbes may have a generalized relationship with lepidopterans or, more broadly, with sugar-
858 feeding insects.
859
860 *5b. Fungi*
861 Most of the fungal inhabitants of butterfly guts were saccharomycete yeasts. Members of
862 the genus *Hanseniaspora* are early colonizers of decaying fruits and are also commonly isolated
863 from drosophilid flies. They are dispersed by the flies, which also use them as a food source
864 (Kurtzman et al 2011). The genera *Candida* and *Pichia* are globally widespread and present in
865 many different environments. Some are pathogenic, but many are animal commensals or
866 mutualists. *Candida* and *Pichia* species have been detected in the guts of both vertebrates and

867 invertebrates including humans, gorillas, beetles, and larval lepidopterans (Rosa et al 1992; Suh
868 et al 2005; Marchesi 2010; Dematheis et al 2012; Hamad et al 2014). All reported isolates of
869 *Candida quercitrusa*, the species most closely related to the OTUs in our dataset, derive from
870 flowers or insects (Kurtzman et al 2011). Not much is known about the ecology of *Kazachstania*
871 *exigua*, but most strains have been isolated from human food products (Kurtzman et al 2011).
872 While *Kazachstania* was present at artificially high abundance in our raw data due to its
873 dominance in the trap baits, it was also detected on wild fruits and in the guts of butterflies that
874 had not fed on trap baits, suggesting that it does naturally occur in butterfly guts at low
875 abundance.

876 Many of the abundant fungal OTUs outside of the Saccharomycetales were too divergent
877 from known species' sequences to identify them with fine taxonomic resolution. One OTU was a
878 member of the class Letiomycetes, a class of plant pathogens. Another belonged to the order
879 Pleosporales, which predominantly consists of plant-decaying fungi, though some species
880 associate with live plants (Zhang et al 2009). These fungi may have arrived in the adult gut via
881 spores in fruits and nectars, or they could have colonized the larval gut via ingestion of leaf
882 tissue and then persisted through pupation. In summary, the fungal flora of the butterfly gut
883 included many sugar-loving yeasts and plant associates, some of which may pass through insect
884 guts as accidental vagrants or hitchhikers, as well as potential insect gut-associated commensals
885 or mutualists.

886

887 **Conclusions**

888 Gut microbiota are a ubiquitous feature of animal life, yet we still understand little about
889 the structure and function of these symbiotic communities. Here we developed adult Neotropical

890 butterflies as a study system to partition variation in gut community membership at the levels of
891 the host individual, species, and feeding guild; to compare the microbial species in the gut to
892 those in host foods; and to investigate the catabolic potential of the gut flora. All samples were
893 collected from a single site, thus eliminating the potential for large-scale geographic variation
894 that is often present in, and can potentially confound, studies of similar scale.

895 We found that the majority of variation in the butterfly gut flora is expressed at the level
896 of the individual host, followed by host species, with the least amount of variation explained by
897 dietary guild. The consistency of these patterns across both bacterial and fungal communities
898 lends support to the observed pattern. The pattern itself demonstrates the need for future work to
899 expand from the current focus on broad diet categories and host taxonomy to additional, species-
900 level or individual-level traits— for example, gut chemistry— in order to better understand the
901 causes and consequences of the gut community structure in nature.

902 Butterfly gut flora had little resemblance to food microbial communities, suggesting that
903 conditions in the gut exert a strong ecological filter on colonists. The initial colonists of the
904 butterfly gut may be derived largely from butterfly foods, but once established, the butterfly gut
905 community appears to be resistant to perturbation via ingestion of microbes. However, foods do
906 influence the gut flora via their chemical composition: culture-based assays suggested that the
907 chemical composition of the diet selects for gut microbes that are good at digesting common
908 chemical components of that diet.

909 Butterfly guts host a rich community of bacteria, as well as yeasts, which are woefully
910 understudied in contemporary analyses of the gut flora. By characterizing variation of the gut
911 microbiota within and among butterfly species we have laid the foundation for a mechanistic
912 understanding of how this hidden symbiosis affects and is affected by its host.

913

914 **Acknowledgments**

915 Sincere thanks to Luke Frishkoff, Meredith Blackwell, and Tadashi Fukami's 2013 Ecological
916 Statistics class for insightful discussion and feedback. Jon Sanders provided the 16s rRNA
917 standard for our qPCR standard curve. Carlos de la Rosa, Ronald Vargas, Bernal Matarrita
918 Carranza, Danilo Brenes Madrigal and the staffs of OTS and La Selva Biological Station offered
919 invaluable logistical assistance and support in the field. This work was supported by an NSF
920 Graduate Research Fellowship, a Stanford Center for Computational, Evolutionary and Human
921 Genomics Graduate Fellowship, and an NIH PERT postdoctoral fellowship to AR, funds from
922 Stanford University, and grants from Stanford University's Biology Department, Center for
923 Latin American Studies, and Biosciences Office of Graduate Education.

924

925 **Data Accessibility**

926 Raw Illumina sequences will be made available in the NCBI Short Read Archive. All other data
927 (e.g. the OTU table, bacterial abundances calculated from qPCR, butterfly fecundity, egg
928 chemical composition, etc) will be uploaded to Dryad.

929

930 **Literature Cited**

931 Adler P, Pearson D (1982) Why do male butterflies visit mud puddles? Can J Zool 60:322–325.

932 Anderson MJ (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecol 26:32–46.
933 doi: 10.1111/j.1442-9993.2001.01070.pp.x

934 Anderson MJ, Walsh DCI (2013) PERMANOVA, ANOSIM, and the Mantel test in the face of heterogeneous
935 dispersions: What null hypothesis are you testing? Ecol Monogr 83:557–574. doi: 10.1890/12-2010.1

936 Beasley DE, Koltz AM, Lambert JE, et al (2015) The evolution of stomach acidity and its relevance to the human
937 microbiome. PLoS One 10:1–12. doi: 10.1371/journal.pone.0134116

938 Belda E, Pedrola L, Peretó J (2011) Microbial diversity in the midguts of field and lab-reared populations of the
939 European Corn Borer *Ostrinia nubilalis*. PLoS One 6:e21751. doi: 10.1371/journal.pone.0021751

940 Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: A practical and powerful approach to
941 multiple testing. J R Stat Soc Ser B 57:289–300. doi: 10.2307/2346101

942 Berasategui A, Axelsson K, Norlander G, et al (2016) The gut microbiota of the pine weevil (*Hylobius abietis*,
943 Coleoptera: Curculionidae) is similar across Europe and resembles that of other confer-feeding beetles. Mol
944 Ecol. doi: 10.1111/mec.13702

945 Bergmann GT, Bates ST, Eilers KG, et al (2011) The under-recognized dominance of Verrucomicrobia in soil
946 bacterial communities. Soil Biol Biochem 43:1450–1455. doi: 10.1016/j.soilbio.2011.03.012

947 Blackwell M, Jones K (1997) Taxonomic diversity and interactions of insect-associated ascomycetes. Biodivers
948 Conserv 6:689–699. doi: 10.1023/A:1018366203181

949 Boissière A, Tchioffo MT, Bachar D, et al (2012) Midgut microbiota of the malaria mosquito vector *Anopheles*
950 *gambiae* and interactions with *Plasmodium falciparum* infection. PLoS Pathog 8:e1002742. doi:
951 10.1371/journal.ppat.1002742

952 Brinkmann N, Martens R, Tebbe CC (2008) Origin and diversity of metabolically active gut bacteria from
953 laboratory-bred larvae of *Manduca sexta* (Sphingidae, Lepidoptera, Insecta). Appl Environ Microbiol
954 74:7189–7196. doi: 10.1128/AEM.01464-08

955 Broderick NA, Raffa KF, Goodman RM, Handelsman J (2004) Census of the bacterial community of the gypsy
956 moth larval midgut by using culturing and culture-independent methods. Appl Environ Microbiol 70:293–300.
957 doi: 10.1128/AEM.70.1.293

