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 organism 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 evolutionary history of a host and its associated microbial communities, altogether 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 the host spiders and their microbial constituents have responded in similar ways to the dynamic abiotic environment of the volcanic archipelago. The expectation was that each component of the holobiont (the spider hosts, intracellular endosymbionts, and gut microbiota) should show a similar pattern of 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. Results showed that the host *A. waikula* is strongly structured by isolation, suggesting sequential colonization from older to younger volcanoes. Similarly, the endosymbiont communities were markedly different between *Ariamnes* species on different islands, but more homogenized among *A. waikula* populations. In contrast, the gut microbiota was largely conserved across all populations and species, and probably mostly environmentally derived. Our results highlight the different evolutionary trajectories of the distinct components of the holobiont, showing the necessity of understanding the interplay between components in order to assess any role of the microbial communities in host diversification.

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

52 **Introduction**

53 Patterns of biodiversity are influenced by both ecological and evolutionary processes
54 operating within the dynamic context of a community (Weber et al. 2017). The external
55 environment can serve to isolate populations for various periods, and select for traits that influence
56 the evolutionary trajectory. At the same time, a given organism also represents a community by
57 hosting a diverse array of microbial species, many of which perform essential functions for their
58 host. Among arthropods, associated microbial communities are often highly diverse assemblages,
59 accounting for an extensive range of interactions with their host (Engel & Moran 2013). Many
60 arthropods host different microbial communities occupying various niches such as the gut
61 microbiota or intracellular endosymbionts (Hansen & Moran 2014). The importance of microbial
62 communities for promoting the isolation of their hosts (Sharon et al. 2010) and facilitating their
63 adaptation to novel ecological niches (O'Connor et al. 2014) has been increasingly recognized. It
64 is thus assumed that a species' response to the dynamic changes in the environment can be
65 dictated by the "holobiont" of host and microbial associates (Margulis & Fester 1991). Therefore,
66 understanding the nature and the interplay between different components of the holobiont – the
67 host and the different communities of microbes - in response to external drivers, is essential for
68 understanding potential drivers of evolution (McFall-Ngai et al. 2013).

69 First considering the gut microbiota, its composition is often determined by complex
70 interactions of environment, diet, developmental stage, and host evolutionary history (Yun et al.
71 2014), contributing to various functions such as host nutrition or protection against pathogens
72 (Engel & Moran 2013). However, for some arthropod taxa, recent work also suggests that a large
73 proportion of the arthropod gut microbiota is purely environmentally derived, highly transient, and
74 does not always have an apparent functional relevance (Hammer et al. 2017). For example,
75 predators may have a microbiota derived from their prey items (Kennedy et al. 2020). In contrast,
76 functional reliance of the host on its microbial communities could warrant more stable and
77 predictable gut microbial communities, which may otherwise be less deterministic. In such a case

(*i.e.* host dependence), the observed microbial communities may even co-evolve with their host (Engel & Moran 2013). That may lead to a co-diversification of microbial communities and host taxa. On the other hand, host and microbial evolutionary histories may not be tightly coupled if the environment of the host dictates microbial assemblage on short time scales.

In contrast to the gut microbiota, endosymbionts are mostly 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 and thus enhance genetic isolation (Shropshire & Bordenstein 2016). Some endosymbionts can also affect dispersal ability (Goodacre et al. 2006; Pekár & Šobotník 2007, 2008), which can further impact their host's diversification. Considering their strong effect on the reproductive system, endosymbionts often evolve in concert with their host. The dominant endosymbiont taxon in a lineage of arthropods is often stable, and the endosymbiont's phylogeny commonly reflects that of their host, with major endosymbiont switching events being infrequent (Bailly-Bechet et al. 2017). Recent evolutionary divergence in the host may thus be mirrored by differentiation among associated endosymbionts.

In summary, various environmental and evolutionary factors can differentially influence a microbial assemblage depending on the nature of the host/microbe relationship. Some microbes may be purely environmentally sourced, while others may closely track their host's adaptation and diversification. A key point of interest is then dissecting the extent, conditions, and mechanisms under which hosts and their microbial communities influence one another's evolutionary trajectories. We pursue this task 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, a hotspot volcanic chain with the islands showing a geological chronosequence of increasing age from southeast to northwest, provides an ideal system for tracking the interplay between a host and the different components of its microbial community, within a constrained setting (Shaw & Gillespie 2016). Even on the youngest island of Hawaiʻi, 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. Blankers et al. 2018; Eldon et al. 2019; Goodman et al. 2019).

Stick-spiders in the genus *Ariamnes* (Theridiidae) have diversified rapidly across the landscapes in the Hawaiian archipelago (Gillespie & Rivera 2007) and exhibit repeated diversification into ecomorphs adapted to specific microhabitats (Gillespie et al. 2018). However, their diet is conserved and they are specialized consumers of other spiders (Kennedy et al. 2018). While the activity of *Ariamnes* is exclusively nocturnal, the ecomorphs are defined by the microhabitat with which they are associated during the day, with the “gold” ecomorph on the underside of leaves, the “dark” ecomorph on dark vegetation and rocks, and the “matte white” ecomorph on white lichen (Gillespie & Rivera 2007; Gillespie *et al.* 2018). The ecomorphs are entirely cryptic on their daytime microhabitat, suggesting that the primary selective agent for morph differences is diurnal predators, most likely birds (Gillespie et al. 2018).

The current study focuses on a single species of the Hawaiian *Ariamnes* (*A. waikula*), endemic to the youngest island of Hawaiʻi, to understand how the early differentiation of the host might be linked to the different components of its holobiont. We aim to determine whether this highly specialized spider lineage and its microbial associates have responded in the same way to recurrent colonization events across volcanoes within the island. We hypothesize that the population structure of *A. waikula* reflects a stepping stone colonization from older to younger volcanoes, and that populations from geologically older sites will show increased differentiation and higher within population diversity compared to younger sites. If microbial associates are

129 largely conserved (because of transmissions or host-filtering (Moran & Sloan 2015)), we predict
130 that the microbiota of the *A. waikula* holobiont will closely mirror the population structure of the
131 host, including a diversity bottleneck in younger sites (Minard et al. 2015; Brooks et al. 2016). In
132 addition, we do not expect significant taxonomic changes in the endosymbionts across populations
133 given their typically vertical transmission.

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

Material and Methods

Sampling

We sampled *Ariamnes* across Hawai‘i Island, focusing on individuals of *A. waikula*, from 6 populations, while including 2 individuals of the related *A. hiwa* (brown ecomorph). We also sampled individuals from two other species: *A. melekalikimaka* on West Maui and *A. n. sp.* Moloka‘i (Gillespie et al. 2018). We included *A. hiwa*, *A. melekalikimaka*, and *A. n. sp.* Moloka‘i to confirm monophyly of the clade on the Hawai‘i Island (as outgroups) and to compare the diversity of the microbial communities of other species between and within islands. Individuals were collected by hand and immediately preserved in 90% EtOH. We collected a total of 133 individuals for sequencing (Table 1; Supplementary Tables S1 & S2). Only adults were collected for this study, to decrease the likelihood of capturing differences driven by age.

Table 1: Individual, population, and sampling locality information for *Ariamnes* spiders in this study. Substrate age (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	Kohala	Kohala, 0.43	300,000	18	6
A. waikula	Hawai‘i	Saddle	Mauna Loa, Kea 0.01-0.38	4,000	6	3
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	Thurston	Kīlauea, 0.004	600	13	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
A. melekalikimaka	Maui	Pu‘u Kukui	W. Maui, 1.3	1,500,000	16	9
A. n. sp	Moloka‘i	Kamakou	E. Moloka‘i, 1.8	1,400,000	16	7
					Total: 123	**Total: 71**

162 *Ariamnes - ddRAD library preparation and sequencing*

163 To examine the population structure of the spider host, we used ddRAD to obtain reduced
164 representation genome-wide SNP data. Genomic DNA was extracted from spider legs with several
165 modifications to the Qiagen DNeasy kit protocol. Legs were first removed from each specimen
166 using sterile tweezers so that the abdomen remained intact for the microbial DNA analysis. DNA
167 was then extracted by placing the tissue in Proteinase K and lysis buffer and grinding them with a
168 sterile pestle to break up the exoskeleton. We then added 4uL RNase A (100 mg/ml) and
169 incubated the extractions for two minutes at room temperature. Tubes with tissue and extraction
170 solution were then placed overnight in a heat block at 56°C. The remainder of the extraction
171 protocol was performed following the manufacturer's instructions. We built ddRAD libraries
172 following an adapted protocol of Peterson *et al.* (2012) (Saarman & Pogson, 2015; see Maas et
173 al. 2018 for protocol optimization steps). Briefly, we started the ddRAD protocol with a total of 100
174 nanograms of DNA per sample. The DNA was digested using SphI-HF (rare-cutting) and MluCI
175 (frequent-cutting) restriction enzymes. We assessed fragmentation with a Bioanalyzer High
176 Sensitivity chip (Agilent). We multiplexed 15-20 individuals per library for a total of eight ddRAD
177 libraries. We used a Sage Science Pippen Prep to size select 451-551bp (including internal
178 adapters) fragments, and confirmed the sizes using a Bioanalyzer. Ten indexing polymerase chain
179 reaction cycles (PCRs) were run on each library to enrich for double-digested fragments and to
180 incorporate a unique external index for each library pool. The eight libraries were sequenced using
181 100bp paired-end sequencing on one Illumina HiSeq 2500 lane at the Vincent J. Coates Genomic
182 Sequencing Facility at UC Berkeley.

183 *Ariamnes - ddRAD data filtering and processing*

184 We used a custom perl script invoking a variety of external programs to filter and process
185 the ddRAD data (RADTOOLKIT v0.13.10; https://github.com/CGRL-QB3-UCBerkeley/RAD). Briefly,

186 raw fastq reads were first de-multiplexed based on the sequence composition of internal barcodes
187 with tolerance of one mismatch. De-multiplexed reads were removed if the expected cutting site
188 was not found at the beginning of the 5'-end of the sequences. The reads were then filtered using
189 cutadapt (Martin 2011) and Trimmomatic using default parameters (Bolger et al. 2014) to trim off
190 Illumina adapter contaminations and low-quality reads. The resulting cleaned forward reads of
191 each individual were first clustered using cd-hit (Fu et al. 2012; Li & Godzik 2006), keeping only
192 those clusters with at least two reads. In each cluster, we pulled out the corresponding reverse
193 reads based on the identifiers. Both forward and reverse clusters at the same time were kept only
194 if the corresponding reverse reads also formed one cluster. If reverse reads were grouped into
195 more than one cluster, then only the forward read cluster was kept. For each paired cluster, the
196 representative sequences for each forward and reverse cluster determined by cd-hit were
197 retained. We then merged the forward and reverse sequences using FLASH (Magoč & Salzberg
198 2011). If they could not be merged then they were joined by placing "N"s between the two
199 sequences. The resulting loci were then masked for putative repetitive and low complexity
200 elements, and short repeats using RepeatMasker (Smit et al. 2004) with "spider" as a database.
201 After masking, we eliminated loci if more than 60% of the nucleotides were Ns. The resulting
202 ddRAD loci from each individual were combined and clustered for all individuals. Contigs that were
203 at least 40 nucleotides in length as well as shared by at least 60% of all the individuals served as
204 a reference. Cleaned sequence reads from each individual were then aligned to the reference
205 using Novoalign (http://www.novocraft.com) and reads that mapped uniquely to the reference
206 were kept. We used Picard (http://picard.sourceforge.net) to add read groups and GATK
207 (McKenna et al. 2010) to perform realignment around indels in BAM format generated by
208 SAMtools (Li et al. 2009). We then used SAMtools/BCFtools to generate data quality control
209 information in VCF format. These data were then further filtered using a custom perl script,
210 SNPcleaner (Bi et al. 2013). We filtered out any loci with more than two called alleles. We masked
211 sites within 10 bp upstream and downstream of an indel. We discarded sites with a total depth

212 outside of the genome-wide 1st and 99th percentile. To avoid excessive heterogeneity in sample
213 representation among sites, we also removed alignments if more than 40% of the samples had
214 less than 3X coverage. The resulting sites passing all of the above filters were analyzed using
215 ANGSD (Korneliussen *et al.* 2014).

