

1 **Caterpillars lack a resident gut microbiome**

2

3 Tobin J. Hammer^{1,2}, Daniel H. Janzen^{3,4}, Winnifred Hallwachs^{3,4}, Samuel L. Jaffe⁵, Noah
4 Fierer^{1,2}

5

6 1. Department of Ecology and Evolutionary Biology, University of Colorado Boulder, Boulder,
7 CO, USA

8 2. Cooperative Institute for Research in Environmental Sciences, University of Colorado
9 Boulder, Boulder, CO, USA

10 3. Department of Biology, University of Pennsylvania, Philadelphia, PA, USA

11 4. Área de Conservación Guanacaste, Liberia, Guanacaste, Costa Rica

12 5. The Caterpillar Lab, Keene, NH, USA

13

14 Corresponding author: Tobin J. Hammer, tobin.hammer@colorado.edu

15

16 **Keywords:** insects, herbivory, Lepidoptera, symbiosis, mutualism

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47 **Abstract**

48

49 Many animals are inhabited by microbial symbionts that influence their hosts’
50 development, physiology, ecological interactions, and evolutionary diversification. However,
51 firm evidence for the existence and functional importance of resident microbiomes in larval
52 Lepidoptera (caterpillars) is lacking, despite the fact that these insects are enormously diverse,
53 major agricultural pests, and dominant herbivores in many ecosystems. Using 16S rRNA gene
54 sequencing and quantitative PCR, we characterized the gut microbiomes of wild leaf-feeding
55 caterpillars in the United States and Costa Rica, representing 124 species from 16 families.
56 Compared with other insects and vertebrates assayed using the same methods, the microbes we
57 detected in caterpillar guts were unusually low-density and highly variable among individuals.
58 Furthermore, the abundance and composition of leaf-associated microbes were reflected in the
59 feces of caterpillars consuming the same plants. Thus, microbes ingested with food are present
60 (though possibly dead or dormant) in the caterpillar gut, but host-specific, resident symbionts are
61 largely absent. To test whether transient microbes might still contribute to feeding and
62 development, we conducted an experiment on field-collected caterpillars of the model species
63 *Manduca sexta*. Antibiotic suppression of gut bacterial activity did not significantly affect
64 caterpillar weight gain, development, or survival. The high pH, simple gut structure, and fast
65 transit times that typify caterpillar digestive physiology may prevent microbial colonization.
66 Moreover, host-encoded digestive and detoxification mechanisms likely render microbes
67 unnecessary for caterpillar herbivory. Caterpillars illustrate the potential ecological and
68 evolutionary benefits of independence from symbionts, a lifestyle which may be widespread
69 among animals.

70

71 **Introduction**

72

73 Many animals are colonized by microbial symbionts that have beneficial and
74 fundamentally important impacts on host biology. Microbes can regulate animal development,
75 immunity and metabolism, mediate ecological interactions, and facilitate the evolutionary origin
76 and diversification of animal clades (1–7). These integral host-microbe relationships have led to
77 the notion that all animals can be conceptualized as “holobionts” (8–10), superorganism-like
78 entities composed of the host plus its microbiome—defined here as the entire assemblage of
79 commensal, pathogenic, and mutualistic microorganisms (11). Furthermore, the recent
80 proliferation of microbiome surveys supports a widely held assumption that microbial symbioses
81 are universal across animals (12, 13).

82 The Lepidoptera (butterflies, moths, and their caterpillar larvae), despite being key
83 components of most terrestrial foodwebs and extraordinarily diverse (14), are one group in which
84 the role of microbes remains ambiguous. Here we focus on caterpillars, which are the main—and
85 in some Lepidoptera, the exclusive—feeding stage, and which have long been intensively studied
86 in many fields (15). The vast majority of caterpillars are herbivores, and some insect herbivores
87 rely on microbes to supplement missing nutrients, neutralize toxins, or digest plant cell walls (16,
88 17). However, considering caterpillars’ simple gut morphology and rapid digestive throughput, it
89 has been speculated that microbes may be unable to persist in the caterpillar gut and do not
90 contribute to digestion (18, 19). Indeed, microscopy-based studies report no, or minimal,
91 microbial growth within the caterpillar gut (20–22).

92 DNA- and culture-based investigations of caterpillar gut microbiomes have produced
93 mixed findings, with conflicting implications for microbial involvement in caterpillar biology.
94 Some studies report a highly abundant and consistent bacterial community (23–25),
95 characteristics that may indicate a functional association with the host. Others report high
96 intraspecific variability in composition, and similarity between diet- and gut-associated microbes
97 (26–29). Inconsistencies may arise from methodological factors such as contamination of low-
98 biomass samples (30), starvation prior to sampling, sequencing of extracellular DNA (31), and
99 the use of laboratory-raised insects or artificial diets (27, 32, 33). Regardless of these potential
100 biases, when studying gut microbiomes, it is often difficult to distinguish between dead or
101 dormant passengers (“transients” (34)) and persistent, living populations (“residents” (34) or
102 “symbionts” *sensu* (35)). Further, microbes in the latter category may be parasitic or pathogenic,
103 as well as beneficial. While microbes were known to cause disease in caterpillars as early as
104 Louis Pasteur’s experiments on silkworms (36), their potential importance as mutualists remains
105 unclear.

106 Do caterpillars depend on gut microbes for feeding and development? To answer this
107 question, we surveyed microbiomes of a taxonomically and geographically broad array of wild
108 caterpillars and conducted a field-based experiment on the model species *Manduca sexta*
109 (Sphingidae). Our analyses are focused on the digestive tract, the most likely habitat for
110 microbial colonization, as abundant microbes have not been observed elsewhere in the caterpillar
111 body (32, 37). First, we characterized gut microbial abundance and composition across 124
112 species of actively feeding caterpillars in Costa Rica and the United States. We applied the same
113 methods to 24 additional insect, bird, and mammal species that we expected to have functional
114 microbiomes, to assess the reliability of our protocol and to contextualize our findings. Second,
115 we experimentally tested whether gut bacteria impact the growth and survival of *M. sexta*. Our
116 findings question the generality of animal-microbe symbioses, and may inform a multitude of
117 research programs based on caterpillar herbivory in both natural and managed ecosystems (e.g.,
118 (38–41)).

119

120 **Results**

121

122 *Survey of caterpillar gut microbiomes*

123

124 Using quantitative PCR and sequencing of the 16S rRNA gene, we found that wild
125 caterpillars representing a broad diversity of Lepidoptera had gut bacterial densities multiple
126 orders of magnitude lower than the whole-body microbiomes of other insects and vertebrate
127 feces measured using identical methods ($p < 0.0001$, Fig. 1A, Table S1). Some animals host
128 symbiotic fungi (42), but fungal biomass was also lower in caterpillar guts relative to other
129 insects and vertebrates (median 6.1×10^2 vs. 9.5×10^4 rRNA gene copies per gram, $p < 0.0001$).
130 As another line of evidence of low microbial biomass, sequence libraries from caterpillar fecal
131 samples were dominated by plant DNA. Though there was extensive variability, typically more
132 than 80% of 16S rRNA gene sequences in caterpillar feces were from plant chloroplasts or
133 mitochondria, versus 0.1% for other herbivores or omnivores with plant-rich diets ($p < 0.0001$,
134 Fig. 1B). In a subset of caterpillars from which we sampled whole, homogenized midgut and
135 hindgut tissue, plant DNA represented an even higher proportion of sequences in guts than in
136 feces (Fig. S1A). This pattern is more likely a function of plant DNA degradation during

137 intestinal transit than of bacterial proliferation, as bacterial density remained similar or decreased
138 slightly from midgut to feces, depending on the caterpillar species (Fig. S1B).

139 Caterpillar gut bacterial assemblages also exhibited a high degree of intraspecific
140 variability, as shown by higher beta diversity within caterpillar species relative to other insects
141 and vertebrates ($p = 0.0002$). Such variability could indicate that the microbes found in
142 caterpillar guts are generally transient, as animals with functionally important, resident
143 microbiomes tend to host a high abundance of microbial taxa shared among conspecific
144 individuals (e.g., (43–45)). In agreement with this expectation, within most species of the other
145 animals analyzed here, microbiomes were largely made up of a common set of bacterial
146 phylotypes. For example, >99% of sequences in any one honeybee belonged to phylotypes found
147 in all honeybees included in the analysis. In contrast, even when raised on the same species of
148 food plant under identical conditions, caterpillars had a much lower proportion of their gut
149 bacterial assemblage belonging to core phylotypes (median 19.5%, $p < 0.0001$; Fig. 1C). In
150 *Schausiella santarosensis*, which among caterpillars had the highest median core size of ~50%,
151 four of its six core phylotypes belong to *Methylobacterium*, a typical inhabitant of leaf surfaces
152 (46). This observation hints that many of the core phylotypes which were found in caterpillars
153 may be transient, food-derived microbes.

154 In addition to low total abundance and high inter-individual variability, caterpillar gut
155 bacterial assemblages are dominated by leaf-associated taxa, further suggesting that resident,
156 host-specific symbionts are sparse or absent. The bacterial phylotypes present in the feces of at
157 least half of the sampled caterpillar individuals are *Staphylococcus*, *Escherichia*,
158 *Methylobacterium*, *Klebsiella/Enterobacter*, *Enterococcus*, and *Sphingomonas* (Table S2). In
159 Colorado and Costa Rica, we sampled leaf-associated bacteria from the same plant individuals
160 consumed by the sampled caterpillars to examine whether leaves are a potential source of these
161 taxa. Of the aforementioned phylotypes, all but *Staphylococcus*—a potential caterpillar pathogen
162 (47) or, like *Corynebacterium*, a transient from human skin (48)—are also among the ten most
163 common phylotypes found in leaf samples (Table S2). Across caterpillar individuals, a median
164 89.6% (interquartile range: 80.2–99.0%) of fecal bacterial sequences belonged to phylotypes
165 detected on leaves. However, bacterial assemblages were not identical between leaves and
166 caterpillar feces ($p = 0.001$). Besides the potential growth of parasites and/or mutualists in the
167 gut, this difference could arise from digestion filtering out subsets of the leaf bacterial
168 community.

169 Transient input of leaf-associated microbes could explain the substantial variation we
170 observed in caterpillar gut bacterial loads (Fig. 1A). Leaf bacterial densities were highly variable
171 within (tomato) and between (milkweed, eggplant, tomato) plant species, and this variation was
172 reflected in the feces of monarch (*Danaus plexxipus*) and *M. sexta* caterpillars feeding on them
173 ($R^2 = 0.24$, $p = 0.03$; Fig. 2A). Furthermore, bacterial densities dropped by a median of 214-fold
174 from leaves to feces, suggesting that any potential bacterial growth within the gut is relatively
175 minor (Fig. 2A). The extent of this reduction varied widely (from 5 to 8400-fold, Fig. 2A),
176 possibly because of inter-individual or interspecific differences in physiological traits that may
177 eliminate leaf microbes, such as gut pH. As with patterns in total abundance, variation in
178 bacterial taxonomic composition among leaves and caterpillar feces was correlated (Mantel $r =$
179 0.28 , $p = 0.001$; Fig. 2B). In other words, caterpillars consuming leaves with more distinct
180 bacterial assemblages produce more distinct bacterial assemblages in their feces, as would be
181 expected from a digestive system in which microbes are diet-derived and only transiently
182 present. Moreover, this process could explain a potential relationship between host relatedness

183 and microbiome structure, a pattern sometimes interpreted to indicate functional host-symbiont
184 interactions (49). Specifically, although confamilial caterpillars in Costa Rica had marginally
185 more similar gut bacterial assemblages than did caterpillars in different families ($p = 0.053$), they
186 had also been feeding on plants with especially similar leaf microbiomes ($p = 0.005$).