958 Caporaso JG, Lauber CL, Walters WA, et al (2012) Ultra-high-throughput microbial community analysis on the
959 Illumina HiSeq and MiSeq platforms. ISME J 6:1621–1624. doi: 10.1038/ismej.2012.8
960 Chandler JA, Lang J, Bhatnagar S, et al (2011) Bacterial communities of diverse *Drosophila* species: Ecological
961 context of a host-microbe model system. PLoS Genet 7:e1002272. doi: 10.1371/journal.pgen.1002272
962 Chouaia B, Rossi P, Epis S, et al (2012) Delayed larval development in *Anopheles* mosquitoes deprived of *Asaia*
963 bacterial symbionts. BMC Microbiol 12:S2. doi: 10.1186/1471-2180-12-S1-S2
964 Colman DR, Toolson EC, Takacs-Vesbach CD (2012) Do diet and taxonomy influence insect gut bacterial
965 communities? Mol Ecol 21:5124–5137. doi: 10.1111/j.1365-294X.2012.05752.x
966 Crotti E, Rizzi A, Chouaia B, et al (2010) Acetic acid bacteria, newly emerging symbionts of insects. Appl Environ
967 Microbiol 76:6963–6970. doi: 10.1128/AEM.01336-10
968 Daily GC, Ehrlich PR (1995) Preservation of biodiversity in small rainforest patches: rapid evaluations using
969 butterfly trapping. Biodivers Conserv 4:35–55. doi: 10.1007/BF00115313
970 Delsuc F, Metcalf JL, Wegener Parfrey L, et al (2014) Convergence of gut microbiomes in myrmecophagous
971 mammals. Mol Ecol 23:1301–1317. doi: 10.1111/mec.12501
972 Dematheis F, Kurtz B, Vidal S, Smalla K (2012) Microbial communities associated with the larval gut and eggs of
973 the western corn rootworm. PLoS One 7:e44685. doi: 10.1371/journal.pone.0044685
974 Deshpande V, Wang Q, Greenfield P, et al (2016) Fungal identification using a Bayesian classifier and the Warcup
975 training set of internal transcribed spacer sequences. Mycologia 108:1–5. doi: 10.3852/14-293
976 Dillon RJ, Dillon VM (2004) The gut bacteria of insects: nonpathogenic interactions. Annu Rev Entomol 49:71–92.
977 doi: 10.1146/annurev.ento.49.061802.123416
978 Dillon RJ, Vennard CT, Charnley AK (2002) A note: Gut bacteria produce components of a locust cohesion
979 pheromone. J Appl Microbiol 92:759–763. doi: 10.1046/j.1365-2672.2002.01581.x
980 Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nat Methods 10:996–
981 8. doi: 10.1038/nmeth.2604
982 Engel P, Moran NA (2013) The gut microbiota of insects - diversity in structure and function. FEMS Microbiol Rev
983 37:699–735. doi: 10.1111/1574-6976.12025
984 Freitak D, Schmidtberg H, Dickel F, et al (2014) The maternal transfer of bacteria can mediate trans-generational
985 immune priming in insects. Virulence 5:547–554. doi: 10.4161/viru.28367

Frost CL, Pollock SW, Smith JE, Hughes WOH (2014) *Wolbachia* in the flesh: Symbiont intensities in germ-line and somatic tissues challenge the conventional view of *Wolbachia* transmission routes. PLoS One 9:e95122. doi: 10.1371/journal.pone.0095122

Gardes M, Bruns TD (1993) ITS primers with enhanced specificity for basidiomycetes- application to the identification of mycorrhizae and rusts. Mol Ecol 2:113–118. doi: Doi 10.1111/J.1365-294x.1993.Tb00005.X

Garland JL, Mills AL, Young JS (2001) Relative effectiveness of kinetic analysis vs single point readings for classifying environmental samples based on community-level physiological profiles (CLPP). Soil Biol Biochem 33:1059–66.

Gilbert JA, Jansson JK, Knight R (2014) The Earth Microbiome project: successes and aspirations. BMC Biol 12:69. doi: 10.1186/s12915-014-0069-1

Gordon GL, Phillips MW (1998) The role of anaerobic gut fungi in ruminants. Nutr Res Rev 11:133–68. doi: 10.1079/NRR19980009

Green GL, Brostoff J, Hudspith B, et al (2006) Molecular characterization of the bacteria adherent to human colorectal mucosa. J Appl Microbiol 100:460–469. doi: 10.1111/j.1365-2672.2005.02783.x

Grimont F, Grimont PAD (2006) The Genus *Serratia*. In: Balows A, Trüper H, Dworkin M, et al (eds) Prokaryotes. Springer-Verlag, New York, pp 219–244

Gross EM, Brune A, Walenciak O (2008) Gut pH, redox conditions and oxygen levels in an aquatic caterpillar: Potential effects on the fate of ingested tannins. J Insect Physiol 54:462–471. doi: 10.1016/j.jinsphys.2007.11.005

Hakim RS, Baldwin K, Smagghe G (2010) Regulation of midgut growth, development, and metamorphosis. Annu Rev Entomol 55:593–608. doi: 10.1146/annurev-ento-112408-085450

Halos L, Jamal T, Maillard R, et al (2004) Role of Hippoboscidae flies as potential vectors of *Bartonella* spp. infecting wild and domestic ruminants. Appl Environ Microbiol 70:6302–6305. doi: 10.1128/AEM.70.10.6302-6305.2004

Hamad I, Keita MB, Peeters M, et al (2014) Pathogenic eukaryotes in gut microbiota of western lowland gorillas as revealed by molecular survey. Sci Rep 4:1–9. doi: 10.1038/srep06417

Hammer TJ, Dickerson JC, Fierer N (2015) Evidence-based recommendations on storing and handling specimens for analyses of insect microbiota. PeerJ e1190. doi: 10.7717/peerj.1190

1014 Hammer TJ, McMillan WO, Fierer N (2014) Metamorphosis of a butterfly-associated bacterial community. PLoS
1015 One 9:e86995. doi: 10.1371/journal.pone.0086995
1016 Hedges LM, Brownlie JC, O'Neill SL, Johnson KN (2008) *Wolbachia* and virus protection in insects. Science
1017 322:702. doi: 10.1126/science.1162418
1018 Hosokawa T, Koga R, Kikuchi Y, et al (2010) *Wolbachia* as a bacteriocyte-associated nutritional mutualist. Proc
1019 Natl Acad Sci 107:769–774. doi: 10.1073/pnas.0911476107
1020 Jaenike J, Unckless R, Cockburn SN, et al (2010) Adaptation via symbiosis: recent spread of a *Drosophila* defensive
1021 symbiont. Science 329:212–215. doi: 10.1126/science.1188235
1022 Jeyaprakash A, Hoy MA, Allsopp MH (2003) Bacterial diversity in worker adults of *Apis mellifera capensis* and
1023 *Apis mellifera scutellata* (Insecta: Hymenoptera) assessed using 16S rRNA sequences. J Invertebr Pathol
1024 84:96–103. doi: 10.1016/j.jip.2003.08.007
1025 Johnson KS, Barbehenn R V. (2000) Oxygen levels in the gut lumens of herbivorous insects. J Insect Physiol
1026 46:897–903. doi: 10.1016/S0022-1910(99)00196-1
1027 Johnston PR, Rolff J (2015) Host and symbiont jointly control gut microbiota during complete metamorphosis.
1028 PLoS Pathog 11:e1005246. doi: 10.1371/journal.ppat.1005246
1029 Kaltenpoth M, Steiger S (2014) Unearthing carrion beetles' microbiome: Characterization of bacterial and fungal
1030 hindgut communities across the Silphidae. Mol Ecol 23:1251–1267. doi: 10.1111/mec.12469
1031 Kieft TL, Simmons KA (2015) Allometry of animal-microbe interactions and global census of animal-associated
1032 microbes. Proc R Soc B Biol Sci 282:20150702. doi: 10.1098/rspb.2015.0702
1033 Kim JY, Lee J, Shin N-R, et al (2013) *Orbus sasakiae* sp. nov., a bacterium isolated from the gut of the butterfly
1034 *Sasakia charonda*, and emended description of the genus *Orbus*. Int J Syst Evol Microbiol 63:1766–1770. doi:
1035 10.1099/ijs.0.041871-0
1036 Kingsley V (1972) Persistence of intestinal bacteria in the developmental stages of the Monarch butterfly (*Danaus*
1037 *plexippus*). J Invertebr Pathol 20:51–58.
1038 Kirk P, Cannon P, Minter D, Stalpers J (eds) (2008) Dictionary of the Fungi, 10th edn. CABI, Wallingford, UK
1039 Komagata K, Iino T, Yamada Y (2014) The Family Acetobacteraceae. In: Rosenberg E, DeLong EF, Lory S, et al
1040 (eds) Prokaryotes Vol. 8 Alphaproteobacteria Betaproteobacteria, 4th edn. pp 3–78
1041 Kozich JJ, Westcott SL, Baxter NT, et al (2013) Development of a dual-index sequencing strategy and curation