216 Since most of the individuals were expected to have low coverage (between 5-15x) we
217 used ANGSD (Korneliussen et al. 2014) to calculate genotype likelihoods and genotypes for the
218 analyses. This tool was specifically developed for population and evolutionary genomic analyses
219 of low coverage data. We used genotype likelihoods whenever the downstream tools allowed us
220 to, since genotypes called from low coverage sites involve a non-negligible amount of uncertainty
221 arising from randomness in allele sampling as well as sequencing or mapping errors (Crawford &
222 Lazzaro 2012)).

223 *Ariamnes - Phylogenetic Analyses*

224 First, we investigated the phylogenetic relationships between populations of *A. waikula*
225 and the other *Ariamnes* species to better understand the colonization patterns. We used the
226 Stacks pipelines (Catchen et al. 2013) to group reads into homologous loci across all individuals
227 and extract phylogenetically informative sites (i.e. fixed within individuals but variable between
228 individuals). Next, we obtained an alignment composed of 58,899 sites and performed
229 phylogenetic reconstruction using IQtree (Nguyen et al. 2015) combining model selection with
230 ModelFinder Plus (Kalyaanamoorthy et al. 2017) and assessing branch supports with 1,000
231 ultrafast bootstrap (Hoang et al. 2018). Finally, we rooted the tree using the *A. hiwa* individual and
232 calibrated it with r8s (Sanderson 2003) without specifying any absolute dating (*i.e.* setting the root
233 age to 1). In addition, we also reconstructed the phylogeny using pairwise genetic distances
234 calculated from genotype likelihoods using the software ngsDist (Viera et al. 2015) and balanced
235 minimum evolution using FastME (Lefort et al. 2015). To do so, we first calculated genotype
236 likelihoods using ANGSD, which were used to construct a pairwise distance matrix between all

237 individuals using ngsDist. The phylogeny was then inferred from this matrix using FastME and
238 bootstrapping was carried out using RaxML (Stamatakis 2014, 100 replicates).

239 *A. waikula - Population Genetic Analyses*

240 We explored *A. waikula* population structure using Principal Components Analysis (PCA).
241 We used the previously generated genotype likelihoods to calculate genotype posterior
242 probabilities under an allele frequency prior (-doPost 1) in binary format (-doGeno 32) for use with
243 ngsCovar (part of the ngsTools package; (Fumagalli et al. 2014). We used ngsCovar to calculate
244 a genetic covariance matrix among individuals from the genotype probabilities at sites having a
245 minor allele frequency of at least 0.004 to avoid noise from very rare alleles (which could be due
246 to sequencing errors). Comparisons between the first three principal component axes were then
247 plotted using R. Pairwise F_{ST} values were calculated in ANGSD from allele frequency likelihoods
248 (-doSaf 1) using the respective pair's genome-wide, unfolded, joint SFS as a prior for jointly
249 observing any combination of allele frequencies between the two populations. We used unfolded
250 allele frequencies by supplying the reference sequences as a pseudo-ancestral sequence, as
251 ANGSD is only able to accurately estimate F_{ST} using unfolded data. In order to visualize F_{ST}
252 distances, we performed multidimensional scaling (to 2 dimensions) with R (R Core Team, 2020;
253 using the *cmdscale* function).

254 A crucial factor underlying population structure is gene flow between populations. In order
255 to investigate signatures of connectivity between populations of *A. waikula*, we used two different
256 approaches: ngsAdmix (Skotte *et al.* 2013) and EEMS (Petkova et al. 2016). ngsAdmix is a
257 genotype likelihood based-tool for estimating individual admixture proportions, while EEMS uses
258 genotype calls to infer effective migration surfaces. We inferred ancestry proportions indicative of
259 admixture for different values of K (number of ancestral populations; ranging from two to six) using
260 ngsAdmix and R for visualization. We then generated genotype calls for the EEMS analysis using
261 ANGSD with the following flags: '-doMaf2', '-doGeno 2', '-doPost 1', '-doSaf 1', '-fold 1', a SNP p-

262 value cut-off of 1e-6, and a postCutoff value of 0.75. We only considered sites with a minimum
263 genotype depth of 3. We then converted the output into the adegenet (R package) input format
264 using PopGenTools (https://github.com/CGRL-QB3-UCBerkeley/PopGenTools) to calculate
265 genetic distances between all individuals. We then divided Hawai'i Island into grids of 10km
266 (Supplementary Fig. S1). We then added the samples to the grid using one GPS coordinate per
267 population (Supplementary Table S1). Using a stepping stone model, we then calculated migration
268 rates between the demes. Convergence of the MCMC runs was assessed by plotting and
269 inspecting the traces by eye (Supplementary Fig. S2). We performed 10 independent runs for
270 each of the analyses.

271 Next, we calculated population genetic statistics such as nucleotide diversity (Pi),
272 Watterson's Theta, and Tajima's D. To do so, we generated a folded SFS for each population
273 using ANGSD and realSFS. We then ran ANGSD using *–doTheta 1* and *–pest* (which provides
274 the genome-wide SFS prior to ANGSD) and the respective output formats were converted to bed
275 format using *thetaStat make_bed*. Subsequently, we calculated the per population statistics using
276 *thetaStat do_stat*. The average, min, and max Tajima's D were then calculated and we further
277 extracted the genome-wide average Watterson's theta and Pi values. To assess whether
278 populations show genetic signatures indicative of serial founder events and expansions, we used
279 linear models in R to test whether genetic diversity (Pi or Watterson's theta) was positively related
280 to the age of the youngest lava flow of each sampling site (referred to as the volcano age), as
281 more time would allow for genetic diversity to recover in a large population.

282 *Microbial Communities*

283 To characterize the microbial community within the *A. waikula* hosts, a subset of 71
284 individuals from the eight populations on Hawai'i and three additional islands (Table 1;
285 Supplementary Tables S1 & S2) were selected for analysis. We focused on the mid and hindgut,
286 both located in the spider's opisthosoma. The preservation in ethanol led to considerable

287 shrinkage of the opisthosoma and thus did not allow us to separately dissect out the gut. Instead,
288 we used the whole opisthosoma to extract DNA, as described in Kennedy et al. (2020). Specimens
289 which did not have the opisthosoma intact were not used. The digestive tract comprises the
290 majority of the opisthosoma's cavity. In addition, it contains silk glands, the heart, lungs, and
291 gonads. The opisthosoma was removed with a sterile razor blade and then washed in ethanol to
292 remove external bacteria (Hammer et al. 2015). We considered it to be representative of the "gut
293 microbiota", even if it technically consists of the "opisthosoma microbiota", but previous studies
294 have shown that the gut microbiota dominates in the opisthosoma (Sheffer et al. 2020, Kennedy
295 et al. 2020). The tissue was then transferred into lysis buffer and finely ground with a sterile pestle.
296 DNA was extracted using the Gentra Puregene Tissue Kit (Qiagen, Hilden, Germany) according
297 to the manufacturer's protocol. Spider abdominal tissue can contain PCR inhibitors (Schrader et
298 al. 2012), thus we cleaned the DNA extract with 0.9X AmPure Beads XP.

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

311 We used the 16S profiling analysis pipeline for Illumina paired-end sequences of the
312 Brazilian Microbiome Project (Pylro et al. 2014), including QIIME 1.8.0 (Caporaso et al. 2010) and

313 USEARCH 11 (Edgar 2013). We modified some steps of these pipelines, using our own Bash and
314 R scripts (R Core Team, 2020; see Data Accessibility section). The QIIME script
315 join_paired_ends.py was used to merge paired reads. The fastq_filter command in USEARCH
316 was used for quality filtering the assemblies (below a base calling error probability of 0.5, using
317 an average Q-score for each read). We used the Stream EDitor in UNIX to remove PCR primers
318 from all assembled sequences. Sequences were de-replicated using USEARCH, removing all
319 singletons. OTUs were generated at a similarity cutoff of 3 % or 0% (0 radius OTUs, or "Z-OTU")
320 from the de-replicated sequences and chimera removed de novo using USEARCH. The following
321 analyses were thus independently applied on the two distinct sets of OTUs. We assigned
322 taxonomy to the resulting OTUs using the assign_taxonomy.py script based on the Greengenes
323 database (http://greengenes.secondgenome.com). We removed sequences corresponding to
324 OTUs found in high prevalence and abundance in the different negative controls from all samples,
325 as these could represent contaminants.

326 Spiders are known to carry various endosymbiotic bacteria (White et al. 2020). These can
327 be vastly overrepresented in microbial analyses and thus may completely dominate the microbial
328 community structure. Since we did not extract the gut from individuals, endosymbionts from
329 outside of the gut could be particularly prevalent in our analysis. We thus separated known
330 endosymbionts from remaining bacterial sequences (referred to as the "gut microbiota"), resulting
331 in two OTU sequence files. Both these datasets were used for the following analyses separately.

332 OTU tables were prepared by mapping sequences back to the filtered OTU sequence files
333 using USEARCH. The OTU tables were rarefied to an even coverage (from 400 to 8,000 reads
334 with 20 replications per rarefaction depth) using the multiple_rarefactions.py script in QIIME. Given
335 the rarefaction curves (Supplementary Fig. S3), we chose rarefied depths at 3,200, 110 and 3,000
336 reads for the whole microbiota, the endosymbionts, and the gut microbiota respectively at
337 replicated 20 times this rarefaction, and these rarefied OTU tables were used in all the following

338 analyses. We plotted the relative abundances of different microbial taxa at the genus and order
339 levels for the different *Ariamnes* populations.

340 We then investigated whether microbial associates showed similar diversity patterns
341 compared to their *Ariamnes* hosts along the chronosequence. Chao1 indexes of alpha diversity
342 were calculated from the rarefied OTU tables using QIIME (alpha_diversity.py), and to evaluate
343 the presence of a bottleneck in microbial diversity, linear mixed models were used to test for an
344 effect of the volcano age and the genetic diversity of *Ariamnes* host populations (Pi and
345 Watterson's theta) on the bacterial alpha diversity. Homoscedasticity and normality of the model
346 residuals were verified.