187
188 *Test of microbiome function in Manduca sexta*

189
190 Supporting our claim that caterpillars lack resident, functional gut microbiomes, we show
191 experimentally that the growth and survival of field-collected *Manduca sexta* caterpillars are not
192 dependent on gut bacterial activity. As measured by qPCR, wild *M. sexta* contain ~61,000-fold
193 lower bacterial loads than expected from allometric scaling relationships based on animals with
194 resident microbiomes ((50), Fig. S2). Feeding *M. sexta* antibiotics reduced this already low
195 number of gut bacteria by 14- to 365-fold (range of medians across dosages), as measured using
196 culture-dependent methods ($R^2 = 0.13$, $p = 0.003$, Fig. S3A). These colony counts were
197 positively correlated with the number of 16S rRNA gene copies ($r = 0.38$, $p = 0.003$; Fig. S3B).
198 Suppression of viable bacteria had no effect on pupal weight (antibiotics: $p = 0.45$; sex: $p =$
199 0.014 ; interaction: $p = 0.70$; Fig. 3), which is correlated with fecundity in insects (51), nor on
200 development time (antibiotics: $p = 0.19$; sex: $p = 0.023$; interaction: $p = 0.63$; Fig. S4A).
201 Likewise, antibiotic treatment did not affect survival from larval hatching to adult emergence (p
202 $= 0.19$, Fig. S4B), nor generally impact total feces production, which is an integrated measure of
203 leaf consumption and assimilation efficiency (antibiotics: $p = 0.07$; sex: $p = 0.002$; interaction: p
204 $= 0.048$). As expected with *M. sexta* (52) we found clear sexual size dimorphism, suggesting our
205 experimental design had sufficient power to detect biologically meaningful differences. Given
206 that antibiotics reduced fecal bacteria to a variable extent within and among treatments (Fig.
207 S3A), we repeated the aforementioned analyses using gut bacterial abundance as the predictor
208 variable. In all cases there was no significant relationship with host performance ($p > 0.1$),
209 further indicating that reducing or eliminating gut bacteria from caterpillars does not negatively
210 impact *M. sexta* fitness.

211
212 **Discussion**

213
214 Consistent with previous microscopy-based (20–22, 53) and molecular studies (26–29),
215 we found that resident microbial symbionts are generally absent or present only in low numbers
216 in caterpillar guts. As expected for herbivores consuming microbe-rich leaf tissue, diet-derived
217 microbes are transiently present in caterpillar guts, wherein they may be dead or inactive. That
218 the microbial biomass in caterpillar guts is far lower than in the guts or whole bodies of many
219 other animals (Fig. 1A), and also lower than in their food (Fig. 2A), suggests a lack of persistent
220 microbial growth within the gut. Moreover, any potential microbial metabolism might be too
221 limited to substantially affect digestive processes, as illustrated by our observation that *Manduca*
222 *sexta* caterpillars contain microbial loads orders of magnitude lower than comparably sized
223 animals with resident microbiomes (Fig. S2). In addition to low abundance, the composition of
224 microbes detected in caterpillar guts is highly variable among conspecific individuals (Fig. 1C).
225 Lacking stable populations of core microbial taxa, caterpillar gut microbiomes may be easily
226 influenced by the idiosyncrasies of which microbes are present on a given leaf and in what
227 abundance, and which leaf microbes can survive transit through the digestive tract. Ingested
228 microbes which die within the host may still be beneficial as a food source or by stimulating the

229 immune system, but are not themselves symbionts (following the original definition of symbiosis
230 as the “living together of different species” (referenced in (35)).

231 Based on the experiment with *M. sexta*, it is unlikely that microbes have cryptic, but
232 essential, functions in caterpillar guts. Antibiotic suppression of viable gut bacterial loads in *M.*
233 *sexta* had no apparent negative consequences, contrasting sharply with the many examples of
234 major reductions in host growth or survival upon removal of beneficial symbionts (e.g., (54–
235 56)). If anything, caterpillars treated with antibiotics showed slight (but not statistically
236 significant) increases in performance (Fig. 3, Fig. S4B). Antibiotics increase the weight gain of
237 laboratory-bred caterpillars (57–59), and commercially made caterpillar diets often contain
238 antibiotics. This effect, also observed in livestock (60), might reflect microbial parasitism
239 occurring in even apparently healthy caterpillars, and/or costly immune responses to the presence
240 of pathogens (61). Aside from known leaf-specialists, some of the most frequently detected
241 bacterial genera in this study (Table S2), including *Acinetobacter*, *Clostridium*, *Enterobacter*,
242 *Enterococcus*, *Escherichia*, and *Staphylococcus*, have been reported to cause disease in
243 caterpillars under some circumstances (37, 47, 62, 63).

244 The lack of a resident gut microbiome in caterpillars may directly result from a digestive
245 physiology that is particularly unfavorable to microbial growth (18). The midgut, the largest
246 section of the digestive tract wherein caterpillars digest leaf material and absorb the resulting
247 nutrients (64), is a hostile environment for microbes (24). It is highly alkaline, with pH values
248 often >10 (65) and as high as 12 (66), and contains host-encoded antimicrobial peptides (67).
249 Additional attributes of the caterpillar gut that may hinder microbial colonization include a
250 simple tube-like morphology without obvious microbe-housing structures (18), a continually
251 replaced lining (the peritrophic matrix) covering the midgut epithelium (68) which may prevent
252 biofilm formation, and short retention times (food transit takes ~2 hours in *M. sexta* (69)).
253 Although some insects harbor symbionts in specialized organs (53), to our knowledge, similar
254 structures have not been reported in caterpillars. Buchner’s foundational survey of animal
255 endosymbiosis describes Lepidoptera only as “a group in which no symbiont bearers have been
256 discovered” ((53), p. 817). Moreover, previous studies did not find abundant microbes outside of
257 the gut (32, 37).

258 Without the aid of microbial symbionts, how are caterpillars able to overcome the dietary
259 challenges posed by herbivory? First, caterpillars use a combination of mechanical disruption,
260 endogenously produced digestive enzymes, and high pH to extract easily solubilized nutrients,
261 primarily from the contents of plant cells (18, 70, 71). Although this method of processing leaves
262 is relatively inefficient, essential nutrients are not totally absent, so that caterpillars can
263 compensate by simply eating more (18, 64). Some insects likely require microbes for
264 detoxification (16), but many caterpillars possess host-encoded mechanisms for degrading or
265 resisting plant allelochemicals (72). However, there may be a vestigial role for microbes in these
266 processes, as genomes of many Lepidoptera contain microbial genes encoding enzymes with
267 related functions (73, 74). These gene acquisitions may have enabled a symbiont-free feeding
268 strategy.

269 The caterpillars surveyed here are likely to be representative of most externally leaf-
270 feeding Lepidoptera, as we included a range of families, habitat types, and diet breadths from
271 monophagous to highly generalist. However, a lack of resident gut microbiome in the caterpillar
272 may not apply to the adult butterfly or moth. Compared with larvae, adult butterflies host distinct
273 bacterial communities (32) and high gut microbial loads (75). Many other Lepidoptera lack
274 functioning mouthparts or digestive tracts as adults, and in these groups microbes may be

275 altogether irrelevant to digestion or nutrition. However, we cannot exclude the possibility that
276 microbial symbionts may influence host fitness by their potential activities in eggs or pupae.

277 The extraordinary diversity and abundance of Lepidoptera (14) indicates that a symbiont-
278 independent feeding strategy can be highly successful. Perhaps such success reflects a release
279 from constraints imposed on other animals that do host and depend on symbionts. There are costs
280 to engaging in mutualisms (e.g., (76–78)), and in a gut microbiome context one cost includes
281 nutrient competition between host and microbes (60). A high availability of food allows
282 caterpillars to “skim the cream” (64), assimilating simple nutrients that might otherwise be used
283 by gut microbes and excreting recalcitrant material. In other words, “Why not do the digestion
284 yourself rather than pay someone else to do it?” ((79), p. 53). Other costs include the risk of gut
285 microbes becoming pathogenic (80, 81), and the potential for pathogens to exploit a gut
286 environment that is hospitable to microbial mutualists. The extreme conditions in the caterpillar
287 midgut may instead exclude all microbial growth, providing some degree of protection against
288 disease.

289 Dependence on microbes with different physiological tolerances than the host constrains
290 overall niche breadth (7, 77). As compared with groups lacking functional microbiomes, animals
291 whose biology is heavily influenced by microbial mutualists may be less able to switch to new
292 food plants or new habitats over evolutionary time. Indeed, it has been argued that while
293 microbial symbioses can provide novel ecological functions, they may also increase the
294 extinction risk of host lineages (7, 82). As Lepidoptera represent one of the most speciose animal
295 radiations (83), a conspicuous question is whether independence from microbes may, in some
296 cases, facilitate animal diversification.

297 Caterpillars do not appear to be unique in lacking a resident microbiome that is important
298 for feeding and development. Microbiomes of walking sticks (84), sawfly larvae (85, 86), a
299 saprophagous fly (87), a parasitic horsehair worm (88), a leaf beetle (89, 90), and certain ants
300 (91) display features similar to those we observed in caterpillars. In fact, our data suggest that
301 some vertebrates also have minimal gut microbiomes, and these species may feed relatively
302 autonomously. Feces of the herbivorous brent goose (*Branta bernicla*) had low bacterial loads
303 and a high proportion of plant DNA, and the insectivorous little brown bat (*Myotis lucifugus*) had
304 similarly low fecal bacterial loads (Figs. 1A,B; Table S1). These species exhibit caterpillar-like
305 physiological traits such as a relatively short gut and rapid digestive transit (92, 93). Additional
306 examples in the microbiome literature might be obscured by contaminants masquerading as
307 mutualists (94), a frequent lack of quantitative information (91) and experimental validation of
308 microbial function *in vivo*, and publication bias against “negative results.”

309 While recent literature has documented extraordinary variation in the types of services
310 provided by microbial symbionts, less explored is variation in the degree to which animals
311 require any such services. Animals likely exist on a spectrum from tightly integrated host-
312 microbe holobionts to simply animals, *sensu stricto*, in which a microbial presence is only
313 relictual (i.e. mitochondria and horizontally transferred genes). Documenting the existence of
314 microbially independent animals as well as their ecological, physiological and phylogenetic
315 contexts, is a first step toward understanding the causes and consequences of evolutionary
316 transitions along this continuum.

317
318
319
320

321 **Methods**

322

323 *Sampling and Sequencing.* Caterpillar fecal samples (N=185) were obtained from actively
324 feeding, field-collected individuals in AZ, CO, MA, and NH, USA, and Área de Conservación
325 Guanacaste, Costa Rica. To sample plant microbiomes, we collected leaves from the same
326 branch used to feed caterpillars prior to fecal or gut sampling. Sequence composition was
327 corelated between feces and guts (*Supplemental Methods*), in line with a previous finding (32).
328 All samples were preserved in 95% ethanol or dry at -20°C (90). We extracted DNA, PCR-
329 amplified the 16S rRNA V4-V5 gene region and sequenced amplicons on an Illumina MiSeq in
330 the same manner as previous insect microbiome studies (32, 90). These DNA extracts and
331 primers were also used for quantitative PCR, which provides microbial biomass estimates
332 concordant with those from microscopy (91) and culturing (Fig. S3B). We did not find evidence
333 that low amplification of caterpillar fecal bacteria is due to primer bias, PCR inhibitors, or
334 storage methods (*Supplemental Methods*).

