A holobiont view of island biogeography: unraveling patterns driving the nascent diversification of a Hawaiian spider and its microbial associates

Ellie E. Armstrong*[,1], Benoît Perez-Lamarque*[,2,3], Ke Bi[4,5,6], Cerise Chen[7,8], Leontine E. Becking[9,10], Jun Ying Lim[11], Tyler Linderoth[12], Henrik Krehenwinkel[7,13], Rosemary Gillespie[7]

[1] Department of Biology, Stanford University, Stanford, CA, USA
[2] Institut de Biologie de l'ENS (IBENS), Département de biologie, École normale supérieure, CNRS, INSERM, Université PSL, Paris, France
[3] Institut de Systématique, Évolution, Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, UA, Paris, France
[4] Computational Genomics Resource Laboratory, California Institute for Quantitative Biosciences, University of California, Berkeley, CA, USA 94720
[5] Museum of Vertebrate Zoology, University of California, Berkeley, CA, USA 94720
[6] Ancestry, 153 Townsend St., Ste. 800 San Francisco, CA, USA 94107
[7] Department of Environmental Science, Policy and Management, University of California, Berkeley, CA, USA
[8] Long Marine Laboratory, University of California, Santa Cruz, CA, USA
[9] Marine Animal Ecology Group, Wageningen University & Research, Wageningen, The Netherlands
[10] Wageningen Marine Research, Den Helder, The Netherlands
[11] School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 637551
[12] Department of Genetics, University of Cambridge, UK
[13] Department of Biogeography, Trier University, Trier, Germany

* Contributed equally

Corresponding Author: elliea@stanford.edu, benoit.perez@ens.psl.eu

Abstract (250 words)

The diversification of a host lineage can be influenced by both the external environment and its assemblage of microbes. Here, we use a young lineage of spiders, distributed along a chronologically arranged series of volcanic mountains, to determine the parallels between the evolutionary histories of the host spiders and their associated microbial communities, together forming the "holobiont". Using the stick spider *Ariamnes waikula* (Araneae, Theridiidae) on the island of Hawai'i, and outgroup taxa on older islands, we tested whether each component of the holobiont (the spider hosts, the intracellular endosymbionts, and the gut microbial communities) showed correlated signatures of diversity due to sequential colonization from older to younger volcanoes. In order to investigate this, we generated ddRAD data for the host spiders and 16S rRNA gene amplicon data from their microbiota. We expected sequential colonizations to result in a (phylo)genetic structuring of the host spiders and in a diversity gradient in microbial communities. Results showed that the host *A. waikula* is indeed structured by geographic isolation, suggesting sequential colonization from older to younger volcanoes. Similarly, the endosymbiont communities were markedly different between *Ariamnes* species on different islands, but more homogeneous among *A. waikula* populations on the island of Hawai'i. Conversely, the gut microbiota was largely conserved across all populations and species, which we suspect are generally environmentally derived. Our results highlight that the different components of the holobiont have responded in distinct ways to the dynamic environment of the volcanic archipelago, showing the necessity of understanding the interplay between components to better characterize holobiont evolution.

Keywords: Host-associated microbes, endosymbiont, speciation, population structure, adaptive radiation, Ariamnes, Hawaiian Islands

53 **Introduction**

54 Patterns of biodiversity are influenced by both ecological and evolutionary processes
55 operating within the dynamic context of a community (Weber et al. 2017). The external
56 environment can serve to isolate populations for various periods, and select for traits that influence
57 their evolutionary trajectory. At the same time, a given organism also represents a community
58 itself by hosting a diverse array of microbial species, many of which perform essential functions
59 for their host (Koskella & Vos, 2015). The importance of microbial communities for promoting the
60 isolation of their hosts (Bordenstein et al. 2001) and facilitating their adaptation to novel ecological
61 niches (O'Connor et al. 2014) has been increasingly recognized. A species' response to dynamic
62 changes in the environment can frequently be modulated by the "holobiont" of host and microbial
63 associates (Margulis & Fester 1991, Moran et al. 2019). In some cases, this holobiont may even
64 act as a unit of selection (Guerrero et al. 2013). Therefore, understanding whether host organisms
65 and their associated microbial communities respond in a similar manner to external forces, is
66 essential for identifying potential drivers of evolution (McFall-Ngai et al. 2013).

67 Among arthropods, associated microbial communities are often highly diverse
68 assemblages, accounting for an extensive range of interactions with their host (Engel & Moran
69 2013). Many arthropods host different microbial communities in different body parts that occupy
70 various niches such as the gut microbiota or intracellular endosymbionts (Hansen & Moran 2014).
71 The composition of the gut microbiota is often determined by complex interactions of environment,
72 diet, developmental stage, and host evolutionary history (Yun et al. 2014), contributing to various
73 functions for the host ranging from nutrition to protection against pathogens (Engel & Moran 2013,
74 Moran et al. 2019, Hammer et al. 2019). However, for some arthropod taxa, recent work also
75 suggests that a large proportion of the gut microbiota is purely environmentally derived, highly
76 transient, and does not always have apparent functional relevance (Hammer et al. 2017; Kennedy
77 et al. 2020). In contrast, a functional reliance of the host on its microbial communities could result

in more stable and predictable gut microbial communities, that may be conserved and co-evolve with their host (Engel & Moran 2013, Kwong et al. 2017).

In contrast to the gut microbiota, most endosymbionts are vertically-transmitted intracellular bacteria. They can comprise tightly coevolved taxa, supplying their host with essential nutrients, such as bacteria of the genus *Buchnera* in aphids (Koga et al. 2003). Many other endosymbionts manipulate the reproduction of their host, such as species in the genera *Wolbachia*, *Rickettsia*, *Rickettsiella,* and *Cardinium* (Duron et al. 2017; Hoy & Jeyaprakash 2005; Vanthournout & Hendrickx 2015; White et al. 2020; Zhang et al. 2017). These taxa can promote cytoplasmic incompatibilities between hosts, enhancing genetic isolation (Shropshire et al. 2020). Some endosymbionts have also been associated with behavioral tendencies for long-distance dispersal (Goodacre et al. 2006), translating into higher gene flow and suppressing divergence between lineages otherwise isolated by distance.

Considering this background, a key point of interest is dissecting to what extent the nature of the host-microbe relationship affects the assembly and composition of host-associated microbial communities as the host adapts and evolves in its environment. Here, we examine the assembly and composition of different microbial communities within an evolving host by focusing on a lineage of spiders that shows recent divergence between populations on the youngest island of Hawai‘i (Gillespie et al. 2018).

The Hawaiian archipelago is comprised of a chain of volcanic islands that increase in age from southeast to northwest, providing an ideal system for tracking the interplay between a host and the different components of its microbial community within an isolated setting (Shaw & Gillespie 2016). Even on the youngest island of Hawai‘i, the volcanoes are arranged chronologically, from approximately 430,000 years old (Kohala), to the active flows of Kīlauea, with population structure of local flora and fauna generally shaped by progressive colonization of the newly emerged volcanoes (e.g. Simon 1987; Wagner & Funk, 1995; Blankers et al. 2018; Eldon et al. 2019; Goodman et al. 2019). Stick-spiders in the genus *Ariamnes* (Theridiidae) have

104 diversified rapidly across the landscapes in the Hawaiian archipelago (Gillespie & Rivera 2007)
105 and exhibit repeated diversification into ecomorphs adapted to specific microhabitats such as the
106 "gold" ecomorph that is are entirely cryptic on the underside of leaves (Gillespie & Rivera 2007;
107 Gillespie *et al.* 2018). However, their diet is conserved and they are specialized consumers of
108 other spiders (Kennedy et al. 2018). Compared to their mainland counterparts in which dispersal
109 by ballooning has been well documented (e.g. Larrivee & Buddle 2011), Hawaiian *Ariamnes* (as
110 with other Hawaiian spiders) show little propensity for dispersal; such a loss of dispersal, and
111 associated high levels of local endemism, are a common feature of organisms on remote islands
112 (Gillespie et al 2012).

113 The current study focuses on a single species of the Hawaiian *Ariamnes* (*A. waikula*),
114 endemic to the youngest island of Hawai'i. We aim to determine whether this highly specialized
115 spider lineage shows parallel signatures of recurrent colonization events across volcanoes within
116 the island in the same way as its microbial associates. We hypothesize that the population
117 structure of *A. waikula* reflects a stepping stone colonization from older to younger volcanoes, and
118 that populations from geologically older sites will show higher differentiation and higher within
119 population diversity compared to younger sites, as they generate novel diversity subsequent to
120 the founder event(s) (Roderick et al. 2012). For the microbial communities within the spider hosts,
121 patterns of composition may or may not reflect the patterns of host differentiation depending on
122 the nature of the host/microbe relationship. First, we predict that microbial associates that are
123 vertically transmitted, including the intracellular endosymbionts, (i) will show patterns of community
124 composition that closely mirror the population structure of the host (i.e. strong correlation between
125 the host genetic structure and the microbial beta diversities) and (ii) will present lower microbial
126 alpha diversities at the younger sites due to stochastic losses of some microbial symbionts during
127 colonization events (Minard et al. 2015). Second, if microbial associates are mainly
128 environmentally acquired and filtered by phylogenetically-conserved host traits (Moran & Sloan
129 2015), we can expect a correlation between the host genetic structure and the microbial beta

130 diversities (Mazel et al. 2018) and we do not expect any reduction of the microbial alpha diversities
131 at the younger sites. Third, if host-associated microbial communities are passively assembled
132 without any filtering by the host, we expect no particular patterns in terms of microbial alpha
133 diversity or relating host genetic structure and the microbial beta diversities. In the latter case,
134 microbial beta diversities might rather be linked to geographical distances or other environmental
135 factors.

136 To test these predictions, we examined the population genetic structure of *A. waikula* on
137 Hawai‘i Island, along with several outgroup species from other islands, using genome-wide single
138 nucleotide polymorphism (SNP) data generated with double digest RAD sequencing (ddRAD). We
139 then investigated how different components of their microbiota have changed as the spiders
140 colonized new locations. To do so, we compared the community structure of microbial populations
141 to that of their host individual using 16S rRNA gene amplicon sequencing, capturing the diversity
142 of both the endosymbionts and the gut microbiota. Finally, we complemented these analyses with
143 simulations to assess the robustness of our findings.

144 **Materials and Methods**

145 *Sampling*

146 We sampled *Ariamnes* across Hawai'i Island, focusing on individuals of *A. waikula* (gold
147 ecomorph), from 6 populations. In addition, we included 2 individuals of the related *A. hiwa* (brown
148 ecomorph) and 32 gold ecomorph individuals from two other species: *A. melekalikimaka* on West
149 Maui and *A. n. sp.* Moloka'i (Gillespie et al. 2018). Using *A. hiwa*, *A. melekalikimaka*, and *A. n. sp.*
150 Moloka'i we were able to confirm monophyly of the Hawai'i Island clade and compare the diversity
151 of the microbial communities of other species between and within islands. Sampling sites were all
152 located within the wet forest (rainfall 2500-3000mm) and between elevations of 1000m and
153 1300m, with the exception of the "Saddle" population which was slightly higher (Table S1). Our
154 sites were all chosen to standardize variables (temperature, precipitation, elevation) to the extent
155 possible. Individuals were collected by hand and immediately preserved in 90% EtOH. We
156 collected a total of 133 individuals for sequencing (Supplementary Tables S1 & S2). Only adults
157 were collected for this study, to decrease the likelihood of capturing differences driven by age.

158 *Ariamnes - ddRAD library preparation and sequencing*

159 To examine the population structure of the spider host, we used ddRAD to obtain reduced
160 representation genome-wide SNP data. Genomic DNA was extracted from spider legs with several
161 modifications to the Qiagen DNeasy kit protocol. Legs were first removed from each specimen
162 using sterile tweezers so that the abdomen remained intact for the microbial DNA analysis. DNA
163 was then extracted by placing the tissue in Proteinase K and lysis buffer and grinding them with a
164 sterile pestle to break up the exoskeleton. We then added 4uL RNase A (100mg/ml) and incubated
165 the extractions for two minutes at room temperature. Tubes with tissue and extraction solution
166 were then placed overnight in a heat block at 56°C. The remainder of the extraction protocol was
167 performed following the manufacturer's instructions. We built ddRAD libraries following an

168 adapted protocol of Peterson *et al.* (2012) (Saarman & Pogson, 2015; see Maas et al. 2018 for
169 protocol optimization steps). Briefly, we started the ddRAD protocol with a total of 100 nanograms
170 of DNA per sample. The DNA was digested using SphI-HF (rare-cutting) and MluCI (frequent-
171 cutting) restriction enzymes. We assessed fragmentation with a Bioanalyzer High Sensitivity chip
172 (Agilent). We multiplexed 15-20 individuals per library for a total of eight ddRAD libraries, using
173 oligo sequences of 42bp including a barcode of 5bp (Peterson *et al.* 2012). We used a Sage
174 Science Pippen Prep to size select 451-551bp (including internal adapters) fragments, and
175 confirmed the sizes using a Bioanalyzer. Ten indexing polymerase chain reaction (PCR) cycle
176 were run on each library to enrich for double-digested fragments and to incorporate a unique
177 external index for each library pool. The eight libraries were sequenced using 100bp paired-end
178 sequencing on one Illumina HiSeq2500 lane at the Vincent J. Coates Genomic Sequencing
179 Facility at UC Berkeley.

180 *Ariamnes - ddRAD data filtering and processing*

181 We used a custom perl script invoking a variety of external programs to filter and process
182 the ddRAD data (RADTOOLKITv0.13.10; https://github.com/CGRL-QB3-UCBerkeley/RAD). Briefly,
183 raw fastq reads were first de-multiplexed based on the sequence composition of internal barcodes
184 with tolerance of one mismatch. De-multiplexed reads were removed if the expected cutting site
185 was not found at the beginning of the 5'-end of the sequences. The reads were then filtered using
186 cutadapt (Martin 2011) and Trimmomatic using default parameters (Bolger et al. 2014) to trim off
187 Illumina adapter contaminations and low-quality reads. The resulting cleaned forward reads of
188 each individual were first clustered using cd-hit with 97% as the clustering threshold (Fu et al.
189 2012; Li & Godzik 2006), keeping only those clusters with at least two reads. In each cluster, we
190 pulled out the corresponding reverse reads based on the identifiers. Both forward and reverse
191 clusters were kept only if the corresponding reverse reads also formed one cluster. If reverse
192 reads were grouped into more than one cluster, then only the forward read cluster was kept. For

193 each paired cluster, the representative sequences for each forward and reverse cluster
194 determined by cd-hit were retained. We then merged the forward and reverse sequences using
195 FLASH (Magoč & Salzberg 2011). If they could not be merged then they were joined by placing
196 "N"s between the two sequences. The resulting loci were then masked for putative repetitive and
197 low-complexity elements, and short repeats using RepeatMasker (Smit et al. 2004) with "spider"
198 as the database. After masking, we eliminated loci if more than 60% of the nucleotides were Ns.
199 The resulting ddRAD loci from each individual were combined and clustered for all individuals.
200 Contigs that were at least 40 nucleotides in length and shared by at least 60% of all the individuals
201 served as a reference. Cleaned sequencing reads from each individual were then aligned to the
202 reference using Novoalign (http://www.novocraft.com), discarding any ambiguously mapped
203 reads. We used Picard (http://picard.sourceforge.net) to add read groups and GATK (McKenna et
204 al. 2010) to perform realignment around indels in BAM format generated by SAMtools (Li et al.
205 2009). We then used SAMtools/BCFtools (Li et al. 2009) to generate data quality control
206 information in VCF format. These data were then further filtered using a custom perl script,
207 SNPcleaner (Bi et al. 2013). We filtered out any loci with more than two called alleles. We masked
208 sites within 10 bp upstream and downstream of an indel. We discarded sites with a total depth
209 outside of the genome-wide 1st and 99th percentile. To avoid excessive heterogeneity in sample
210 representation among sites, we also removed alignments if more than 40% of the samples had
211 less than 3X coverage.

212 Since most of the individuals were expected to have low coverage (between 5-15x) we
213 used ANGSD (Korneliussen et al. 2014) to calculate genotype likelihoods and genotypes for the
214 analyses. This tool was specifically developed for population and evolutionary genomic analyses
215 of low coverage data. We used genotype likelihoods whenever the downstream tools allowed us
216 to, randomness in allele sampling as well as sequencing or mapping errors lead to uncertainty in
217 genotype calling (Crawford & Lazzaro 2012)).

218 *Ariamnes - Phylogenetic Analyses*

219 We investigated the phylogenetic relationships between populations of *A. waikula* and the
220 other *Ariamnes* species to better understand colonization patterns. To do so, we used the Stacks
221 pipelines (Catchen et al. 2013) to group reads into homologous loci across all individuals and
222 extract phylogenetically informative sites (i.e. fixed within individuals but variable between
223 individuals). This yielded an alignment of 58,899 sites with which we performed phylogenetic
224 reconstruction using IQ-TREE (Nguyen et al. 2015) by combining model selection from
225 ModelFinder Plus (Kalyaanamoorthy et al. 2017) and assessing branch supports with 1,000
226 ultrafast bootstrap (Hoang et al. 2018). Finally, we rooted the tree using the *A. hiwa* individuals
227 and calibrated it with r8s (Sanderson 2003) without specifying any absolute dating (*i.e.* setting the
228 root age to 1).

229 *A. waikula - Population Genetic Analyses*

230 We first explored *A. waikula* population structure using Principal Components Analysis
231 (PCA). PCA, as opposed to structure analyses which force data into a specific number of clusters,
232 is an unsupervised method which attempts to find orthogonal directions of maximum variance
233 through a cloud of data. Therefore, it is a useful way to explore data structure prior to using
234 supervised methods for inferring population structure. We used the previously generated genotype
235 likelihoods to calculate genotype posterior probabilities under an allele frequvvency prior (-doPost
236 1) in binary format (-doGeno 32) for use with ngsCovar (part of the ngsTools package; (Fumagalli
237 et al. 2014). We used ngsCovar to calculate a genetic covariance matrix among individuals from
238 the genotype probabilities at sites having a minor allele frequency of 5% to avoid noise from very
239 rare alleles (which could be due to sequencing errors). Comparisons between the first three
240 principal component axes were then plotted using R (R Core Team, 2020). Pairwise F_{ST} values
241 were calculated in ANGSD from allele frequency likelihoods (-doSaf 1) and the respective pair's

genome-wide, unfolded, joint site frequency spectrum (SFS) as a prior for jointly observing any combination of allele frequencies between the two populations. We used unfolded allele frequencies by supplying the reference sequences as a pseudo-ancestral sequence since ANGSDv0.930 is only able to accurately estimate F_{ST} using unfolded data. In order to visualize F_{ST} distances, we performed multidimensional scaling (to 2 dimensions) with R (R Core Team, 2020; using the *cmdscale* function). Last, we estimated pairwise genetic distances calculated from genotype likelihoods using the software ngsDist (Viera et al. 2015) and balanced minimum evolution using FastME (Lefort et al. 2015). To do so, we first calculated genotype likelihoods using ANGSD (see above), which were used to construct a pairwise distance matrix between all individuals using ngsDist.

A crucial factor affecting population structure is gene flow between populations. In order to determine levels of connectivity between populations of *A. waikula*, we used two different approaches: ngsAdmix (Skotte *et al.* 2013) and EEMS (Petkova et al. 2016). ngsAdmix is a genotype likelihood based-tool for estimating individual admixture proportions, while EEMS uses genotype calls to infer effective migration surfaces. We inferred ancestry proportions indicative of admixture for different values of K (number of ancestral populations; ranging from two to six, which is the number of populations sampled) using R for visualization. We then generated genotype calls for the EEMS analysis using ANGSD with the following flags: '-doMaf2', '-doGeno 2', '-doPost 1', '-doSaf 1', '-fold 1', a SNP p-value cut-off of 1e-6, and a posterior probability for calling genotypes of 0.75. We only considered sites with a minimum depth of 3. We converted the output into the adegenet (R-package) input format using PopGenTools (https://github.com/CGRL-QB3-UCBerkeley/PopGenTools) to calculate genetic distances between all individuals. EEMS also requires a geographic map of the area in question, so we used the program to divide Hawai'i Island into grids of 10km (Supplementary Fig. S1) and subsequently added the samples to the grid using one GPS coordinate per population (Supplementary Table S1). Using a stepping stone model, we calculated migration rates between the demes. Convergence of the MCMC runs was

268 assessed by plotting and inspecting the traces by eye (Supplementary Fig. S2). We performed 10
269 independent runs for each of the analyses.