1042 pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. Appl Environ
1043 Microbiol 79:5112–5120. doi: 10.1128/AEM.01043-13
1044 Kurtzman CP, Fell JW, Boekhout T (eds) (2011) The Yeasts, A Taxonomic Study, Volume 2, 5th edn. Elsevier, San
1045 Diego, CA
1046 Kwong WK, Moran NA (2013) Cultivation and characterization of the gut symbionts of honey bees and bumble
1047 bees: Description of *Snodgrassella alvi* gen. nov., sp. nov., a member of the family Neisseriaceae of the
1048 betaproteobacteria, and *Gilliamella apicola* gen. nov., sp. Int J Syst Evol Microbiol 63:2008–2018. doi:
1049 10.1099/ijs.0.044875-0
1050 Lenth R V (2016) Least-Squares Means: The R Package lsmeans. J Stat Softw 69:1–33. doi: 10.18637/jss.v069.i01
1051 Ley RE, Hamady M, Lozupone C, et al (2008a) Evolution of mammals and their gut microbes. Science 320:1647–
1052 1651. doi: 10.1126/science.1155725
1053 Ley RE, Lozupone CA, Hamady M, et al (2008b) Worlds within worlds: evolution of the vertebrate gut microbiota.
1054 Nat Rev Microbiol 6:776–788. doi: 10.1038/nrmicro1978
1055 Ley RE, Peterson DA, Gordon JI (2006) Ecological and evolutionary forces shaping microbial diversity in the
1056 human intestine. Cell 124:837–848. doi: 10.1016/j.cell.2006.02.017
1057 Lundberg DS, Yourstone S, Mieczkowski P, et al (2013) Practical innovations for high-throughput amplicon
1058 sequencing. Nat Methods 10:999–1002. doi: 10.1038/nmeth.2634
1059 Luttge U (1977) Nectar composition and membrane transport of sugars and amino acids: a review on the present
1060 state of nectar research. Apidologie 8:305–319. doi: 10.1051/apido:19770402
1061 Marchesi JR (2010) Prokaryotic and eukaryotic diversity of the human gut. Adv. Appl. Microbiol., 1st edn. Elsevier
1062 Inc., pp 43–62
1063 Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal
1064 17:10. doi: 10.14806/ej.17.1.200
1065 McDade LA, Bawa KS, Hespenheide HA, Hartshorn GS (1994) La Selva: ecology and natural history of a
1066 Neotropical rain forest. University of Chicago Press, Chicago
1067 McDonald D, Price MN, Goodrich J, et al (2012) An improved Greengenes taxonomy with explicit ranks for
1068 ecological and evolutionary analyses of bacteria and archaea. ISME J 6:610–618. doi: 10.1038/ismej.2011.139
1069 McFall-Ngai M (2007) Adaptive immunity: care for the community. Nature 445:153. doi: 10.1038/445153a

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

McNulty NP, Yatsunenko T, Hsiao A, et al (2011) The impact of a consortium of fermented milk strains on the gut microbiome of gnotobiotic mice and monozygotic twins. Sci Transl Med 3:1–14. doi: 10.1126/scitranslmed.3002701

Minnick MF, Anderson BE (2000) *Bartonella* interactions with host cells. In: Oelschlaeger TA, Hacker JH (eds) Subcell. Biochem. Bact. Invasion into Eukaryot. Cells. Springer, New York, NY, pp 97–118

Mitraka E, Stathopoulos S, Siden-Kiamos I, et al (2013) *Asaia* accelerates larval development of *Anopheles gambiae*. Pathog Glob Health 107:305–11. doi: 10.1179/2047773213Y.0000000106

Muegge BD, Kuczynski J, Knights D, et al (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332:970–4. doi: 10.1126/science.1198719

Nardi JB, Mackie RI, Dawson JO (2002) Could microbial symbionts of arthropod guts contribute significantly to nitrogen fixation in terrestrial ecosystems? J Insect Physiol 48:751–763.

Oksanen J, Blanchet FG, Kindt R, et al (2015) vegan: Community Ecology Package. R package version 2.3-0.

Peay KG, Bruns TD, Kennedy PG, et al (2007) A strong species-area relationship for eukaryotic soil microbes: Island size matters for ectomycorrhizal fungi. Ecol Lett 10:470–480. doi: 10.1111/j.1461-0248.2007.01035.x

Peay KG, Kennedy PG, Talbot JM (2016) Dimensions of biodiversity in the Earth mycobiome. Nat Rev Microbiol 14:434–447. doi: 10.1038/nrmicro.2016.59

Pinto-Tomás A, Sittenfeld A (2011) Comparison of midgut bacterial diversity in tropical caterpillars (Lepidoptera: Saturniidae) fed on different diets. Environ Entomol 40:1111–1122.

R Core Team (2015) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ravenscraft A, Boggs CL (2016) Nutrient acquisition across a dietary shift: fruit feeding butterflies crave amino acids, nectivores seek salt. Oecologia 181:1–12. doi: 10.1007/s00442-015-3403-6

Rawls JF, Mahowald MA, Ley RE, Gordon JI (2006) Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. Cell 127:423–433. doi: 10.1016/j.cell.2006.08.043

Robinson CJ, Bohannan BJM, Young VB (2010a) From structure to function: the ecology of host-associated microbial communities. Microbiol Mol Biol Rev 74:453–76. doi: 10.1128/MMBR.00014-10

1098 Robinson CJ, Schloss P, Ramos Y, et al (2010b) Robustness of the bacterial community in the cabbage white
1099 butterfly larval midgut. Microb Ecol 59:199–211. doi: 10.1007/s00248-009-9595-8
1100 Roh SW, Nam Y-D, Chang H-W, et al (2008) Phylogenetic characterization of two novel commensal bacteria
1101 involved with innate immune homeostasis in *Drosophila melanogaster*. Appl Environ Microbiol 74:6171–
1102 6177. doi: 10.1128/AEM.00301-08
1103 Rosa CA, Hagler AN, Mendonca-Hagler LCS, et al (1992) *Clavispora opuntiae* and other yeasts associated with the
1104 moth *Sigelgaita* sp. in the cactus *Pilosocereus arrabidae* of Rio de Janeiro, Brazil. Antonie Van Leeuwenhoek
1105 62:267–272. doi: 10.1007/BF00572594
1106 Rosa CA, Peter G (eds) (2006) Biodiversity and Ecophysiology of Yeasts. Springer-Verlag
1107 Russell V, Dunn PE (1996) Antibacterial proteins in the midgut of *Manduca sexta* during metamorphosis. J Insect
1108 Physiol 42:65–71. doi: 10.1016/0022-1910(95)00083-6
1109 Ryu JH, Kim SH, Lee HY, et al (2008) Innate immune homeostasis by the homeobox gene caudal and commensal-
1110 gut mutualism in *Drosophila*. Science 319:777–782. doi: 10.1126/science.1149357
1111 Salter SJ, Cox MJ, Turek EM, et al (2014) Reagent and laboratory contamination can critically impact sequence-
1112 based microbiome analyses. BMC Biol 12:87. doi: 10.1186/s12915-014-0087-z
1113 Salzman NH, Hung K, Haribhai D, et al (2009) Enteric defensins are essential regulators of intestinal microbial
1114 ecology. Nat Immunol 11:76–82. doi: 10.1038/ni.1825
1115 Sculley CE, Boggs CL (1996) Mating systems and sexual division of foraging effort affect puddling behaviour by
1116 butterflies. Ecol Entomol 21:193–197.
1117 Shin SC, Kim S-H, You H, et al (2011) *Drosophila* microbiome modulates host developmental and metabolic
1118 homeostasis via insulin signaling. Science 334:670–4. doi: 10.1126/science.1212782
1119 Smith DP, Peay KG (2014) Sequence depth, not PCR replication, improves ecological inference from next
1120 generation DNA sequencing. PLoS One 9:e90234. doi: 10.1371/journal.pone.0090234
1121 Solomon K V., Haitjema CH, Henske JK, et al (2016) Early-branching gut fungi possess a large, comprehensive
1122 array of biomass-degrading enzymes. Science 351:1192–1195. doi: 10.1126/science.aad1431
1123 Starmer WT, Lachance MA (2011) Yeast Ecology. In: Kurtzman CP, Fell JW, Boekhout T (eds) The Yeasts, 5th
1124 edn. Elsevier, San Diego, CA, pp 65–83
1125 Steinhaus E (1941) A study of the bacteria associated with thirty species of insects. J Bacteriol 42:757–90.