335

336 *Antibiotic Experiment.* We collected *Manduca sexta* eggs from *Datura wrightii* plants near
337 Portal, AZ, USA. 72 newly hatched larvae were randomly and evenly divided among six
338 treatments varying from 0-1.68 mg total antibiotics per ml distilled water, and reared in separate
339 unused plastic bags on *D. wrightii* foliage at the Southwestern Research Station. Water with or
340 without antibiotics was sprayed onto leaves, which were briefly dried prior to feeding. The
341 compounds used (rifampicin, tetracycline, streptomycin, in a 1:2:4 ratio) suppressed bacterial
342 symbionts in other insect herbivores (55, 95). We collected a fresh fecal pellet from each
343 caterpillar midway through the final instar, from which one subsample was cultured on LB
344 media, and another used for qPCR and sequencing with the aforementioned protocol. Pupae were
345 weighed six days after pupation and monitored daily for adult eclosion.

346

347 *Data Analysis.* Statistical analyses were conducted in R (96). Differences in bacterial loads, core
348 sizes, and *M. sexta* performance variables were tested using linear models; residuals were
349 visually inspected for Gaussian structure. The betadisper function in the vegan package was used
350 to compare intraspecific beta diversity. *M. sexta* survival was analyzed using logistic regression.
351 We used a Mantel test to estimate the rank correlation between leaf and fecal microbiome
352 dissimilarities. A Wilcoxon test was used for proportions of plant DNA. Differences in
353 community composition were analyzed using PERMANOVA. DNA sequences, metadata, and R
354 code available at doi:10.6084/m9.figshare.4955648.

355

356 **Acknowledgments**

357

358 We thank the parataxonomists and staff of ACG and The Caterpillar Lab, without whom
359 this project would not have been possible. J Dickerson, K Vaccarello, and H Layton provided
360 invaluable help in the field and laboratory. HA Woods and MD Bowers aided in the planning
361 and interpretation of this study. T Sharpton provided feedback on an earlier version of the
362 manuscript. We also acknowledge administrative support from ACG, the Southwestern Research
363 Station, and A Dietz. TJH was supported by the American Philosophical Society's Lewis and
364 Clark Fund, the National Science Foundation (NSF) Graduate Research Fellowship Program
365 (1144083) and an NSF Doctoral Dissertation Improvement Grant (1601787). DHJ and WH were

366 supported by the University of Pennsylvania, Wege Foundation, Permian Global, government of
367 Costa Rica, and Guanacaste Dry Forest Conservation Fund.

368

369 References

370

- 371 1. McFall-Ngai M, et al. (2013) Animals in a bacterial world, a new imperative for the life
372 sciences. *Proc Natl Acad Sci* 110(9):3229–3236.
- 373 2. Sudakaran S, Kost C, Kaltenpoth M (2017) Symbiont acquisition and replacement as a
374 source of ecological innovation. *Trends Microbiol* 25(5):375–390.
- 375 3. Janson EM, Stireman JO, Singer MS, Abbot P (2008) Phytophagous insect-microbe
376 mutualisms and adaptive evolutionary diversification. *Evolution* 62(5):997–1012.
- 377 4. Frago E, Dicke M, Godfray HCJ (2012) Insect symbionts as hidden players in insect–plant
378 interactions. *Trends Ecol Evol* 27(12):705–711.
- 379 5. Sommer F, Bäckhed F (2013) The gut microbiota — masters of host development and
380 physiology. *Nat Rev Microbiol* 11(4):227–38.
- 381 6. Douglas AE (2014) Symbiosis as a general principle in eukaryotic evolution. *Cold Spring*
382 *Harb Perspect Biol* 6(2):1–14.
- 383 7. Moran NA (2002) The ubiquitous and varied role of infection in the lives of animals and
384 plants. *Am Nat* 160:S1-8.
- 385 8. Zilber-Rosenberg I, Rosenberg E (2008) Role of microorganisms in the evolution of
386 animals and plants: The hologenome theory of evolution. *FEMS Microbiol Rev*
387 32(5):723–735.
- 388 9. Bordenstein SR, Theis KR (2015) Host biology in light of the microbiome: Ten principles
389 of holobionts and hologenomes. *PLoS Biol* 13(8):e1002226.
- 390 10. Gilbert SF, Sapp J, Tauber AI (2012) A symbiotic view of life: We have never been
391 individuals. *Q Rev Biol* 87(4):325–341.
- 392 11. Lederberg J, McCray AT (2001) 'Ome Sweet 'Omics—a genealogical treasury of words.
393 *Sci*:8.
- 394 12. Russell JA, Dubilier N, Rudgers JA (2014) Nature's microbiome: Introduction. *Mol Ecol*
395 23(6):1225–1237.
- 396 13. Vavre F, Kremer N (2014) Microbial impacts on insect evolutionary diversification: from
397 patterns to mechanisms. *Curr Opin Insect Sci* 4(1):29–34.
- 398 14. Scoble MJ (1992) *The Lepidoptera* (Oxford University Press).
- 399 15. Stamp NE, Casey TM (1993) *Caterpillars: ecological and evolutionary constraints on*
400 *foraging* (Chapman & Hall Ltd.).
- 401 16. Hammer TJ, Bowers MD (2015) Gut microbes may facilitate insect herbivory of
402 chemically defended plants. *Oecologia* 179(1):1–14.
- 403 17. Douglas AE (2009) The microbial dimension in insect nutritional ecology. *Funct Ecol*
404 23(1):38–47.
- 405 18. Appel H (1994) The chewing herbivore gut lumen: physicochemical conditions and their
406 impact on plant nutrients, allelochemicals, and insect pathogens. *Insect–Plant*
407 *Interactions*, ed Bernays EA, pp 209–223.
- 408 19. Bernays EA, Janzen DH (1988) Saturniid and sphingid caterpillars: Two ways to eat
409 leaves. *Ecology* 69(4):1153–1160.
- 410 20. Shannon AL, Attwood G, Hopcroft DH, Christeller JT (2001) Characterization of lactic
411 acid bacteria in the larval midgut of the keratinophagous lepidopteran, *Hofmannophila*

- 412 *pseudospretella*. *Lett Appl Microbiol* 32(1):36–41.
- 413 21. Kukul O, Dawson TE (1989) Temperature and food quality influences feeding behavior,
414 assimilation efficiency and growth rate of arctic woolly-bear caterpillars. *Oecologia*
415 79(4):526–532.
- 416 22. Vilanova C, Latorre A, Porcar M (2016) The generalist inside the specialist: Gut bacterial
417 communities of two insect species feeding on toxic plants are dominated by *Enterococcus*
418 sp. *Front Microbiol* 7(June):1–8.
- 419 23. Anand AAP, et al. (2010) Isolation and characterization of bacteria from the gut of
420 *Bombyx mori* that degrade cellulose, xylan, pectin and starch and their impact on
421 digestion. *J Insect Sci* 10(107):1–20.
- 422 24. Broderick NA, Raffa KF, Goodman RM, Handelsman J (2004) Census of the bacterial
423 community of the gypsy moth larval midgut by using culturing and culture-independent
424 methods. *Appl Environ Microbiol* 70(1):293–300.
- 425 25. Kingsley V V. (1972) Persistence of intestinal bacteria in the developmental stages of the
426 monarch butterfly (*Danaus plexippus*). *J Invertebr Pathol* 20:51–58.
- 427 26. Priya NG, Ojha A, Kajla MK, Raj A, Rajagopal R (2012) Host plant induced variation in
428 gut bacteria of *Helicoverpa armigera*. *PLoS One* 7(1):e30768.
- 429 27. Staudacher H, et al. (2016) Variability of bacterial communities in the moth *Heliothis*
430 *virescens* indicates transient association with the host. *PLoS One* 11(5):1–21.
- 431 28. Mason CJ, Raffa KF (2014) Acquisition and structuring of midgut bacterial communities
432 in gypsy moth (Lepidoptera: Erebidae) larvae. *Environ Entomol* 43(3):595–604.
- 433 29. Whitaker M, Pierce N, Salzman S, Kaltenpoth M, Sanders J (2016) Microbial
434 communities of lycaenid butterflies do not correlate with larval diet. *Front Microbiol*
435 7:1920.
- 436 30. Salter SJ, et al. (2014) Reagent and laboratory contamination can critically impact
437 sequence-based microbiome analyses. *BMC Biol* 12(1):87.
- 438 31. Carini P, et al. (2016) Relic DNA is abundant in soil and obscures estimates of soil
439 microbial diversity. *Nat Microbiol* 5(9):680840.
- 440 32. Hammer TJ, McMillan WO, Fierer N (2014) Metamorphosis of a butterfly-associated
441 bacterial community. *PLoS One* 9(1):e86995.
- 442 33. Lighthart B (1988) Some changes in gut bacterial flora of field-grown *Peridroma saucia*
443 (Lepidoptera: Noctuidae) when brought into the laboratory. *Appl Environ Microbiol*
444 54(7):1896–8.
- 445 34. Berg RD (1996) The indigenous gastrointestinal microflora. *Trends Microbiol* 4(11):430–
446 435.
- 447 35. Douglas AE (1994) *Symbiotic Interactions* (Oxford University Press).
- 448 36. Steinhaus EA (1956) *Microbial control—the emergence of an idea*
449 doi:10.1017/CBO9781107415324.004.
- 450 37. Bucher GE (1967) Pathogens of tobacco and tomato hornworms. *J Invertebr Pathol* 9:82–
451 89.
- 452 38. Ehrlich P, Raven P (1964) Butterflies and plants: a study in coevolution. *Evolution*
453 18(4):586–608.
- 454 39. I-M-Arnold A, et al. (2016) Forest defoliator pests alter carbon and nitrogen cycles. *R Soc*
455 *Open Sci* 3(10):160361.
- 456 40. Pearse IS, Altermatt F (2013) Predicting novel trophic interactions in a non-native world.
457 *Ecol Lett* 16(8):1088–94.