270 Next, we calculated population genetic statistics such as nucleotide diversity (Pi),
271 Watterson's Theta, and Tajima's D. To do so, we generated a folded SFS for each population
272 using ANGSD and realSFS (part of the ANGSD package). We then ran ANGSD *-doTheta 1* and -
273 *pest* (which provides the genome-wide SFS prior to ANGSD) and the respective output formats
274 were converted to bed format using *thetaStat make_bed*. Subsequently, we calculated the per
275 population statistics using *thetaStat do_stat*. The average, min, and max Tajima's D were then
276 calculated and we further extracted the genome-wide average Watterson's theta and Pi values.
277 While Watterson's theta and Pi are both measures of genetic diversity, Watterson's theta is based
278 only on the number of segregating sties and so is sensitive to novel, rare diversity, whereas Pi is
279 most influenced by sites with alleles at intermediate frequency more typical ancestral diversity.
280 We measured both Pi and theta, as well as Tajima's D to gain a wholistic understanding of the
281 genetic signatures of colonization events and possible subsequent expansions. We then also used
282 linear models in R to test whether genetic diversity (Pi or Watterson's theta) was positively related
283 to the age of a given volcano.

284 *Microbial Communities*

285 To characterize the microbial community within the *A. waikula* hosts, a subset of 71
286 individuals from the eight populations on Hawai'i and two additional islands (Table 1;
287 Supplementary Tables S1 & S2) were selected for analysis. We focused on the mid and hindgut,
288 both located in the spider's opisthosoma. The preservation in ethanol led to considerable
289 shrinkage of the opisthosoma and thus did not allow us to separately dissect out the gut. Instead,
290 we used the whole opisthosoma to extract DNA, as described in Kennedy et al. (2020). Specimens
291 which did not have the opisthosoma intact were not used. The digestive tract comprises the
292 majority of the opisthosoma's cavity. In addition, it contains silk glands, the heart, lungs, and

293 gonads. The opisthosoma was removed with a sterile razor blade and then washed in ethanol to
294 remove external bacteria (Hammer et al. 2015). Besides the intracellular endosymbionts, we
295 considered the remaining bacteria to be representative of the “gut microbiota”, even if it technically
296 consists of the “opisthosoma microbiota”, but previous studies have shown that the gut microbiota
297 dominates in the opisthosoma (Sheffer et al. 2020, Kennedy et al. 2020). The tissue was then
298 transferred into lysis buffer and finely ground with a sterile pestle. DNA was extracted using the
299 Gentra Puregene Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer’s protocol.
300 Spider abdominal tissue can contain PCR inhibitors (Schrader et al. 2012), thus we cleaned the
301 DNA extract with 0.9X AmPure Beads XP.

302 We next amplified a ~300 bp fragment of the V1-V2 region of the bacterial 16S rRNA using
303 the Qiagen Multiplex PCR kit according to the manufacturer’s protocols and using the primer pair
304 MS-27F (AGAGTTTGATCCTGGCTCAG) and MS-338R (TGCTGCCTCCCGTAGGAGT) (Gibson
305 et al. 2014). PCRs were run with 20ng of template DNA and 30 cycles at an annealing temperature
306 of 55°C. PCR products were separated from leftover primer by 1X AmPure Beads XP. A six-cycle
307 indexing PCR was performed on the cleaned products, adding dual indexes to every sample using
308 the Qiagen Multiplex PCR kit. Indexing was performed according to (Lange et al. 2014). The dual
309 indexed libraries were isolated from leftover primer as described above, quantified using a Qubit
310 fluorometer, and pooled in equal amounts into a single tube. The library was sequenced on an
311 Illumina MiSeq using V3 chemistry and 300 bp paired reads. In order to discard contaminants from
312 our final dataset, we also performed blank extraction controls and negative PCR controls (without
313 DNA template), which were sequenced along the other samples.

314 We used the 16S profiling analysis pipeline for Illumina paired-end sequences of the
315 Brazilian Microbiome Project (Pylro et al. 2014), including QIIME 1.8.0 (Caporaso et al. 2010) and
316 USEARCH 11 (Edgar 2013). We modified some steps of these pipelines, using our own Bash and
317 R scripts (R Core Team, 2020; see Data Accessibility section). The QIIME script
318 join_paired_ends.py was used to merge paired reads. The fastq_filter command in USEARCH

319 was used for quality filtering the assemblies (below a base calling error probability of 0.5, using
320 an average Q-score for each read). We used the Stream EDitor in UNIX to remove PCR primers
321 from all assembled sequences. Sequences were de-replicated using USEARCH, removing all
322 singletons. OTUs were generated at a similarity cutoff of 3% or 0% (0 radius OTUs, or "Z-OTU")
323 from the de-replicated sequences and chimera removed de novo using USEARCH. The following
324 analyses were thus independently applied on the two distinct sets of OTUs. We assigned
325 taxonomy to the resulting OTUs using the assign_taxonomy.py script based on the Greengenes
326 database (http://greengenes.secondgenome.com). We removed sequences corresponding to
327 OTUs found in high prevalence and abundance in the different negative controls from all samples,
328 as these could represent contaminants (Hornung et al. 2019).

329 Spiders are known to carry various endosymbiotic bacteria (Goodacre et al. 2006,
330 Vanthournout et al. 2015, White et al. 2020). These can be vastly overrepresented in microbial
331 analyses and thus may completely dominate the microbial community structure. Since we did not
332 extract the gut from individuals, endosymbionts from outside of the gut could be particularly
333 prevalent in our analysis. We thus separated known endosymbionts in our samples (*i.e. Rickettsia,*
334 *Rickettsiella, Wolbachia,* and *Canditatus Cardinium*) from the remaining bacterial sequences
335 (referred to as the "gut microbiota"), resulting in two OTU sequence files. Both of these datasets
336 were used for the following analyses separately.

337 OTU tables were prepared by mapping sequences back to the filtered OTU sequence files
338 using USEARCH. The OTU tables were rarefied to an even coverage (from 400 to 8,000 reads
339 with 20 replications per rarefaction depth) using the multiple_rarefactions.py script in QIIME. Given
340 the rarefaction curves (Supplementary Fig. S3), we chose rarefied depths at 3,200, 110, and 3,000
341 reads for the whole microbiota, the endosymbionts, and the gut microbiota respectively and
342 replicated 20 times this rarefaction; these rarefied OTU tables were used in all the following
343 analyses. Rarefying at 110 reads for the endosymbionts enabled us to keep all the samples for
344 the following analyses, but using a higher threshold did not qualitatively affect our results (not

345 shown) given that most samples were dominated by a single endosymbiotic OTU (see Results).
346 We plotted the relative abundances of different microbial taxa at the genus and order levels for
347 the different *Ariamnes* populations.

348 We then investigated whether microbial associates showed similar diversity patterns
349 compared to their *Ariamnes* hosts along the chronosequence. Chao1 indexes of alpha diversity
350 were calculated from the rarefied OTU tables using QIIME (alpha_diversity.py), and to evaluate
351 the presence of a bottleneck in microbial diversity, linear mixed models were used to test for an
352 effect of the volcano age and the genetic diversity of *Ariamnes* host populations (Pi and
353 Watterson's theta) on the bacterial alpha diversity, with sampling site as a random effect. The
354 different variables were log-transformed prior to analyses, and homoscedasticity and normality of
355 the model residuals were verified.

356 To measure microbiota differentiation across *Ariamnes* species and populations, we
357 computed beta diversity between microbial communities of each individual using QIIME
358 (beta_diversity.py with Bray-Curtis dissimilarities). Beta diversity of microbial populations was
359 visualized with a Principal Coordinate Analysis (PCoA) and as dendrograms using a neighbor
360 joining reconstruction with the R-package ape (Paradis et al., 2004). To test whether individuals
361 from the same population tend to host similar bacterial communities across the different islands
362 (*i.e.* different *Ariamnes* species) or within Hawai'i Island only (*A. waikula*), we performed a
363 Permutational analysis of variance (PERMANOVA; *adonis* function, vegan R-package) on the
364 beta diversity matrices with 10,000 permutations, after having verified the homogeneity of the
365 variances (*betadispers* function). Next, we analyzed whether the microbiota of the *Ariamnes*
366 holobiont mirror the host's phylogeny by testing the correlations between microbial Bray-Curtis
367 dissimilarities and host genetic distances (ngsDist distances) or between microbiota differentiation
368 and host phylogenetic distances, using Mantel tests with 10,000 permutations (vegan R-package).
369 These analyses were performed between populations of *A. waikula* in Hawai'i Island and were

370 compared to the analyses performed between populations of different *Ariamnes* species (*i.e.*
371 including *A. melekalikimaka* on West Maui and *A. n. sp.* Moloka‘i). Finally, we tested whether the
372 microbiota differentiation could be explained by geography alone, by performing a Mantel test
373 between the microbial Bray-Curtis dissimilarities and the geographical distances between
374 sampling sites.

375 In addition, we replicated all diversity analyses of the microbial communities using
376 phylogenetic diversity indexes: alpha diversities were measured using Faith's phylogenetic
377 diversity, and beta-diversities were measured using weighted UniFrac distances (Lozupone &
378 Knight, 2005).

379 During all analyses of diversity, we also controlled for batch effects during the PCR steps
380 by assessing the correlations between proximal samples in the PCR plates.

381

382 *Simulations*

383 Our data for the microbial communities relied on 16S rRNA, raising the question as to
384 whether the rate of evolution of this particular molecular marker was sufficient to capture changes
385 at the very short timescales (<2 Myr) involved in the current study. The 16S rRNA gene of
386 symbiotic bacteria is generally assumed to accumulate only 1-2% of sequence divergences per
387 50 Myr on average (Moran et al. 1993; Ochman et al. 1999). To investigate whether such a slow
388 substitution rate is sufficient to detect microbial differentiation on <2Myr timescales we performed
389 simulations. We directly modeled the evolution of the DNA sequences of transmitted OTUs along
390 the *Ariamnes* phylogeny so as to simulate their evolution on a phylogenetic tree independent from
391 the *Ariamnes* tree (following Perez-Lamarque & Morlon 2019). We considered DNA sequences of
392 300 bp, with 10% variable sites, and a substitution rate of 0.001 events per site per million years
393 (such that two sequences would accumulate 1% divergence in 50 Myr on average). The *Ariamnes*
394 microbiota were composed of 100 OTUs with a certain percentage of vertically transmitted OTUs

395 (either 0, 10, 25, 50, 75, or 100%) and the remainder acquired from the environment. For each
396 percentage of transmitted OTUs, we performed 30 simulations of mock microbiota using the
397 function *simulate_alignment* from the R-package HOME (Perez-Lamarque & Morlon 2019). Then,
398 we considered that each OTU had an abundance of 100 and we added a sampling process of
399 50% of the OTUs in each microbiota (representing the sampling hazards of a metabarcoding
400 sequencing). Finally, we computed the Bray-Curtis beta diversities of the mock *Ariamnes*
401 microbiota and tested the presence of microbial differentiation by performing a Mantel test with
402 the host phylogenetic distances.

403 We then tested whether the presence of a bottleneck in the pool of available microbes
404 along the island chronosequence could be detected in the alpha diversities of the *Ariamnes*
405 microbiota. We considered 100 OTUs (each OTU having an abundance of 100) and simulated
406 bottlenecks of the microbial OTUs between the old sites (Moloka‘i and Maui islands) and the
407 younger Hawai‘i Island. We considered different strengths of bottlenecks (from 0% to 50% of the
408 microbial OTUs) and simulated mock microbiota for each *Ariamnes* by randomly picking 100
409 microbes from the local pools. We then tested the relationship between microbial alpha diversities
410 measured using Chao1 index and the volcano age using linear mixed models. Note that here we
411 only considered microbial bottlenecks as the stochastic losses of some microbial OTUs during
412 island colonization, but the signal of the microbial bottlenecks can also be reinforced by the
413 reduction of the genetic diversity within each microbial lineage at colonization events.

414 **Results**

415 *Ariamnes - Phylogenetic Analyses*

416 We obtained approximately 210 million paired-end reads from Illumina sequencing of the
417 ddRAD libraries. After filtering for contaminants and low-quality reads, there were approximately
418 83 million remaining across the 123 demultiplexed samples. From these, we retained a total of
419 2,957,301 sites passing filters out of a possible 7,378,384 across all individuals that were used in
420 downstream analyses and resulting in 123 individuals for population genetic analysis. After quality
421 filtering, the average sequencing depth of samples was 12x (range of 3-42x across 123 samples).
422 Phylogenetic analysis confirms that the *A. melekalikimaka* (West Maui), *A. n. sp.* (Molokaʻi)
423 and *A. waikula* (Hawaiʻi Island) are monophyletic groups (Fig. 1). Among *A. waikula* individuals,
424 the population from Kohala (the oldest volcano on Hawaiʻi Island) is sister to the clade that contains
425 all other individuals. Except for one individual from Saddle and two from Alili, these populations
426 from Saddle (between Mauna Loa and Mauna Kea) and Alili form monophyletic clades, suggesting
427 that there may be infrequent dispersal between some populations. The populations from Olaʻa,
428 Puʻu Makaʻala and Thurston, all sites that are close together in the saddle between the volcanoes
429 of Mauna Loa and Kilauea (MLKS), form one mixed clade.

430 *A. waikula - Population genetics*

431 We analyzed population differentiation and structure for the full data set, and a subset
432 including only *A. waikula* from Hawaiʻi Island. We did not include the two outgroup *A. hiwa*
433 individuals. Divergence (F_{ST}) was highest and similar between the different species on separate
434 islands: 0.52-0.54 for *A. melekalikimaka* (Maui) compared to *A. waikula* (Hawaiʻi Island); 0.46-0.49
435 for *A. n. sp.* (Molokaʻi) compared to *A. waikula (*Hawai'i Island); 0.47 for *A. n. sp.* (Molokai)
436 compared to *A. melekalikimaka* (Maui) (Table S3). These levels of divergence are close to those
437 reported for COI data between *Ariamnes* populations on Maui compared to Hawaii Island

438 (Roderick et al. 2012). *A. waikula* populations on Hawai‘i Island were structured primarily
439 according to locality, with Kohala, Alili, and Saddle being distinct from each other and from the
440 MLKS sites (pairwise F_{ST} range 0.03-0.15; Pu‘u Maka‘ala, Thurston and Ola‘a; Fig. 1A).

441 Clustering analyses (PCA and ngsAdmix) of *A. waikula* broadly showed clustering of
442 groups according to locality/volcano, except at MLKS sites (Fig. 2a, Supplementary Fig. S5). PC1
443 explained approximately 16% of the variation and separated Kohala and the younger sites, placing
444 Saddle in the middle (Fig. 1A). Similarly, admixture analysis showed Kohala first splitting off at
445 K=2. PC2 explained approximately 4% of the variation and primarily separated Alili from the MLKS
446 sites, while PC3 (3%) separates Saddle, and to a lesser degree, Alili from the other populations
447 (Supplementary Fig. S5). Admixture analyses show the same pattern, with individuals from Alili
448 clustering separately at K=3 and Saddle at K=4. At K=6 we see slightly more similar ancestry
449 between individuals from Thurston and Ola‘a, than either of these two site with Pu‘u Maka‘ala (Fig.
450 2B). For the ngsAdmix analyses, we found consistent results from the 50 independent runs for
451 higher values of K for the Hawai‘i only sampling (Supplementary Fig. S4). Lastly, the EEMS
452 analyses with *A. waikula* individuals indicate potential gene flow between the MLKS populations:
453 Pu‘u Maka‘ala, Thurston, and Ola‘a (Fig. 2C).

454 Contrary to expectations, we found that Watterson’s theta was fairly even across localities
455 on the youngest island of Hawai’i (Fig. 3). However, this correlation between Watterson’s theta
456 and volcano age was not significant (F=0.023, p-value = 0.87) for *A. waikula* populations within
457 Hawai’i (Figure 3) and no other correlations were found between theta and volcano age or for
458 *Ariamnes* genetic diversity (Pi or Theta).

459 We also investigated Tajima’s D, which reflects population size changes as a skew in the
460 site frequency spectrum (Supplementary Fig. S6). We found that most sampling sites showed only
461 marginally negative values consistent with population expansion (average: -1.0 to -0.06;
462 Supplementary Table S4).

463

464 *Microbial Communities*

465 We obtained a total of 4,932,236 bacterial reads from the 16S rRNA sequencing. After the
466 quality filtering steps, demultiplexing, and removal of contaminants based on the negative controls
467 (including the genera *Brachybacterium*, *Veillonella*, or *Brevibacterium*), we obtained a total of
468 1,315,469 sequences, which range from 5,000 sequences to 45,000 per individual *Ariamnes*
469 sample. Due to the largely unexplored nature of the microbial communities of Hawaiian
470 arthropods, we could not assign genus level taxonomy to the majority of microbial OTUs (only
471 241OTUs out of the 571 97% OTUs and 742 OTUs out of the 1,357 Z-OTUs), but were able to
472 assign order for the majority of them (528 OTUs 97% and 1,290 Z-OTUs).

473 We did not find any significant differences in the alpha diversities of the whole microbiota
474 among *Ariamnes* populations or species (Supplementary Table S5). Using linear mixed models,
475 we also did not find a significant association of the microbial alpha diversity with volcano age or
476 the genetic diversity of the host population (Fig. 3), even though the alpha diversities of the gut
477 microbiota seemed to be slightly higher for host populations present on old volcanos (Fig. 3D).
478 The results were similar for both Z-OTUs and 97% OTUs.