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

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

Results

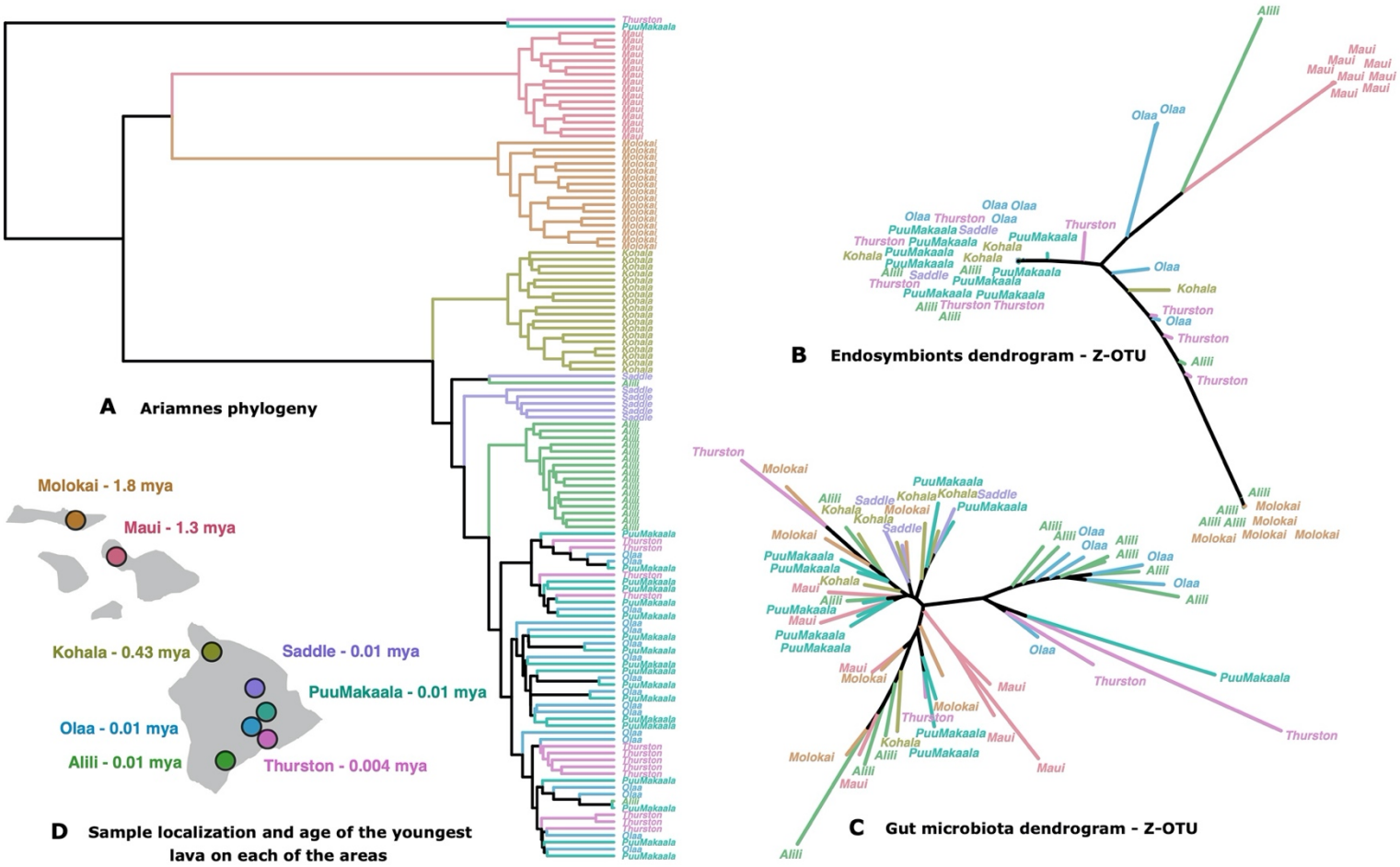
Ariamnes - Phylogenetic Analyses

We obtained approximately 210 million paired-end reads from the Illumina sequencing. After filtering for contaminants and low-quality reads, there were approximately 83 million remaining across the 123 demultiplexed samples. From these reads, we retained a total of 2,957,301 sites passing filters out of a possible 7,378,384 that were used in downstream analyses and resulting in 123 individuals for population genetic analysis. On average, samples had a coverage of 12x (3-42x) across all loci.

All phylogenetic analyses (maximum likelihood and genetic distance based) confirm that the *A. melekalikimaka* (West Maui), *A. n. sp.* (Moloka‘i) and *A. waikula* (Hawai‘i Island) are monophyletic groups (Fig. 1). Among *A. waikula* individuals, the population from Kohala (the oldest volcano on Hawai‘i Island) is sister to the clade that contains all other individuals. Except for one individual from Saddle and two from Alili, the populations from the Saddle (between Mauna Loa and Mauna Kea) and Alili form monophyletic clades. The populations from Ola‘a, Pu‘u Maka‘ala and Thurston, all sites that are close together in the saddle between the volcanoes of Mauna Loa and Kilauea (MLKS), form one mixed clade.

Figure 1: Host phylogenetic history partially recapitulates microbiota differentiation

Phylogenetic tree of the *Ariamnes waikula* individuals across the island of Hawai'i (A). Two specimens of *A. hiwa* (brown ecomorph) were used as outgroup taxa. Microbiota dendrograms reconstructed from the endosymbiont community (B) and the gut microbiota (C) for the Z-OTU. (D) map of the sampled host populations and the corresponding age of the youngest lava on each of the areas.



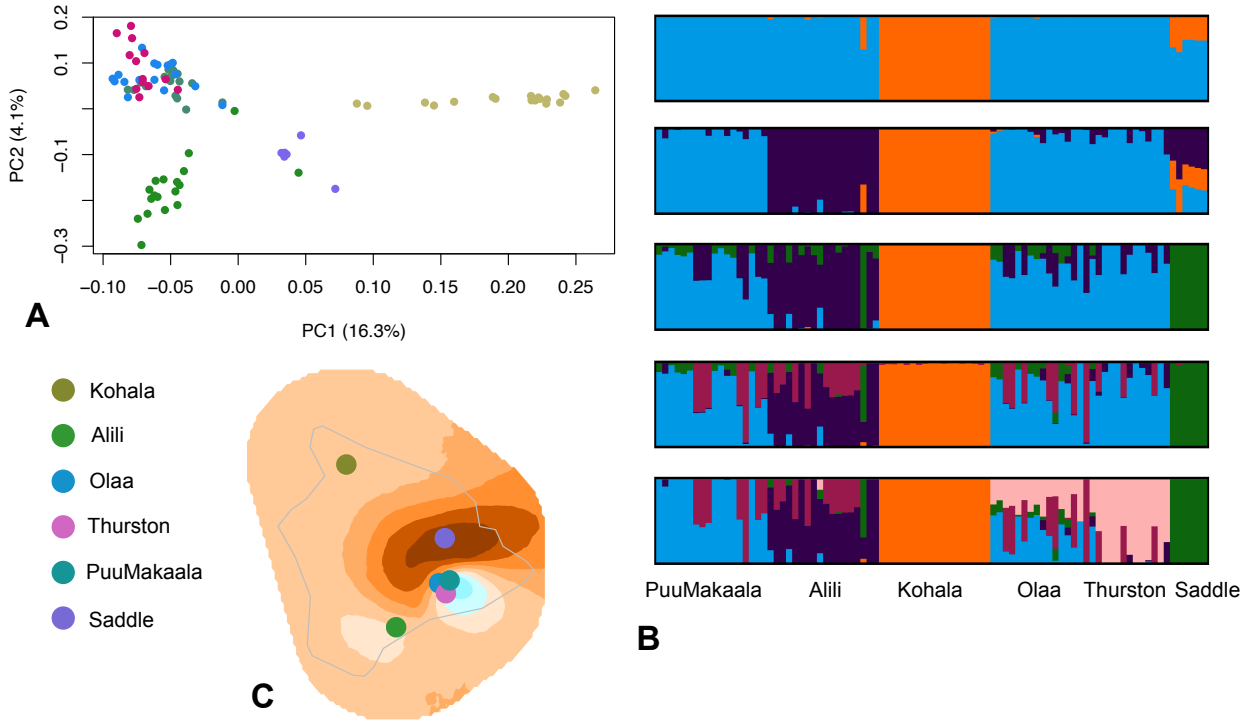
388 *A. waikula - Population genetics*

389 We analyzed population differentiation and structure for the full data set, and a subset
390 including only *A. waikula* from Hawai‘i Island. We did not include the two *A. hiwa* individuals in the
391 population genetic analyses since they represent a different species used as outgroup for the
392 phylogenetic analyses.

393 We found the highest genetic differentiation (F_{ST}) between the different species on these
394 three islands with similar pairwise levels (0.54 between Maui and Hawai‘i Island, 0.48 for Moloka‘i
395 and Hawai’i Island and 0.48 between Moloka‘i and Maui; Table S3). This closely matches the
396 results from COI data (Roderick et al. 2012). *A. waikula* on Hawai‘i Island were primarily structured
397 according to locality with Kohala, Alili, and Saddle being distinct from the MLKS sites (pairwise F_{ST}
398 range 0.03-0.15; Pu‘u Maka‘ala, Thurston and Ola‘a; Fig. 1A).

399 Regarding potential admixture between populations, using ngsAdmix first, we found good
400 convergence between the 50 independent runs for higher values of K for the Hawai‘i only sampling
401 (Supplementary Fig. S4). Kohala forms a separate group at K=2, next is Allili at K=3 and Saddle
402 at K=4. At K=6 we see a slightly closer grouping of Thurston and Ola‘a, then either of the two with
403 Pu‘u Maka‘ala, but overall the MLKS populations make up one group (Fig. 2B). Second, EEMS
404 analyses with *A. waikula* individuals indicate potential gene flow between the MLKS populations:
405 Pu‘u Maka‘ala, Thurston, and Ola‘a (Fig. 2C). These patterns were reiterated using PCA (Fig. 1A,
406 Supplementary Fig. S5). PC1 explained approximately 16% of the variation and separated Kohala
407 and the younger sites, placing Saddle in the middle (Fig. 1A). PC2 explained approximately 4% of
408 the variation and primarily separated Alili from the MLKS sites, while PC3 (3%) interestingly
409 clustered Kohala with Pu‘u Maka‘ala, Thurston, and Ola‘a, while separating the Alili and Saddle
410 populations (Supplementary Fig. S5).
411

Figure 2: Population genetic analyses of *Ariamnes waikula*. A) ngsAdmix results for K=2 to K=8 for all *A. waikula* specimens. B) ngsAdmix results for K=2 to K=6 (from top to bottom) for all *A. waikula* from Hawai'i Island only. C) EEMS analysis of all *A. waikula* specimens. Brown color indicates barriers for gene flow (the stronger the darker), and cyan indicates gene flow between populations (the stronger the darker).