- 458 41. Thomas CD, et al. (1987) Incorporation of a European weed into the diet of a North
459 American herbivore. *Evolution* 41(4):892–901.
- 460 42. Gibson CM, Hunter MS (2010) Extraordinarily widespread and fantastically complex:
461 comparative biology of endosymbiotic bacterial and fungal mutualists of insects. *Ecol Lett*
462 13(2):223–34.
- 463 43. Falony G, et al. (2016) Population-level analysis of gut microbiome variation. *Science*
464 352(6285):560–564.
- 465 44. Moran NA, Hansen AK, Powell JE, Sabree ZL (2012) Distinctive gut microbiota of honey
466 bees assessed using deep sampling from individual worker bees. *PLoS One* 7(4):e36393.
- 467 45. Sudakaran S, Salem H, Kost C, Kaltenpoth M (2012) Geographical and ecological
468 stability of the symbiotic mid-gut microbiota in European firebugs, *Pyrrhocoris apterus*
469 (Hemiptera, Pyrrhocoridae). *Mol Ecol* 21:6134–6151.
- 470 46. Vorholt JA (2012) Microbial life in the phyllosphere. *Nat Rev Microbiol* 10(12):828–840.
- 471 47. Haloi K, Kankana M, Nath R, Devi D (2016) Characterization and pathogenicity
472 assessment of gut-associated microbes of muga silkworm *Antheraea assamensis* Helfer
473 (Lepidoptera: Saturniidae). *J Invertebr Pathol* 138:73–85.
- 474 48. Grice EA, et al. (2009) Topographical and temporal diversity of the human skin
475 microbiome. *Science* 324(5931):1190–2.
- 476 49. Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016)
477 Phylosymbiosis: Relationships and functional effects of microbial communities across
478 host evolutionary history. *PLoS Biol* 14(11):e2000225.
- 479 50. Kieft TL, Simmons KA (2015) Allometry of animal-microbe interactions and global
480 census of animal-associated microbes. *Proc R Soc B* 282:20150702.
- 481 51. Honek A (1993) Intraspecific variation in body size and fecundity in insects: a general
482 relationship. *Oikos* 66(3):483–492.
- 483 52. Stillwell RC, Davidowitz G (2010) A developmental perspective on the evolution of
484 sexual size dimorphism of a moth. *Proc R Soc B* 277(1690):2069–2074.
- 485 53. Buchner P (1965) *Endosymbiosis of animals with plant microorganisms* (John Wiley &
486 Sons).
- 487 54. Salem H, Kreutzer E, Sudakaran S, Kaltenpoth M (2012) Actinobacteria as essential
488 symbionts in firebugs and cotton stainers (Hemiptera, Pyrrhocoridae). *Environ Microbiol*
489 15:1956–1968.
- 490 55. Ceja-Navarro JA, et al. (2015) Gut microbiota mediate caffeine detoxification in the
491 primary insect pest of coffee. *Nat Commun* 6:7618.
- 492 56. Raymann K, Shaffer Z, Moran NA (2017) Antibiotic exposure perturbs the gut microbiota
493 and elevates mortality in honeybees. *PLoS Biol* 15(3):e2001861.
- 494 57. van der Hoeven R, Betrabet G, Forst S (2008) Characterization of the gut bacterial
495 community in *Manduca sexta* and effect of antibiotics on bacterial diversity and nematode
496 reproduction. *FEMS Microbiol Lett* 286(2):249–56.
- 497 58. Murthy MR, Sreenivasaya M (1953) Effect of antibiotics on the growth of the silkworm
498 *Bombyx mori* L. *Nature* 172(4380):684–685.
- 499 59. Visôto LE, Oliveira MGA, Guedes RNC, Ribon AOB, Good-God PIV (2009)
500 Contribution of gut bacteria to digestion and development of the velvetbean caterpillar,
501 *Anticarsia gemmatalis*. *J Insect Physiol* 55(3):185–91.
- 502 60. Gaskins HR, Collier CT, Anderson DB (2002) Antibiotics as growth promotants: mode of
503 action. *Anim Biotechnol* 13:29–42.

- 504 61. Rolff J, Siva-Jothy MT, Rolff M. T. J. S-J (2003) Invertebrate ecological immunology.
505 *Science* 301(5632):472–475.
- 506 62. Bulla LA, Rhodes RA, St. Julian G (1975) Bacteria as insect pathogens. *Annu Rev*
507 *Microbiol* 29:163–190.
- 508 63. Peleg AY, et al. (2009) *Galleria mellonella* as a model system to study *Acinetobacter*
509 *baumannii* pathogenesis and therapeutics. *Antimicrob Agents Chemother* 53(6):2605–
510 2609.
- 511 64. Dow JAT (1986) Insect midgut function. *Adv In Insect Phys* 19:187–328.
- 512 65. Johnson K, Felton G (1996) Potential influence of midgut pH and redox potential on
513 protein utilization in insect herbivores. *Arch Insect Biochem Physiol* 32:85–105.
- 514 66. Dow JAT (1984) Extremely high pH in biological systems: a model for carbonate
515 transport. *Am J Physiol* 246:R633–R636.
- 516 67. Jiang H, Vilcinskas A, Kanost MR (2010) Immunity in Lepidopteran insects. *Invertebrate*
517 *Immunity*, ed Söderhäll K (Springer), pp 181–204.
- 518 68. Hegedus D, Erlandson M, Gillott C, Toprak U (2009) New insights into peritrophic matrix
519 synthesis, architecture, and function. *Annu Rev Entomol* 54:285–302.
- 520 69. Brinkmann N, Tebbe CC (2007) Leaf-feeding larvae of *Manduca sexta* (Insecta,
521 Lepidoptera) drastically reduce copy numbers of *aadA* antibiotic resistance genes from
522 transplastomic tobacco but maintain intact *aadA* genes in their feces. *Environ Biosafety*
523 *Res* 6:121–133.
- 524 70. Santos CD, Ferreira C, Terra WR (1983) Consumption of food and spatial organization of
525 digestion in the cassava hornworm, *Erinnyis ello*. *J Insect Physiol* 29(9):707–714.
- 526 71. Barbehenn R V. (1992) Digestion of uncrushed leaf tissues by leaf-snipping larval
527 Lepidoptera. *Oecologia* 89(2):229–235.
- 528 72. Després L, David J-P, Gallet C (2007) The evolutionary ecology of insect resistance to
529 plant chemicals. *Trends Ecol Evol* 22(6):298–307.
- 530 73. Sun BF, et al. (2013) Multiple ancient horizontal gene transfers and duplications in
531 lepidopteran species. *Insect Mol Biol* 22(1):72–87.
- 532 74. Wybouw N, et al. (2014) A gene horizontally transferred from bacteria protects arthropods
533 from host plant cyanide poisoning. *eLife* 3:e02365.
- 534 75. Ravenscraft A, Berry M, Hammer T, Peay K, Boggs C (2017) Structure and function of
535 the bacterial and fungal gut flora of Neotropical butterflies. *bioRxiv*. doi:10.1101/128884.
- 536 76. Frederickson ME, et al. (2012) The direct and ecological costs of an ant-plant symbiosis.
537 *Am Nat* 179(6):768–78.
- 538 77. Nougé O, Gallet R, Chevin L-M, Lenormand T (2015) Niche limits of symbiotic gut
539 microbiota constrain the salinity tolerance of brine shrimp. *Am Nat* 186(3):390–403.
- 540 78. Simonsen AK, Dinnage R, Barrett LG, Prober SM, Thrall PH (2017) Symbiosis limits
541 establishment of legumes outside their native range at a global scale. *Nat Commun*
542 8:14790.
- 543 79. Janzen DH (1985) The natural history of mutualisms. *The Biology of Mutualism*, ed
544 Boucher DH (Oxford University Press), pp 40–99.
- 545 80. Broderick NA, et al. (2009) Contributions of gut bacteria to *Bacillus thuringiensis*-
546 induced mortality vary across a range of Lepidoptera. *BMC Biol* 7:11.
- 547 81. Young BC, et al. (2017) Severe infections emerge from the microbiome by adaptive
548 evolution. *bioRxiv*. doi:10.1101/116681.
- 549 82. Bennett GM, Moran NA (2015) Heritable symbiosis: The advantages and perils of an

- 550 evolutionary rabbit hole. *Proc Natl Acad Sci* 112(33):10169–10176.
- 551 83. Wiens JJ, Lapoint RT, Whiteman NK (2015) Herbivory increases diversification across
552 insect clades. *Nat Commun* 6:1–7.
- 553 84. Shelomi M, Lo W-S, Kimsey LS, Kuo C-H (2013) Analysis of the gut microbiota of
554 walking sticks (Phasmatodea). *BMC Res Notes* 6(1):368.
- 555 85. Lucarotti CJ, Whittome-Waygood BH, Levin DB (2011) Histology of the larval
556 *Neodiprion abietis* (Hymenoptera: Diprionidae) digestive tract. *Psyche*:1–10.
- 557 86. Whittome B, Graham RI, Levin DB (2007) Preliminary examination of gut bacteria from
558 *Neodiprion abietis* (Hymenoptera: Diprionidae) larvae. *J Entomol Soc Ontario* 138:49–63.
- 559 87. Šustr V, Stingl U, Brune A (2014) Microprofiles of oxygen, redox potential, and pH, and
560 microbial fermentation products in the highly alkaline gut of the saprophagous larva of
561 *Penthetria holosericea* (Diptera: Bibionidae). *J Insect Physiol* 67:64–69.
- 562 88. Hudson AJ, Floate KD (2009) Further evidence for the absence of bacteria in horsehair
563 worms (Nematomorpha: Gordiidae). *J Parasitol* 95(6):1545–1547.
- 564 89. Taylor EC (1985) Cellulose digestion in a leaf eating insect, the mexican bean beetle,
565 *Epilachna varivestis*. *Insect Biochem* 15(2):315–320.
- 566 90. Hammer TJ, Dickerson JC, Fierer N (2015) Evidence-based recommendations on storing
567 and handling specimens for analyses of insect microbiota. *PeerJ* 3:e1190.
- 568 91. Sanders J, et al. (2017) Dramatic differences in gut bacterial densities help to explain the
569 relationship between diet and habitat in rainforest ants. *bioRxiv*. doi:10.1101/114512.
- 570 92. Buchler ER (1975) Food transit time in *Myotis lucifugus* (Chiroptera: Vespertilionidae). *J*
571 *Mammal* 56(1):252–255.
- 572 93. Buchsbaum R, Wilson J, Valiela I (1986) Digestibility of plant constituents by Canada
573 Geese and Atlantic Brant. *Ecology* 67(2):386–393.
- 574 94. Lauder AP, et al. (2016) Comparison of placenta samples with contamination controls
575 does not provide evidence for a distinct placenta microbiota. *Microbiome* 4(1):29.
- 576 95. Wilkinson TL (1998) The elimination of intracellular microorganisms from insects: An
577 analysis of antibiotic-treatment in the pea aphid (*Acyrtosiphon pisum*). *Comp Biochem*
578 *Physiol Part A* 119:871–881.
- 579 96. R Core Team (2016) R: A language and environment for statistical computing. *R Found*
580 *Stat Comput*. Available at: <http://www.r-project.org>.

581 582 **Figure Legends**

583
584 **Figure 1.** Comparisons of bacterial densities, proportions of plant DNA in sequence libraries,
585 and intraspecific variability between caterpillars and other animals expected to host functional
586 microbiomes. Medians are indicated by black dashed lines, and points are horizontally jittered.
587 Data for each species are listed in Table S1. One caterpillar species yielding <100 total
588 sequences was excluded. A) The density of bacterial 16S rRNA gene copies in caterpillar feces
589 versus fecal (vertebrates) or whole-body homogenate (other insect) samples of other animals
590 (N=121 caterpillar species, 24 other species). Two caterpillar species with lower amplification
591 than DNA extraction blanks are not shown. For species with multiple replicates, the median is
592 plotted. B) The proportion of sequence libraries assigned to plant chloroplast or mitochondrial
593 16S rRNA (N=123 caterpillars, 21 other herbivores). C) The proportion of sequences belonging
594 to core phylotypes, defined as those present in all conspecific individuals. Included are species
595 with at least three replicates with >100 bacterial sequences each (N=7 caterpillars, 19 other

596 animals). For species with more than three, points show the median core size across all
597 combinations of three individuals, and error bars show the interquartile range.

598
599 **Figure 2.** The abundance and composition of bacteria present in caterpillar fecal samples, as
600 compared with paired diet (leaf) samples. A) The density of bacterial 16S rRNA gene copies in
601 ground leaves versus feces, for 16 individuals collected in Colorado. Parallel lines indicate the
602 association observed between plant and fecal bacterial abundances across paired samples. B) The
603 correlation between beta diversity (Bray-Curtis dissimilarity metric) across caterpillar fecal
604 samples collected in Costa Rica, and their paired leaf surface samples (N=24 caterpillar species,
605 19 plant species; 26 individuals each). Here only samples with >2,000 sequences are shown, to
606 facilitate visualization.

607
608 **Figure 3.** Increasing concentration of an antibiotic cocktail, delivered by a spray applied to
609 *Datura wrightii* leaves prior to feeding, does not reduce *Manduca sexta* growth (N=62). Males
610 and females are plotted separately, as they were expected to differ in size. Fresh weight was
611 measured six days after pupation. Pupal weight correlates with adult fecundity and is often used
612 as a proxy of insect fitness.