479 *Intracellular endosymbiont community:* When only selecting the endosymbionts, the
480 community showed considerable variation between the different species of *Ariamnes* (Fig. 1B).
481 *Ariamnes melekalikimaka* from West Maui mostly carry *Wolbachia*, while *Rickettsia* dominate the
482 endosymbiont community of *A. n. sp.* Moloka‘i, and *Rickettsiella* dominates the populations of *A.*
483 *waikula* on Hawai‘i Island (Fig. 4A & Supplementary Fig. S7). This suggests that each spider
484 population was colonized by a single endosymbiont strain (single OTU; Supplementary Fig. S8).
485 The exception to this was several individuals of *A. waikula* spiders from Alili, Ola‘a, and Thurston
486 that appear to have been simultaneously colonized by *Rickettsiella* and *Rickettsia.* We observed
487 a large variability in the relative abundance of the endosymbionts within each sample, from
488 occupying the majority of the microbial reads (in most samples) to less than 5% of the reads in a
489 few samples (Supplementary Fig. S8). The pronounced turnover of endosymbiont taxa between

490 host populations was mirrored in the Principal coordinate analyses (PCoA): Individuals of the
491 different species from different islands formed separated clusters in PCoA plots and cluster
492 dendrograms for the endosymbiont communities (Fig. 1B, Supplementary Figs. S9 & S10). PCoA
493 plots and cluster dendrograms exhibit a clear pattern of phylosymbiosis, where the divergences
494 between microbiota reflect the phylogenetic distances between the *Ariamnes* hosts. Mantel tests
495 showed a significant correlation between the endosymbiont community beta diversity and the
496 *Ariamnes* (phylo)genetic distances ($R>0.6$, $P<0.001$; Table 2 & Supplementary Table S7). Mantel
497 tests similarly showed a significant correlation between the endosymbiont community beta
498 diversity and the geographical distances ($R=0.56$, $p<0.001$, Table 2 & Supplementary Table S7),
499 but their correlation coefficients were generally lower that those found with the *Ariamnes*
500 (phylo)genetic distances, suggesting that the different *Ariamnes* host lineages better explain the
501 endosymbiont variation than the geography alone. Within Hawai‘i Island, endosymbiont
502 communities of *A. waikula* from the same population tend to be more similar than between
503 populations (Supplementary Table S6; *e.g.* individuals from Ola‘a tend to cluster together in Fig.
504 1B), suggesting a slight population differentiation. However, we found no strong correlation
505 between microbial beta diversity and *A. waikula* (phylo)genetic distances using Mantel tests (Table
506 2 & Supplementary Table S7).
507
508 *Gut microbiota*: In contrast to the endosymbionts, the gut microbiota showed a
509 homogeneous taxonomic composition of microbial taxa at the genus level across different
510 *Ariamnes* populations or species (Fig. 4B & Supplementary Fig. S7). This stability of the
511 composition of the gut microbiota was also confirmed at the order level (Fig. 4B & Supplementary
512 Fig. S7) and at the level of the individual spiders (Supplementary Fig. S8). No clear differentiation
513 according to the different species (across islands) or different *A. waikula* populations from Hawai’i
514 Island could be visually detected for the beta diversity of the gut microbiota, either from PCoA
515 decomposition (Supplementary Fig. S9) or hierarchical clustering (dendrogram; Fig. 1C &

516 Supplementary Fig. S10), though there are some slight patterns of clustering by host population.
517 This slight differentiation was confirmed using PERMANOVA: The gut microbiota of individuals
518 from the same population were significantly more similar than the gut microbiota from different
519 populations or species, even within *A. waikula* populations from Hawai'i (R^2=0.29, P<0.001,
520 Supplementary Table S6). These results indicate that, although the composition of the gut
521 microbiota seems taxonomically similar, there is a slight trend toward microbiota differentiation
522 according to the host populations. However, using Mantel tests we found no significant
523 correlations between the gut microbial community beta diversities and the *Ariamnes* genetic
524 distances, the *Ariamnes* phylogenetic distances, nor the geographical distances (Table 2 &
525 Supplementary Table S7), suggesting that there is likely no pattern of sequential colonization from
526 older to younger volcanoes (stepping stone colonization) in these gut microbial communities in
527 opposition to their spider hosts. We also noticed that besides the slight effect of host populations
528 on the PCoA decomposition (Supplementary Fig. S9) or on the dendrogram of the gut microbiota
529 (Fig. 1C), these gut bacterial communities tend to group into two clusters (Supplementary Fig. S9
530 and Fig. 1C), with one cluster ("*cluster 1*") being composed of 14 individual spiders from Alili, Ola'a,
531 Thurston, and Pu'u Maka'ala. Examination of the details of the differences in gut microbial
532 composition of these two clusters, we found that spider individuals from *cluster 1* were additionally
533 colonized by a single abundant OTU from the order Entomoplasmatales (Supplementary Fig.
534 S11).
535 Finally, the diversity analyses using phylogenetic metrics showed similar patterns to that obtained
536 with non-phylogenetic metrics (results not shown).
537
538 *Simulations*
539 Simulations to assess whether the slow substitutions rates of 16S rRNA can indeed
540 capture signatures of change among host-associated microbes in recently evolving lineages
541 (<2MY) showed that we can recover a significant phylosymbiotic pattern, i.e. a significant

542 correlation between the microbial beta diversity and the *Ariamnes* phylogenetic distances, even
543 when only 10% of the bacterial OTUs were transmitted and 90% of the remaining OTUs were
544 evolving independently from the *Ariamnes* phylogeny (Figure S12). When >=75% of the bacterial
545 OTUs were transmitted, we always recovered a strong and significant phylosymbiosis (Figure
546 S12). This suggested that, despite the slow evolution of the 16S rRNA gene, we should be able
547 to detect microbiota differentiation if some of the *Ariamnes*-associated bacteria diverged
548 separately in the different populations, whether because of vertical transmission or because of
549 vicariance events in the environmental pools of available bacteria. Moreover, simulations of the
550 bottlenecks in microbial diversity during island colonizations showed that we can recover a signal
551 in the microbial alpha diversity (a decrease of diversity in the youngest sites) even if the bottleneck
552 has impacted only a small proportion of the bacterial OTUs (<10%; Figure S13).

553 **Discussion**

554 In this study, we examined the early stages of diversification of Hawaiian *Ariamnes* spiders
555 in concert with the different components of their associated microbial community on the youngest
556 island of the Hawaiian archipelago. Our results show a strong effect of isolation by distance in the
557 host and contrasting patterns of community composition among both the endosymbionts and the
558 gut microbiota they harbor, suggesting eco-evolutionary distinctiveness of the host and each of
559 the two microbial components.

560 *Isolation by distance explains Ariamnes diversification*

561 Our analyses suggest that the nascent diversification of *A. waikula* is mainly the result of
562 isolation by distance. The different populations of *A. waikula* show a predictable pattern of
563 colonization across Hawaiʻi Island from the oldest (Kohala) to the youngest (MLKS) sites, with little
564 genetic structure between the geographically proximal MLKS sites of Olaʻa, Puʻu Makaʻala, and
565 Thurston (PCA, Supplementary Fig. S5; FST, Supplementary Table S3; Admixture, Fig. 2B). Our
566 analyses suggest either that gene flow is still occurring between these MLKS populations, or that
567 the recency of colonization means that there has been insufficient time for differentiation in this
568 area (Roderick et al. 2012). Interestingly, we did not find a significant increase of genetic diversity
569 with volcano age, as expected with successive founding events. These patterns may become
570 more apparent with additional loci or more extensive sampling.

571 Patterns of genetic differentiation suggest a potentially complicated course of colonization
572 across Hawaiʻi Island for *A. waikula*. Although Kohala, which represents the oldest volcano, was
573 consistently characterized as distinct from other populations in PCA, phylogenetic, and admixture
574 analyses, both Saddle and Alili were also largely distinct. These patterns are potentially explained
575 by a shifting mosaic population structure (Carson et al 1990) with multiple colonization events,
576 historical connectivity, or ongoing admixture. However, sampling of additional sites, individuals,

577 and species will be required to disentangle the historical sequence of events explaining the
578 relationship between these sites. Our results are consistent with other population genetic studies
579 on Hawai‘i Island showing strong differentiation and potential speciation of lineages on the
580 different volcanoes within the island (Funk & Wagner 1995; Holland & Cowie 2007; Muir & Price
581 2008; Croucher et al. 2012; Goodman et al. 2019; Eldon et al. 2019; Blankers et al. 2018).
582 Interestingly, spiders on the mainland would not show this isolation by volcano pattern, because
583 they do efficiently disperse by ballooning; here, our data, in contrast, suggest that the species
584 does not efficiently balloon anymore in this island system (Gillespie et al 2012).
585

586 *Major endosymbiont changes after potential colonization event*

587 We found major shifts among the dominant endosymbiont genera across the different
588 islands and *Ariamnes* species: Species of *Ariamnes* from the two geologically oldest sites included
589 in the study (Moloka‘i and Maui) each harbor a unique endosymbiont genus (*Rickettsia* and
590 *Wolbachia* respectively) while the different populations of *A. waikula* on the youngest Hawaii Island
591 all have a mix of endosymbionts dominated by the genus *Rickettsiella*. Most spider individuals
592 were colonized by a single endosymbiont OTU as previously observed in *Argiope* spiders (Sheffer
593 et al. 2020), but there were also a few individuals that were co-colonized by both *Rickettsiella* and
594 *Rickettsia*. Although such endosymbiont coexistence has been previously observed (*e.g.* Mouton
595 et al. (2004) found up to 3 coexisting *Wolbachia* strains in some wasps), we cannot exclude the
596 possibility that one of these endosymbionts may have come from a prey item that had just been
597 eaten before sampling (Kennedy et al. 2020). The overall shifts in endosymbionts across *Ariamnes*
598 species likely contribute to the congruence between the host phylogeny and the dendrogram of
599 microbiota dissimilarity - a pattern referred as phylosymbiosis (Brooks et al. 2016; Lim &
600 Bordenstein, 2020). The strong conservatism of the dominant endosymbiont within each island is
601 consistent with the maternal inheritance and vertical transmission of the endosymbionts (Duron et

602 al. 2008) and the fact that it generates phylosymbiosis has been similarly observed in other
603 arthropods (e.g. a *Camponotus* ant system and its associated endosymbiont *Candidatus*
604 *Blochmania* (Degnan et al. 2004)). Therefore, the endosymbiont component of the microbiota
605 reflects the overall evolutionary relationships of the *Ariamnes* host. However, from our study, we
606 cannot say whether endosymbionts played a role in promoting the genetic isolation of *A. waikula*
607 populations; this hypothesis would require experimental manipulation of microbial communities
608 within different host populations.

609 The observed turnover of dominant endosymbiont genera between islands and host
610 populations suggests that endosymbiont horizontal transmission happens quickly. Previous
611 studies have also shown frequent horizontal transfers in the *Wolbachia* strains infecting
612 arthropods, this pattern being more common than co-speciation events (Baldo et al., 2008;
613 Stahlhut et al. 2010). In particular, Stahlhut et al. (2010) showed that hosts sharing the same
614 ecology were likely to horizontally exchange *Wolbachia* strains. In contrast, a more recent study
615 reported lower turnover rates over time (7 Myr on average) for the bacterial genus *Wolbachia*
616 colonizing arthropods in French Polynesia and found that *Wolbachia* colonization did not seem
617 related to recent events of isolation mediated by island structure (Bailly-Bechet et al. 2017). Given
618 that the islands used in the current study, and thus the spider populations studied here, are
619 separated by <2 Myr (Gillespie et al. 2018), the more rapid turnover in *Ariamnes* endosymbiotic
620 communities (compared to *Wolbachia* in the French Polynesian arthropods) appears to be linked
621 to the geological formation of islands. Therefore, it seems to exist a range of endosymbiont
622 dynamics, from conserved and stable host associations (Degnan et al. 2004) to labile and
623 horizontally acquired (Baldo et al., 2008; Stahlhut et al. 2010; this study). For *Ariamnes* spiders,
624 including *A. waikula* (Kennedy et al. 2018), predation on other arthropods and cannibalism are
625 particularly important (Whitehouse et al. 2002), providing a potential avenue for endosymbiont
626 horizontal transfer (Su et al. 2019) and may explain the high endosymbiont turnover observed in
627 this system. However, from our current analyses, because we did not study the phylogenetic

628 history of individual microbial lineages, we cannot date the events of endosymbiont acquisition.
629 In particular, we cannot separate a scenario of recent horizontal shifts of the endosymbionts,
630 where endosymbionts were very recently acquired from the environment, from a scenario of older
631 shifts, where endosymbionts were acquired during colonization events and were thus
632 synchronous with speciation events or subject to priority effects. Testing this will require modeling
633 host-microbiota evolution, which is currently challenging on such short evolutionary timescales
634 (Groussin et al. 2017; Perez-Lamarque & Morlon 2019).

635 Several previous studies have suggested that endosymbionts such as *Rickettsia* reduce
636 the behavioral propensity for ballooning dispersal in spiders (Goodacre et al. 2009). Hawaiian
637 *Ariamnes* have never been shown to balloon and exhibit high local endemism, a common
638 phenomenon in organisms endemic to remote islands, generally attributed to distinct local
639 selective pressures (Gillespie et al 2012). The possibility that such a reduction in dispersal may
640 be associated with bacterial infection is intriguing. Experimental studies will be needed to test the
641 hypothesis.

642 *Similarity of the gut microbiota*

643 Across all spider populations, we observed a conserved gut microbiota: Over
644 approximately one million years of divergence, the composition seems to have remained stable
645 and does not reflect the *Ariamnes* genetic structure. The core lineages at the generic level are the
646 same as those typically found in the gut microbiota of other arthropods (Engel & Moran 2013),
647 including spiders (Kennedy et al. 2020, Sheffer et al. 2020). In addition, although we detected
648 slight but significant differences in the gut microbiota between the different spider species and
649 between the *A. waikula* populations within Hawai'i Island, these patterns of differentiation were
650 not associated with *Ariamnes* (phylo)genetic structure, volcano age, nor the geographical
651 distances between the sampling sites. Most of the variation in the gut microbiota observed using
652 PCoA or cluster dendrograms (Fig. 3C and Supplementary Fig. S9) was actually explained by a

653 single OTU from the order Entomoplasmatales that was present in high abundance in many *A.*
654 *waikula* individuals across different populations (Supplementary Fig. S8). Given that
655 Entomoplasmatales are arthropod pathogens (Gasparich, 2014), the observed variation across
656 *Ariamnes* gut microbiota may simply reflect the local acquisition from the environment,
657 decorrelated from the host (phylo)genetic structure.

658 We found that both alpha and beta diversity of gut microbes was similar across the different
659 islands and populations, decoupled from the spider's phylogeny. This absence of a correlation
660 with the host's phylogenetic structure suggests that the gut microbiota is likely acquired from the
661 spider's environment, most probably through the spider's diet (Kennedy et al. 2020; Zhang et al.
662 2018). The pattern of slight differentiation according to the host populations could either be due to
663 gradual changes in the host-filtering process (Mazel et al. 2018; see Introduction), or only reflect
664 the local variations of the bacterial pool available in the host environment.

665 One could argue that the absence of strong differentiation in the microbial communities
666 could come from the slow evolution of the 16S rRNA marker which may not provide the needed
667 resolution to show differences at short timescales (<2 Myr). However, given that we used Z-OTU
668 clustering (0% radius OTUs) that can detect any variation in the 16S sequences and that the 16S
669 gene of symbiotic bacteria is estimated to diverge at least 1-2% per 50 Myr on average (Ochman
670 et al., 1999), our simulations demonstrated that we should expect to have some variation in a
671 portion of the 16S sequences of the transmitted bacteria. Thus, we should expect to detect
672 microbial differentiation (by comparing their beta diversity) and a resulting pattern of
673 phylosymbiosis (Sanders et al. 2014, Perez-Lamarque & Morlon, 2019). In particular, our
674 simulations demonstrated that even 10% transmission of the microbial OTUs can be sufficient to
675 generate a significant pattern of phylosymbiosis within such a short amount of time. In addition, if
676 the *Ariamnes*-associated microbial communities experienced bottlenecks at colonization events,
677 we expect filtering of the preexisting (genetic or strain) diversity that has accumulated for more
678 than 2 Myr, and this should not be affected by the slow evolution of the 16S rRNA gene. Here

again, this expectation was supported by the simulations which showed that diversity bottlenecks in the microbial communities resulting from a stepping stone colonization, could readily be detected and was not sensitive to the short timescales of the *A. waikula* divergences.

The overall similarity of the gut microbiota at the generic level has been shown in other systems, such as recently diversified *Anolis* lizards (Ren et al. 2016) and *Cephalotes* turtle ants (Sanders et al. 2014). Such similarity likely conserves the functional properties of the gut microbiota (Muegge et al. 2011), which could be particularly advantageous for arthropods which often rely on gut symbionts to complement imbalanced nutrition (Engel & Moran 2013) or detoxify toxin-rich diets (Adams et al. 2013). It is unclear whether *Ariamnes* spiders, and predatory arthropods in general, also rely on their microbiota for various functions. A recent study suggested that, the spider microbiota can be mainly derived from the prey items, without any apparent functions (Kennedy et al., 2020). If the same applies for the Hawaiian *Ariamnes*, the conserved gut microbiota of *A. waikula* may reflect its conserved niche, with the diet presumably similar despite changes in geographical range.

Conclusion

Organisms host unique and large microbial diversities, and it is likely that the different components of the microbial community experience different ecological and evolutionary dynamics. In this study, we demonstrated a dichotomy between the endosymbiont communities and the gut microbiota: For the host Hawaiian *Ariamnes* spiders, abiotic factors, in particular isolation by distance along the volcano chronosequence, appear to have the strongest influence in explaining divergence patterns. The evolutionary history of the hosts also strongly impacts the colonization of endosymbionts, which show clear shifts in their composition between species on different islands. However, we find little differentiation in endosymbionts across populations of *A. waikula* within Hawa‘i Island, despite the strong genetic structure of the host. In contrast to the differences we find between taxa in the endosymbiont composition, the composition of the gut

705 microbiota is overall similar across both species and populations despite strong geographic
706 isolation of their hosts. Our results stress the high heterogeneity within the different components
707 of the *Ariamnes* holobiont that likely did not act as a single homogenous unit of selection.
708 Investigating the composition of microbiota through a broader sampling across different
709 ecomorphs and islands would significantly contribute to our understanding of the dynamics of
710 host-microbe interactions as it pertains to rapid speciation events.

711 **Acknowledgements**

712 We would like to acknowledge support and assistance from the following: The permit processing
713 and access to different reserves and private land was possible thanks to Steve Bergfeld (DOFAW
714 Big Island), Pat Bily (TNC Maui), Tabetha Block (HETF), Pomaikaʻi Kaniaupio-Crozier (Maui Land
715 and Pineapple), Lance DaSilva (DOFAW Maui), Charmian Dang (NAR), Melissa Dean (HETF),
716 Betsy Gagne (NAR), Lisa Hadway (DOFAW Big Island), Cynthia King (DLNR), Russell Kallstrom
717 (TNC Molokaʻi), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Molokaʻi). For support and
718 advice in the lab, we are very grateful to Lydia Smith (Evolutionary Genetics Lab, Museum of
719 Vertebrate Zoology, UC Berkeley), Shana McDevitt (Vincent J. Coates Genomics Sequencing
720 Laboratory, QB3, UC Berkeley), and Anna Sellas (California Academy of Sciences, San
721 Francisco). Samples were provided by Susan Kennedy and Andrew Rominger. We thank
722 Benjamin Peter, Thorfinn S. Korneliussen, and Line Skotte for advice on the analyses. L.E.B was
723 supported by the Netherlands Organisation for Scientific Research VENI #863.14.020. B.P.L was
724 supported by a master fellowship from the École Normale Supérieure of Paris. H.K. was supported
725 by a postdoctoral fellowship by the German Research Foundation (DFG). Support for the project
726 was provided by the NSF DEB 1241253 to R.G.G.

Data Accessibility

The pipelines used for processing ddRAD seq data are available in https://github.com/CGRL-QB3-UCBerkeley/RAD. All scripts used to analyze the Ariamnes microbiota are available in https://github.com/BPerezLamarque/Scripts. The raw data can be found on dryad (https://datadryad.org) DOI: https://doi.org/10.5061/dryad.nzs7h44qj

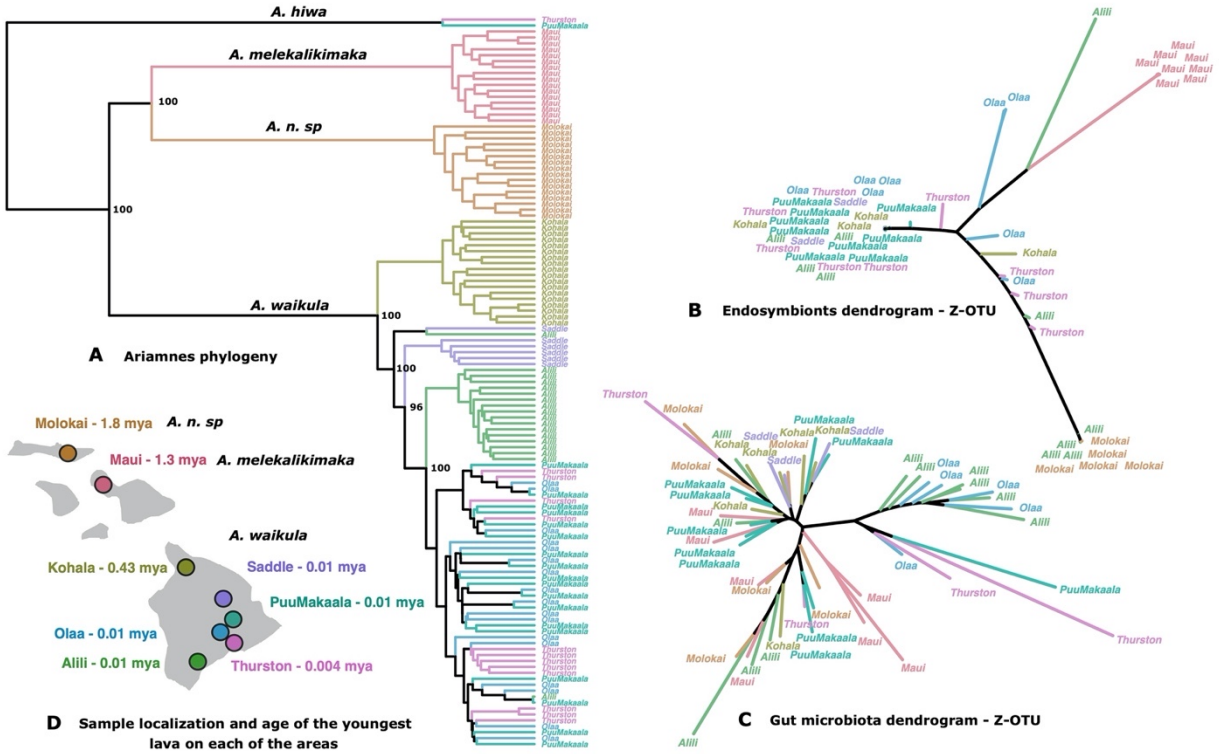
Author Contributions

R.G.G and H.K. conceived of the study. E.A. and B.P-L. performed DNA extractions. C.C. and L.E.B. performed ddRAD lab work. E.A., K.B., and T.L. performed ddRAD genetic analyses. B.P-L. and H.K. performed lab work and analyses for the microbial component of the study. E.A., B.P-L., H.K., and R.G.G wrote the manuscript. All authors edited and approved the manuscript before final submission.