1126 Stoll S, Gadau J, Gross R, Feldhaar H (2007) Bacterial microbiota associated with ants of the genus *Tetraponera*.
1127 Biol J Linn Soc 90:399–412. doi: 10.1111/j.1095-8312.2006.00730.x
1128 Suh S-O, McHugh J V, Pollock DD, Blackwell M (2005) The beetle gut: a hyperdiverse source of novel yeasts.
1129 Mycol Res 109:261–265. doi: 10.1017/S0953756205002388
1130 Sullam KE, Essinger SD, Lozupone CA, et al (2012) Environmental and ecological factors that shape the gut
1131 bacterial communities of fish: a meta-analysis. Mol Ecol 21:3363–3378. doi: 10.1111/j.1365-
1132 294X.2012.05552.x
1133 Tap J, Mondot S, Levenez F, et al (2009) Towards the human intestinal microbiota phylogenetic core. Environ
1134 Microbiol 11:2574–2584. doi: 10.1111/j.1462-2920.2009.01982.x
1135 Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naïve Bayesian classifier for rapid assignment of rRNA sequences
1136 into the new bacterial taxonomy. Appl Environ Microbiol 73:5261–5267. doi: 10.1128/AEM.00062-07
1137 Warnecke F, Luginbühl P, Ivanova N, et al (2007) Metagenomic and functional analysis of hindgut microbiota of a
1138 wood-feeding higher termite. Nature 450:560–565. doi: 10.1038/nature06269
1139 White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for
1140 phylogenetics. PCR Protoc A Guid to Methods Appl. doi: http://dx.doi.org/10.1016/B978-0-12-372180-
1141 8.50042-1
1142 Xiang H, Wei G, Jia S (2006) Microbial communities in the larval midgut of laboratory and field populations of
1143 cotton bollworm (*Helicoverpa armigera*). Can J Microbiol 52:1085–1092. doi: 10.1139/W06-064
1144 Zhang Y, Schoch CL, Fournier J, et al (2009) Multi-locus phylogeny of Pleosporales: A taxonomic, ecological and
1145 evolutionary re-evaluation. Stud Mycol 64:85–102. doi: 10.3114/sim.2009.64.04
1146 Zuur AF, Ieno EN, Walker NJ, et al (2009) Mixed Effects Models and Extensions in Ecology with R. Springer, New
1147 York
1148

Table 1. Sample sizes

Number of butterfly or food samples ("individuals") and species remaining in the sequencing datasets after rarefying.

	Bacterial dataset		Fungal dataset	
	Individuals	**Species**	**Individuals**	**Species**
Frugivores	148	24	92	15
Nectivores	142	28	69	19
Fruits	12	1	13	1
Nectars	46	8	39	8
Trap baits	25	n/a	14	n/a

Table 2. Taxonomic identities of the 20 most abundant bacterial OTUs

The 20 most abundant bacterial OTUs. "% Total" is the percentage of sequences (pooled across all butterflies) that were assigned to the OTU. "% Butterflies", "% Frugivores", and "% Nectivores" are the percentages of all butterflies, frugivores, and nectivores in which the OTU was detected. Taxonomy was assigned by the RDP classifier using the Greengenes training set, except as otherwise noted. Confidences measure the degree of certainty of the taxonomic assignment.

OTU ID	% Reads	% Butterflies	% Frugivores	% Nectivores	Frug:Nect	Phylum	Class	Order	Family	Genus	Species	Confidence	Notes
OTU_1	9.8	73	84	63	1.3	Proteobacteria	Gammaproteobacteria	Orbales	Orbaceae	*Orbus*		100	a
OTU_4	8.7	82	81	84	1.0	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Klebsiella/Enterobacter*		100	b
OTU_3	8.3	56	42	70	0.6	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	*Swaminathania/Asaia*		97	c
OTU_6	6.0	83	86	80	1.1	Firmicutes	Bacilli	Lactobacillales	Enterococcaceae	*Vagococcus/Enterococcus*		62	d
OTU_2	4.8	44	41	48	0.9	Bacteroidetes	Flavobacteriia	Flavobacteriales	[Weeksellaceae]			96	
OTU_19	3.2	46	34	59	0.6	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Serratia*	*marcescens*	94	
OTU_5	2.7	28	27	29	0.9	Tenericutes	Mollicutes	Entomoplasmatales		*Spiroplasma/Entomoplasma*		100	e
OTU_10	2.6	28	36	20	1.8	Bacteroidetes	Bacteroidia	Bacteroidales	Porphyromonadaceae			68	
OTU_9	2.5	30	27	34	0.8	Proteobacteria	Alphaproteobacteria	Rhizobiales	Bartonellaceae	*Bartonella*		92	f
OTU_12	2.5	44	40	49	0.8	Firmicutes	Bacilli	Lactobacillales	Streptococcaceae	*Lactococcus*		100	
OTU_18	2.3	46	72	19	3.8	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	*Acetobacter*		100	g
OTU_11	2.2	34	46	22	2.1	Proteobacteria	Gammaproteobacteria	Orbales	Orbaceae	*Gilliamella/Orbus*		91	h
OTU_17	2.2	27	36	18	2.0	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	*Commensalibacter*	*intestini*	95	
OTU_14	1.7	19	25	13	1.9	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Providencia/Serratia*		73	i
OTU_13	1.7	18	22	13	1.7	Proteobacteria	Alphaproteobacteria	Rickettsiales	Rickettsiaceae	*Wolbachia*		100	
OTU_15	1.6	26	31	20	1.5	Proteobacteria	Gammaproteobacteria	Pseudomonadales	Pseudomonadaceae			73	
OTU_23	1.4	23	26	19	1.4	Proteobacteria	Gammaproteobacteria	Pasteurellales				100	
OTU_569	1.4	27	18	37	0.5	Proteobacteria	Gammaproteobacteria	Enterobacteriales	Enterobacteriaceae	*Erwinia/Pantoea*		56	j
OTU_22	1.4	25	23	27	0.9	Proteobacteria	Alphaproteobacteria	Rhodospirillales	Acetobacteraceae	*Commensalibacter*	*intestini*	93	
OTU_1106	1.2	19	16	21	0.8	Bacteroidetes	Flavobacteriia	Flavobacteriales	[Weeksellaceae]			68	

Notes

a. RDP assigned to the order Pasteurellales with no finer resolution. Closest NCBI hit was Orbales:Orbaceae:*Orbus*.

b. In a blast search against NCBI,matched equally closely to the genera *Klebsiella* and *Enterobacter.*

c. RDP assigned to genus *Swaminanthia;* closest NCBI hit was *Asaia.*

d. Closest NCBI hit was *Enterococcus*

e. In the blast search against NCBI, matched equally closely to the genera *Spiroplasma* and *Entomoplasma*

f. RDP assigned to family with no finer resolution. Closest NCBI match was the genus *Bartonella*.

g. Closest NCBI species matches were *indonesiensis* or *okinawensis.*

h. RDP assigned to the order Pasteurellales with no finer resolution. Closest NCBI hit was Orbales:Orbaceae:*Gilliamella/Orbus*.

i. RDP assigned to family with no finer resolution. Closest NCBI matches were the genera *Providencia* and *Serratia*.

j. Closest NCBI match was the genus *Pantoea.*

Table 3. Taxonomic identities of the 20 most abundant fungal OTUs

The 20 abundant fungal OTUs detected in at least 10% of the butterfly samples. "% Total" is the percentage of sequences (pooled across all butterflies) that were assigned to the OTU. "% Butterflies", "% Frugivores", and "% Nectivores" are the percentages of all butterflies, frugivores, and nectivores in which the OTU was detected. Taxonomy was assigned by the RDP classifier using the Warcup training set. Where a rank is *incertae sedis*, the closest available rank is written in brackets. Confidences measure the degree of certainty of the taxonomic assignment and are cumulative from higher to lower taxonomic ranks. Assignments with particularly low confidence are greyed out. (One additional fungal OTU, a *Torulaspora* species, was observed at high abundance in the pooled data but was detected in less than 10% of frugivores and less than 10% of nectivores, so is omitted here.)