613
614 **Figure S1.** Community composition, bacterial density, and midgut pH in five caterpillar species
615 (Saturniidae) from Área de Conservación Guanacaste (ACG), Costa Rica. A) The composition of
616 sequence libraries from the leaf surface, midgut, hindgut, and feces. The median across five
617 replicate individuals is displayed. The food plant species is indicated in parentheses. Note that
618 one species, *Eacles imperialis*, was reared separately on two plant species. Only plant chloroplast
619 or mitochondrial sequences, reagent contaminants, and the top 10 bacterial genera (among the
620 dissected individuals only) are shown; the remainder of the community (summing to 1)
621 represents sequences from a variety of low-abundance taxa. B) The number of bacterial 16S
622 rRNA gene copies per gram (fresh weight) in homogenized midgut or hindgut tissue and feces
623 (N=5 individuals per species, except *E. imperialis* with 10 individuals (5 each on two plant
624 species)). C) Photographs of each species taken in ACG.

625
626 **Figure S2.** Allometric scaling of whole-individual microbial loads with body size. Triangles and
627 the solid line show data replotted from (50), which were originally measured using microscopy
628 or culturing. Circles show data generated in this study, using quantitative PCR. The dashed
629 regression line is calculated from a model only including non-caterpillar species analyzed in this
630 study, limited to those species with bacterial densities not less than 1/100 of the group median.
631 The red horizontal dotted line indicates the median per-caterpillar bacterial load for 17 *Manduca*
632 *sexta* individuals collected in Colorado (N=15) or Arizona (N=2). The photograph is *M. sexta*
633 feeding on *D. wrightii*.

634
635 **Figure S3.** The relationships between antibiotic dose, the number of bacterial colony-forming
636 units cultured on LB media, and the number of bacterial 16S rRNA gene copies measured by
637 qPCR. Fecal samples that yielded no cultured colonies are plotted at 10^0 on \log_{10} axes; for these,
638 nonzero estimates of 16S rRNA gene copies are likely due, in large part, to amplification of
639 DNA from dead or nonviable cells (see *Supplemental Methods*). A) Effect of antibiotic treatment
640 on the number of culturable bacteria in caterpillar feces. Points are individual caterpillars (N=60)
641 and are horizontally jittered for clarity. Dashed lines are medians for each treatment. B)

642 Correlation of bacterial density as measured by culturing versus by DNA quantification. The 1:1
643 line between the two variables is shown.

644

645 **Figure S4.** The relationship between antibiotic treatment and other components of *M. sexta*
646 fitness. 12 individuals, randomly selected from the population, were used to initiate each group.
647 A) Number of days from larval hatching from eggs to the cessation of feeding, which marks the
648 beginning of the prepupal stage. Shown are the 64 individuals that survived to this point. We
649 were unable to identify the sex of two individuals. B) The proportion of individuals surviving
650 from larval hatch to adult eclosion, for the control group and each antibiotic treatment.

651

652 **Figure S5.** Two tests for PCR inhibitory substances in caterpillar feces. A) Fecal DNA from
653 eight *M. sexta* individuals, arranged left-right by decreasing 16S rRNA gene copy number in
654 original extracts. For each individual, \log_{10} (16S rRNA gene copies) is shown for the original
655 sample, and for and extracts diluted 1:10 and 1:100 in pure water. Copy number estimates are
656 standardized per μ l of original DNA extract. Note that variability between technical replicates
657 increases with low concentrations of template DNA. One sample, D-0.01, had less amplification
658 than negative controls and is not shown. B) Amplification (arbitrary units) of rDNA ITS of *B.*
659 *dendrobatidis*, a chytrid fungus of amphibians, showing 12 replicate controls (PCR-grade water
660 only) versus 12 reactions to which 5 μ l of caterpillar fecal DNA was substituted for water.
661 Means of triplicate reactions are shown. The twelve caterpillar species with the lowest total 16S
662 rRNA gene copy number were used for this test. Dashed lines show medians for each group.

Figures

Figure 1.

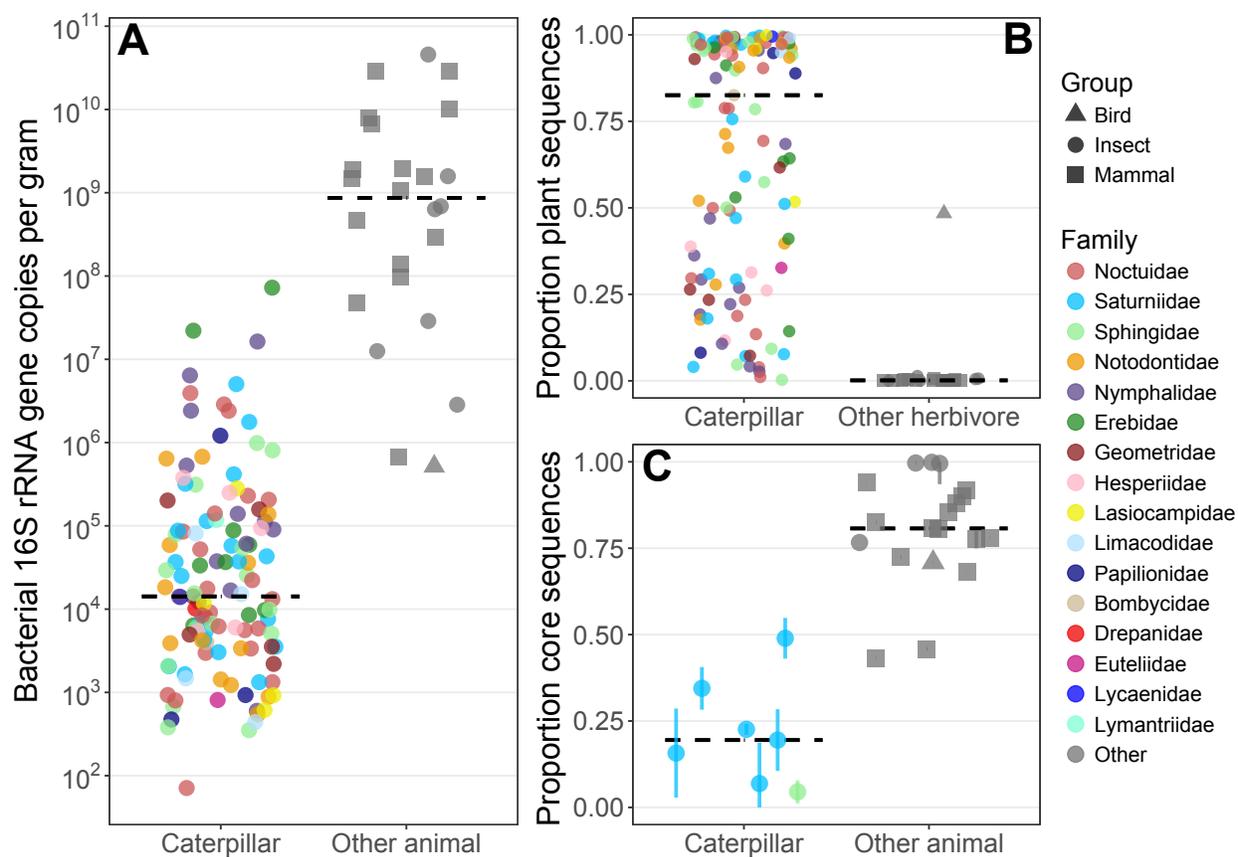


Figure 2.

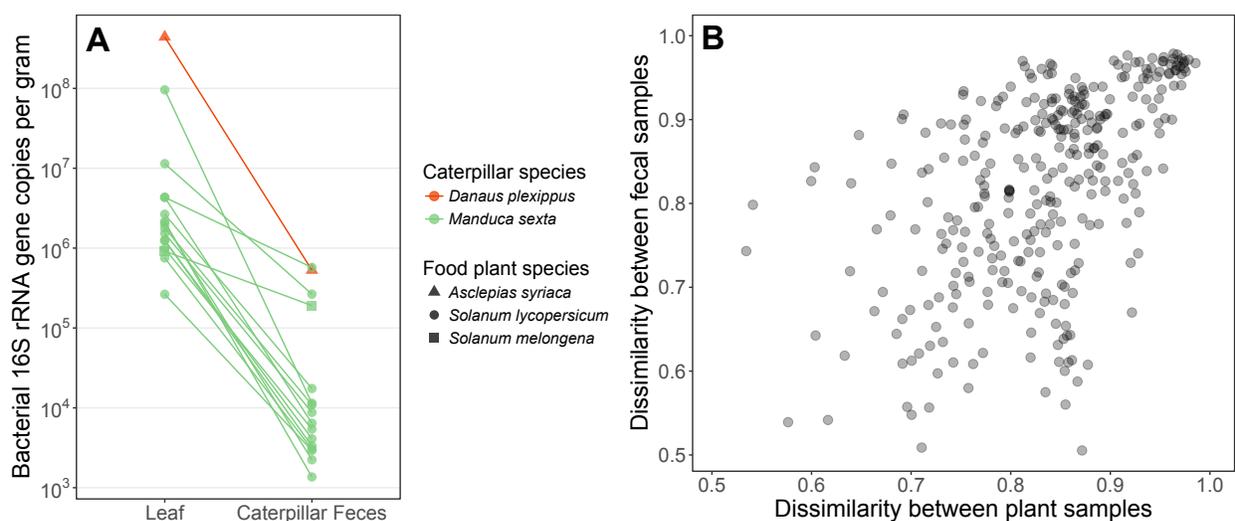
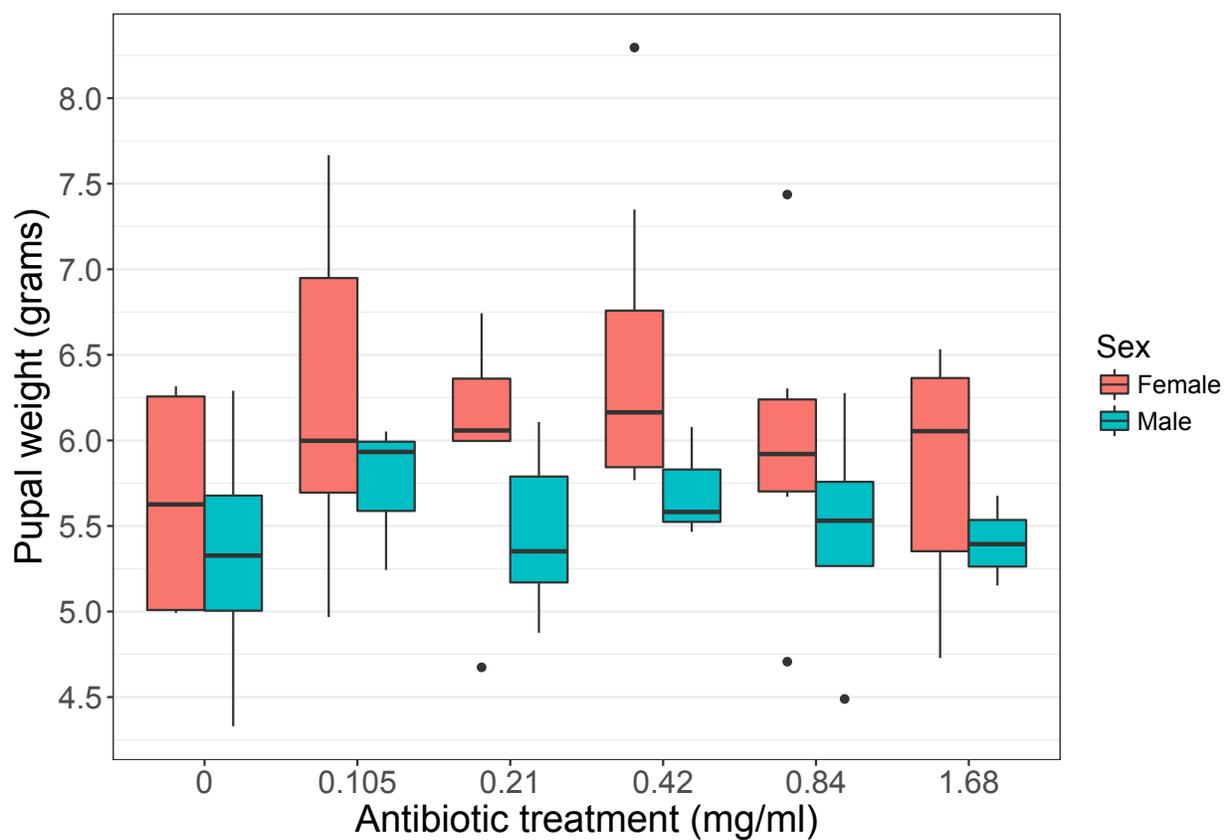


Figure 3.