Figures

Figure 1: Host phylogenetic history partially recapitulates microbiota differentiation

(A) Phylogenetic tree of the *Ariamnes waikula* individuals across the island of Hawai'i obtained using IQ-TREE smoothed with r8s. Two specimens of *A. hiwa* (brown ecomorph) and 32 specimens of *A. melekalikimaka* (from Maui; gold ecomorph) and *A. n. sp.* (from Moloka‘i; gold ecomorph) were used as outgroup taxa**.** The bootstrap support values are indicated on the side of the main phylogenetic nodes. (B) Microbiota dendrograms reconstructed from the endosymbiont community and (C) the gut microbiota for the Z-OTU. (D) map of the sampled host populations and the corresponding volcano age on each of the areas.



752 **Figure 2: Population genetic analyses of *Ariamnes waikula*.** A) ngsAdmix results for K=2 to

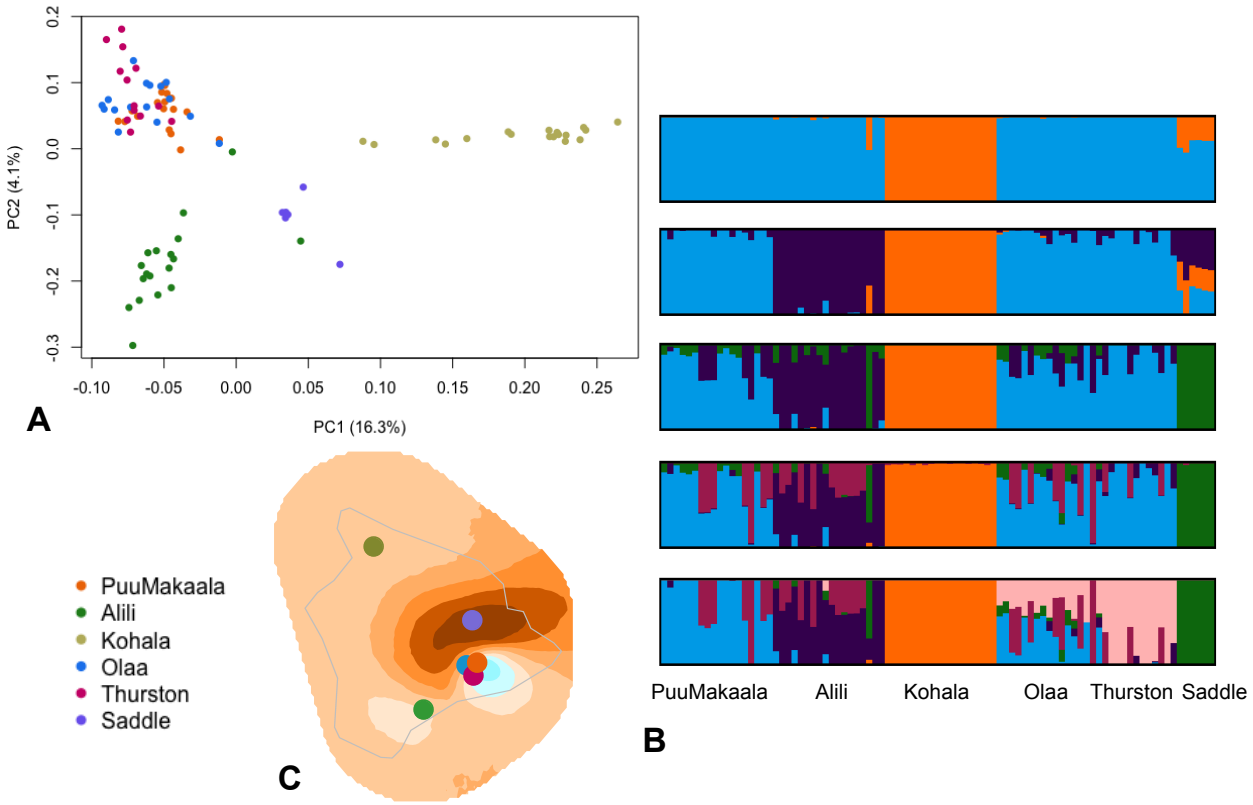
753 K=8 for all *A. waikula* specimens. B) ngsAdmix results f

754 or K=2 to K=6 (from top to bottom) for all *A. waikula* from Hawai'i Island only. C) EEMS analysis

755 of all *A. waikula* specimens. Brown color indicates barriers for gene flow (the stronger the darker),

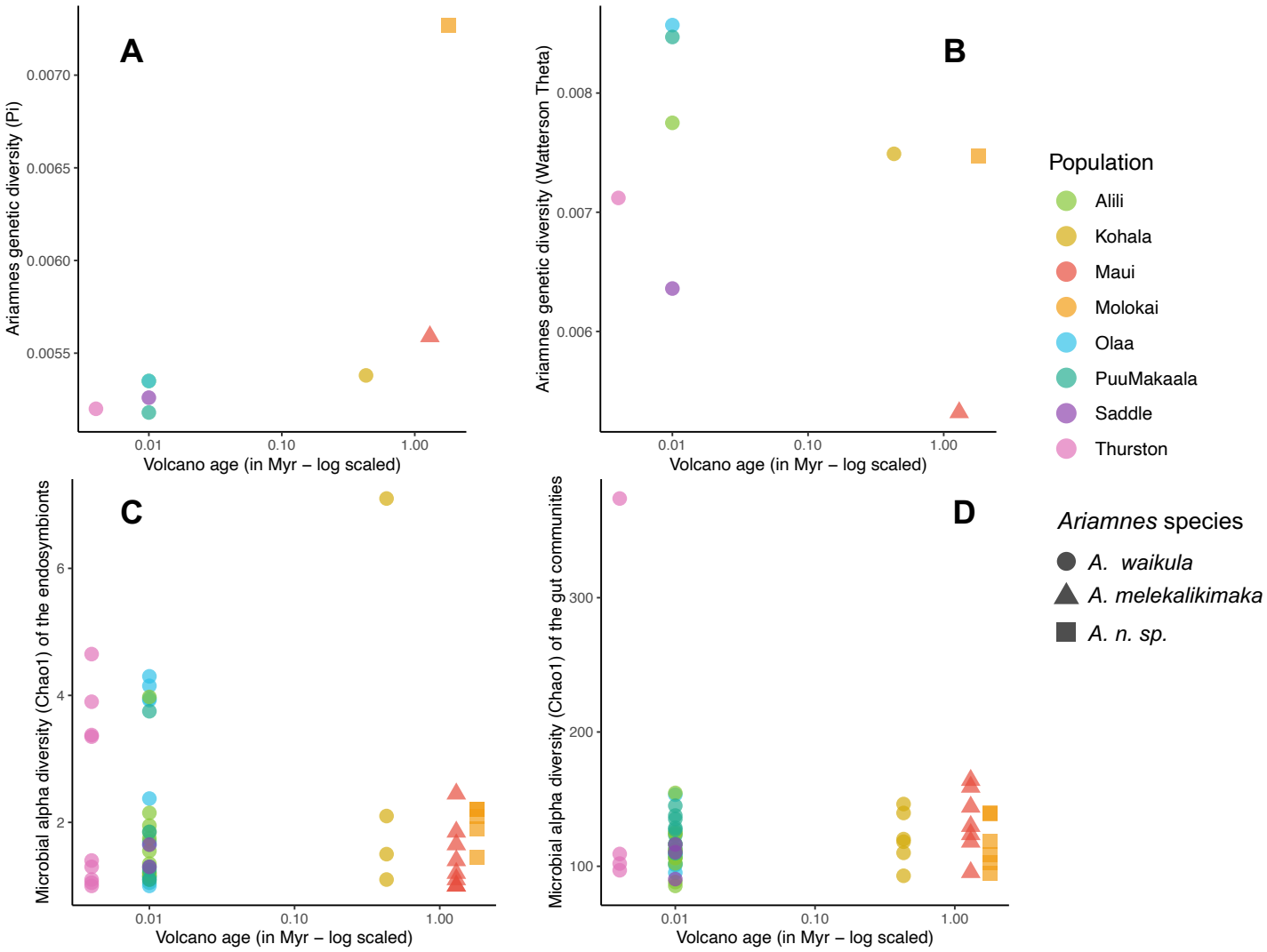
756 and cyan indicates gene flow between populations (the stronger the darker).

757



758
759

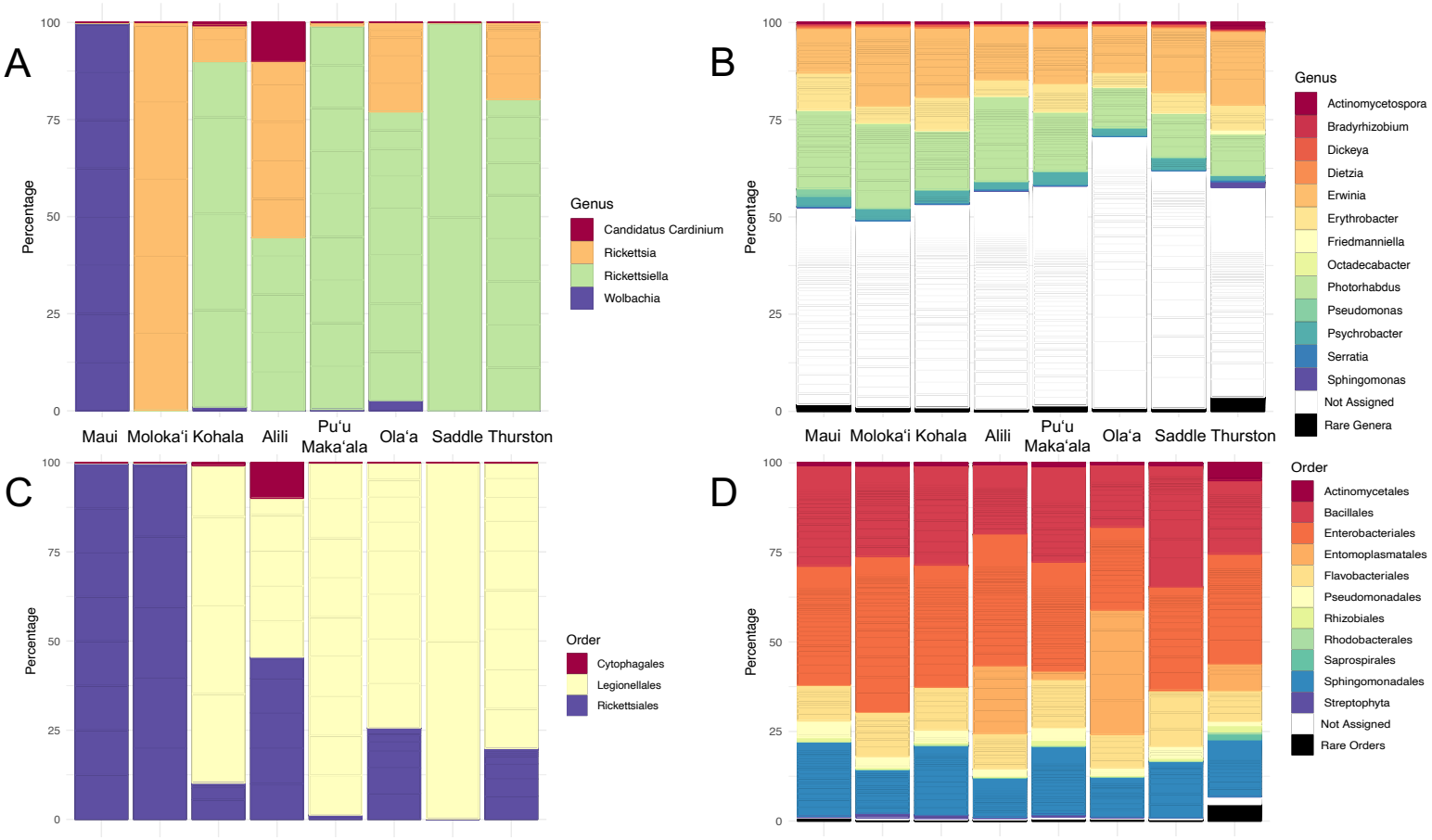
760 **Figure 3: The bottleneck in diversity along the volcano age tends to affect the *Ariamnes***
761 **hosts, but not their microbial associates**
762 Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A.*
763 *n. sp respectively*) and are only represented as a reference here: the following linear model testing
764 the effect of volcano age on host (A-B) or microbial (C-D) diversities were only performed on the
765 six *A. waikula* populations in Hawai'i.
766 (A) *Ariamnes* genetic diversity (Pi) estimated per population as a function of the volcano age (in
767 Myr). We found no significant relationship between *A. waikula* genetic diversity (Pi) and the
768 volcano age (linear model: F=1.7, df=4, p-value=0.27), despite the observed trend, probably
769 because of the small number of sampled populations in Hawai'i (6 only).
770 (B) *Ariamnes* genetic diversity (Watterson's theta) estimated per population as a function of the
771 volcano ago (in Myr). We found no significant relationship between *A. waikula* genetic diversity
772 (Watterson's theta) and the volcano age (linear model: F=0.023, df=4, p-value=0.87).
773 (C) Microbial alpha diversity (Chao1 index on Z-OTUs) estimated for the endosymbionts as a
774 function of the volcano ago (in Myr). Each dot corresponds to the average of the 20 rarefactions
775 performed for each individual. We found no significant relationship between endosymbionts alpha
776 diversity and the volcano age (linear mixed model: t=1.10, p-value=0.24).
777 (D) Microbial alpha diversity (Chao1 index on Z-OTUs) estimated for the gut community as a
778 function of the volcano ago (in Myr). Each dot corresponds to the average of the 20 rarefactions
779 performed for each individual. We found no significant relationship between gut microbial alpha
780 diversity and the volcano age (linear mixed model: t=-0.08, p-value=0.94). The absence of a
781 significant correlation was not influenced by the presence of an outlier sample in Thurston with
782 very high microbial diversity.
783


A
Ariamnes genetic diversity (Pi)
0.0070
0.0065
0.0060
0.0055
0.01
0.10
1.00
Volcano age (in Myr – log scaled)
B
Ariamnes genetic diversity (Watterson Theta)
0.008
0.007
0.006
0.01
0.10
1.00
Volcano age (in Myr – log scaled)
C
Microbial alpha diversity (Chao1) of the endosymbionts
6
4
2
0.01
0.10
1.00
Volcano age (in Myr – log scaled)
D
Microbial alpha diversity (Chao1) of the gut communities
300
200
100
0.01
0.10
1.00
Volcano age (in Myr – log scaled)
Population
Alili
Kohala
Maui
Molokai
Olaa
PuuMakaala
Saddle
Thurston
Ariamnes species
A. waikula
A. melekalikimaka
A. n. sp.

784

Figure 4: Relative abundances of endosymbionts (A and C) and gut symbionts (B and D) per spider population defined from rarefied Z-OTU table.

The panel A-B represents the taxonomic assignations at the genus level, whereas the panel C-D represents the taxonomic assignations at the order level. The relative abundance of each OTU is delimited using the horizontal grey lines and OTUs are colored according to genus. Rare taxonomic assignations (representing less than 1% of the abundance) are merged together. Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island. See Table 1 for the number of individual spiders per sampling site.



Tables

Table 1: Individual, population, and sampling locality information for *Ariamnes* spiders in this study. Volcano ages are given based on K-Ar dating estimates of the maximum time available for occupancy (Langenheim & Clague 1987). Substrate ages (in years) are approximate and based on geologic estimates of the youngest lava flow making up each sampling locality. * Samples with sufficient coverage for microbiota analysis. ¶ Approximate ages and estimates of time available for colonization (Carson & Clague 1995). For full details see Supplementary Tables S1 and S2.