OTU ID	% Reads	% Butterflies	% Frugivores	% Nectivores	Frug:Nect	Phylum	Pconf	Class	Cconf	Order	Oconf	Family	Fconf	Genus	Gconf	Species	Sconf
OTU_7	5.0	22	7	37	0.2	Zygomycota	20	[Mucoromycotina]	15	Mucorales	15	Mucoraceae	7	Rhizopus	6	caespitosus	6
OTU_10	3.7	13	9	16	0.6	Basidiomycota	68	Agaricostilbomycetes	29	Agaricostilbales	29	Agaricostilbaceae	27	Bensingtonia	27	yuccicola	9
OTU_20	3.2	9	12	3	4.0	Ascomycota	68	Saccharomycetes	13	Saccharomycetales	13	Saccharomycodaceae	3	Hanseniaspora	3	uvarum	3
OTU_17	2.9	45	56	25	2.2	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycodaceae	100	Hanseniaspora	100	uvarum	94
OTU_2469	2.6	42	57	16	3.6	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycodaceae	100	Hanseniaspora	100	opuntiae	62
OTU_25	2.4	35	51	11	4.6	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Pichiaceae	99	Pichia	99	fermentans	70
OTU_2439	2.2	21	28	9	3.1	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycetaceae	100	Kazachstania	100	exigua	100
OTU_11	1.9	29	34	18	1.9	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycetales_Incertae sedis	97	Candida	97	quercitrusa	96
OTU_2306	1.9	40	53	18	2.9	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycodaceae	100	Hanseniaspora	100	guilliermondii	42
OTU_2503	1.5	15	6	24	0.2	Ascomycota	35	Saccharomycetes	7	Saccharomycetales	7	Saccharomycodaceae	5	Hanseniaspora	5	uvarum	5
OTU_78	1.5	21	16	25	0.6	Ascomycota	97	Leotiomycetes	90	[Leotiomycetes]	32	[Leotiomycetes]	32	Leohumicola	17	levissima	14
OTU_24	1.5	13	15	9	1.7	Ascomycota	99	Saccharomycetes	98	Saccharomycetales	98	Saccharomycetaceae	98	Kazachstania	96	hellenica	96
OTU_41	1.4	8	3	13	0.2	Basidiomycota	46	Microbotryomycetes	8	Sporidiobolales	7	[Sporidiobolales]	7	Rhodotorula	1	cresolica	0
OTU_127	1.2	14	7	20	0.4	Ascomycota	94	Dothideomycetes	80	Pleosporales	75	[Pleosporales]	38	Clavariopsis	37	aquatica	37
OTU_997	1.2	13	22	0	n/a	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Saccharomycetaceae	100	Kazachstania	100	exigua	100
OTU_151	1.2	5	0	11	0	Ascomycota	43	Saccharomycetes	10	Saccharomycetales	10	Saccharomycetaceae	5	Kazachstania	2	hellenica	2
OTU_1735	1.1	9	10	7	1.4	Ascomycota	78	Saccharomycetes	40	Saccharomycetales	40	Pichiaceae	17	Pichia	10	haplophila	7
OTU_3	1.1	13	6	20	0.3	Ascomycota	22	Saccharomycetes	18	Saccharomycetales	18	[Saccharomycetales]	11	Cyberlindnera	5	fabianii	5
OTU_96	1.0	11	6	14	0.4	Ascomycota	100	Saccharomycetes	100	Saccharomycetales	100	Pichiaceae	100	Pichia	100	jadinii	100
OTU_73	1.0	16	11	20	0.5	Ascomycota	100	[Pezizomycotina]	19	[Pezizomycotina]	19	[Pezizomycotina]	19	Fontanospora	17	fusiramosa	17

Figure 1. Total bacterial load per butterfly by species.

Total bacterial 16s counts per butterfly were estimated via qPCR. Host species are arranged in ascending size (wing length) along the x-axis. Dots indicate copy number for individual frugivores (red) or nectivores (blue), and boxplots depict medians and interquartile ranges of the data. Whiskers are placed at 1.5 times the interquartile range or, if all data fall within this range, they are placed at most extreme value measured. Note that although feeding guild is displayed, it was not a significant predictor of total bacterial load.

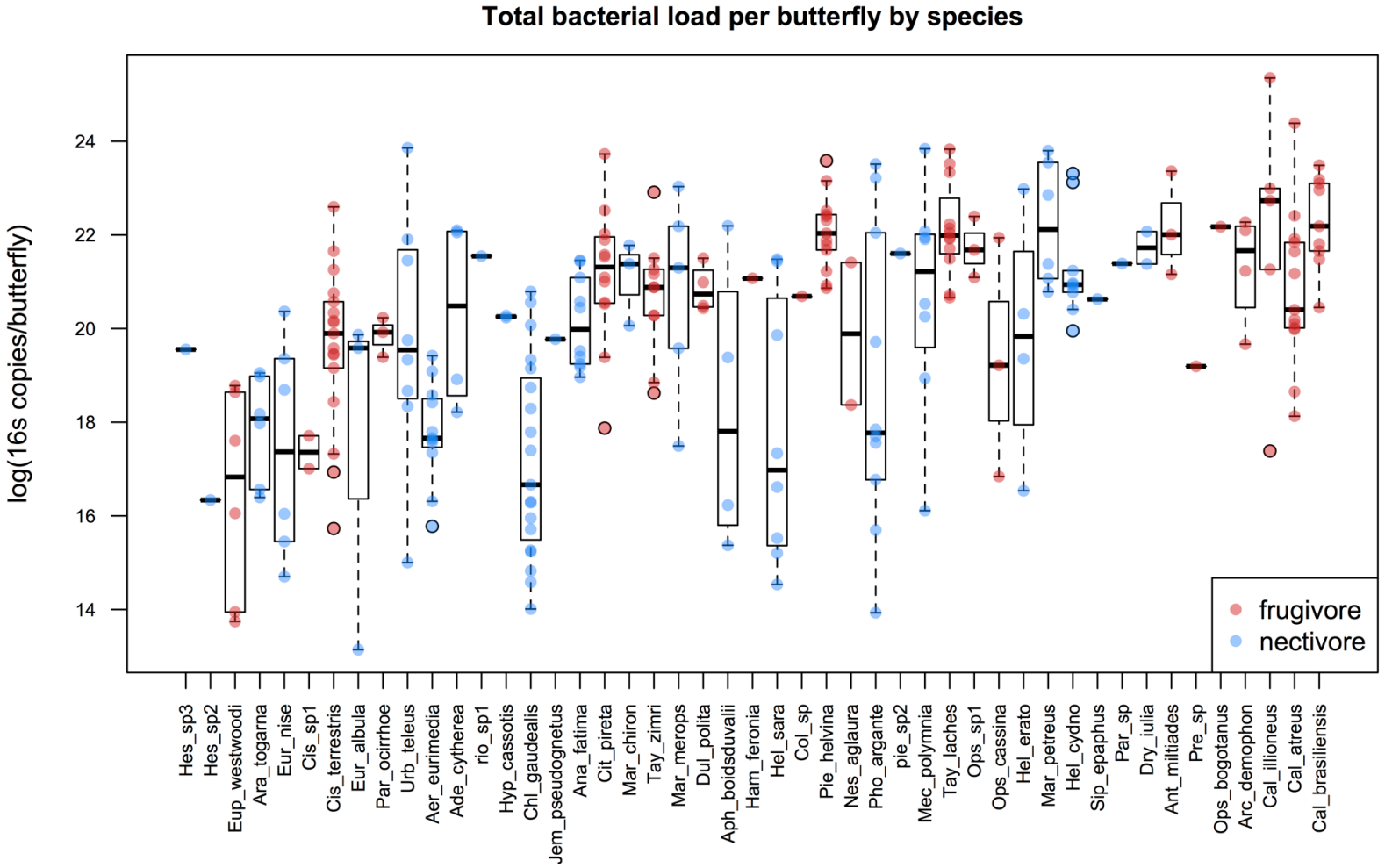

Figure 2. Microbial species richness in frugivores and nectivores.

Bacterial (a) and fungal (b) observed species richness per butterfly did not differ between frugivores and nectivores. P-values are the result of model comparisons and control for host species (see text). Boxplot features are as described in Figure 1. One outlier of 162 bacterial OTUs in a nectivore is not shown in panel (a). Note that since bacteria and fungi were rarefied to different cutoffs, species richness estimates are not comparable between the two.