We found that Watterson's theta was fairly even across localities on the youngest island of Hawai'i, with the populations on younger volcanoes possessing slightly higher thetas (e.g. Pu'u Maka'ala: 0.0085, Ola'a: 0.0086; Table 2). The Maui population had the overall lowest theta (0.0053). Pi estimates showed a similar pattern, with little differentiation between values on the youngest volcanic sites and Maui (0.0051-0.0056). However, Moloka'i populations had the highest value of Pi (0.0073). Despite low variance in Pi, the positive correlation between Pi genetic diversities and the volcano age (linear regression: t=4.6, p-value=0.06) suggested that *Ariamnes*

427 likely experienced successive bottlenecks along the island chronosequence during the sequential
428 colonizations from older to younger volcanoes. However, this correlation was of lower significance
429 for *A. waikula* populations within Hawai'i (Figure 3) and no other correlations between theta and
430 volcano age or between *Ariamnes* genetic diversity (Pi or Theta) and *Ariamnes* microbial diversity
431 were found.
432
433 **Table 2**: Genetic diversity estimates for *Ariamnes* spiders across the different sampling sites
434 (excluding *A. hiwa*). The average volcano age of each site used for the statistical analyses is
435 indicated.
436

Sites	Volcano age (in Myr)	Average Pi	Average Watterson's theta
Pu'u Maka'ala	0.01	0.00518	0.00847
Alili	0.01	0.00535	0.00775
Kohala	0.43	0.00538	0.00749
Ola'a	0.01	0.00535	0.00857
Thurston	0.004	0.00520	0.00712
Saddle	0.19	0.00526	0.00636
Moloka'i	1.8	0.00727	0.00747
West Maui	1.3	0.00559	0.00532

437
438 We also investigated Tajima's D, which reflects population size changes as a skew in the
439 site frequency spectrum (Supplementary Fig. S6). This statistic compares the number of pairwise
440 differences to the number of segregating sites. We found that most sampling sites showed only
441 marginally negative values, but show no significant indications of population bottlenecks or
442 expansions (average: -1.0 to -0.06; Supplementary Table S4).
443

Figure 3: The bottleneck in diversity along the volcano age tends to affect the *Ariamnes* hosts, but not their microbial associates

Note that Maui and Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*) and are only represented as a reference here: the following linear model testing the effect of volcano age on host (A-B) or microbial (C-D) diversities were only performed on *A. waikula* populations in Hawai'i.

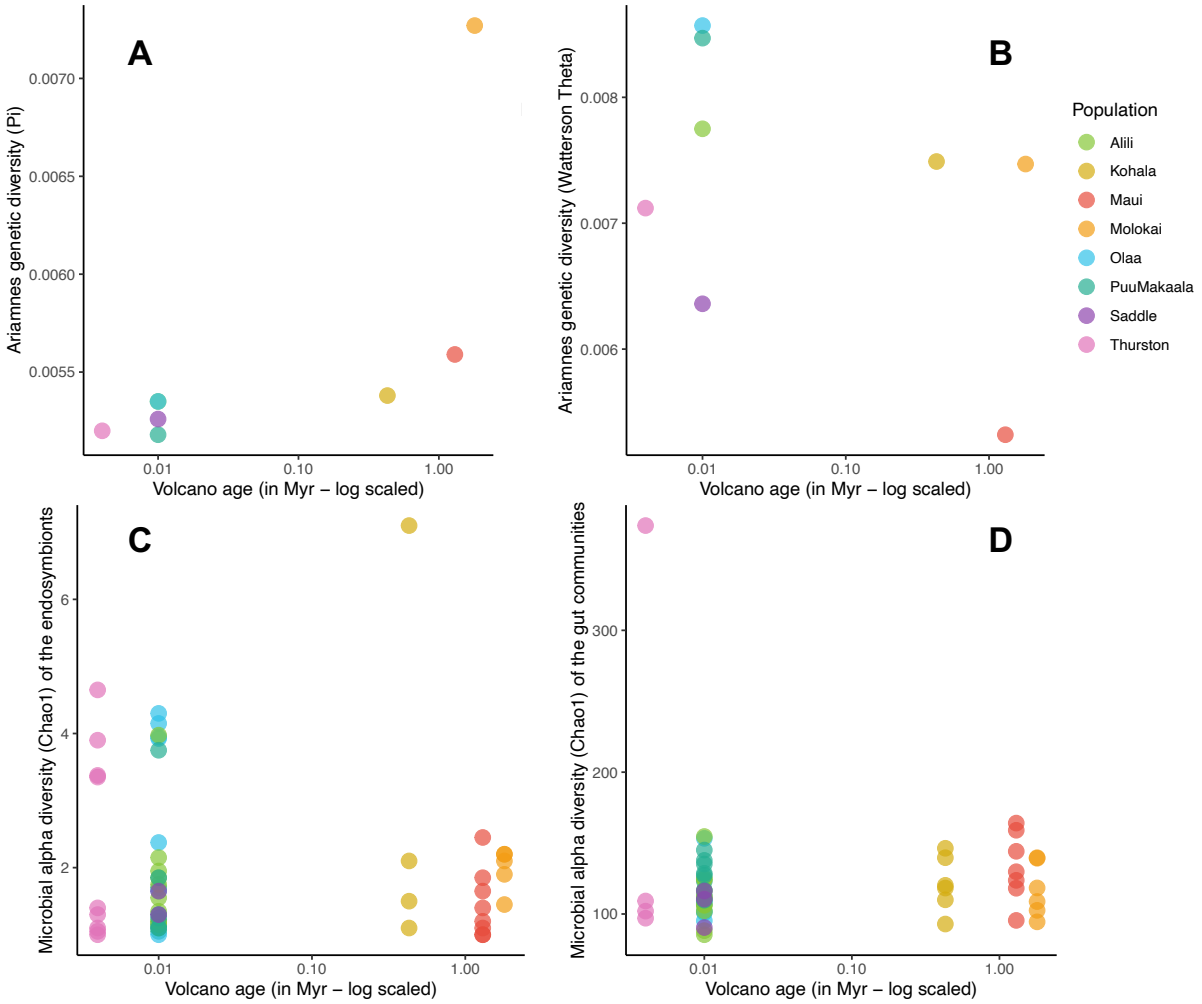
(A) *Ariamnes* genetic diversity (Pi) estimated per population as a function of the volcano age (in Myr). We found no significant relationship between *A. waikula* genetic diversity (Pi) and the volcano age (linear model: F=1.7, df=4, p-value=0.27), despite the observed trend, probably because of the small number of sampled populations in Hawai'i (6 only).

(B) *Ariamnes* genetic diversity (Watterson's theta) estimated per population as a function of the volcano ago (in Myr). We found no significant relationship between *A. waikula* genetic diversity (Watterson's theta) and the volcano age (linear model: F=0.023, df=4, p-value=0.87).

(C) Microbial alpha diversity (Chao1 index on Z-OTUs) estimated for the endosymbionts as a function of the volcano ago (in Myr). Each dot corresponds to the average of the 20 rarefactions performed for each individual. We found no significant relationship between endosymbionts alpha diversity and the volcano age (linear mixed model: t=1.10, p-value=0.24).

(D) Microbial alpha diversity (Chao1 index on Z-OTUs) estimated for the gut community as a function of the volcano ago (in Myr). Each dot corresponds to the average of the 20 rarefactions performed for each individual. We found no significant relationship between gut microbial alpha diversity and the volcano age (linear mixed model: t=-0.08, p-value=0.94).

In addition, we did not find any correlation between the *Ariamnes* genetic diversities and the microbial alpha diversities (not shown; p-value>0.05). Similar results were obtained with 97% OTUs.



468

469

470

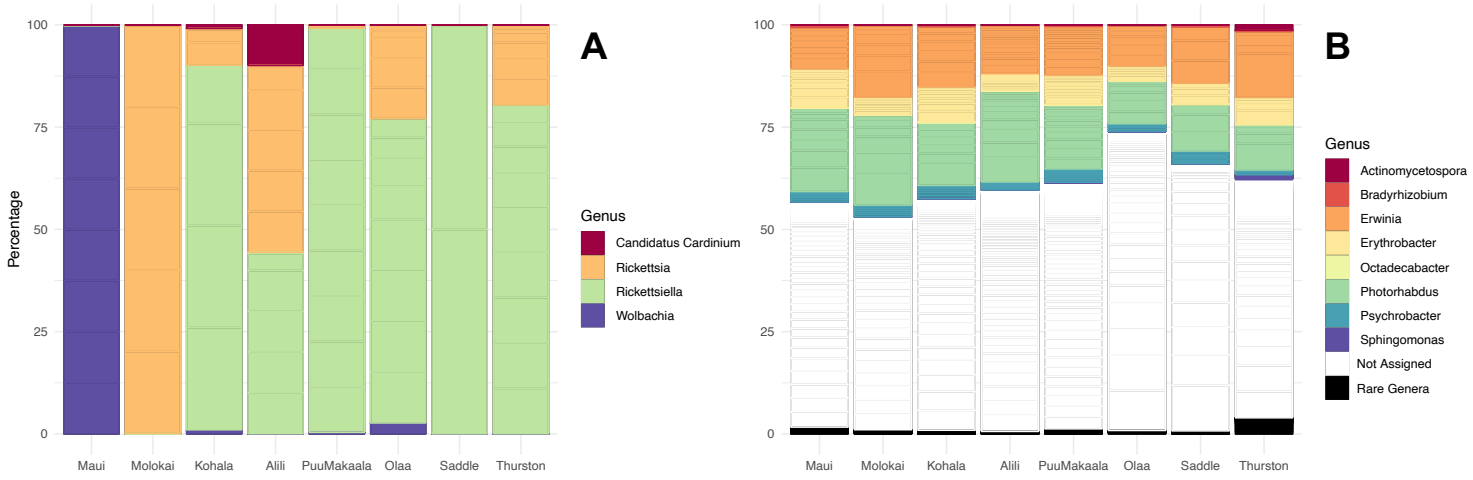
471 *Microbial Communities*

472 We obtained a total of 4,932,236 bacterial reads. After the quality filtering steps,
473 demultiplexing, and removal of contaminants based on the negative controls, we obtained a total
474 of 1,315,469 sequences, which range from 5,000 sequences to 45,000 per individual *Ariamnes*
475 sample. Due to the largely unexplored nature of the microbial communities of Hawaiian
476 arthropods, we could not assign species level taxonomy to the majority of microbial OTUs (538
477 OTUs of 571 at 97%, 615 OTUs over 1,357 Z-OTUs), but a higher proportion at the genus level
478 (330 at 97%, and 1,193 Z-OTUs).

479 We did not find any significant differences in the alpha diversities of the whole microbiota
480 among *Ariamnes* populations or species (Supplementary Table S5). We also did not find a
481 significant association of the microbial alpha diversity with volcano age or the genetic diversity of
482 the host population (Fig. 3), even if the alpha diversities of the gut microbiota seem to be slightly
483 higher for host populations present on old volcanos (Fig. 3D).