Supplemental Figures

Figure S1.

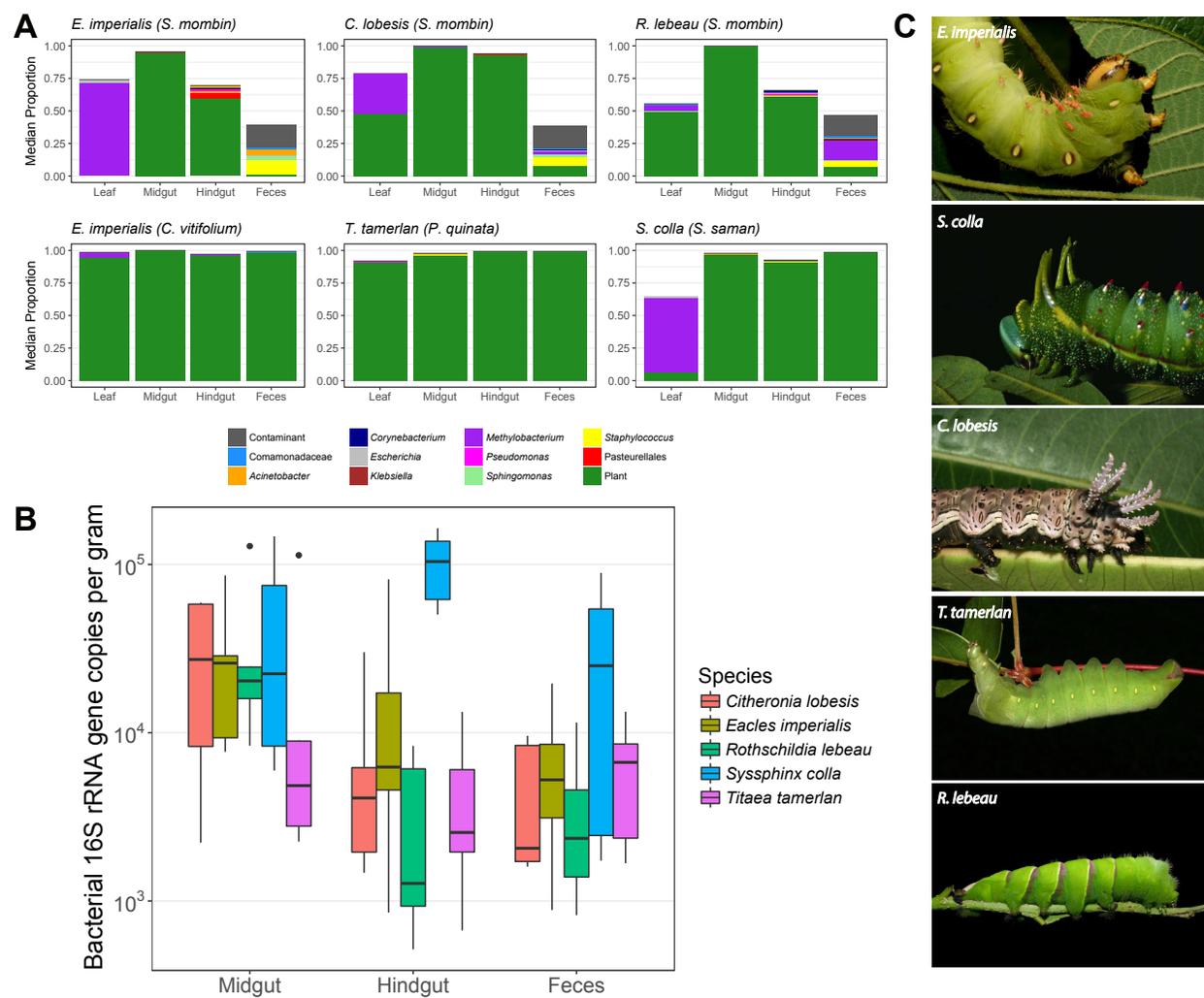


Figure S2.

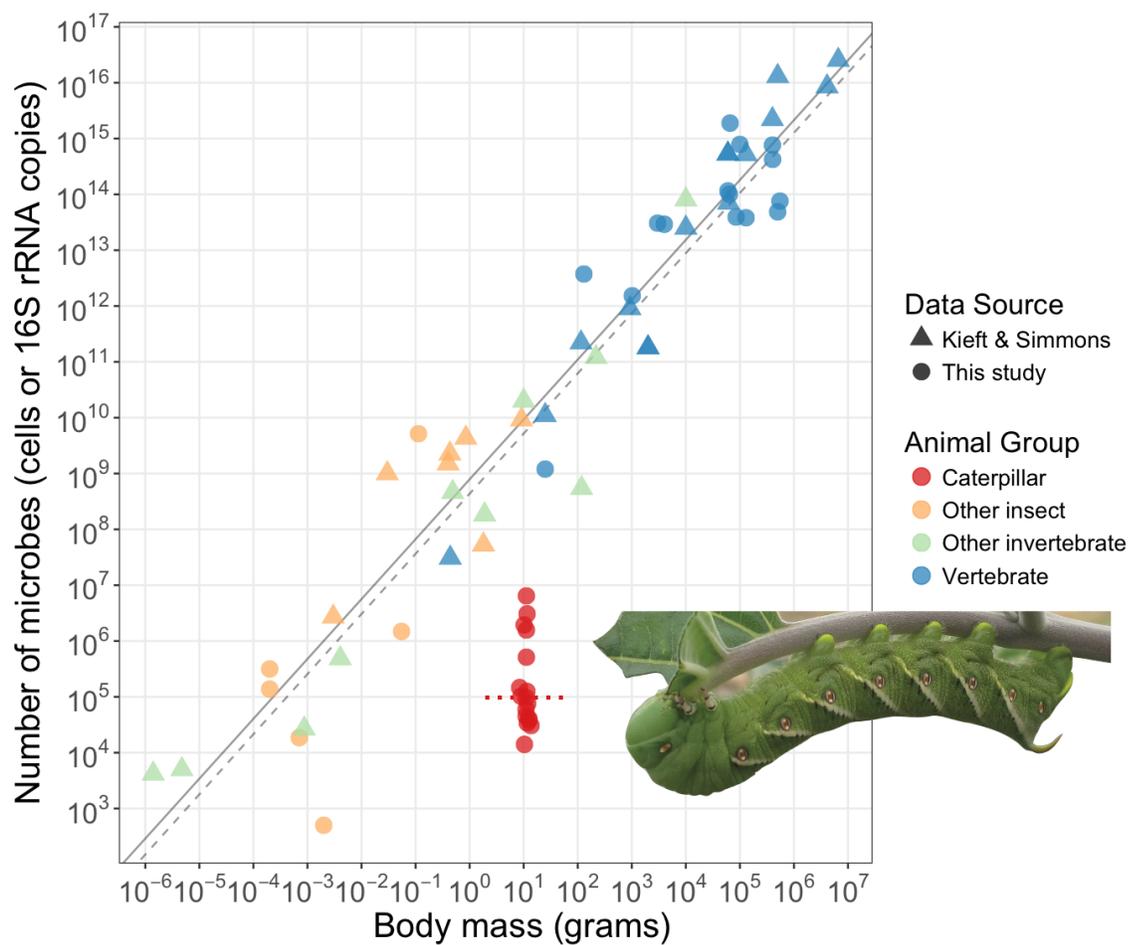


Figure S3.

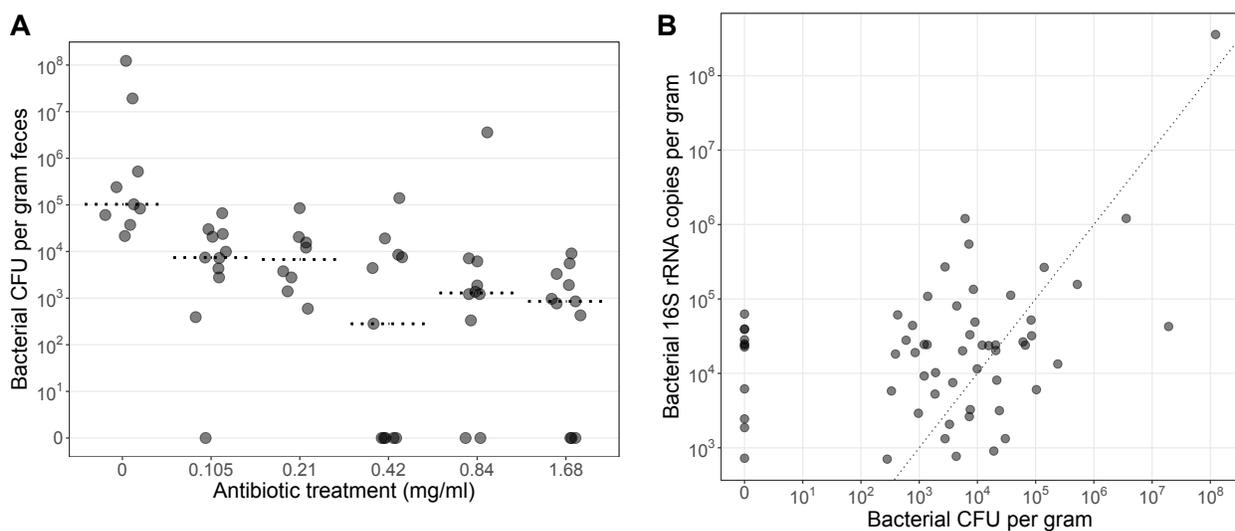


Figure S4.