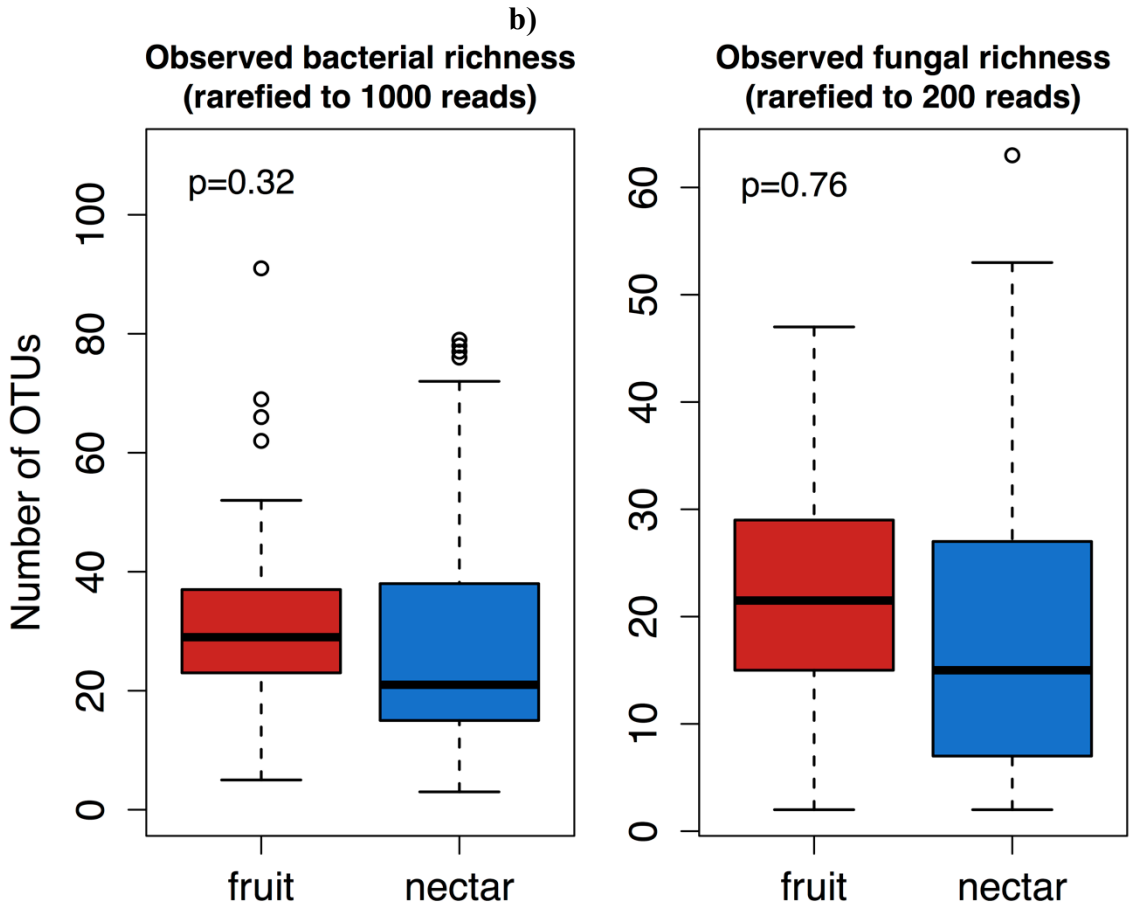

Figure 3. Ordinations of bacterial and fungal community composition.

NMDS plots of the Bray-Curtis dissimilarities between butterflies. PerMANOVA tests confirmed that the gut flora of frugivores and nectivores differed in bacterial and fungal OTU composition (p-values displayed on plots). (a) Bacterial gut flora of 148 frugivores and 142 nectivores. (b) Fungal gut flora of 92 frugivores and 69 nectivores. One outlier is not displayed in panel (b).

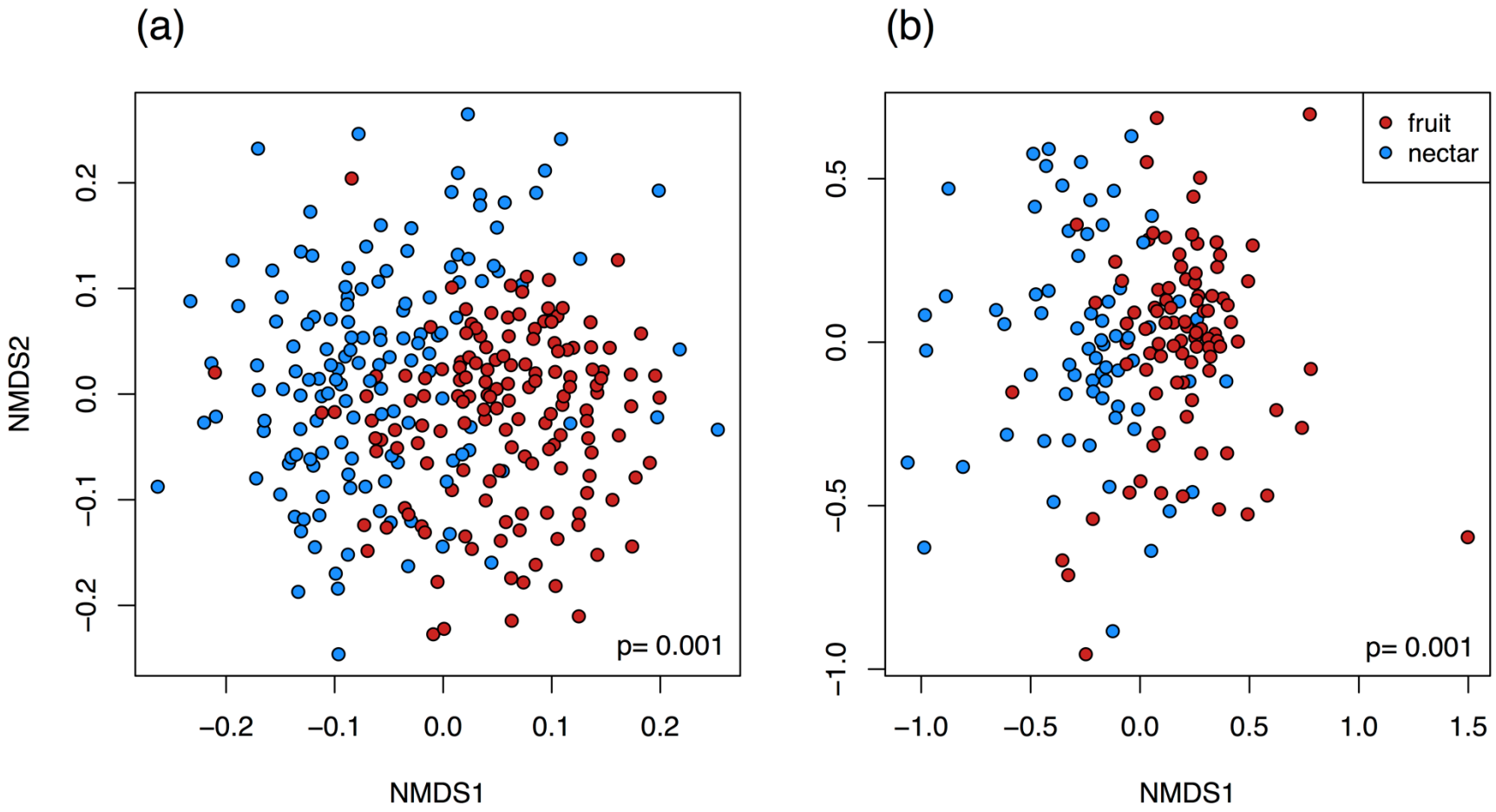

Figure 4. Effect of host diet (a) and host species (b) on bacterial community composition.

(a) Model-predicted estimates of the mean relative abundance per butterfly (counts out of 1000) of the 20 most abundant bacterial OTUs in the dataset. Relative abundance was modeled as a function of the interaction between host diet and OTU identity, with a random intercept of host species. Points indicate the model estimated mean and lines indicate one standard error. OTUs are arranged along the x-axis according to the magnitude and direction of the difference in their abundance between frugivores and nectivores. Each OTU is designated by genus or by the finest taxonomic resolution available, followed by OTU number in parentheses. The shaded gray box indicates OTUs that did not differ in abundance between the feeding guilds.
(b) Mean relative abundances of the 20 most abundant bacterial OTUs per host species. Darker shading indicates higher mean relative abundance.

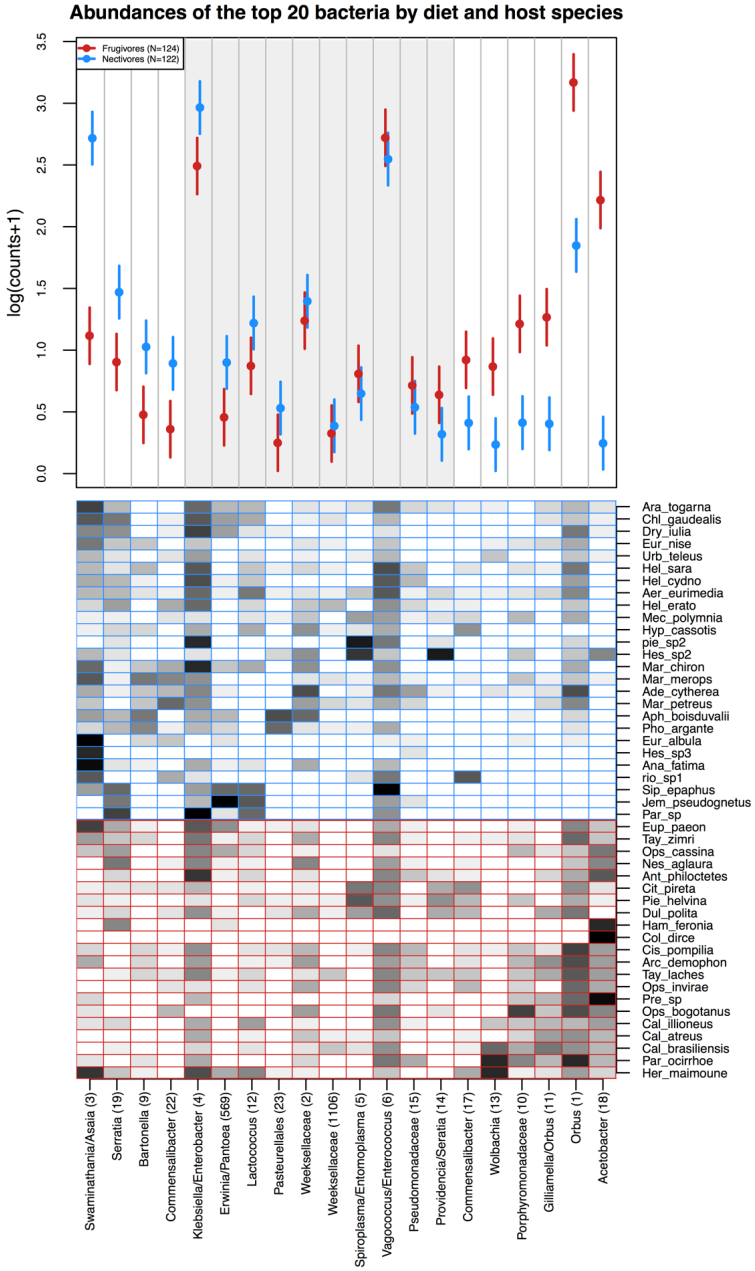

Figure 5. Effect of host diet (a) and host species (b) on fungal community composition.