484 *Gut microbiota*: After removing endosymbionts, the remaining gut microbiota showed a
485 homogeneous taxonomic composition of microbial taxa at the genus level across different
486 *Ariamnes* populations or species (Fig. 4B & Supplementary Fig. S7). This stability of the
487 composition of the gut microbiota was also confirmed at the order level (Supplementary Fig. S8).
488 No clear differentiation according to the islands or the host populations could be detected for the
489 beta diversity of the gut microbiota, either from PCoA decomposition (Supplementary Fig. S9) or
490 hierarchical clustering (dendrogram; Fig. 1C & Supplementary Fig. S10). However, the gut
491 microbiota of individuals from the same population were significantly more similar than gut
492 microbiota from different populations or species, even within *A. waikula* populations from Hawai'i
493 Island (PERMANOVA; Supplementary Table S6). Such slight patterns of clustering by host
494 populations were also visually detectable for some individuals from the same populations on the
495 PCoA decomposition (Supplementary Fig. S9), or on the dendrogram of the gut microbiota (Fig.
496 1C). These results indicate that, although the composition of the gut microbiota seems
497 taxonomically similar, there is a slight trend toward microbiota differentiation according to the host
498 populations. However, using Mantel tests we found no significant correlations between the gut
499 microbial community beta diversities and the *Ariamnes* genetic distances nor the *Ariamnes*
500 phylogenetic distances (Table 3 & Supplementary Table S7), suggesting that there is likely no
501 pattern of sequential colonization from older to younger volcanoes (stepping stone colonization)
502 in these gut microbial communities in opposition to their spider hosts.
503

504 **Figure 4**: Relative abundances of endosymbiont genera per population (A) and gut symbionts per
505 spider population (B) defined from rarefied Z-OTU table. The relative abundance of each OTU is
506 delimited using the horizontal grey lines and are colored according to the OTU genera. Rare
507 genera (representing less than 1% of the abundance are merged together). Note that Maui and
508 Moloka‘i correspond to different *Ariamnes* species (*A. melekalikimaka* and *A. n. sp respectively*)
509 whereas the other populations in Hawai'i correspond to *A. waikula*.



510

511

512 *Endosymbiont community:* In contrast to the gut microbiota, the endosymbiont community
513 showed considerable variation between the different species of *Ariamnes* (Fig. 1B). *Ariamnes*
514 *melekalikimaka* from West Maui mostly carry *Wolbachia*, while *Rickettsia* dominate the
515 endosymbiont community on *A. n. sp.* Moloka‘i, and *Rickettsiella* the populations of *A. waikula* on
516 Hawai‘i Island (Fig. 4A & Supplementary Fig. S7). The pronounced turnover of endosymbiont taxa
517 between host populations is also mirrored in differences of beta diversity. Individuals of the
518 different species from different islands form clearly separated clusters in our PCoA plots and
519 cluster dendrograms for the endosymbiont communities (Fig. 1B, Supplementary Figs. S9 & S10).
520 PCoA plots and cluster dendrograms exhibit a strong pattern of phylosymbiosis, where the
521 divergences between microbiota reflect the phylogenetic distances between the *Ariamnes* hosts.

522 Mantel tests showed a significant correlation between the endosymbiont community beta diversity
523 and the *Ariamnes* genetic distances (P<0.0001; Table 3 & Supplementary Table S7). Within
524 Hawai'i Island, endosymbiont communities of *A. waikula* from the same population tend to be more
525 similar than between populations (Supplementary Table S6; *e.g.* individuals from Ola'a tend to
526 cluster together in Fig. 1B), suggesting a slight population differentiation. However, we found no
527 strong correlation between microbial beta diversity and *A. waikula* genetic distances using Mantel
528 tests (Table 3 & Supplementary Table S7).
529
530 **Table 3:** Mantel tests comparing the microbial beta diversity dis-similarities to the genetic
531 distances or the phylogenetic distances of their associated *Ariamnes*. Beta diversity dissimilarities
532 were computed on rarefied Z-OTU tables using the Bray-Curtis dissimilarities. Bold values
533 represent significant correlations. Mantel tests were either performed on all populations (between
534 *Ariamnes* species and *A. waikula* populations) or only on *A. waikula* populations.
535

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 phylo-genetic distances	**0.334**	**9.9e-05**	0.012	0.40
Endosymbionts beta diversity distances	*Ariamnes* genetic distances	**0.619**	**9.9e-05**	0.015	0.49
	Ariamnes phylo-genetic distances	**0.642**	**9.9e-05**	0.047	0.32
Gut microbiota beta diversity distances	*Ariamnes* genetic distances	-0.048	0.66	-0.052	0.10
	Ariamnes phylo-genetic distances	-0.038	0.63	-0.139	0.93

536

537 **Discussion**

538 In this study, we examined the early stages of diversification of Hawaiian *Ariamnes* spiders
539 in concert with the different components of their associated microbiota on the youngest island of
540 the Hawaiian archipelago. Our results show a strong effect of isolation by distance in the host and
541 contrasting phylogenetic patterns among both the endosymbiont community and the gut
542 microbiota they harbor, suggesting eco-evolutionary distinctiveness of their host and each of the
543 two microbial components.

544 *Isolation by distance explains Ariamnes diversification*

545 Our analyses suggest that the nascent diversification of *A. waikula* is mainly associated
546 with isolation by distance, where island structure or isolation (on separate volcanoes) is the main
547 barrier to gene flow. The different populations of *A. waikula* show a predictable pattern of
548 colonization across Hawai‘i Island from the oldest (Kohala) to the youngest (MLKS) sites (Fig. 1),
549 with little genetic structure between the geographically close MLKS sites of Ola‘a, Pu‘u Maka‘ala,
550 and Thurston (PCA, Supplementary Fig. S5; FST, Supplementary Table S3; Admixture, Fig. 2B).
551 Our analyses suggest that gene flow is still occurring between these three populations, and that
552 recent colonization events have not yet led to differentiation in this area (Roderick et al. 2012).
553 Interestingly, the nucleotide diversity (Pi) tends to increase with volcano age, but their relationship
554 is not significant for *A. waikula* populations. In addition, the pattern is less clear for Watterson’s
555 theta, likely due to loss of rare alleles during colonization and the short time frame for new
556 mutations to arise subsequent to the colonization event. Notably, despite the popular hypothesis
557 that sequential colonizations result in genetic bottlenecks, we show little detectable genetic
558 evidence for such events here. Larger sample sizes and additional genetic data should be
559 obtained to investigate these patterns further.

560 Patterns of genetic differentiation suggest a potentially complicated course of colonization
561 across Hawai‘i Island for *A. waikula*. Although Kohala, which represents the oldest volcano, was
562 repeatedly characterized as distinct from other populations in PCA, phylogenetic, and admixture
563 analyses, both Saddle and Alili also presented as distinct. These patterns are potentially explained
564 by a shifting mosaic population structure (Carson et al 1990) with multiple colonization events,
565 historical connectivity, or ongoing admixture. Sampling of additional sites, individuals, and species
566 will be required to disentangle the historical sequence of events explaining the relationship
567 between these sites. Our results are consistent with other population genetic studies on Hawai‘i
568 Island showing strong differentiation and potential speciation of lineages on the different volcanoes
569 within the island (planthoppers, Goodman et al. 2019; flies, Eldon et al. 2019; crickets, Blankers
570 et al. 2018).

571 *Major endosymbiont changes after potential colonization event*

572 We found major shifts among the dominant endosymbiont genera across the different
573 islands and *Ariamnes* species: Species of *Ariamnes* from the two geologically oldest sites
574 (Moloka‘i and Maui) each harbor a unique endosymbiont genus (*Rickettsia* and *Wolbachia*
575 respectively) while the different populations of *A. waikula* all have a mix of endosymbionts
576 dominated by the genus *Rickettsiella*. These shifts in endosymbionts likely contribute to the
577 congruence between the host phylogeny and the dendrogram of microbiota dissimilarity (a pattern
578 referred as phylosymbiosis (Brooks et al. 2016). These patterns are consistent with the maternal
579 inheritance and vertical transmission of the endosymbionts (Duron et al. 2008) and such pattern
580 have been similarly observed in other arthropods (e.g. a *Camponotus* ant system and its
581 associated endosymbiont *Candidatus Blochmania* (Degnan et al. 2004). Therefore, the overall
582 evolutionary patterns of the *Ariamnes* host are reflected in the endosymbiont component of the
583 microbiota. However, from our study, we cannot say whether endosymbionts played a role in

584 promoting the genetic isolation of *A. waikula* populations; this hypothesis would require
585 experimental manipulation of microbial communities within different host populations.

586 The observed turnover of dominant endosymbiont genera between islands and host
587 populations suggests that the endosymbiont vertical transmission happens unexpectedly fast. A
588 previous study reported turnover rates over time for the bacterial genus *Wolbachia* colonizing
589 arthropods from French Polynesia of 7 Myr on average and found that *Wolbachia* colonization did
590 not seem related to recent events of isolation mediated by island structure (Bailly-Bechet et al.
591 2017). Given that the islands and thus the spider populations studied here are separated by <2
592 Myr (Gillespie et al. 2018), this suggests that the turnover in *Ariamnes* endosymbiotic
593 communities, strongly linked to the Hawaiian archipelago structure, is much faster than the
594 extinction dynamics of *Wolbachia* in the French Polynesian arthropods. Predation on other
595 arthropods and cannibalism, which are particularly important for *Ariamnes* spiders (Whitehouse et
596 al. 2002), including *A. waikula* (Kennedy et al. 2018), are a potential avenue for increasing the
597 chance of endosymbiont horizontal transfer (Su et al. 2019). Such a mechanism may explain the
598 high endosymbiont turnover observed in this system. However, from our current analyses, it is
599 impossible to separate a scenario of recent horizontal shifts of the endosymbionts, where
600 endosymbionts are acquired from the environment, from a scenario of older shifts, where
601 endosymbionts are acquired primarily during colonization events. Testing this will require
602 modeling host-microbiota evolution, which is currently challenging on such short evolutionary
603 timescales (Groussin et al. 2017; Perez-Lamarque & Morlon 2019).

604 Several previous studies have suggested that endosymbionts such as *Rickettsia* enhance
605 spider dispersal (Goodacre et al. 2009; Pekár & Šobotník 2007, 2008). However, although
606 colonized by various endosymbionts including *Rickettsia*, most populations of *A. waikula* are highly
607 locally endemic which suggests that species within a given genus of endosymbiont are not
608 associated with the loss of the host's dispersal here.

609 *Similarity of the gut microbiota*

610 Across all spider populations, we observe a conserved gut microbiota: Over approximately
611 one million years, the composition seems to have remained stable and does not reflect the
612 *Ariamnes* genetic structure. The core lineages at the generic level reflect typical gut microbiota of
613 arthropods (Engel & Moran 2013). However, slight but significant differences in the gut microbiota
614 are detectable between the different spider species and between the *A. waikula* populations within
615 Hawai'i Island.

616 We would expect transmission of the gut microbes to generate a strong pattern of
617 phylosymbiosis and potentially evidence of bottlenecks in the microbial diversity of the youngest
618 populations (Minard et al. 2015). Instead, we find a strong similarity in terms of microbial diversity
619 across the different islands and populations, decoupled from the spider's phylogeny. This absence
620 of a correlation with the host's phylogenetic structure suggests that the gut microbiota is likely
621 acquired from the spider's environment, most probably through the spider's diet (Kennedy et al.
622 2020; Zhang et al. 2018). The pattern of slight differentiation according to the host populations
623 could either be due to gradual changes in the host-filtering process (Mazel et al. 2018), or only
624 reflect the local variations of the bacterial pool available in the host environment.