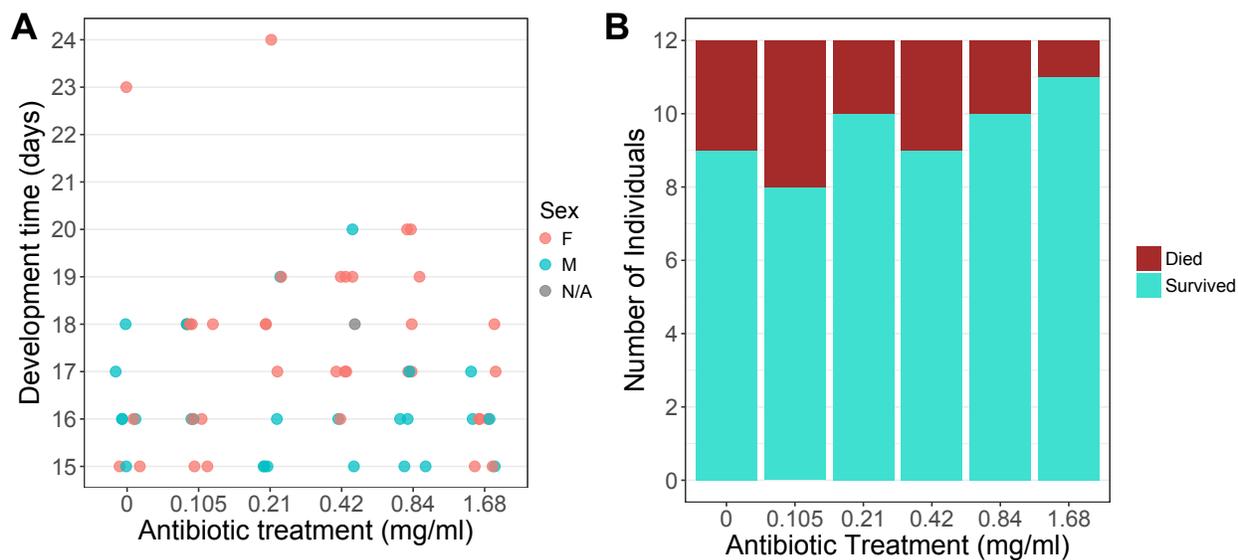
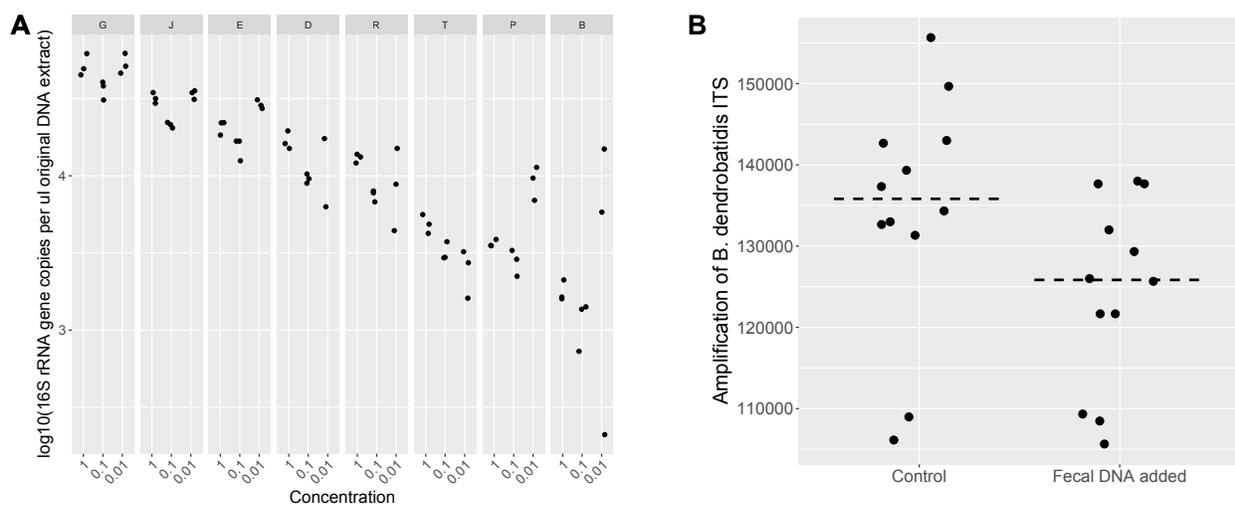


Figure S5.



664 Supplemental Methods

665

666 *Sampling*

667

668 Fecal samples were obtained from wild populations of caterpillars in four regions: Área
669 de Conservación Guanacaste (Costa Rica), New Hampshire and Massachusetts (USA), Boulder
670 County, Colorado (USA), and Portal, Arizona (USA). Caterpillars were collected in ACG under
671 permit #ACG-PI-027-2015 and in Arizona under a Scientific Use Permit from the United States
672 Forest Service. For more details about the ACG landscape and collection, rearing, and
673 identification protocols, see (1–3). Most species were collected as caterpillars, but some ACG
674 specimens were reared from eggs either found on foliage or laid by females caught at light traps
675 (see data file “Metadata_ACG2015”). For some caterpillars we had information on whether they
676 died of parasitoids or disease after sampling, and these samples were discarded in order to focus
677 on apparently healthy individuals. Most caterpillars were sampled in the final or penultimate
678 instar.

679 All samples were preserved within 30 minutes of defecation, as preliminary evidence
680 suggested rapid (by 6-12 hours) bacterial and fungal growth in excreted fecal pellets, which
681 would render old feces unsuitable as a proxy for gut microbial communities. In five caterpillar
682 species, we did not find evidence for abundant bacterial populations in the midgut (including
683 both ecto- and endoperitrophic spaces) or hindgut that were not captured in feces (Fig. S1A),
684 supporting a previous finding that caterpillar feces approximates the whole-body microbial
685 community (4). Further supporting the use of fresh feces to sample microbes in the caterpillar
686 gut, we found that the inter-individual variation in sequence composition (including nonbacterial
687 DNA) was reflected in fecal samples (Mantel test: midgut $r = 0.33$, $p = 0.001$; hindgut $r = 0.39$, p
688 $= 0.001$).

689 We preserved gut and fecal samples using either dry storage at -20°C or 95% ethanol
690 (Table S1); both methods are suitable for storing insect microbiome samples and do not
691 substantially alter community composition (5). Approximately 50 mg (fresh weight) of sample
692 was used for DNA extraction. Prior to DNA extraction, ethanol-preserved samples were dried in
693 a vacuum centrifuge; since this also evaporated water, their fresh weight equivalent was
694 estimated using percent water content calculated from *M. sexta* guts or feces. To test whether
695 microbial biomass estimates may have been biased by ethanol storage, we compared PCR
696 amplification for paired ethanol-stored and frozen fecal pellets from eight *M. sexta* individuals.
697 From a collection of pellets defecated by each individual during a 1-2 hour window, separate
698 pellets were randomly chosen for each storage type (note that pre-storage inter-pellet microbial
699 variation is possible even under these relatively controlled conditions). As assessed by a linear
700 mixed-effects model treating individual as a random effect, there was no significant influence of
701 storage method on 16S rRNA gene copy number ($\chi^2(1) = 1.09$, $p = 0.30$).

702 For caterpillars in Costa Rica and Colorado, we also sampled microbes from leaves of the
703 same branch as that fed to the caterpillar prior to feces collection. With this strategy we aimed to
704 maximize microbial similarity between the leaves that were sampled and those consumed by the
705 caterpillar, although leaf microbiomes can also vary substantially within a branch (6). These
706 leaves appeared clean and had not, to our knowledge, come into contact with any caterpillars
707 prior to sampling. Leaves from Colorado plants were frozen dry at -20°C and ground under
708 liquid N_2 with a mortar and pestle prior to DNA extraction (thus including endophytes as well as
709 surface-associated microbes). Leaves from Costa Rican plants were stored in 95% ethanol, and

710 surface-associated microbes were concentrated in a vacuum centrifuge and resuspended in
711 molecular grade water prior to DNA extraction. As this sampling method was not quantitative,
712 we did not perform qPCR on plant samples from Costa Rica and used them only for analyses of
713 microbial composition.

714 Non-lepidopteran animals were sampled using the same procedures outlined above, with
715 five species preserved in ethanol and 19 preserved dry at -20°C (Table S1). With the exception
716 of two dung beetles feeding on herbivore dung, and the insectivorous bat *M. lucifugus*, these
717 species are either predominantly or exclusively herbivorous, although the type of plant matter
718 consumed (leaves, seeds, fruit, pollen, etc.) varies. We extracted DNA from feces for vertebrates
719 and from subsamples of homogenized whole bodies for insects (as some insects house the
720 majority of symbionts in organs outside the gut). By including all tissue from these insects, we
721 may have underestimated bacterial densities in the particular organs where microbes are housed
722 (Fig. 1A).

723

724 ***DNA extraction, PCR and sequencing***

725

726 Following previous studies of insect microbiomes (4, 5, 7), we used the MoBio Powersoil
727 kit to extract DNA (100 µl eluate) from measured amounts of sample material. We then PCR-
728 amplified a portion of the 16S rRNA gene with barcoded 515f/806r primers (8). PCR products
729 were cleaned and normalized (up to 25 ng DNA/sample) using the SequalPrep Normalization kit
730 (Thermo Fisher Scientific), and then sequenced on an Illumina MiSeq. Paired-end sequences of
731 16S rRNA amplicons were merged, quality-filtered, and clustered into operational taxonomic
732 units (“phylotypes”) at the 97% sequence similarity level using UPARSE (9), and classified
733 using the RDP classifier and Greengenes (10, 11) as previously described (12). The
734 representative sequences of phylotypes unclassified at this stage, and mitochondrial rRNA
735 phylotypes (which could be from plant, insect, fungal or other mitochondria) were aligned to the
736 NCBI nonredundant nucleotide database (nt) using BLAST for taxonomic identification. (Many
737 universal 16S primers amplify rRNA genes of chloroplasts and mitochondria as well as bacteria
738 and archaea (13)).

739 As bacterial DNA is ubiquitous in laboratory reagents used for DNA extraction and PCR,
740 and especially problematic with low-biomass samples (such as caterpillar feces) (14), we
741 removed contaminants from our samples using information from the 22 DNA extraction blanks
742 and PCR no-template controls that yielded >100 bacterial sequences. Importantly, phylotypes
743 detected in these blanks are not exclusively composed of reagent contaminants, because they
744 receive some input from sample DNA during laboratory processing (15). As high-biomass
745 samples are both least likely to experience reagent contamination (14), and themselves most
746 likely to be the source of “real” sample phylotypes identified in blanks, they can be used to
747 distinguish between laboratory contaminants and true sample sequences (15). We classified
748 contaminants as phylotypes present at $\geq 1\%$ abundance in one or more blank samples, excepting
749 phylotypes present at $\geq 1\%$ abundance in one or more of the best-amplifying samples (the top
750 third in 16S rRNA gene copy number as measured by qPCR). These 25 phylotypes were
751 removed from the dataset prior to analyses of bacterial abundance and composition (they are
752 retained only in Fig. S1A). This approach does not include other types of contaminants
753 introduced prior to DNA extraction, such as those from human skin. Finally, we note that the
754 high relatedness between microbes commonly present in laboratory reagents (listed in (14)) and

755 those present in soil, water and leaves—all possible genuine microbial inputs to the caterpillar
756 gut—precludes a taxonomy-based approach to removing contaminants.

757

758 ***Sequence Data Analysis***

759

760 All analyses were conducted in R version 3.3.2 and are available in the file
761 “Hammer2017_Rcode.R”. Analyses involving bacterial composition were limited to samples
762 with at least 100 bacterial sequences. To calculate phylotype-level overlap between fecal and
763 plant samples, “phylotypes detected on leaves” are defined as those present at any abundance in
764 any plant sample in our dataset. New England and Arizona fecal samples which lack paired plant
765 samples were excluded from this comparison. In measuring intraspecific beta diversity and core
766 microbiome size in caterpillars and other animals, we excluded species with fewer than three
767 replicate individuals. Further, to be conservative, only caterpillars sampled from the same
768 location, and feeding on the same species of plant were compared. As the number of replicates
769 could affect these metrics, and varied among species, we iterated these analyses over multiple
770 combinations of only three replicates per species.

771

772 ***Quantitative PCR***

773

774 We measured 16S rRNA gene copy number using quantitative PCR with the same
775 primers and DNA extracts as above. Reaction conditions and other details are specified in (16).
776 Each sample was run in triplicate (except 11 non-caterpillar species for which limited DNA was
777 available, which were run singly) and the mean of these technical replicates was used for
778 subsequent analyses. Standard curves were calculated using purified genomic DNA from *E. coli*
779 DH10B, which has seven 16S rRNA operons per genome (17). The median copy number of 31
780 qPCR'd DNA extraction blanks was subtracted from sample copy numbers. Resulting counts of
781 total 16S rRNA genes in samples were then multiplied by the proportion of noncontaminant
782 bacterial sequences identified from the same DNA extract, resulting in estimates of bacterial 16S
783 copy numbers.