(a) Model-predicted estimates of the mean relative abundance per butterfly (counts out of 1000) of the 20 most abundant fungal OTUs in the dataset. Relative abundance was modeled as a function of the interaction between host diet and OTU identity, with a random effect of host species. Points indicate the predicted mean and lines indicate one standard error. OTUs are arranged along the x-axis according to the magnitude and direction of the difference in their abundance between frugivores and nectivores. Each OTU is designated by genus or by the finest taxonomic resolution available, followed by OTU number in parentheses. The shaded gray box indicates OTUs that did not differ in abundance between the feeding guilds.
(b) Mean relative abundances of the 20 most abundant fungal OTUs per host species. Darker shading indicates higher mean relative abundance.

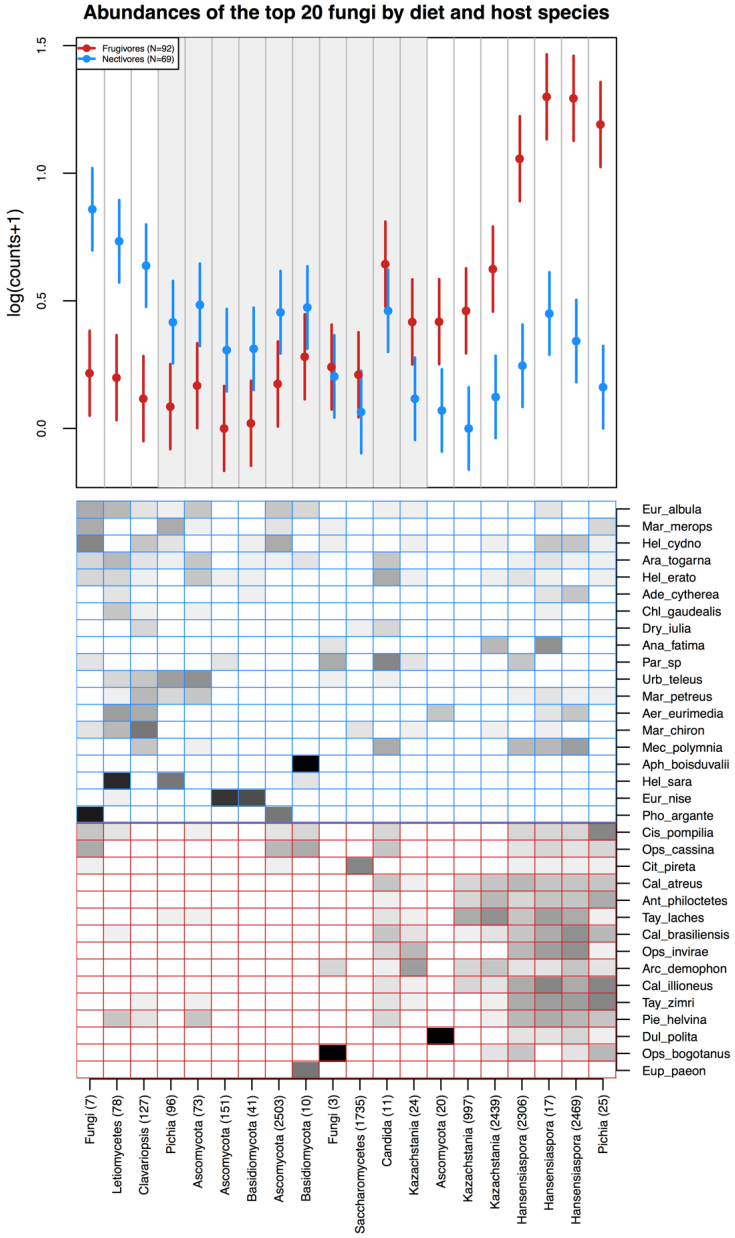

Figure 6. Dissimilarity of gut flora within versus between host species.

Every point represents the average Bray-Curtis dissimilarity between of a given individual's gut community and the gut communities of all other conspecific hosts (green) or all heterospecific hosts (orange). Boxplot features are as described in Figure 1. P-values are the results of t-tests for pairwise differences in dissimilarity.

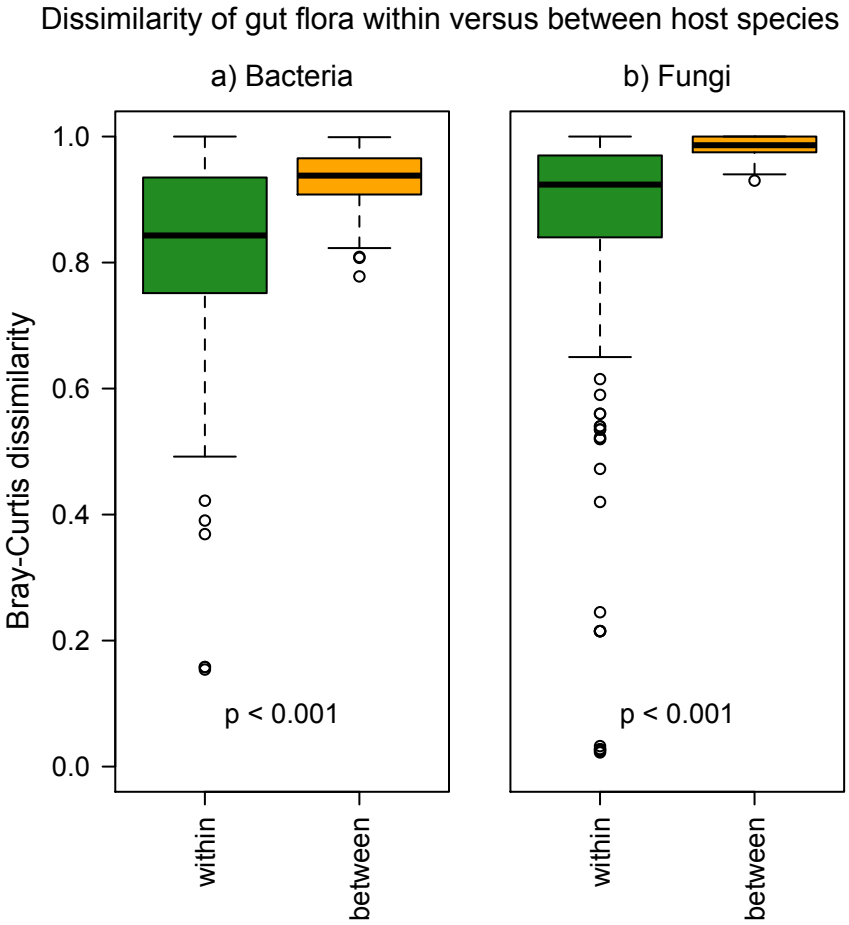

Figure 7. Comparison of frugivore and nectivore gut communities to the microbial communities of butterfly foods.

NMDS plots of the Bray-Curtis dissimilarities between butterflies and their foods. (a) Bacterial flora of 149 frugivores, 149 nectivores, 46 nectars, 25 trap baits, and 12 fruits. (b) Fungal gut flora of 94 frugivores, 70 nectivores, 39 nectars, 14 trap baits, and 13 fruits. Two samples were fungal outliers and are not displayed in panel (b). (Sample sizes differ slightly from Figure 3 because in order to include as many food samples as possible, the combined gut and food sequence data were rarefied to different depths than the gut sequence data alone.)