625 The overall similarity of the gut microbiota in terms of the identity and diversity of genera
626 represented, has been shown in other systems, such as recently diversified *Anolis* lizards (Ren et
627 al. 2016) and *Cephalotes* turtle ants (Sanders et al. 2014). Such similarity likely conserves the
628 functional properties of the gut microbiota (Muegge et al. 2011), which could be particularly
629 advantageous for arthropods which often rely on gut symbionts to complement imbalanced
630 nutrition (Engel & Moran 2013) or detoxify toxin-rich diets (Adams et al. 2013). This gut microbiota
631 conservatism of *A. waikula* may reflect its niche conservatism, despite changes in geographical
632 range.

633

634 ***Conclusion***

635 Host organisms can harbor a large microbial diversity, and it is likely that the different
636 components of the microbial community experience different ecological and evolutionary
637 dynamics. In this work, we demonstrated a dichotomy between the endosymbiont communities
638 and the gut microbiota: For the host Hawaiian *Ariamnes* spiders, abiotic factors, in particular
639 isolation by distance along the volcano chronosequence, appear to have the strongest influence
640 in explaining divergence patterns. The phylogeny of the host strongly impacts the evolution of the
641 transmitted endosymbionts, which show clear shifts in their composition between species on
642 different islands. However, we find little differentiation in endosymbionts across populations of *A.*
643 *waikula* within Hawaʻi Island, despite the strong genetic structure of the host. In contrast to the
644 differences in the endosymbionts between taxa, the composition of the gut microbiota is similar
645 across both species and populations despite strong geographic isolation of their hosts. Our results
646 stress the high heterogeneity within the different components of the microbiota in terms of host-
647 associated evolution and calls for further investigations of the role on these heterogenous
648 microbial symbionts on promoting host diversification. Specifically, the conservation of the
649 microbiota across ecomorph and island in concert with broader sampling of both *Ariamnes* and
650 other arthropods would significantly contribute to our understanding of the dynamics of co-
651 evolution of microbe and host as it pertains to rapid speciation events.

652 **Acknowledgements**

653 We would like to acknowledge support and assistance from the following: The permit processing
654 and access to different reserves and private land was possible thanks to Steve Bergfeld (DOFAW
655 Big Island), Pat Bily (TNC Maui), Tabetha Block (HETF), Pomaikaʻi Kaniaupio-Crozier (Maui Land
656 and Pineapple), Lance DaSilva (DOFAW Maui), Charmian Dang (NAR), Melissa Dean (HETF),
657 Betsy Gagne (NAR), Lisa Hadway (DOFAW Big Island), Cynthia King (DLNR), Russell Kallstrom

(TNC Moloka‘i), Joey Mello (DOFAW Big Island), Ed Misaki (TNC Moloka‘i). For support and advice in the lab, we are very grateful to Lydia Smith (Evolutionary Genetics Lab, Museum of Vertebrate Zoology, UC Berkeley), Shana McDevitt (Vincent J. Coates Genomics Sequencing Laboratory, QB3, UC Berkeley), and Anna Sellas (California Academy of Sciences, San Francisco). Samples were provided by Susan Kennedy and Andrew Rominger. We thank Benjamin Peter, Thorfinn S. Korneliussen, and Line Skotte for advice on the analyses. L.K.B was supported by the Netherlands Organisation for Scientific Research VENI #863.14.020. B.P.L was supported by a master fellowship from the École Normale Supérieure of Paris. H.K. was supported by a postdoctoral fellowship by the German Research Foundation (DFG). Support for the project 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.K.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.

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.

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.

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.

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, 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.

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.

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.

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

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.

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.

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

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.

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.

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

Pekár S, Šobotník J (2007) Comparative study of the femoral organ in Zodarion spiders (Araneae: Zodariidae). *Arthropod structure & development* **36**, 105-112.

Pekár S, Šobotník J (2008) Erratum to "Comparative study of the femoral organ in Zodarion spiders (Araneae: Zodariidae)"[Arthropod Structure & Development 36 (2)(2007) 105–112]. *Arthropod Structure & Development* **37**, 93-94.

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.

Sharon G, Segal D, Ringo JM*, et al.* (2010) Commensal bacteria play a role in mating preference of Drosophila melanogaster. *Proceedings of the National Academy of Sciences* **107**, 20051-20056.

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, Bordenstein SR (2016) Speciation by symbiosis: the microbiome and behavior. *MBio* **7**, e01785-01715.

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.

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**.

Vieira FG, Lassalle F, Korneliussen TS, Fumagalli M (2015) Improving the estimation of genetic distances from Next-Generation Sequencing data. Biological Journal of the Linnean Society **117**(1):139-149

Wang IJ (2013) Examining the full effects of landscape heterogeneity on spatial genetic variation: a multiple matrix regression approach for quantifying geographic and ecological isolation. *Evolution* **67**, 3403-3411.

Weber MG, Wagner CE, Best RJ, Harmon LJ, Matthews B (2017) Evolution in a community context: on integrating ecological interactions and macroevolution. *Trends in ecology & evolution* **32**, 291-304.

White JA, Styer A, Rosenwald LC*, et al.* (2020) Endosymbiotic bacteria are prevalent and diverse in agricultural spiders. *Microbial ecology* **79**, 472-481.

Whitehouse M, Agnarsson I, Miyashita T*, et al.* (2002) *Argyrodes*: phylogeny, sociality and interspecific interactions—a report on the *Argyrodes* symposium, Badplaas 2001. *Journal of Arachnology* **30**, 238-245.

Yun J-H, Roh SW, Whon TW*, et al.* (2014) Insect gut bacterial diversity determined by environmental habitat, diet, developmental stage, and phylogeny of host. *Appl. Environ. Microbiol.* **80**, 5254-5264.

Zhang L, Yun Y, Hu G, Peng Y (2018) Insights into the bacterial symbiont diversity in spiders. *Ecology and evolution* **8**, 4899-4906.

Zhang L, Zhang G, Yun Y, Peng Y (2017) Bacterial community of a spider, Marpiss magister (Salticidae). *3 Biotech* **7**, 371.

Supplementary Figures

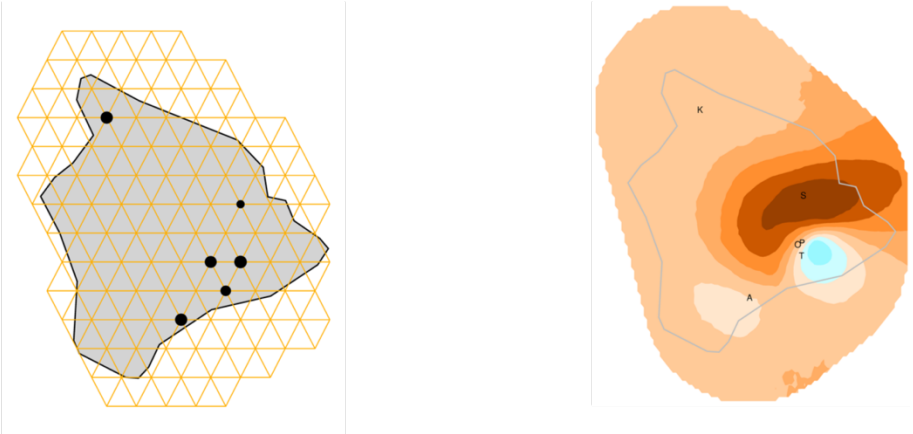


Figure S1: **Grid system used for the EEMS analyses on Hawaii Island (left) and corresponding results (right)**

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

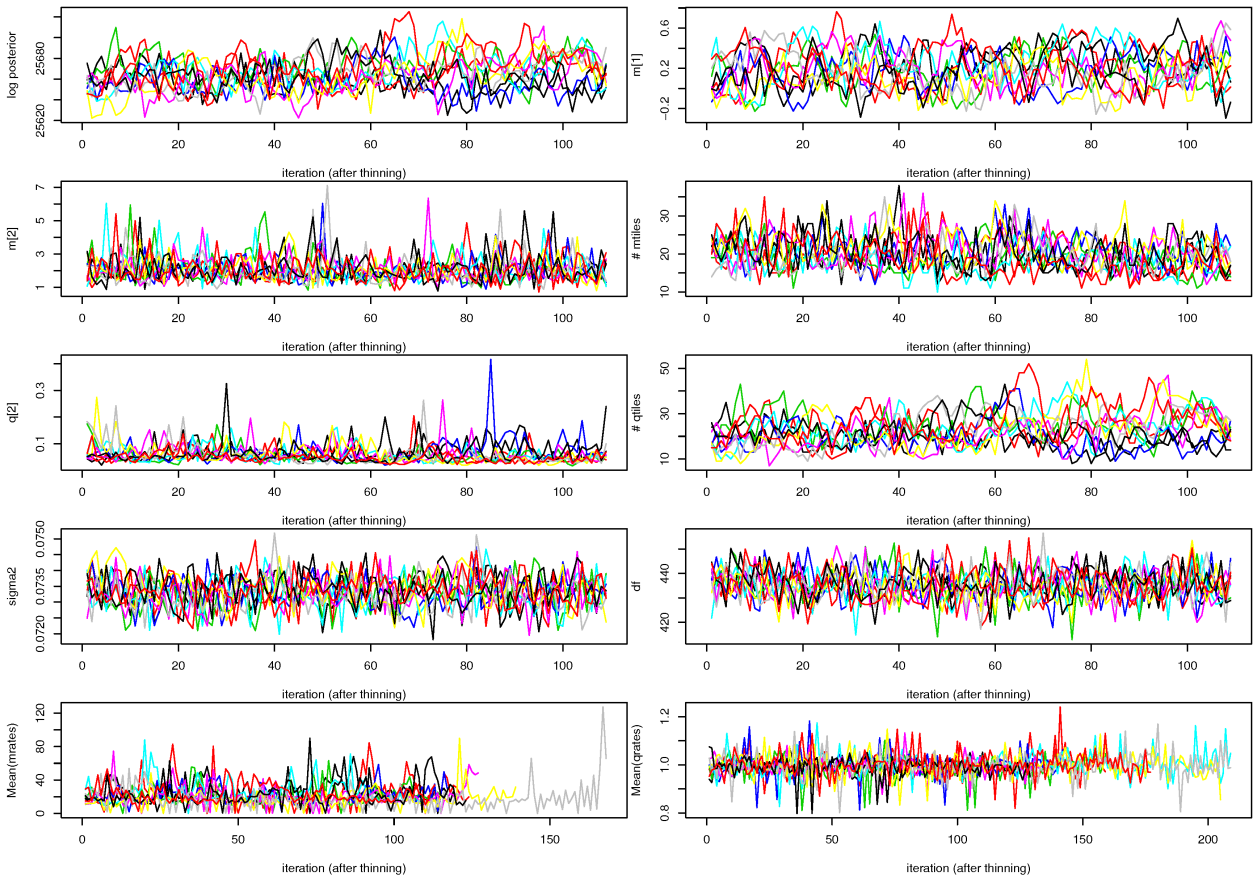


Figure S2. Traces of the MCMC runs of the 10 replicates to assess convergence. Only individuals of the gold ecomorph from Hawaii Island were used here.