784 It is unlikely that the low amplification we found in caterpillar samples results from
785 primer bias against abundant bacterial taxa. First, these primers successfully amplified bacteria in
786 non-lepidopteran animals, even when in some cases (such as aphids, (18)), the dominant
787 symbiont has been strictly vertically transmitted between hosts for tens of millions of years.
788 Even in this case, divergence from free-living relatives has not been so great that its 16S rRNA
789 gene is un-amplifiable using 515f/806r primers. Second, the caterpillar gut-associated microbial
790 taxa we found are similar to those reported as being relatively (i.e., in terms of the proportion of
791 sequence libraries) abundant in metagenomic surveys (19, 20) and amplicon-based studies using
792 different 16S rRNA-targeting primer pairs (e.g., (21–25)).

793 To estimate the relationship between body size and whole-animal microbial loads (Fig.
794 S2), we combined published data from (26) with body mass data we calculated directly or
795 derived from other studies (see supplemental file “Body_mass_data.txt”). To restrict the
796 allometric scaling relationship for noncaterpillar animals to those species likely to harbor
797 resident microbiomes, we removed species that had bacterial densities $< 1/100^{\text{th}}$ of the group
798 median. These species were the goose *Branta bernicla*, the bat *Myotis lucifugus*, and the dung
799 beetle *Geotrupes stercorosus*. The body size of two *M. sexta* individuals from Arizona was not
800 recorded and so we substituted the median from other *M. sexta*. Furthermore, as we only had

801 direct gut mass measurements for *M. sexta* (30-40% of body mass), for species sampled using
802 feces (including *M. sexta*) we calculated total microbial loads by multiplying 16S rRNA gene
803 density in feces by body mass. This procedure is likely to have slightly overestimated the
804 microbial load for these species. Despite the numerous methodological uncertainties, microbial
805 counts from (26) and our qPCR-based data, and their allometric scaling relationship with body
806 size (excepting *M. sexta*) were remarkably similar (compare solid and dashed line in Fig. S2).

807

808 ***PCR inhibition assays***

809

810 To examine whether low 16S rRNA gene copy number estimates in caterpillar samples
811 are an artifact of caterpillar-specific PCR inhibitors, we used two distinct approaches. First, we
812 tested whether diluting extracted DNA improves PCR amplification by minimizing inhibitor
813 effects (27). However, 1:10 and 1:100 dilutions of fecal DNA from eight *M. sexta* individuals did
814 not have this effect (Fig. S5A). Second, we individually added the twelve lowest-amplifying
815 caterpillar fecal samples—which might be especially likely to contain PCR inhibitors—to qPCR
816 reactions with targeted primers and a template highly unlikely to be present in caterpillar feces
817 (rDNA ITS region of *Batrachochytrium dendrobatidis* strain JEL270, a chytrid fungus
818 pathogenic to amphibians). As compared to replicate reactions with pure molecular-grade water,
819 adding caterpillar fecal DNA reduced amplification of *B. dendrobatidis* rDNA by 7.4% (Fig.
820 S5B). This inhibition effect, which is also present in feces of humans (27) and likely many other
821 species, is miniscule relative to the difference in bacterial loads between caterpillars and non-
822 lepidopterans spanning multiple orders of magnitude (Fig. 1A). Therefore, the relatively low
823 PCR amplification of 16S rRNA genes from caterpillar feces is most likely due to low microbial
824 biomass rather than high PCR-inhibitory substances.

825

826 ***Additional information on the antibiotic experiment***

827

828 *M. sexta* larval feces production was measured by collecting, drying (50°C for 24 hours),
829 and weighing all fecal pellets in the final instar. To culture bacteria, we plated a dilution series
830 (in sterilized phosphate-buffered saline) of weighed (10-20 mg) subsamples of feces, incubated
831 in aerobic conditions at 37°C. After 24 hours, visible colonies were counted and then, if present,
832 collected *en masse* from the agar surface for sequencing using a sterile swab. This plate-scrape
833 method produces a list of the most abundant bacterial phylotypes potentially culturable using our
834 approach. It should be noted that the presence of fecal bacteria in culture demonstrates that these
835 taxa were viable, but not necessarily growing or metabolically active, while in the caterpillar gut.

836

837 ***Comparison of biomass estimates and evidence of extracellular DNA***

838

839 Among *M. sexta* fecal samples collected during the antibiotic experiment, we found that
840 qPCR-estimated bacterial abundances were correlated with the number of cultured bacterial
841 colonies (see Results; Fig. S3B). Eleven individuals' fecal pellets did not produce any bacterial
842 colonies whatsoever, but did contain measurable levels of DNA (Fig. S3B), and excluding these
843 “zero-colony” samples yielded a stronger association between bacterial colony counts and 16S
844 rRNA gene copy number ($r = 0.51$, $p = 0.0002$). This result could stem from the presence of
845 bacteria that cannot grow aerobically or on LB. Alternatively, it may be due to PCR
846 amplification of extracellular DNA or DNA from dead or otherwise nonviable cells (16). To

847 evaluate these possibilities, we compared the phylotypes (identified by 16S rRNA gene
848 sequencing) in zero-colony fecal samples to those from other samples that did yield colonies, in
849 which bacterial biomass was swabbed directly from the agar surface and sequenced. Most of the
850 16S rRNA gene sequences in the zero-colony fecal samples (median 84%, interquartile range:
851 74-95%) belong to phylotypes cultured from other samples, suggesting that qPCR may have
852 overestimated viable bacterial loads by amplifying DNA from lysed or nonviable cells. If the
853 fraction of the gut microbiome originating from dead or nonviable cells is disproportionately
854 high in caterpillars in general (e.g., due to their digestive physiology – see Discussion), then the
855 difference in living, active microbial biomass between caterpillars and other animals (Fig. 1A)
856 may have been underestimated.

857

858 **References for Supplemental Methods**

859

- 860 1. Janzen DH, Hallwachs W (2011) Joining inventory by parataxonomists with DNA
861 barcoding of a large complex tropical conserved wildland in northwestern Costa Rica.
862 *PLoS One* 6(8):e18123.
- 863 2. Janzen DH, et al. (2009) Integration of DNA barcoding into an ongoing inventory of
864 complex tropical biodiversity. *Mol Ecol Resour* 9:1–26.
- 865 3. Janzen DH, Hallwachs W (2016) DNA barcoding the Lepidoptera inventory of a large
866 complex tropical conserved wildland, Area de Conservacion Guanacaste, northwestern
867 Costa Rica. *Genome* 59(9):641–660.
- 868 4. Hammer TJ, McMillan WO, Fierer N (2014) Metamorphosis of a butterfly-associated
869 bacterial community. *PLoS One* 9(1):e86995.
- 870 5. Hammer TJ, Dickerson JC, Fierer N (2015) Evidence-based recommendations on storing
871 and handling specimens for analyses of insect microbiota. *PeerJ* 3:e1190.
- 872 6. Espinosa-Garcia FJ, Langenheim JH (1990) The endophytic fungal community in leaves
873 of a coastal redwood population diversity and spatial patterns. *New Phytol* 116(1):89–97.
- 874 7. Hammer TJ, et al. (2016) Treating cattle with antibiotics affects greenhouse gas
875 emissions, and microbiota in dung and dung beetles. *Proc R Soc B* 283(1831):20160150.
- 876 8. Caporaso JG, et al. (2012) Ultra-high-throughput microbial community analysis on the
877 Illumina HiSeq and MiSeq platforms. *ISME J* 6(8):1621–1624.
- 878 9. Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon
879 reads. *Nat Methods* 10(10):996–8.
- 880 10. Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid
881 assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol*
882 73(16):5261–5267.
- 883 11. McDonald D, et al. (2012) An improved Greengenes taxonomy with explicit ranks for
884 ecological and evolutionary analyses of bacteria and archaea. *ISME J* 6(3):610–618.
- 885 12. Ramirez KS, et al. (2014) Biogeographic patterns in below-ground diversity in New York
886 City’s Central Park are similar to those observed globally. *Proc R Soc B Biol Sci*
887 281:20141988.
- 888 13. Rastogi G, Tech JJ, Coaker GL, Leveau JHJ (2010) A PCR-based toolbox for the culture-
889 independent quantification of total bacterial abundances in plant environments. *J*
890 *Microbiol Methods* 83(2):127–132.
- 891 14. Salter SJ, et al. (2014) Reagent and laboratory contamination can critically impact
892 sequence-based microbiome analyses. *BMC Biol* 12(1):87.

- 893 15. Lazarevic V, Gaña N, Girard M, Schrenzel J (2016) Decontamination of 16S rRNA gene
894 amplicon sequence datasets based on bacterial load assessment by qPCR. *BMC Microbiol*
895 16(1):73.
- 896 16. Carini P, et al. (2016) Relic DNA is abundant in soil and obscures estimates of soil
897 microbial diversity. *Nat Microbiol* 53(9):680840.
- 898 17. Durfee T, et al. (2008) The complete genome sequence of *Escherichia coli* DH10B:
899 Insights into the biology of a laboratory workhorse. *J Bacteriol* 190(7):2597–2606.
- 900 18. Clark MA, Moran NA, Baumann P (1999) Sequence evolution in bacterial endosymbionts
901 having extreme base compositions. *Mol Biol Evol* 16:1586–1598.
- 902 19. Belda E, et al. (2011) Microbial diversity in the midguts of field and lab-reared
903 populations of the European corn borer *Ostrinia nubilalis*. *PLoS One* 6(6):e21751.
- 904 20. Xia X, et al. (2017) Metagenomic sequencing of diamondback moth gut microbiome
905 unveils key holobiont adaptations for herbivory. *Front Microbiol*.
906 doi:10.3389/fmicb.2017.00663.
- 907 21. Pinto-Tomás AA, et al. (2011) Comparison of midgut bacterial diversity in tropical
908 caterpillars (Lepidoptera: Saturniidae) fed on different diets. *Environ Entomol*
909 40(5):1111–1122.
- 910 22. Broderick NA, Raffa KF, Goodman RM, Handelsman J (2004) Census of the bacterial
911 community of the gypsy moth larval midgut by using culturing and culture-independent
912 methods. *Appl Environ Microbiol* 70(1):293–300.
- 913 23. Brinkmann N, Martens R, Tebbe CC (2008) Origin and diversity of metabolically active
914 gut bacteria from laboratory-bred larvae of *Manduca sexta* (Sphingidae, Lepidoptera,
915 Insecta). *Appl Environ Microbiol* 74(23):7189–7196.
- 916 24. Staudacher H, et al. (2016) Variability of bacterial communities in the moth *Heliothis*
917 *virescens* indicates transient association with the host. *PLoS One* 11(5):1–21.
- 918 25. Mason CJ, Raffa KF (2014) Acquisition and structuring of midgut bacterial communities
919 in gypsy moth (Lepidoptera: Erebidae) larvae. *Environ Entomol* 43(3):595–604.
- 920 26. Kieft TL, Simmons KA (2015) Allometry of animal-microbe interactions and global
921 census of animal-associated microbes. *Proc R Soc B* 282:20150702.
- 922 27. Nechvatal JM, et al. (2008) Fecal collection, ambient preservation, and DNA extraction
923 for PCR amplification of bacterial and human markers from human feces. *J Microbiol*
924 *Methods* 72(2):124–132.