Species	Island	Population	Volcano, age (mill. years) ¶	Substrate age (years)	Individuals (ddRAD)	Microbiota analysis *
A. waikula	Hawai‘i	Thurston	Kīlauea, 0.004	600	13	9
A. waikula	Hawai‘i	Alili	Mauna Loa, 0.01	20,000	18	13
A. waikula	Hawai‘i	Pu‘u Maka‘ala	Mauna Loa, 0.01	11,000	18	13
A. waikula	Hawai‘i	Ola‘a	Mauna Loa, 0.01	7,500	16	11
A. waikula	Hawai‘i	Saddle	Mauna Loa, Kea 0.01-0.38	4,000	6	3
A. waikula	Hawai‘i	Kohala	Kohala, 0.43	300,000	18	6
A. n. sp	Moloka‘i	Kamakou	E. Moloka‘i, 1.8	1,400,000	16	7
A. melekalikimaka	Maui	Pu‘u Kukui	W. Maui, 1.3	1,500,000	16	9
A. hiwa	Hawai‘i	Pu‘u Maka‘ala	Mauna Loa, 0.01	11,000	1	N/A
A. hiwa	Hawai‘i	Thurston	Kīlauea, 0.004	600	1	N/A
					Total: 123	**Total: 71**

807 **Table 2:** Mantel tests comparing the microbial beta diversity dissimilarities to the genetic
808 distances, the phylogenetic distances of their associated *Ariamnes,* and the geographic distances
809 between the sampling sites. Beta diversity dissimilarities were computed on rarefied Z-OTU tables
810 using the Bray-Curtis dissimilarities. Bold values represent significant correlations. Mantel tests
811 were either performed on all populations (between *Ariamnes* species and *A. waikula* populations)
812 or only on *A. waikula* populations.
813

Matrix 1	Matrix 2	All populations		*A. waikula* populations	
		Correlation	p-value	Correlation	p-value
Whole microbiota beta diversity distances	*Ariamnes* genetic distances	**0.324**	**9.9e-05**	0.063	0.20
	Ariamnes phylogenetic distances	**0.334**	**9.9e-05**	0.012	0.40
	Geographic distances	**0.303**	**9.9e-05**	0.053	0.20
Endosymbionts beta diversity distances	*Ariamnes* genetic distances	**0.619**	**9.9e-05**	0.015	0.49
	Ariamnes phylogenetic distances	**0.642**	**9.9e-05**	0.047	0.32
	Geographic distances	**0.567**	**9.9e-05**	0.108	0.169
Gut microbiota beta diversity distances	*Ariamnes* genetic distances	-0.048	0.66	-0.052	0.70
	Ariamnes phylogenetic distances	-0.038	0.63	-0.139	0.93
	Geographic distances	0.007	0.42	-0.084	0.78

814

Supplementary Figures

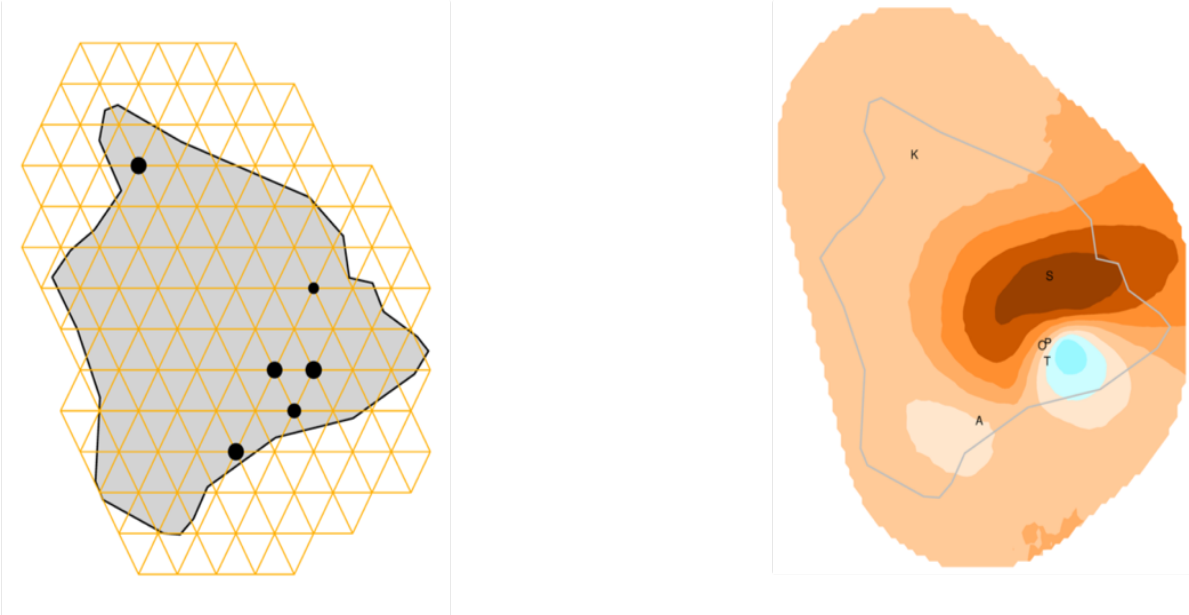


Figure S1: **Grid system used for the EEMS analyses on Hawai‘i Island (left) and corresponding results (right)**

A: Alili, K: Kohala, O: Ola‘a , P: Pu‘u Maka‘ala, S: Saddle, T: Thurston

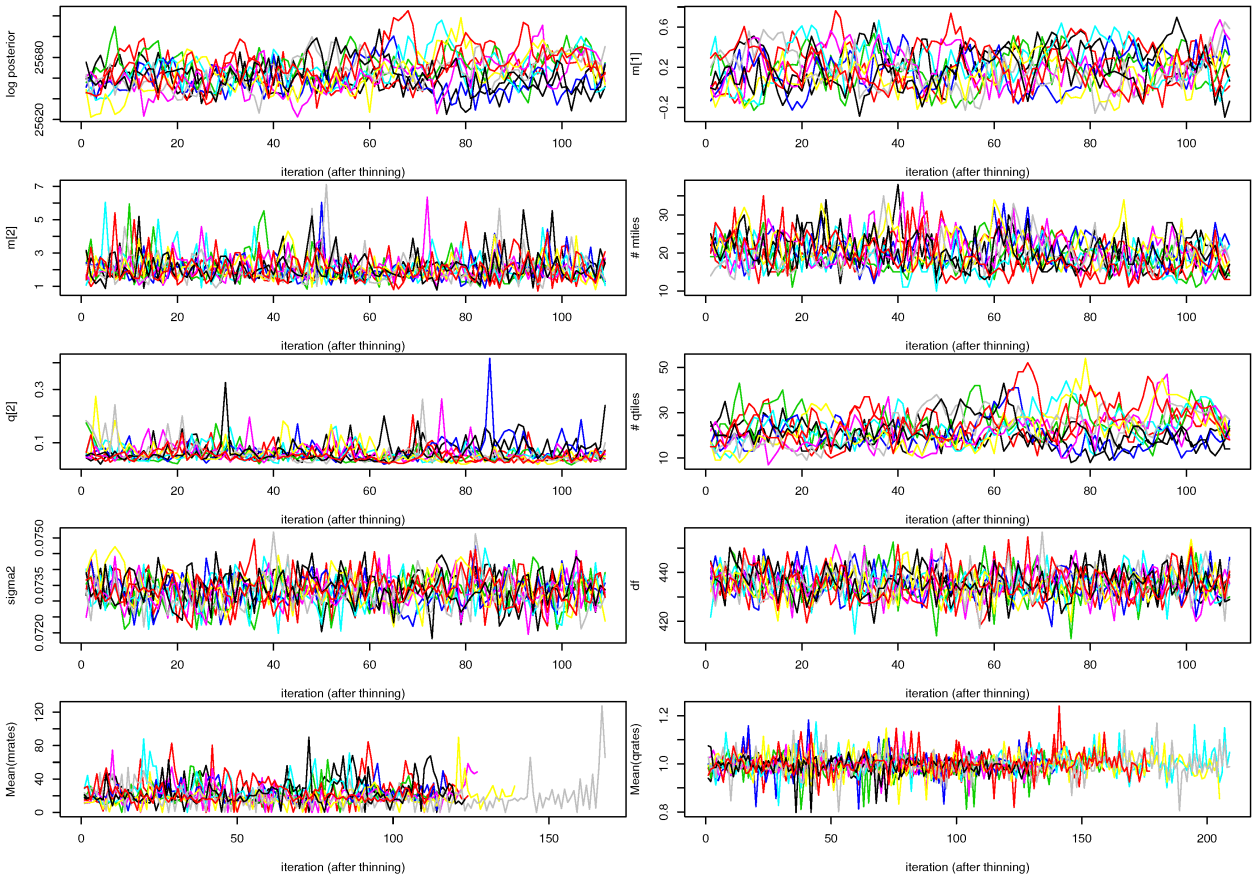


Figure S2. Traces of the MCMC runs of the 10 replicates to assess convergence.

Only individuals of the gold ecomorph from Hawai‘i Island were used here.

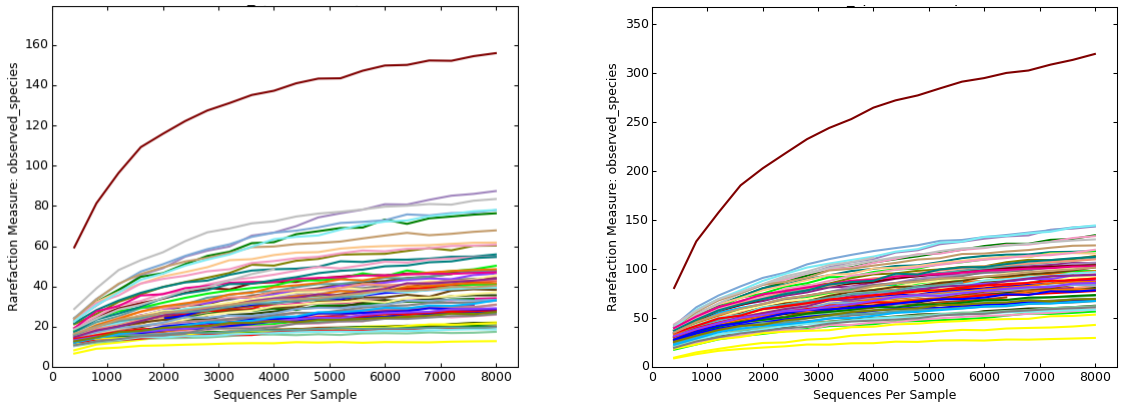
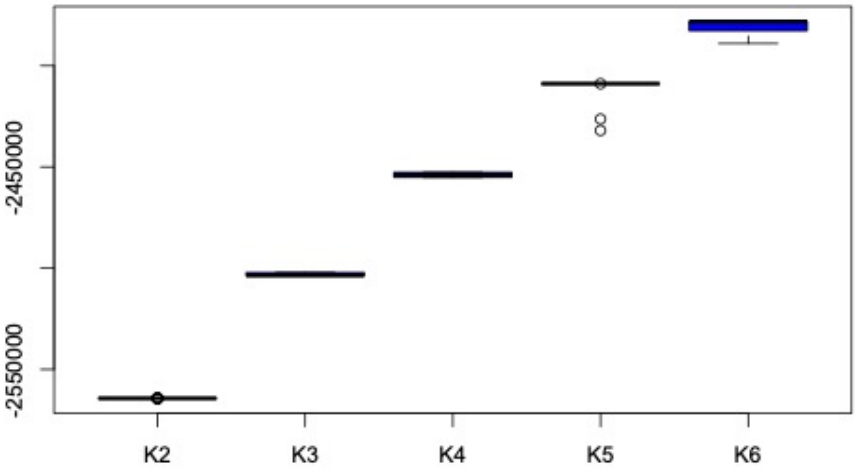


Figure S3: Rarefaction curve of the whole microbiota: number of observed OTU in each sample according to the number of rarefied reads per sample. Results of OTU at 97% are presented on the left, whereas Z-OTU are on the right. Rarefactions are performed 20 times independently and the mean value is plotted here. Note that one sample (from Thurston) present an unexpectedly high diversity compared to other samples.

833



834

835 **Figure S4**: Log likelihood for 50 iterations of NGSAdmix K values 2-6 for *A. waikula* from Hawai'i

836 Island.

.

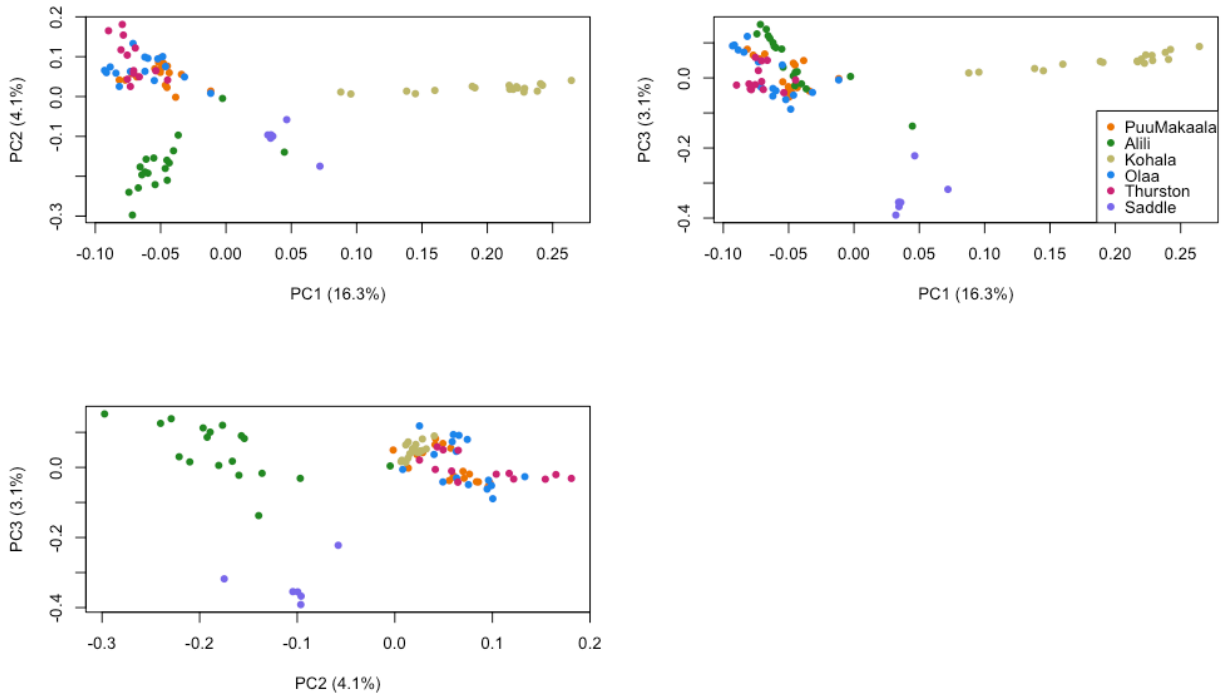
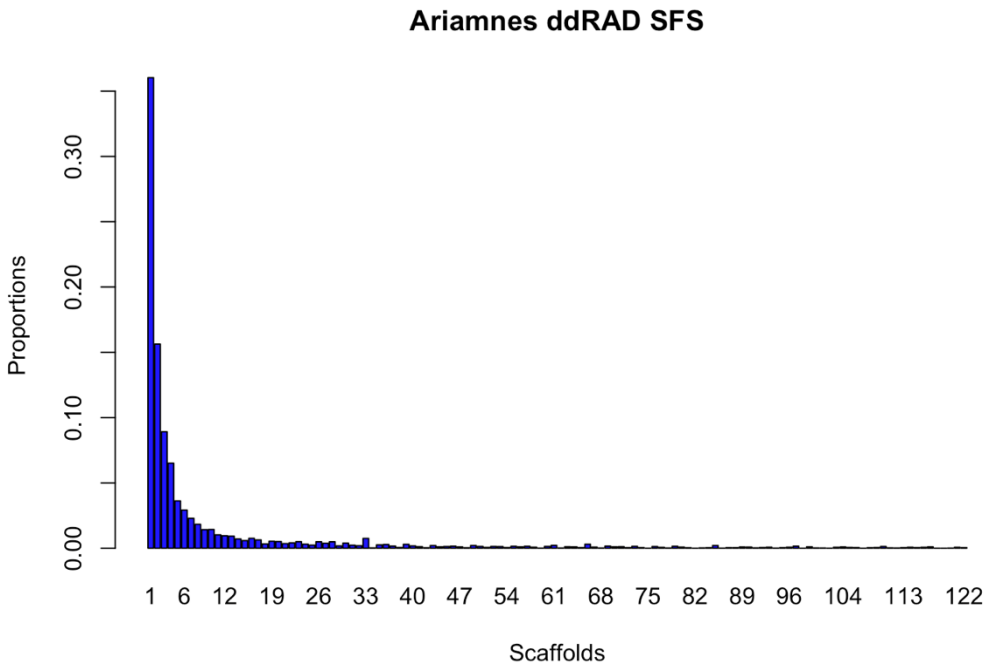


Figure S5: PCA analyses of ddRAD data of *A. waikula* from Hawai'i Island

841



842

843

844 **Figure S6**: Site frequency spectrum of all *Ariamnes* ddRAD data after filtering.

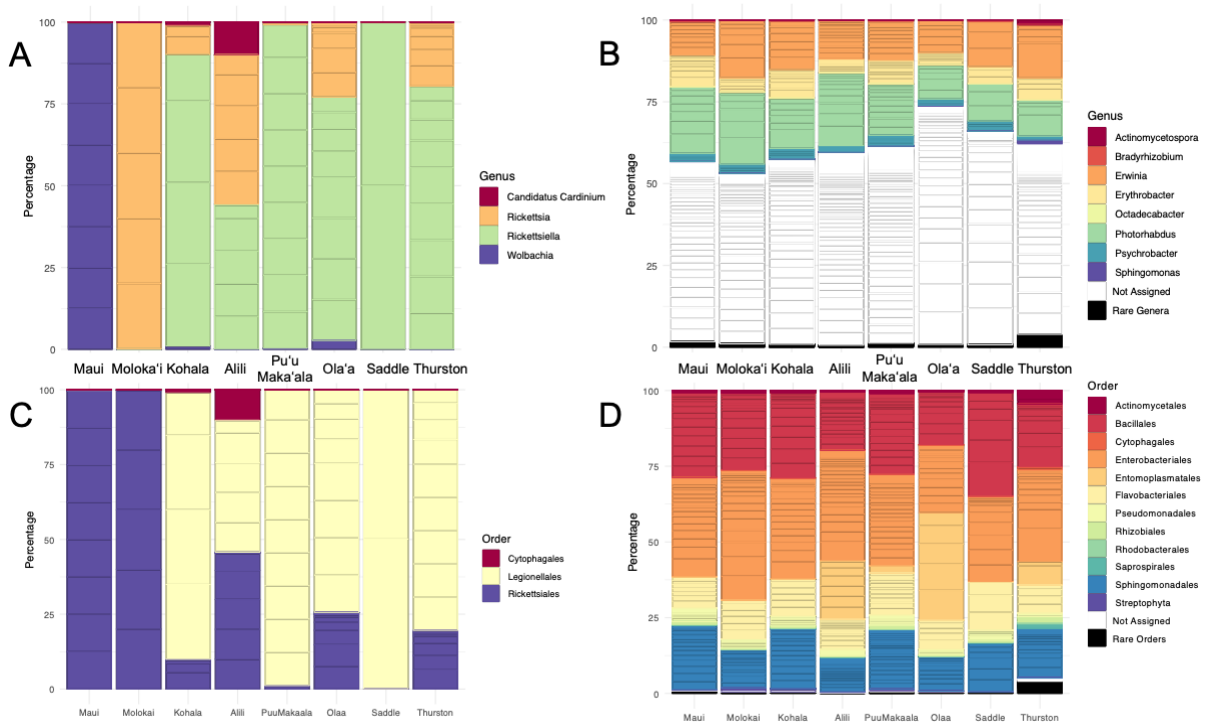


Figure S7: Relative abundances of endosymbionts per population (A and C) and gut symbionts per spider population (B and D) defined from rarefied 97% OTU table.

The panel A-B represents the taxonomic assignations at the genus level, whereas the panel C-D represents the taxonomic assignations at the order level. The relative abundance of each OTU is delimited using the horizontal grey lines and OTUs are colored according to genus/order. Rare taxonomic assignations (representing less than 1% of the abundance) are merged together. Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island. See Table 1 for the number of individual spiders per sampling site.

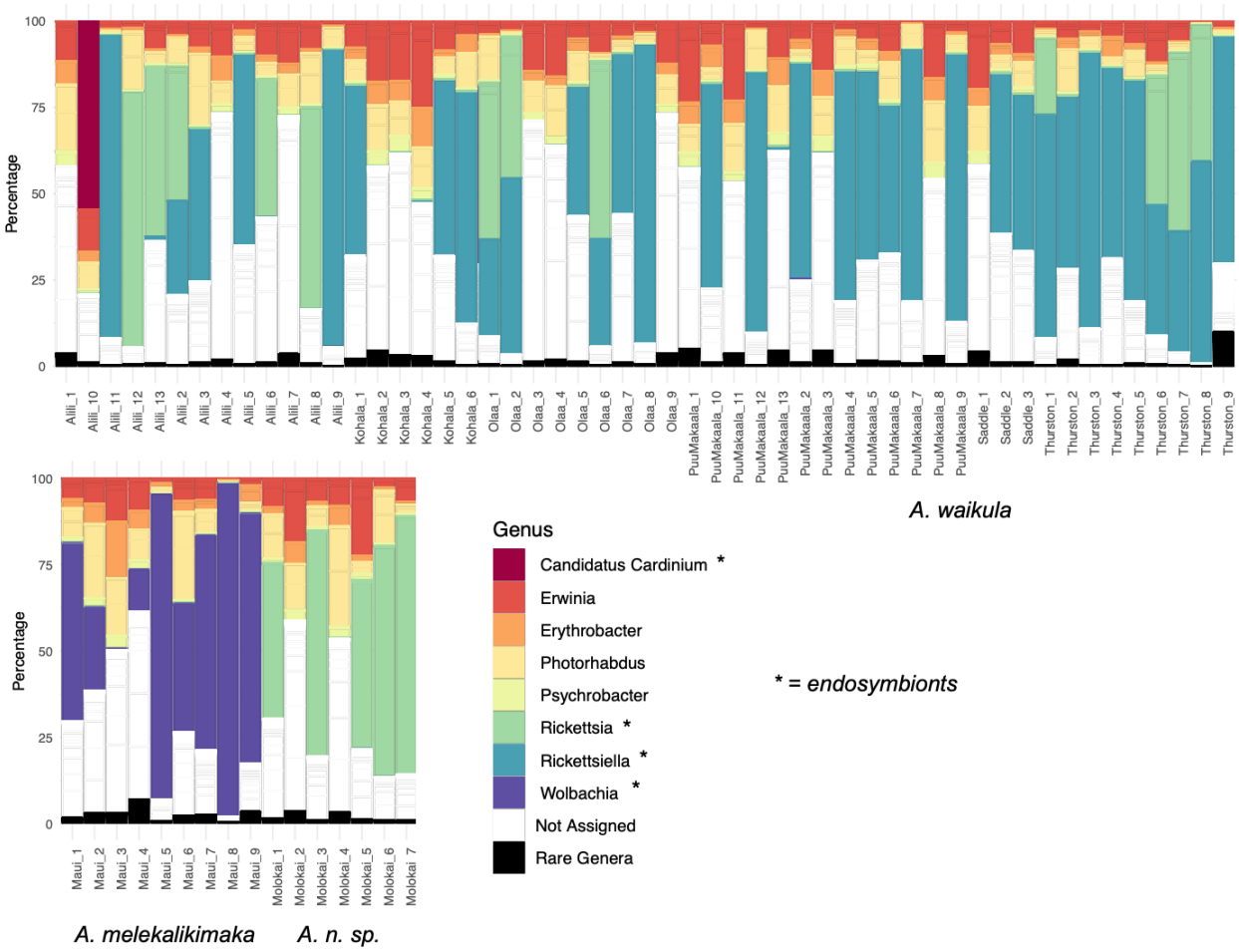


Figure S8: Relative abundances of the microbial community (both endosymbionts and gut symbionts) per spider individual from rarefied Z-OTU table.