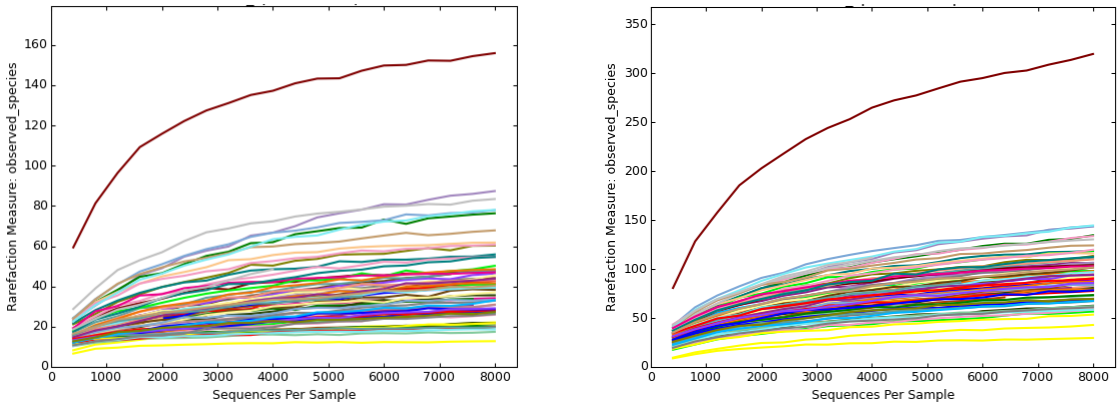


Figure S3: Rarefaction curve of the whole microbiota: number of observed OTU in each sample according to the number of rarefied reads per sample. Results of OTU at 97% are presented on the left, whereas Z-OTU are on the right. Rarefactions are performed 20 times independently and the mean value is plotted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

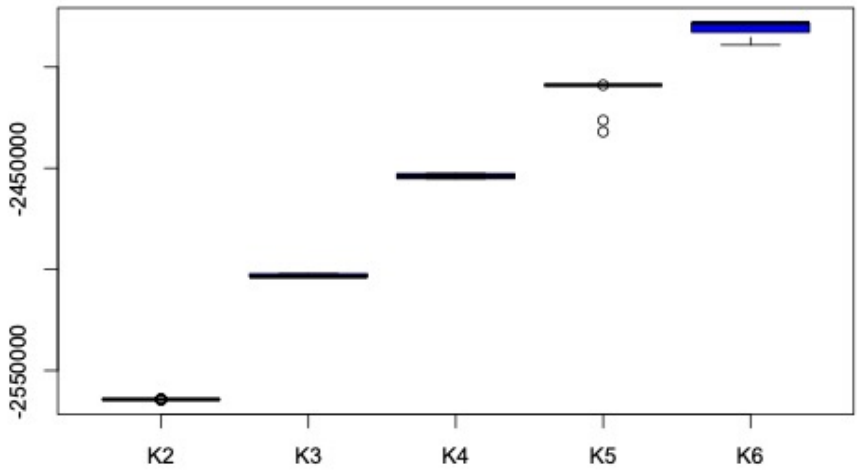
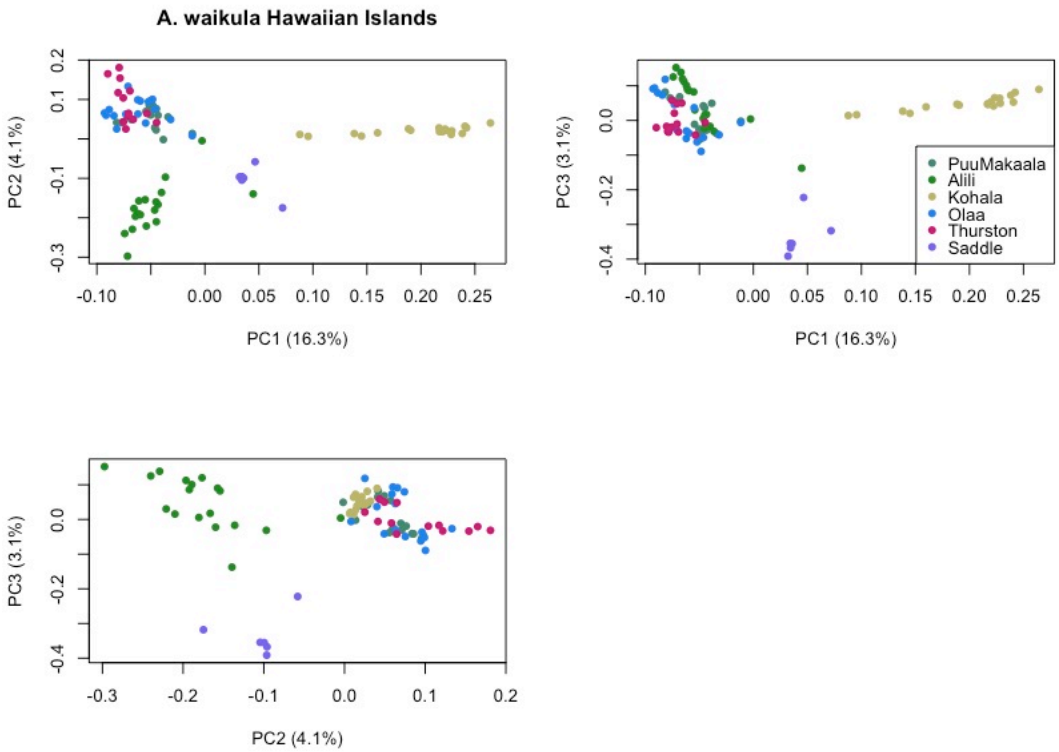


Figure S4: NgsAdmix convergence likelihoods for 50 iterations K values 2-6 for *A. waikula* from Hawai'i Island.

926 .

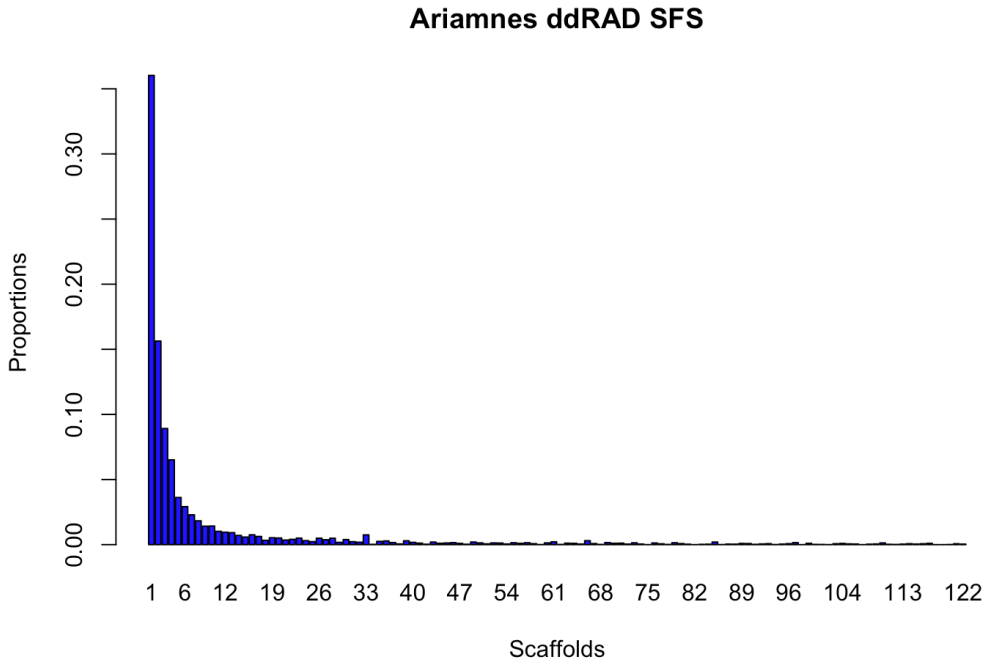


927

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

929

930



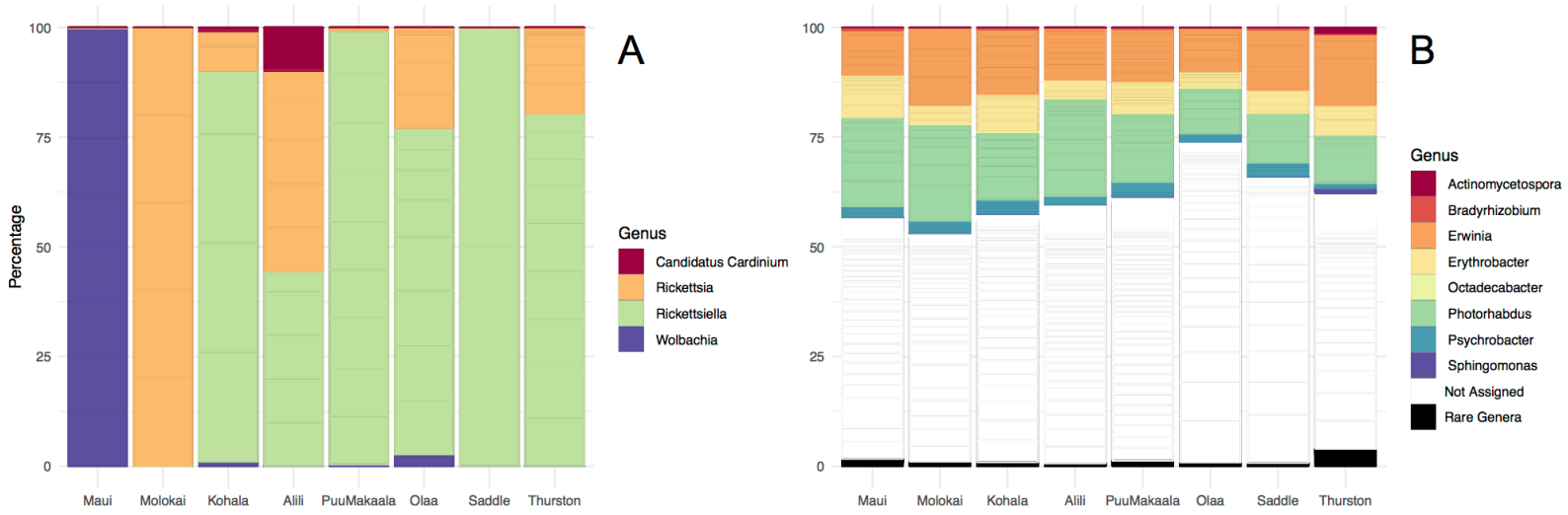
931

932

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

934

935



936 **Figure S7:** Relative abundances of endosymbiont genera per population (A) and gut symbionts

937 per spider population (B) defined from rarefied OTU table at 97%. Rare genera (representing less

938 than 1% of the abundance are merged together).