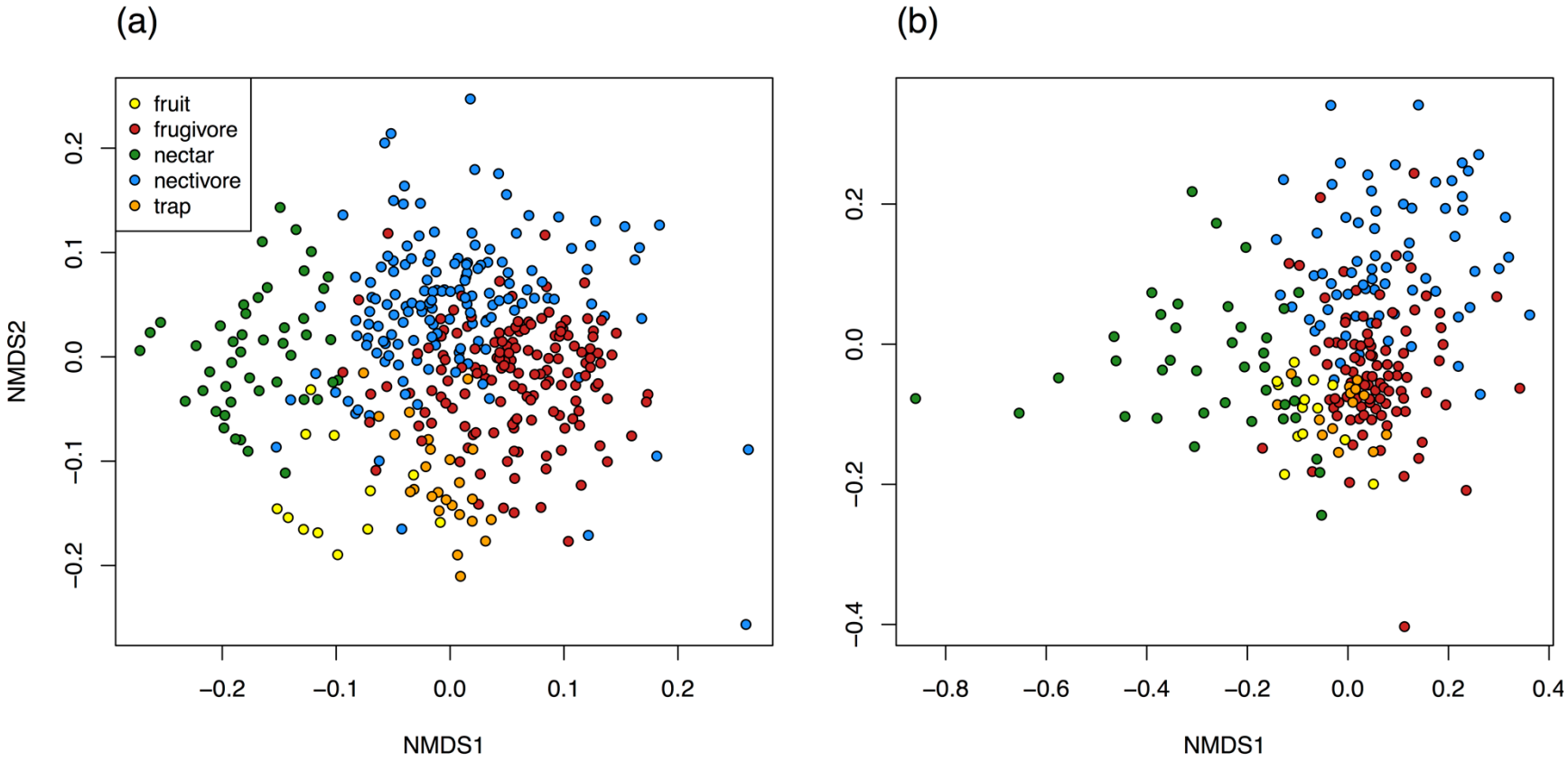

Figure 8. Dissimilarity between gut and food microbial communities**.**

Bray-Curtis dissimilarities between butterfly gut bacterial (a) and fungal (b) communities and the microbial communities in butterfly foods. Every point represents the average distance of a butterfly gut community to all food communities in a given category (fruit, trap bait, or nectar). Comparisons to frugivorous butterflies are colored red; comparisons to nectivores are colored blue. Boxplot features are as described in Figure 1. P-values are the results of t-tests for pairwise differences in dissimilarity; these have been FDR-corrected.
Note that three fungal strains of *Kazachstania exigua* were excluded from the butterfly data but not the food data, since these OTUs were introduced into frugivores' guts via the trap baits. (Described in the main text.) Frugivores would appear more similar to the trap baits if these OTUs had been included.

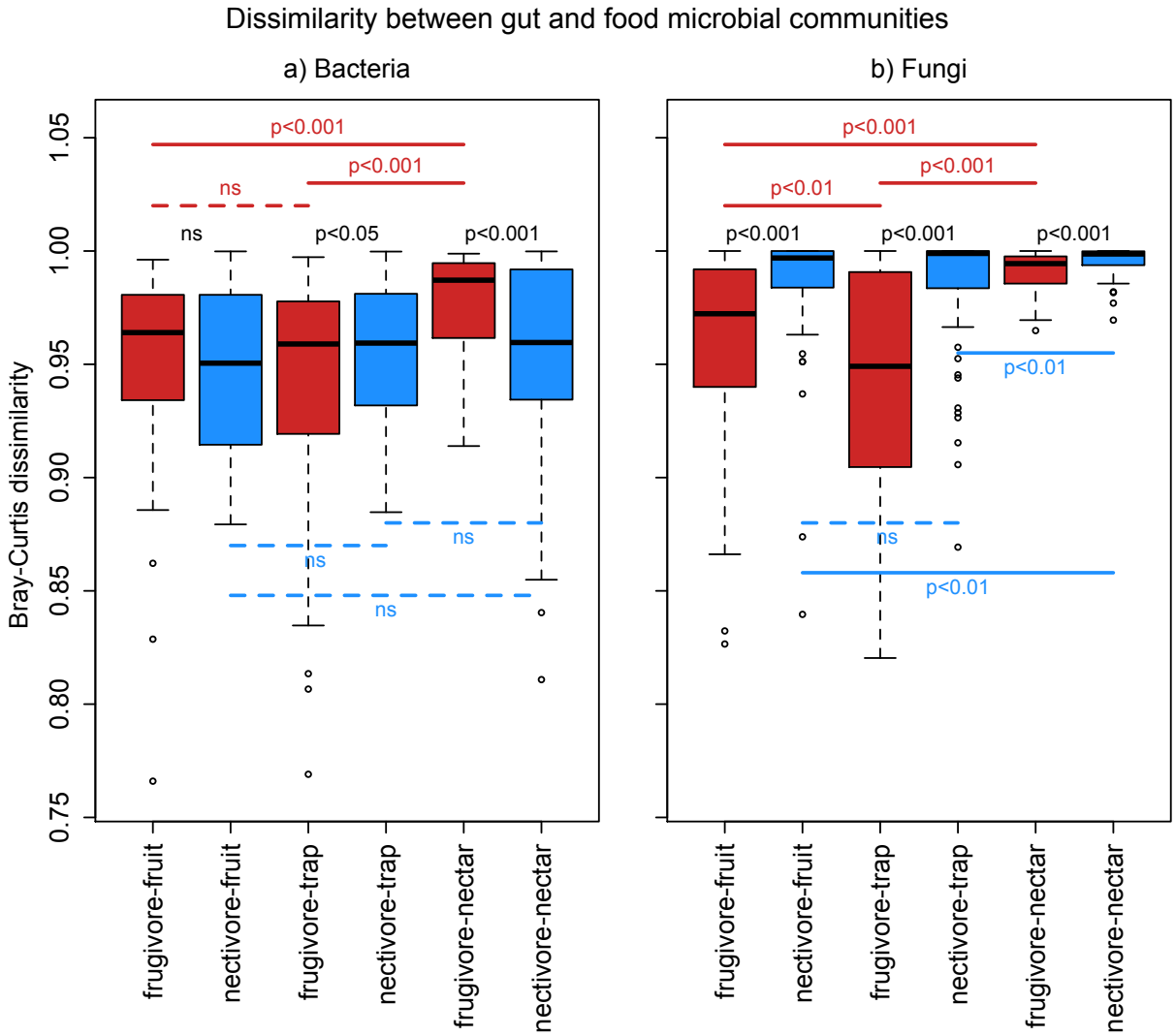

Figure 9. Differences in microbial community catabolism between frugivores and nectivores.

Points indicate the model-estimated mean difference in catabolism of a substrate (units of absorbance at 590 nm, standardized by plate and Hellinger transformed) between frugivores' and nectivores' gut microbial communities. Gut community catabolism of substrates in dark red and dark blue was significantly greater in frugivores or nectivores, respectively, after multiple test correction. Light red and light blue substrates were significantly different between guilds before, but not after multiple test correction. Thick lines and points indicate model-estimated means and standard errors for substrate classes. Red indicates classes that were catabolized more strongly by frugivore gut communities; blue indicates classes that were digested more actively by nectivore gut flora.