The relative abundance of each OTU is delimited using the horizontal grey lines and OTUs are colored according to genus. Rare taxonomic assignations (representing less than 0.5% of the abundance) are merged together. Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island.

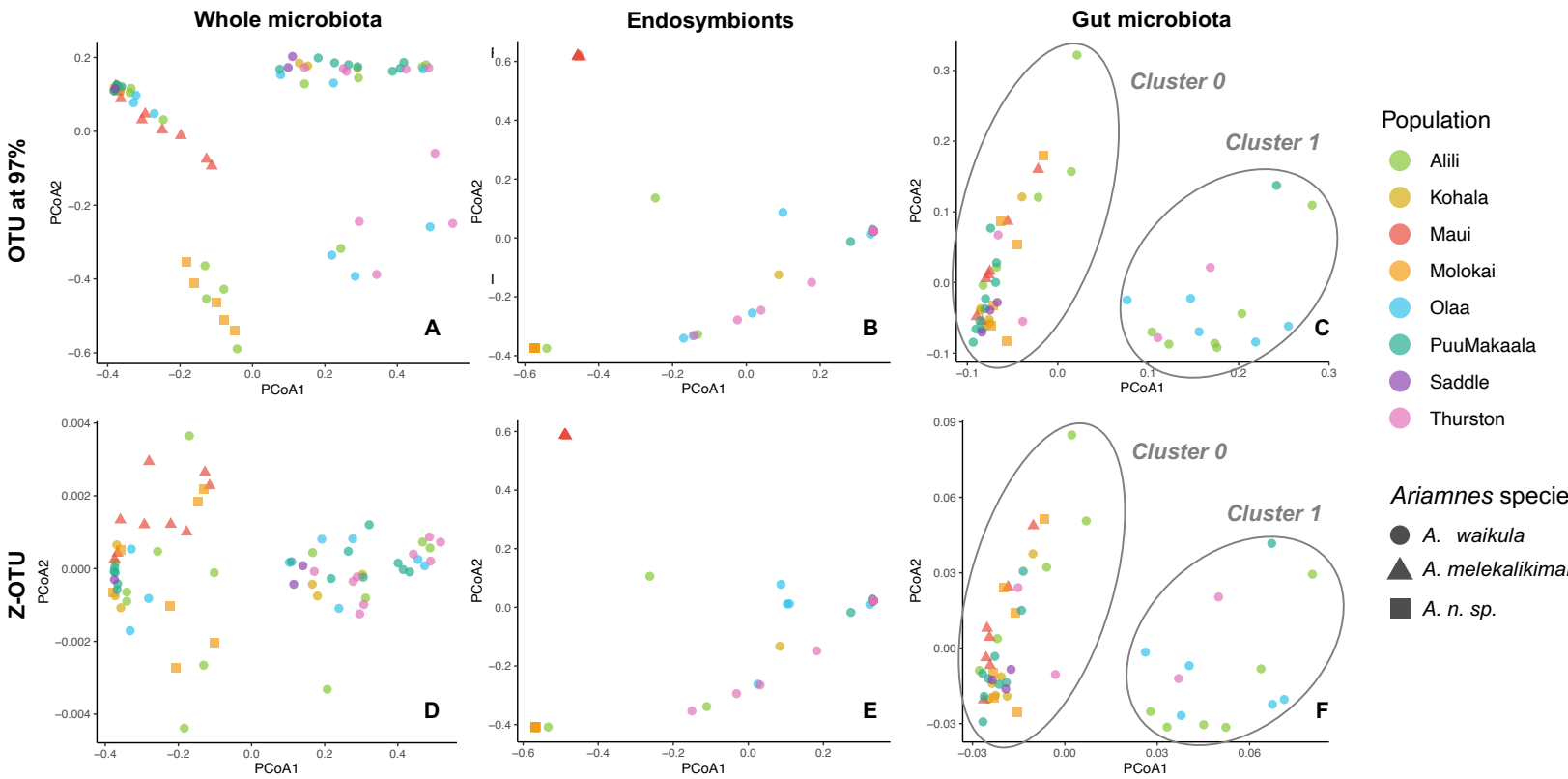


Figure S9: Principal coordinate analyses (PCoA) decomposition of the microbial communities according to the type of OTU clustering (OTU at 97% or Z-OTU) and the type of microbial communities (whole microbiota, endosymbionts, or gut microbiota).

Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island.

We noticed the presence of two clusters in the PCoA of the gut microbiota that are not explained by the identity of the host species or population: these two clusters (cluster 0 and cluster 1) are highlighted in grey.

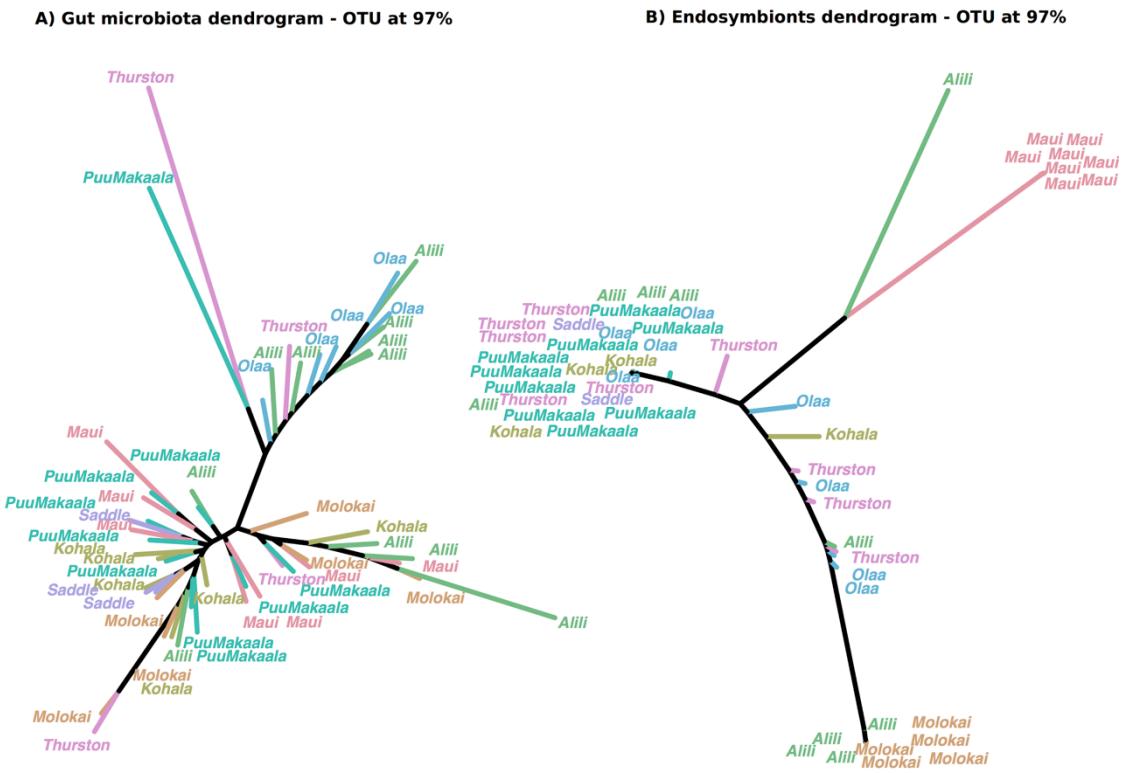


Figure S10: (A) Microbiota dendrograms reconstructed from the gut microbiota and (B) the endosymbiont community, for the OTU at 97%. Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island.

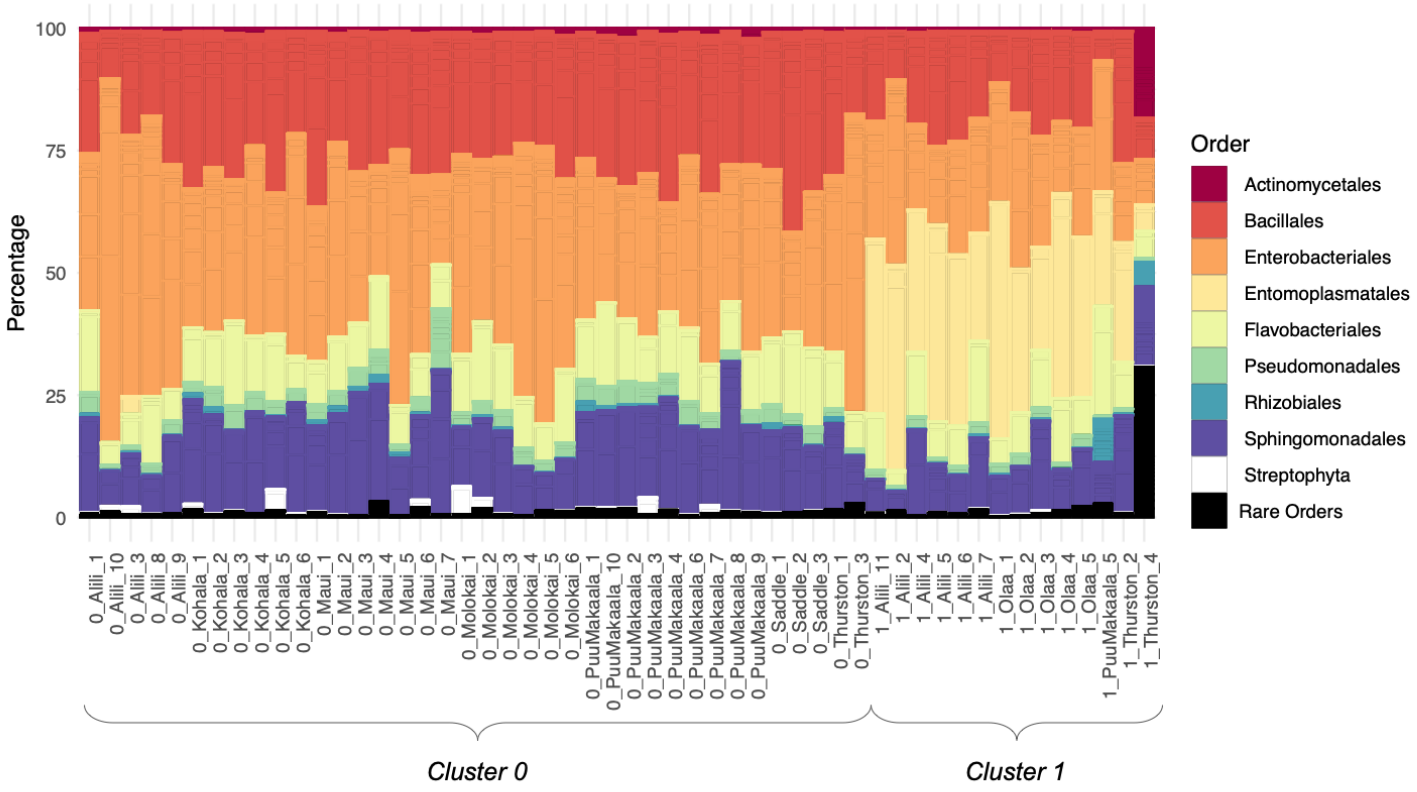


Figure S11: Relative abundances of the gut microbial community per spider individual from the rarefied Z-OTU table.

The relative abundance of each OTU is delimited using the horizontal grey lines and OTUs are colored according to order. Rare taxonomic assignations (representing less than 0.5% of the abundance) are merged together. Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) whereas the other populations correspond to *A. waikula* on Hawai’i Island.

Samples were sorted by cluster (cluster 0 and cluster 1), which emerged in the PCoA decomposition of the gut microbiota (see Fig. S9).

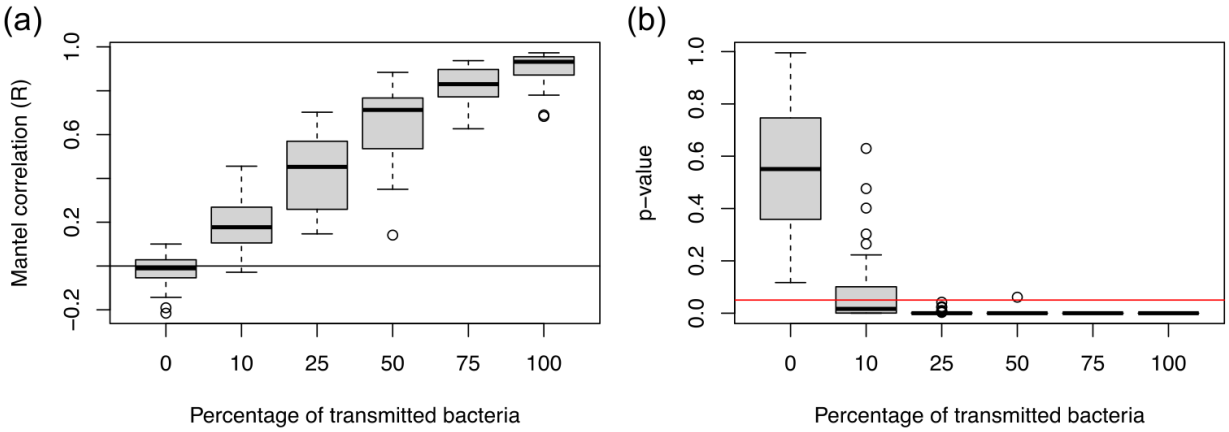


Figure S12: Pattern of phylosymbiosis as a function of the simulated percentage of transmitted bacteria:

For each percentage of vertically transmitted bacteria (0, 10, 25, 50, 75, or 100%), 30 mock microbiota communities, each of 100 OTUs, associated with *Ariamnes* spiders were simulated. Mantel tests between the *Ariamnes* phylogenetic distances and the simulated microbial beta diversity (Bray-Curtis) show (a) the Pearson correlation values (R) of the Mantel tests and (b) the corresponding p-values (obtained with 10,000 permutations – the red horizontal line indicates the significance threshold of 0.05, i.e. below this line there is a significant pattern of phylosymbiosis).

Note that the results are qualitatively similar when using UniFrac distances instead of Bray-Curtis dissimilarities.

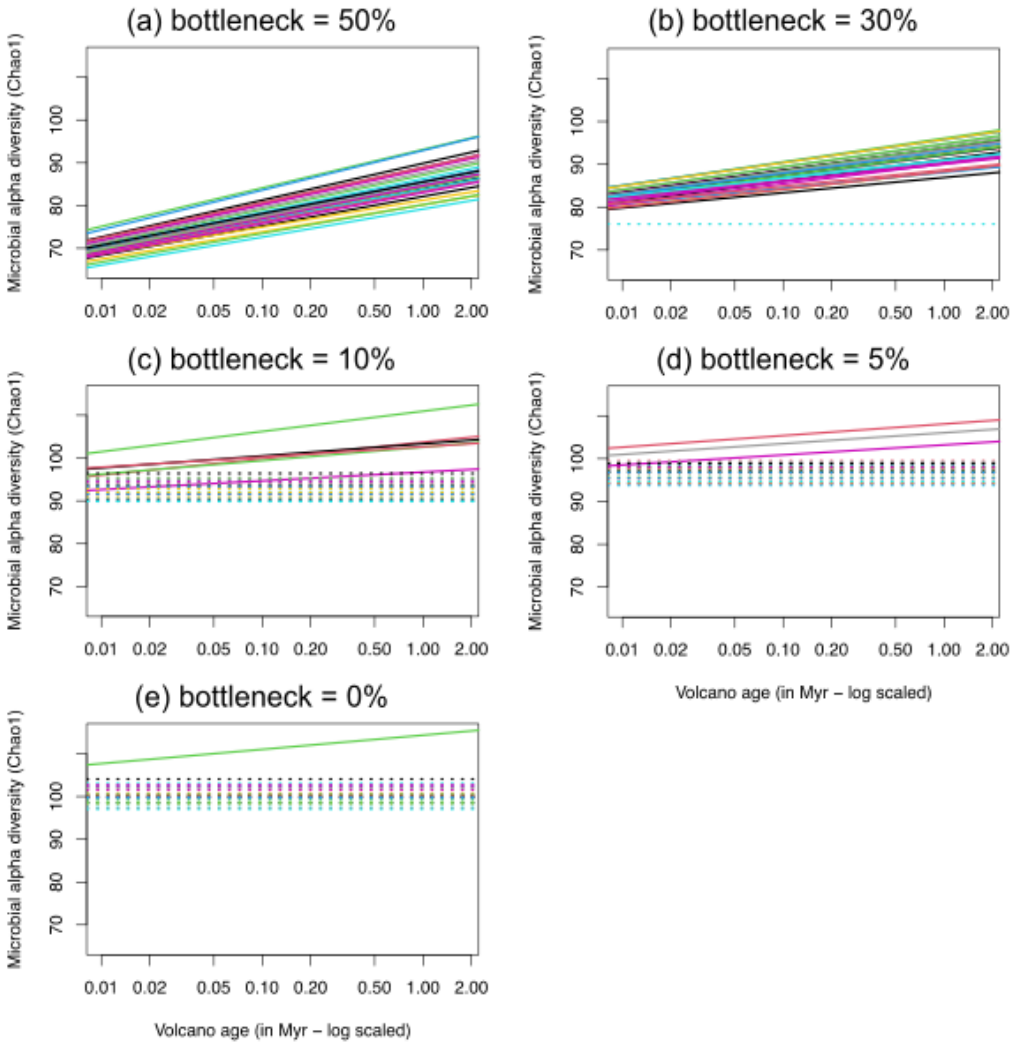


Figure S13: Presence of a bottleneck in *Ariamnes* microbial alpha diversities when simulating a bottleneck in the environmental pools of available microbes.

We simulated a diversity bottleneck along the island chronosequence (between the older Moloka‘i-Maui islands and the recent Hawai‘i Island) and randomly assembled mock *Ariamnes* microbiota from the pool of available microbes (see Methods). For each percentage of microbial OTUs experiencing the bottleneck (0, 5, 10, 30, or 50%), we replicated 30 mock microbiota communities and tested whether the alpha diversities (Chao1 index) of the *Ariamnes* microbiota increase with the volcano age (along the chronosequence) using linear mixed models. Plots show solid lines when the relationship was significant (p-value<0.05), dotted lines when not significant.

915 **Supplementary Tables**

916 **Table S1:** Specimen collection detail for population genetic and microbial samples.