939

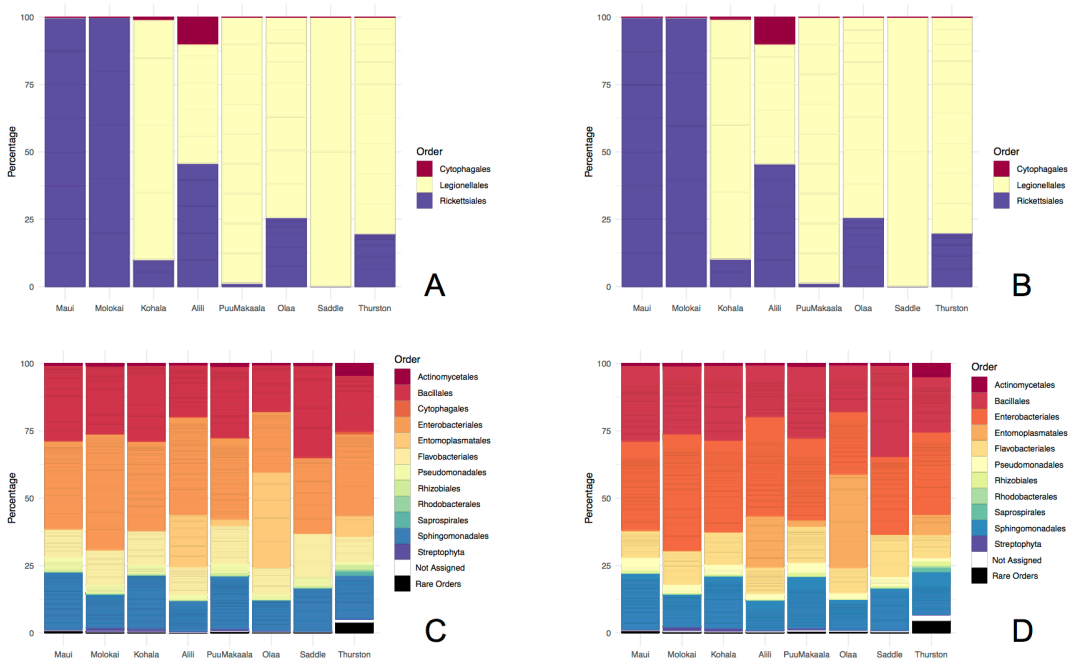
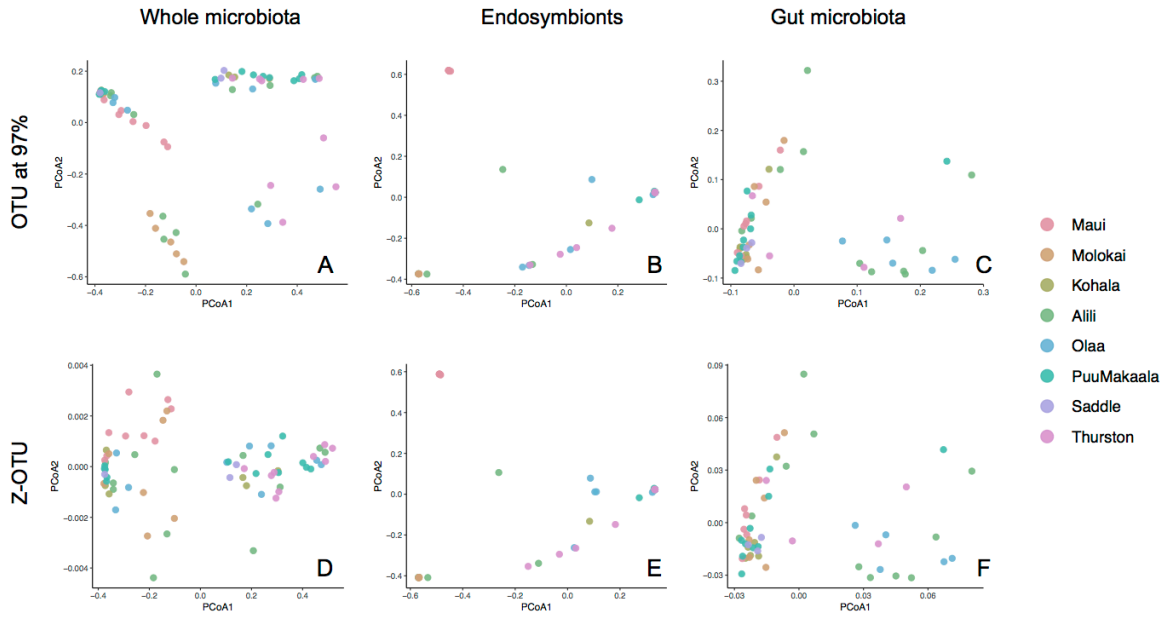


Figure S8: Relative abundances of orders: endosymbiont orders per population (A-B) and gut symbionts per spider population (C-D) defined from rarefied OTU table. Rare orders (representing less than 1% of the abundance are merged together). A and C (resp. B and D) are from 97% OTU (resp. Z-OTU).



947

948

949 **Figure S9:** PCoA decomposition of the microbial communities according to the type of OTU

950 clustering (OTU at 97% or Z-OTU) and the type of microbial communities (whole microbiota,

951 endosymbionts, or gut microbiota).

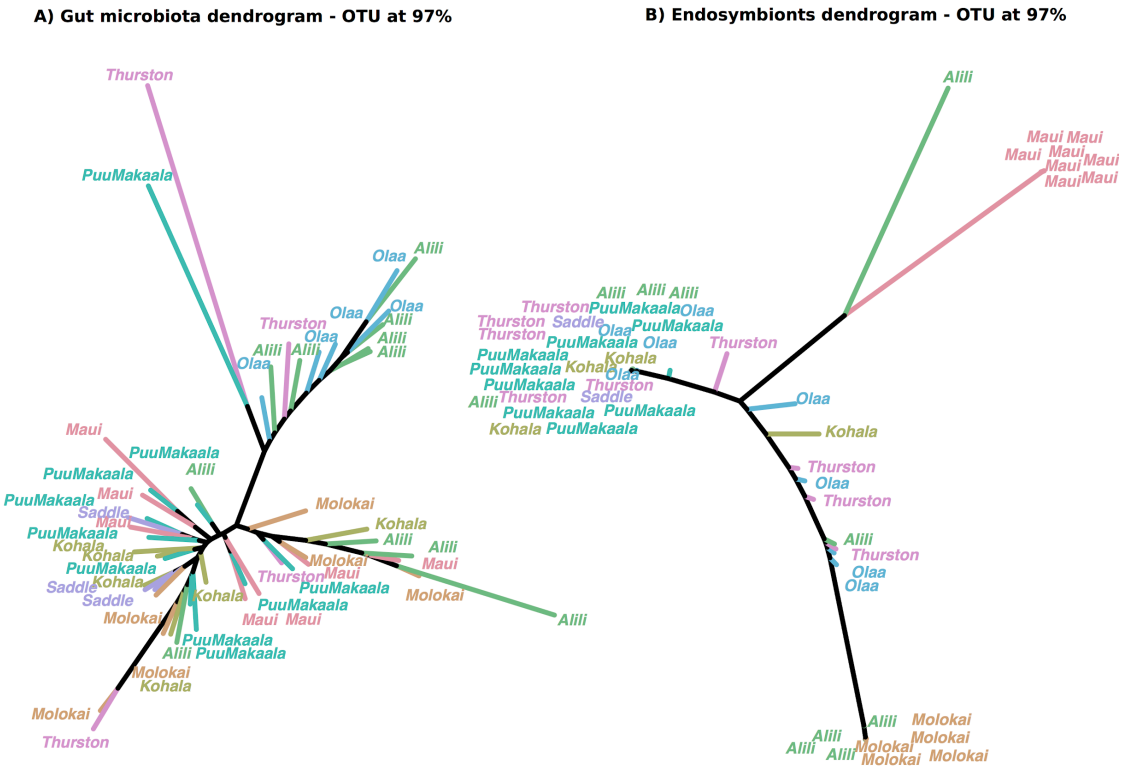


Figure S10: Microbiota dendrograms reconstructed from the gut microbiota (A) or the endosymbiont community (B) for the OTU at 97%.

Supplementary Tables

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

960 **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

961

962 **Table S3**: F_{ST} values for ddRAD data from *Ariamnes* spiders.

963
964
965

	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

966 **Table S4: Average Tajima's D values for each population of *A. waikula* and Maui (*A.***

967 ***melekelikimaka*) and Molokai (*A. n. spp.*).**

968

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
Molokai	-0.23	-2.46	3.66

969

970 **Table S5:** Mean alpha diversity of the microbial communities per host populations. Alpha
971 diversities were computed on rarefied OTU tables defined at 97% or Z-OTUs using the Chaos 1
972 index.
973

Island	Population	Number individuals	OTU at 97%		Z-OTU	
			Alpha diversity (Chaos 1)	Standard deviation	Alpha diversity (Chaos 1)	Standard deviation
Hawai'i	Alili	13	42.4	7.2	95.4	4.0
Hawai'i	Kohala	6	55.8	12.3	114.8	7.8
Hawai'i	Ola'a	9	41.6	3.0	96.3	8.5
Hawai'i	Pu'u Maka'ala	13	49.5	3.2	108.0	3.4
Hawai'i	Saddle	3	44.8	1.6	103.1	5.2
Hawai'i	Thurston	9	40.8	4.6	103.2	5.7
Maui	Puu Kukui	9	46.0	5.7	106.6	4.4
Moloka'i	Kamakou	7	52.2	6.7	106.6	5.3

974
975
976

977

978 **Table S6:** PERMANOVA testing the of host populations on the dissimilarities between microbial
979 communities according to the host populations. We indicated the adjusted R-squared and the
980 associated p-value for each test. Significant relationships indicate a differentiation of the microbial
981 communities according to the host populations.

982

983

Populations	OTU at 97%			Z-OTU		
	Whole microbiota	Gut microbiota	Endo-symbionts	Whole microbiota	Gut microbiota	Endo-symbionts
All populations	0.37905 (0.0001)	0.31109 (0.0001)	0.7146 (0.0001)	0.37709 (0.0001)	0.31739 (0.0001)	0.7115 (0.0001)
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)

984 **Table S7:** Mantel test comparing the microbial beta diversity dis-similarities to the genetic
985 distances or the phylogenetic distances of their associated *Ariamnes*. Beta diversity dissimilarities
986 were computed on rarefied OTU tables defined at 97% using the Bray-Curtis dissimilarities. Bold
987 values represent significant correlations. Mantel tests were either performed on all populations
988 (between *Ariamnes* species and *A. waikula* populations) or only on *A. waikula* populations.
989

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**	0.117	0.066
	Ariamnes phylogenetic distances	**0.323**	**9.9e-05**	0.026	0.321
Endosymbionts beta diversity distances	*Ariamnes* genetic distances	**0.620**	**9.9e-05**	0.011	0.500
	Ariamnes phylogenetic distances	**0.646**	**9.9e-05**	0.061	0.257
Gut microbiota beta diversity distances	*Ariamnes* genetic distances	-0.096	0.837	-0.043	0.668
	Ariamnes phylogenetic distances	-0.095	0.861	-0.130	0.912

990