Species	Eco-morph	Island	Population	Volcano (age)	Substrate age (years)	Latitude	Longitude	Collector	Date	Individuals (ddRAD)	Microbiota
A. waikula	Gold	Hawai'i	Kohala	Kohala, 0.43	300,000	20.1048	-155.7242	Rominger	2014-01-12	18	6
A. waikula	Gold	Hawai'i	Saddle	Mauna Loa, Kea 0.01-0.38	4,000	19.6746	-155.3316	Rominger	2014-01-07	6	3
A. waikula	Gold	Hawai'i	Puu Makaala	Kilauea, 0.004	11,000	19.3748	-155.2202	Rominger	2014-01-05	18	13
A. waikula	Gold	Hawai'i	Thurston	Kilauea, 0.004	600	19.4078	-155.2388	Rominger	2014-01-08	13	9
A. waikula	Gold	Hawai'i	Ola'a	Kilauea, 0.004	7,500	19.4530	-155.2466	Rominger	2014-01-09	16	11
A. waikula	Gold	Hawai'i	Alili	Mauna Loa, 0.01	20,000	19.2309	-155.5134	Rominger	2014-01-04	18	13
A. hiwa	Brown	Hawai'i	Puu Makaala	Kilauea, 0.004	11,000	19.3748	-155.2202	Rominger	2014-01-05	1	N/A
A. hiwa	Brown	Hawai'i	Thurston	Kilauea, 0.004	600	19.4078	-155.2388	Rominger	2014-01-08	1	N/A
A. meleka-likimaka	Gold	W. Maui	Puu Kukui	Puu Kukui, 1.3	1,500,000	20.9194	-156.5972	Gillespie	2012-06-29	16	9
A. n. sp	Gold	Moloka'i	Kamakou	Moloka'i, 1.8	1,400,000	21.1108	-156.8906	Gillespie	2012-06-23	16	7

917
918

919 **Table S2:** Specimen collection detail for population genetic and microbial samples.

Island	Population	Name ddRAD	Name microbiota	SeqID	Pool	Sample ID
Hawai'i	Pu'u Maka'ala	1	83	RGCC001A_S25_AACCA	A2	AJR923
Hawai'i	Pu'u Maka'ala	2	NA	RGCC001A_S25_AAGGA	A8	AJR913
Hawai'i	Pu'u Maka'ala	3	82	RGCC001A_S25_ACACA	A10	AJR948
Hawai'i	Pu'u Maka'ala	4	NA	RGCC001A_S25_ACTGG	A13	AJR956
Hawai'i	Pu'u Maka'ala	5	70	RGCC001A_S25_ACGGT	A12	AJR955
Hawai'i	Pu'u Maka'ala	6	65	RGCC001A_S25_ACTTC	A14	AJR964
Hawai'i	Pu'u Maka'ala	7	47	RGCC001A_S25_AGCTA	A9	AJR915
Hawai'i	Pu'u Maka'ala	8	NA	RGCC001A_S25_ATACG	A15	AJR965
Hawai'i	Pu'u Maka'ala	9	59	RGCC001A_S25_ATGAG	A16	AJR1017
Hawai'i	Pu'u Maka'ala	10	NA	RGCC001A_S25_CAACC	A6	AJR950
Hawai'i	Pu'u Maka'ala	11	85	RGCC001A_S25_CGATC	A3	AJR924
Hawai'i	Pu'u Maka'ala	12	80	RGCC001A_S25_CTGCG	A17	AJR1033
Hawai'i	Pu'u Maka'ala	13	60	RGCC001A_S25_CTGTC	A18	AJR1034
Hawai'i	Pu'u Maka'ala	14	52	RGCC001A_S25_CTTGG	A19	AJR1036
Hawai'i	Pu'u Maka'ala	15	74	RGCC001A_S25_GACAC	A20	AJR1037
Hawai'i	Pu'u Maka'ala	16	NA	RGCC001A_S25_GCATG	A1	AJR922
Hawai'i	Pu'u Maka'ala	17	NA	RGCC001A_S25_GGTTG	A7	AJR951
Hawai'i	Pu'u Maka'ala	18	NA	RGCC001A_S25_TCGAT	A4	AJR946
Hawai'i	Pu'u Maka'ala	19	61	RGCC001A_S25_TGCAT	A5	AJR947
Hawai'i	Alili	20	66	RGCC001B_S26_AACCA	B2	AJR839
Hawai'i	Alili	21	62	RGCC001B_S26_AAGGA	B8	AJR850
Hawai'i	Alili	22	88	RGCC001B_S26_AATTA	B11	AJR855
Hawai'i	Alili	23	100	RGCC001B_S26_ACACA	B10	AJR854
Hawai'i	Alili	24	45	RGCC001B_S26_ACGGT	B12	AJR857
Hawai'i	Alili	25	NA	RGCC001B_S26_ACTTC	B14	AJR859
Hawai'i	Alili	26	78	RGCC001B_S26_AGCTA	B9	AJR852
Hawai'i	Alili	27	118	RGCC001B_S26_ATACG	B15	AJR860
Hawai'i	Alili	28	104	RGCC001B_S26_ATGAG	B16	AJR861
Hawai'i	Alili	29	98	RGCC001B_S26_CAACC	B6	AJR848
Hawai'i	Alili	30	NA	RGCC001B_S26_CGATC	B3	AJR844
Hawai'i	Alili	31	107	RGCC001B_S26_CTGCG	B17	AJR891
Hawai'i	Alili	32	NA	RGCC001B_S26_CTGTC	B18	AJR893

Hawai'i	Alili	33	NA	RGCC001B_S26_GACAC	B20	AJR897
Hawai'i	Alili	34	NA	RGCC001B_S26_GCATG	B1	AJR838
Hawai'i	Alili	35	102	RGCC001B_S26_GGTTG	B7	AJR849
Hawai'i	Alili	36	NA	RGCC001B_S26_TCGAT	B4	AJR846
Hawai'i	Alili	37	NA	RGCC001B_S26_TGCAT	B5	AJR847
Hawai'i	Kohala	38	NA	RGCC001C_S27_AACCA	C2	AJR1413
Hawai'i	Kohala	39	55	RGCC001C_S27_AAGGA	C8	AJR1425
Hawai'i	Kohala	40	120	RGCC001C_S27_AATTA	C11	AJR1428
Hawai'i	Kohala	41	99	RGCC001C_S27_ACACA	C10	AJR1427
Hawai'i	Kohala	42	84	RGCC001C_S27_ACTGG	C13	AJR1430
Hawai'i	Kohala	43	NA	RGCC001C_S27_ACTTC	C14	AJR1476
Hawai'i	Kohala	44	NA	RGCC001C_S27_AGCTA	C9	AJR1426
Hawai'i	Kohala	45	NA	RGCC001C_S27_ATACG	C15	AJR1477
Hawai'i	Kohala	46	NA	RGCC001C_S27_ATGAG	C16	AJR1478
Hawai'i	Kohala	47	NA	RGCC001C_S27_CAACC	C6	AJR1417
Hawai'i	Kohala	48	NA	RGCC001C_S27_CGATC	C3	AJR1414
Hawai'i	Kohala	49	NA	RGCC001C_S27_CTGCG	C17	AJR1518
Hawai'i	Kohala	50	NA	RGCC001C_S27_CTGTC	C18	AJR1519
Hawai'i	Kohala	51	NA	RGCC001C_S27_CTTGG	C19	AJR1520
Hawai'i	Kohala	52	NA	RGCC001C_S27_GCATG	C1	AJR1412
Hawai'i	Kohala	53	NA	RGCC001C_S27_GGTTG	C7	AJR1418
Hawai'i	Kohala	54	117	RGCC001C_S27_TCGAT	C4	AJR1415
Hawai'i	Kohala	55	NA	RGCC001C_S27_TGCAT	C5	AJR1416
Hawai'i	Ola‘a	56	NA	RGCC001D_S28_AACCA	D2	AJR1151
Hawai'i	Ola‘a	57	NA	RGCC001D_S28_AAGGA	D8	AJR1158
Hawai'i	Ola‘a	58	67	RGCC001D_S28_AATTA	D11	AJR1161
Hawai'i	Ola‘a	59	64	RGCC001D_S28_ACACA	D10	AJR1160
Hawai'i	Ola‘a	60	81	RGCC001D_S28_ACTGG	D13	AJR1187
Hawai'i	Ola‘a	61	49	RGCC001D_S28_ACTTC	D14	AJR1188
Hawai'i	Ola‘a	62	58	RGCC001D_S28_AGCTA	D9	AJR1159
Hawai'i	Ola‘a	63	48	RGCC001D_S28_ATACG	D15	AJR1190
Hawai'i	Ola‘a	64	72	RGCC001D_S28_CAACC	D6	AJR1156
Hawai'i	Ola‘a	65	53	RGCC001D_S28_CGATC	D3	AJR1152
Hawai'i	Ola‘a	66	119	RGCC001D_S28_CTGCG	D17	AJR1192

Hawai'i	Ola‘a	67	NA	RGCC001D_S28_CTGTC	D18	AJR1232
Hawai'i	Ola‘a	68	NA	RGCC001D_S28_CTTGG	D19	AJR1233
Hawai'i	Ola‘a	69	NA	RGCC001D_S28_GCATG	D1	AJR1150
Hawai'i	Ola‘a	70	NA	RGCC001D_S28_GGTTG	D7	AJR1157
Hawai'i	Ola‘a	71	NA	RGCC001D_S28_TCGAT	D4	AJR1153
Hawai'i	Ola‘a	72	109	RGCC001D_S28_TGCAT	D5	AJR1155
Hawai'i	Thurston	105	NA	RGCC001H_S31_AACCA	H2	AJR1061
Hawai'i	Thurston	106	44	RGCC001H_S31_AAGGA	H8	AJR1067
Hawai'i	Thurston	107	111	RGCC001H_S31_AATTA	H11	AJR1107
Hawai'i	Thurston	108	87	RGCC001H_S31_ACACA	H10	AJR1106
Hawai'i	Thurston	109	NA	RGCC001H_S31_ACGGT	H12	AJR1108
Hawai'i	Thurston	110	NA	RGCC001H_S31_ACTGG	H13	AJR1109
Hawai'i	Thurston	111	NA	RGCC001H_S31_ACTTC	H14	AJR1110
Hawai'i	Thurston	112	NA	RGCC001H_S31_AGCTA	H9	AJR1068
Hawai'i	Thurston	114	108	RGCC001H_S31_CGATC	H3	AJR1062
Hawai'i	Thurston	115	115	RGCC001H_S31_GCATG	H1	AJR1060
Hawai'i	Thurston	116	79	RGCC001H_S31_GGTTG	H7	AJR1066
Hawai'i	Thurston	117	97	RGCC001H_S31_TGCAT	H5	AJR1064
Hawai'i	Saddle	118	NA	RGCC001I_S32_AACCA	H2	AJRA2
Hawai'i	Saddle	119	NA	RGCC001I_S32_CAACC	H6	AJRA6
Hawai'i	Saddle	120	63	RGCC001I_S32_CGATC	H3	AJRA3
Hawai'i	Saddle	121	103	RGCC001I_S32_GCATG	H1	AJRA1
Hawai'i	Saddle	122	NA	RGCC001I_S32_TCGAT	H4	AJRA4
Hawai'i	Saddle	123	57	RGCC001I_S32_TGCAT	H5	AJRA5
Hawai'i	Thurston	NA	42	RGCC001H_S31_TCGAT	H4	AJR1063
Hawai'i	Kohala	NA	51	RGCC001C_S27_ACGGT	C12	AJR1429
Hawai'i	Thurston	NA	69	RGCC001H_S31_CAACC	H6	AJR1065
Hawai'i	Alili	NA	71	RGCC001B_S26_ACTGG	B13	AJR858
Hawai'i	Alili	NA	77	RGCC001B_S26_CTTGG	B19	AJR894
Hawai'i	Pu‘u Maka‘ala	NA	101	RGCC001A_S25_AATTA	A11	AJR949
Hawai'i	Ola‘a	NA	113	RGCC001D_S28_ACGGT	D12	AJR1183
Hawai'i	Kohala	NA	NA	RGCC001C_S27_GACAC	C20	AJR1521
Hawai'i	Ola‘a	NA	NA	RGCC001D_S28_ATGAG	D16	AJR1191
Moloka‘i	Moloka‘i	73	NA	RGCC001F_S29_AACCA	F2	RGGA10

Moloka‘i	Moloka‘i	74	NA	RGCC001F_S29_AAGGA	F8	RGG43
Moloka‘i	Moloka‘i	75	106	RGCC001F_S29_AATTA	F11	RGG46
Moloka‘i	Moloka‘i	76	86	RGCC001F_S29_ACACA	F10	RGG45
Moloka‘i	Moloka‘i	77	68	RGCC001F_S29_ACGGT	F12	RGG47
Moloka‘i	Moloka‘i	78	73	RGCC001F_S29_ACTGG	F13	RGG48
Moloka‘i	Moloka‘i	79	NA	RGCC001F_S29_ACTTC	F14	RGG5
Moloka‘i	Moloka‘i	80	114	RGCC001F_S29_AGCTA	F9	RGG44
Moloka‘i	Moloka‘i	81	NA	RGCC001F_S29_ATACG	F15	RGG6
Moloka‘i	Moloka‘i	82	NA	RGCC001F_S29_ATGAG	F16	RGG9
Moloka‘i	Moloka‘i	83	105	RGCC001F_S29_CAACC	F6	RGG41
Moloka‘i	Moloka‘i	84	NA	RGCC001F_S29_CGATC	F3	RGGA11
Moloka‘i	Moloka‘i	85	NA	RGCC001F_S29_GCATG	F1	RGGA9
Moloka‘i	Moloka‘i	86	112	RGCC001F_S29_GGTTG	F7	RGG42
Moloka‘i	Moloka‘i	87	NA	RGCC001F_S29_TCGAT	F4	RGG2
Moloka‘i	Moloka‘i	88	NA	RGCC001F_S29_TGCAT	F5	RGG3
Maui	W_Maui	89	56	RGCC001G_S30_AACCA	G2	RGG121
Maui	W_Maui	90	NA	RGCC001G_S30_AAGGA	G8	RGG135
Maui	W_Maui	91	43	RGCC001G_S30_AATTA	G11	RGG138
Maui	W_Maui	92	50	RGCC001G_S30_ACACA	G10	RGG137
Maui	W_Maui	93	116	RGCC001G_S30_ACGGT	G12	RGG151
Maui	W_Maui	94	NA	RGCC001G_S30_ACTGG	G13	RGG152
Maui	W_Maui	95	54	RGCC001G_S30_ACTTC	G14	RGG153
Maui	W_Maui	96	46	RGCC001G_S30_AGCTA	G9	RGG136
Maui	W_Maui	97	76	RGCC001G_S30_ATACG	G15	RGG154
Maui	W_Maui	98	41	RGCC001G_S30_ATGAG	G16	RGG155
Maui	W_Maui	99	NA	RGCC001G_S30_CAACC	G6	RGG133
Maui	W_Maui	100	110	RGCC001G_S30_CGATC	G3	RGG122
Maui	W_Maui	101	NA	RGCC001G_S30_GCATG	G1	RGGA2
Maui	W_Maui	102	NA	RGCC001G_S30_GGTTG	G7	RGG134
Maui	W_Maui	103	NA	RGCC001G_S30_TCGAT	G4	RGG131
Maui	W_Maui	104	NA	RGCC001G_S30_TGCAT	G5	RGG132
Hawai’i	Ola‘a	NA	NA	RGCC001D_S28_GACAC	D20	AJR1186
Hawai’i	Thurston	113	NA	RGCC001H_S31_ATACG	H15	AJR1111

920
921

Table S3: F_{ST} values for ddRAD data from *Ariamnes waikula* spiders from Hawai'i.

	Alili	Kohala	Ola‘a	Thurston	Saddle	Moloka‘i	Maui
Pu‘u Maka‘ala	0.0650	0.1466	0.0291	0.0461	0.1066	0.4870	0.5454
Alili		0.1394	0.0544	0.0629	0.0978	0.4829	0.5407
Kohala			0.1367	0.1496	0.1315	0.4742	0.5264
Ola‘a				0.0334	0.0937	0.4817	0.5404
Thurston					0.1120	0.4815	0.5468
Saddle						0.4629	0.5410
Moloka‘i							0.4781

Table S4: Average Tajima's D values for each population of *A. waikula* (Hawai'i) and *A. melekalikimaka (*Maui) and *A. n. spp.* (Moloka'i).

Population	Average	Min	Max
Alili	-0.77	-2.68	3.33
Ola'a	-0.95	-2.83	3.84
Pu'u Maka'ala	-1.00	-2.77	3.25
Saddle	-0.61	-2.97	3.79
Thurston	-0.74	-2.87	3.73
Kohala	-0.71	-2.82	3.52
Maui	-0.06	-2.63	3.94
Moloka'i	-0.23	-2.46	3.66

931 **Table S5:** Mean alpha diversity of the microbial communities per host populations. Alpha
932 diversities were computed on rarefied OTU tables defined at 97% or Z-OTUs using the Chao 1
933 index.
934

Island / species	Population	Number individuals	OTU at 97%		Z-OTU	
			Alpha diversity	Standard deviation	Alpha diversity	Standard deviation
Hawai'i *A. waikula*	**Alili**	13	42.4	7.2	95.4	4.0
Hawai'i *A. waikula*	**Kohala**	6	55.8	12.3	114.8	7.8
Hawai'i *A. waikula*	**Ola'a**	9	41.6	3.0	96.3	8.5
Hawai'i *A. waikula*	**Pu'u Maka'ala**	13	49.5	3.2	108.0	3.4
Hawai'i *A. waikula*	**Saddle**	3	44.8	1.6	103.1	5.2
Hawai'i *A. waikula*	**Thurston**	9	40.8	4.6	103.2	5.7
Maui *A. melekelikimaka*	**Puu Kukui**	9	46.0	5.7	106.6	4.4
Moloka'i *A. n. sp.*	**Kamakou**	7	52.2	6.7	106.6	5.3

935

936 **Table S6:** PERMANOVA testing the effect of host populations on the dissimilarities between
937 microbial communities. The adjusted R-squared and the associated p-value are shown for each
938 test. Significant relationships indicate a differentiation of the microbial communities according to
939 the host populations.
940

Populations	OTU at 97%			Z-OTU		
	Whole microbiota	Gut microbiota	Endo-symbionts	Whole microbiota	Gut microbiota	Endo-symbionts
All populations (from different species)	0.37905 (0.0001)	0.31109 (0.0001)	0.7146 (0.0001)	0.37709 (0.0001)	0.31739 (0.0001)	0.7115 (0.0001)
A. waikula populations within Hawai'i Island	0.21715 (0.0062)	0.29671 (0.0007)	0.28093 (0.0190)	0.21542 (0.0057)	0.29 (0.0003)	0.29072 (0.0120)

941

942 **Table S7:** Mantel tests comparing the microbial beta diversity dissimilarities to the genetic
943 distances, the phylogenetic distances of their associated *Ariamnes,* and the geographical
944 distances between the sampling sites. Beta diversity dissimilarities were computed on rarefied
945 OTU tables defined at 97% using the Bray-Curtis dissimilarities. Bold values represent significant
946 correlations. Mantel tests were either performed on all populations (between *Ariamnes* species
947 and *A. waikula* populations) or only on *A. waikula* populations.

948

Matrix 1	Matrix 2	All populations		*A. waikula* populations	
		Correlation	p-value	Correlation	p-value
Whole microbiota beta diversity distances	*Ariamnes* genetic distances	**0.312**	**9.9e-05**	0.12	0.07
	Ariamnes phylogenetic distances	**0.323**	**9.9e-05**	0.03	0.32
	Geographic distances	**0.292**	**9.9e-05**	0.06	0.16
Endosymbionts beta diversity distances	*Ariamnes* genetic distances	**0.620**	**9.9e-05**	0.01	0.50
	Ariamnes phylogenetic distances	**0.646**	**9.9e-05**	0.06	0.26
	Geographic distances	**0.565**	**9.9e-05**	0.12	0.14
Gut microbiota beta diversity distances	*Ariamnes* genetic distances	-0.096	0.84	-0.04	0.67
	Ariamnes phylogenetic distances	-0.095	0.86	-0.13	0.91
	Geographic distances	-0.030	0.62	-0.07	0.73

949

References

Adams AS, Aylward FO, Adams SM*, et al.* (2013) Mountain pine beetles colonizing historical and naive host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. *Appl. Environ. Microbiol.* **79**, 3468-3475.

Bailly-Bechet M, Martins-Simões P, Szöllősi GJ*, et al.* (2017) How long does Wolbachia remain on board? *Molecular Biology and Evolution* **34**, 1183-1193.

Baldo L, Ayoub NA, Hayashi CY, Russell JA, Stahlhut JK, Werren JH (2008), Insight into the routes of Wolbachia invasion: high levels of horizontal transfer in the spider genus Agelenopsis revealed by Wolbachia strain and mitochondrial DNA diversity. Mol Ecol.17(2):557-69.

Bi K, Linderoth T, Vanderpool D*, et al.* (2013) Unlocking the vault: next-generation museum population genomics. *Molecular ecology* **22**, 6018-6032.

Blankers T, Oh KP, Bombarely A, Shaw KL (2018) The genomic architecture of a rapid island radiation: recombination rate variation, chromosome structure, and genome assembly of the Hawaiian cricket Laupala. *Genetics* **209**, 1329-1344.

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114-2120.

Bordenstein S, O'Hara F & Werren J (2001), Wolbachia-induced incompatibility precedes other hybrid incompatibilities in Nasonia. Nature 409, 707–71

Brooks AW, Kohl KD, Brucker RM, van Opstal EJ, Bordenstein SR (2016) Phylosymbiosis: relationships and functional effects of microbial communities across host evolutionary history. *PLoS biology* **14**, e2000225.

Brucker RM, Bordenstein SR (2012) Speciation by symbiosis. *Trends in ecology & evolution* **27**, 443-451.

Caporaso JG, Kuczynski J, Stombaugh J*, et al.* (2010) QIIME allows analysis of high-throughput community sequencing data. *Nature methods* **7**, 335.

Carson, H.L., Lockwood, J.P. and Craddock, E.M., 1990. Extinction and recolonization of local populations on a growing shield volcano. *Proceedings of the National Academy of Sciences*, *87*(18), pp.7055-7057.

Carson H, Clague D (1995) Geology and biogeography of the Hawaiian Islands. In Wagner, WL & Funk, VA (eds). Hawaiian Biogeography, Evolution on a Hot Spot Archipelago, Smithsonian Institution Press, Washington: pp. 14–29.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular ecology* **22**, 3124-3140.

Crawford JE, Lazzaro BP (2012) Assessing the accuracy and power of population genetic inference from low-pass next-generation sequencing data. *Frontiers in genetics* **3**.

Croucher PJ, Oxford GS, Lam A, Mody N, Gillespie RG (2012) Colonization history and population genetics of the color-polymorphic Hawaiian happy-face spider Theridion grallator (Araneae, Theridiidae). *Evolution: International Journal of Organic Evolution* **66**, 2815-2833.

Degnan PH, Lazarus AB, Brock CD, Wernegreen JJ (2004) Host–symbiont stability and fast evolutionary rates in an ant–bacterium association: cospeciation of Camponotus species and their endosymbionts, Candidatus Blochmannia. *Systematic biology* **53**, 95-110.

Duron O, Binetruy F, Noël V*, et al.* (2017) Evolutionary changes in symbiont community structure in ticks. *Molecular ecology* **26**, 2905-2921.

Duron O, Bouchon D, Boutin S*, et al.* (2008) The diversity of reproductive parasites among arthropods: Wolbachiado not walk alone. *BMC biology* **6**, 27.

Edgar RC (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods* **10**, 996.

Eldon J, Bellinger MR, Price DK (2019) Hawaiian picture-winged Drosophila exhibit adaptive population divergence along a narrow climatic gradient on Hawai'i Island. *Ecology and evolution* **9**, 2436.

Engel P, Moran NA (2013) The gut microbiota of insects–diversity in structure and function. *FEMS microbiology reviews* **37**, 699-735.

Fu L, Niu B, Zhu Z, Wu S, Li W (2012) CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150-3152.

Fumagalli M, Vieira FG, Linderoth T, Nielsen R (2014) ngsTools: methods for population genetics analyses from next-generation sequencing data. *Bioinformatics* **30**, 1486-1487.

Gasparich GE (2014), The Family Entomoplasmataceae. In: Rosenberg E., DeLong E.F., Lory S., Stackebrandt E., Thompson F. (eds) The Prokaryotes. Springer, Berlin, Heidelberg.

Guerrero R, Margulis L, Berlanga M (2013) Symbiogenesis: the holobiont as a unit of evolution. Int Microbiol. 16(3):133-43

Gibson J, Shokralla S, Porter TM*, et al.* (2014) Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasystematics. *Proceedings of the National Academy of Sciences* **111**, 8007-8012.

Gillespie RG, Baldwin BG, Waters JM, Fraser CI, Nikula R, and Roderick GK (2012) "Long-distance dispersal: a framework for hypothesis testing." Trends in ecology & evolution 27, no. 1: 47-56.

Gillespie RG, Benjamin SP, Brewer MS, Rivera MAJ, Roderick GK (2018) Repeated Diversification Of Ecomorphs In Hawaiian Stick Spiders. *Current Biology* **28**, 941-947.

Gillespie RG, Rivera MA (2007) Free-living spiders of the genus Argyrodes (Araneae: Theridiidae) in Hawai'i. *Journal of Arachnology* **35**, 11-37.

Goodacre SL, Martin OY, Bonte D*, et al.* (2009) Microbial modification of host long-distance dispersal capacity. *Bmc Biology* **7**, 32.

Goodacre SL, Martin OY, Thomas CG, Hewitt GM (2006) Wolbachia and other endosymbiont infections in spiders. *Molecular Ecology* **15**, 517-527.

Goodman KR, Prost S, Bi K, Brewer MS, Gillespie RG (2019) Host and geography together drive early adaptive radiation of Hawaiian planthoppers. *Molecular ecology* **28**, 4513-4528.

Groussin M, Mazel F, Sanders JG, Smillie CS, Lavergne S, Thuiller W, & Alm EJ (2017). Unraveling the processes shaping mammalian gut microbiomes over evolutionary time. Nature Communications, 8(1), 163–185. doi:10.1038/ncomms14319

Hammer TJ, Dickerson JC, Fierer N (2015). Evidence-based recommendations on storing and handling specimens for analyses of insect microbiota. *PeerJ*, *3*, e1190. doi:10.7717/peerj.1190

Hammer TJ, Janzen DH, Hallwachs W, Jaffe SP, Fierer N (2017) Caterpillars lack a resident gut microbiome. *Proceedings of the National Academy of Sciences* **114**, 9641-9646.

Hammer TJ, Sanders JG and Fierer N (2019) Not all animals need a microbiome. FEMS microbiology letters, 366(10): p.fnz117

Hansen AK, Moran NA (2014) The impact of microbial symbionts on host plant utilization by herbivorous insects. *Molecular ecology* **23**, 1473-1496.

Hiller AE, Koo MS, Goodman KR*, et al.* (2019) Niche conservatism predominates in adaptive radiation: comparing the diversification of Hawaiian arthropods using ecological niche modelling. *Biological Journal of the Linnean Society* **127**, 479-492.

Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS (2018) UFBoot2: improving the ultrafast bootstrap approximation. *Molecular biology and evolution* **35**, 518-522.

Holland BS & Cowie RH (2007) A geographic mosaic of passive dispersal: population structure in the endemic Hawaiian amber snail Succinea caduca (Mighels, 1845). Molecular Ecology, 16(12)

Hornung B, Zwittink RD, Kuijper EJ (2019), Issues and current standards of controls in microbiome research, FEMS Microbiology Ecology, 95: fiz045

Hoy MA, Jeyaprakash A (2005) Microbial diversity in the predatory mite Metaseiulus occidentalis (Acari: Phytoseiidae) and its prey, Tetranychus urticae (Acari: Tetranychidae). *Biological Control* **32**, 427-441.

Huson, D. H., & Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Molecular biology and evolution*, *23*(2), 254-267.

Kalyaanamoorthy S, Minh BQ, Wong TK, von Haeseler A, Jermiin LS (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nature methods* **14**, 587.

Kennedy SR, Dawson TE, Gillespie RG (2018) Stable isotopes of Hawaiian spiders reflect substrate properties along a chronosequence. *PeerJ* **6**, e4527.

Kennedy SR, Tsau S, Gillespie R, Krehenwinkel H (2020) Are you what you eat? A highly transient and prey-influenced gut microbiome in the grey house spider Badumna longinqua. *Molecular Ecology*.

Koga R, Tsuchida T, Fukatsu T (2003) Changing partners in an obligate symbiosis: a facultative endosymbiont can compensate for loss of the essential endosymbiont Buchnera in an aphid. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 2543-2550.

Korneliussen TS, Albrechtsen A, Nielsen R (2014) ANGSD: analysis of next generation sequencing data. *BMC bioinformatics* **15**, 356.

Koskella B & Vos M (2015) Adaptation in natural microbial populations. Annual Review of Ecology, Evolution, and Systematics, 46, 503-522.

Lange V, Böhme I, Hofmann J*, et al.* (2014) Cost-efficient high-throughput HLA typing by MiSeq amplicon sequencing. *BMC genomics* **15**, 63.

LangenheimAVM, and Clague DA (1987) The Hawaiian-Emperor volcanic chain, Geologic evolution: Stratigraphic framework of volcanic rocks of the Hawaiian Islands Vol. 1, Chap 1, Part 2 IN: Volcanism in Hawaii: Papers to commemorate the 75th anniversary of the founding of the Hawaiian Volcano Observatory, R. W. Decker , T. L.Wright , and P. H. Stauffer , Eds., USGS Prof. Paper 1350, Vol. 1, 771–780

Larrivee M & Buddle CM (2011) Ballooning propensity of canopy and understorey spiders in a mature temperate hardwood forest. Ecological Entomology, 36(2)

Lefort, V., Desper, R., & Gascuel, O. (2015). FastME 2.0: a comprehensive, accurate, and fast distance-based phylogeny inference program. *Molecular biology and evolution*, **32**(10), 2798-2800.

Kwong WK, Medina LA, Koch H, Sing KW, Soh EJY, Ascher JS, Jaffé R and Moran NA, (2017) Dynamic microbiome evolution in social bees. Science Advances, 3(3): p.e1600513

Li H, Handsaker B, Wysoker A*, et al.* (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078-2079.

Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658-1659.

Lim SJ & Bordenstein SR (2020) An introduction to phylosymbiosi sProc. R. Soc. B. 2872019290020192900

Lozupone C & Knight R (2005) UniFrac: a New Phylogenetic Method for Comparing Microbial Communities Applied and Environmental Microbiology 71 (12): 8228-8235

Maas, D. L., Prost, S., Bi, K., Smith, L. L., Armstrong, E. E., Aji, L. P., Toha, A.H.A., Gillespie, R.G., & Becking, L. E. (2018). Rapid divergence of mussel populations despite incomplete barriers to dispersal. *Molecular ecology*, *27*(7), 1556-1571.

Magoč T, Salzberg SL (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**, 2957-2963.

Margulis L, Fester R (1991) *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis* Mit Press.

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal* **17**, pp. 10-12.

Mazel F, Davis KM, Loudon A*, et al.* (2018) Is host filtering the main driver of phylosymbiosis across the tree of life? *mSystems* **3**, e00097-00018.

McFall-Ngai M, Hadfield MG, Bosch TC*, et al.* (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proceedings of the National Academy of Sciences* **110**, 3229-3236.

McKenna A, Hanna M, Banks E*, et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research* **20**, 1297-1303.

Minard G, TRAN F-H, Goubert C*, et al.* (2015) French invasive Asian tiger mosquito populations harbor reduced bacterial microbiota and genetic diversity compared to Vietnamese autochthonous relatives. *Frontiers in Microbiology* **6**, 970.

Moeller AH, Caro-Quintero A, Mjungu D*, et al.* (2016) Cospeciation of gut microbiota with hominids. *Science* **353**, 380-382.

Moran NA, Munson MA, Baumann P, Ishikawa H. (1993) A molecular clock in endosymbiotic bacteria is calibrated using the insect hosts. Proc R Soc B Biol Sci; 253: 167–171.

Moran NA, Ochman H, and Hammer TJ (2019) Evolutionary and ecological consequences of gut microbial communities. Annual Review of Ecology, Evolution, and Systematics 50: 451-475

Moran NA, Sloan DB (2015) The hologenome concept: helpful or hollow? *PLoS biology* **13**.

Mouton L, Dedeine F, Henri H, Boulétreau M, Profizi N, Vavre F (2004). Virulence, multiple infections and regulation of symbiotic population in the Wolbochia-Asobara tabida symbiosis. Genetics 168:181-189.

Muegge BD, Kuczynski J, Knights D*, et al.* (2011) Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* **332**, 970-974.

Muir C & Price DK (2008) Population structure and genetic diversity in two species of Hawaiian picture-winged Drosophila. Molecular phylogenetics and evolution, 47(3)

Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ (2015) IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular biology and evolution* **32**, 268-274.

O'Connor TK, Humphrey PT, Lapoint RT, Whiteman NK, O'Grady PM (2014) Microbial interactions and the ecology and evolution of Hawaiian Drosophilidae. *Frontiers in microbiology* **5**, 616.

Ochman H, Elwyn S, Moran NA (1999) Calibrating bacterial evolution. Proc Natl Acad Sci; 96: 12638–12643.

Oksanen J, Kindt R, Pierre L, O'Hara B, Simpson GL, Solymos P, Stevens MH. HH, Wagner H, Blanchet FG, Kindt R, et al. 2016. vegan: Community Ecology Package, R package version 2.4-0. R package version 2.2-1.

Papadopulos AS, Kaye M, Devaux C*, et al.* (2014) Evaluation of genetic isolation within an island flora reveals unusually widespread local adaptation and supports sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **369**, 20130342.

Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20*(2), 289–290. doi:10.1093/bioinformatics/btg412

Perez-Lamarque B, Morlon H (2019). Characterizing symbiont inheritance during host–microbiota evolution: Application to the great apes gut microbiota. Molecular Ecology Resources, 1–14. doi:10.1111/1755-0998.13063

Petkova, D., Novembre, J., & Stephens, M. (2016). Visualizing spatial population structure with estimated effective migration surfaces. *Nature genetics*, *48*(1), 94-100.

Pylro VS, Roesch LFW, Morais DK*, et al.* (2014) Data analysis for 16S microbial profiling from different benchtop sequencing platforms. *Journal of microbiological methods* **107**, 30-37.

Ren T, Kahrl AF, Wu M, Cox RM (2016) Does adaptive radiation of a host lineage promote ecological diversity of its bacterial communities? A test using gut microbiota of Anolis lizards. *Molecular ecology* **25**, 4793-4804.

Rennison DJ, Rudman SM, Schluter D (2019) Parallel changes in gut microbiome composition and function during colonization, local adaptation and ecological speciation. *Proceedings of the Royal Society B* **286**, 20191911.

Roderick GK, Croucher PJ, Vandergast AG, Gillespie RG (2012) Species differentiation on a dynamic landscape: shifts in metapopulation genetic structure using the chronology of the Hawaiian archipelago. *Evolutionary Biology* **39**, 192-206.

Sanders JG, Powell S, Kronauer DJ*, et al.* (2014) Stability and phylogenetic correlation in gut microbiota: lessons from ants and apes. *Molecular Ecology* **23**, 1268-1283.

Sanderson MJ (2003) r8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. *Bioinformatics* **19**, 301-302.

Schrader C, Schielke A, Ellerbroek L, Johne R (2012) PCR inhibitors–occurrence, properties and removal. *Journal of applied microbiology* **113**, 1014-1026.

Simon C (1987) Hawaiian evolutionary biology: an introduction. Trends in ecology & evolution 2, no. 7: 175-178

Shaw KL, Gillespie RG (2016) Comparative phylogeography of oceanic archipelagos: Hotspots for inferences of evolutionary process. *Proceedings of the National Academy of Sciences* **113**, 7986-7993.

Sheffer MM, Uhl G, Prost S, Lueders T, Urich T, Bengtsson MM (2019). Tissue- and Population-Level Microbiome Analysis of the Wasp Spider Argiope bruennichi Identified a Novel Dominant Bacterial Symbiont. Microorganisms, 8(1), 8. doi:10.3390/microorganisms8010008

Sherrod DR, Sinton JM, Watkins SE, Brunt KM (2007) Geologic map of the State of Hawai'i. *US geological survey open-file report* **1089**.

Shropshire JD, Leigh B, Bordenstein SR (2020) Symbiont-mediated cytoplasmic incompatibility: what have we learned in 50 years? Elife. 9:e61989.

Skotte L, Korneliussen TS, Albrechtsen A (2013) Estimating individual admixture proportions from next generation sequencing data. *Genetics* **195**, 693-702.

Smit A, Hubley R, Green P (2004) RepeatMasker Open-3.0. 1996–2004. *Institute for Systems Biology*.

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**(9), 1312-1313.

Stahlhut JK, Desjardins CA, Clark ME, Baldo L, Russell JA, Werren JH, Jaenike J (2010). The mushroom habitat as an ecological arena for global exchange of Wolbochia. Molecular Ecology 19:1940-1952.

Su Q, Hu G, Yun Y, Peng Y (2019) Horizontal transmission of Wolbachia in Hylyphantes graminicola is more likely via intraspecies than interspecies transfer. *Symbiosis* **79**, 123-128.

Team RC (2018) R Foundation for Statistical Computing; Vienna, Austria: 2015. *R: A language and environment for statistical computing*, 2013.

Vanthournout B, Hendrickx F (2015) Endosymbiont dominated bacterial communities in a dwarf spider. *PLoS One* **10**.

1193 Vieira FG, Lassalle F, Korneliussen TS, Fumagalli M (2015) Improving the estimation of genetic
1194 distances from Next-Generation Sequencing data. Biological Journal of the Linnean
1195 Society **117**(1):139-149
1196 Wagner WL & Funk VA (1995) Hawaiian Biogeography: Evolution on a Hot Spot Archipelago.
1197 Smithsonian Institution Press, Washington, DC.
1198 Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation:
1199 a multiple matrix regression approach for quantifying geographic and ecological isolation.
1200 *Evolution* **67**, 3403-3411.
1201 Weber MG, Wagner CE, Best RJ, Harmon LJ, Matthews B (2017) Evolution in a community
1202 context: on integrating ecological interactions and macroevolution. *Trends in ecology &*
1203 *evolution* **32**, 291-304.
1204 White JA, Styer A, Rosenwald LC*, et al.* (2020) Endosymbiotic bacteria are prevalent and diverse
1205 in agricultural spiders. *Microbial ecology* **79**, 472-481.
1206 Whitehouse M, Agnarsson I, Miyashita T*, et al.* (2002) *Argyrodes*: phylogeny, sociality and
1207 interspecific interactions—a report on the *Argyrodes* symposium, Badplaas 2001. *Journal*
1208 *of Arachnology* **30**, 238-245.
1209 Yun J-H, Roh SW, Whon TW*, et al.* (2014) Insect gut bacterial diversity determined by
1210 environmental habitat, diet, developmental stage, and phylogeny of host. *Appl. Environ.*
1211 *Microbiol.* **80**, 5254-5264.
1212 Zhang L, Yun Y, Hu G, Peng Y (2018) Insights into the bacterial symbiont diversity in spiders.
1213 *Ecology and evolution* **8**, 4899-4906.
1214 Zhang L, Zhang G, Yun Y, Peng Y (2017) Bacterial community of a spider, Marpiss magister
1215 (Salticidae). *3 Biotech* **7**, 371.