

1   **Classification:** Major - Biological Sciences, Minor - Ecology

2

3

4

5

6   **The interaction networks of plants and symbiotic fungi are not restored**  
7   **following reforestation of a degraded subtropical ecosystem**

8   Cameron P. Egan<sup>1,2</sup>, Sean O.I. Swift<sup>1</sup>, Nicole A. Hynson<sup>1\*</sup>

9   <sup>1</sup>Pacific Biosciences Research Center, University of Hawai‘i at Mānoa, Honolulu, HI, USA

10   <sup>2</sup>Department of Biology, Okanagan College, Kelowna, BC, Canada

11

12   \*Corresponding Author: 3190 Maile Way Rm 101, University of Hawai‘i at Mānoa, Honolulu  
13   HI, 96822. E-mail: nhynson@hawaii.edu

14

15

16

17

18   **Keywords**

19   arbuscular mycorrhizal fungi, ecological networks, ecological restoration, symbiosis, subtropical  
20   native forests

21

22

23

24

25

26

27

28 **Abstract**

29 Through our actions, humans have caused nearly a 50% decline in forest cover globally. To  
30 ameliorate the impacts of deforestation such as habitat loss and increased carbon emissions,  
31 ecological restoration practices such as reforestation are often invoked. However, the success of  
32 these efforts is inconsistent and rarely do they lead to the reestablishment of target ecosystems or  
33 their services. One important factor that is often overlooked during restoration is the synergistic  
34 interactions of organisms with each other, for example plant interactions with microbial  
35 symbionts. This is a considerable oversight, as a multitude of studies have indicated the  
36 importance of symbiotic microbes in plant community assembly and ecological succession.

37 We examined whether restoration practices focused aboveground, were successful in restoring  
38 belowground symbiotic fungal communities and their ecological interactions. We sampled  
39 remnant subtropical montane forests and restored forest patches, and characterized their root-  
40 associated symbiotic fungi to assess their ecological interactions.

41 Plant-fungal network analyses revealed that remnant forest networks harboured significantly  
42 more specialized, and less dense interactions than restored habitat patches. Analysis of fungal co-  
43 occurrence networks revealed that remnant forest keystone taxa were absent in restored sites.  
44 Using classic ecological metrics we found that fungal community composition differed  
45 significantly between habitats, with noted differences in the relative abundance of fungal  
46 families whose functions also differ. Combined, these findings have adverse consequences for  
47 the re-establishment of native plant communities and ecological succession. We suggest that  
48 moving forward, restoration practices must take a multi-guild approach in order to increase their  
49 success.

50

51 **Significance**

52 Reforestation, or the practice of replanting trees and other woody species into degraded  
53 landscapes, has the potential to counter the negative impacts of forest loss. However,  
54 reforestation efforts rarely meet their goals of re-establishing similar biological communities and  
55 ecological functions as primary forests. By focusing on a single group of organisms like plants,  
56 reforestation efforts are potentially over simplifying complex ecosystems that include  
57 interactions among species. Here we show that even *c.* 20 years post reforestation, key  
58 interactions between plants and their associated fungal symbionts are missing. We contend that  
59 this loss of complexity limits the ability of subtropical forests to re-establish. Our results  
60 highlight the importance of considering symbiotic interactions during restoration projects, which  
61 will potentially improve their outcomes.

62

63

64 **Introduction**

65 Through our actions humans have modified >50% of the Earth's land surface (1). While  
66 concerted efforts have been taken to help minimize ecosystem losses through habitat  
67 preservation, there is a recognized need for effectual restoration of already degraded areas (2).  
68 However, recovery rates are often slow and restored ecosystems rarely achieve their targeted pre-  
69 disturbance state (3). A contributing factor could be that restoration methods over-simplify  
70 complex systems (4), and omit key ecosystem components or processes, thereby preventing full  
71 recovery. This is evident by restoration projects often being strongly biased towards single  
72 taxonomic groups (eg. plants, 5). Despite their well-recognized roles in local- and broad-scale  
73 ecosystem processes (6) microbes are often overlooked in the context of ecological restoration  
74 (5). Root symbiotic mycorrhizal fungi are a particularly salient guild of microorganisms to  
75 consider during restoration due to their ubiquity and significant impacts on plant success,  
76 biodiversity and ecosystem functioning (7, 8), as well as the trajectory of community succession  
77 (9). These fungi can potentially serve a valuable role during terrestrial ecological restoration,  
78 acting either as an ecosystem component to be manipulated (10, 11, 12), or as indicators of the  
79 state of recovery to a target ecosystem (13). In the present study, we set out to test the latter.

80 Ecosystem disturbance and secondary succession has been shown to affect plant  
81 symbiotic microbes such as mycorrhizal fungi by changing their abundance (14, 15), altering  
82 their community composition (16) and their functional traits (17). Due to the reliance of the  
83 majority of plants on mycorrhizal fungi to survive in the wild (18), the success of plant  
84 restoration projects and the sequential reestablishment of the food-webs for which plants form  
85 the foundation, may hinge upon whether the diversity and function of mycorrhizal interactions  
86 can also be restored. While previous studies have examined the potential role that mycorrhizal  
87 fungal pre-inoculation can have in steering successful ecological restoration (9, 19, 20, 21),  
88 information is lacking on the reestablishment of mycorrhizal interactions *in situ* and over time,  
89 limiting our understanding of belowground ecosystem development and plant-soil feedbacks  
90 post-restoration. To assess this, we employed traditional community ecology metrics and a  
91 contemporary ecological network approach within a landscape-scale reforestation project on  
92 Hawaii Island.

93 For symbiotic organisms, ecological network theory is an especially powerful tool to  
94 understand the complex interactions among hosts and their symbionts in nature (22). Rather than  
95 more traditional community ecology metrics of diversity, ecological networks measure the  
96 degree, extent and intensity of interactions among species within an ecosystem. Network patterns  
97 also provide valuable information on the stability of biotic interactions in the face of  
98 perturbations, and the degree of specificity among hosts and symbionts. Because they measure  
99 complex interactions, network approaches provide a framework for assessing the system-wide  
100 status of restored areas, relative to a reference such as primary forests. By assessing and  
101 comparing network properties we can target specific ecosystem components to manipulate that  
102 may accelerate or at least increase the success of future restoration efforts.

103 Previous work provides multiple insights about how ecological network properties are  
104 influenced by disturbance, and how networks vary during ecosystem recovery. For instance,  
105 examination of plant-pollinator networks has revealed that networks within intact ecosystems are  
106 more complex, relative to sites that have been disturbed (23). Investigations of networks between  
107 plants and soil biota have revealed similar patterns. In their examination of abandoned arable  
108 lands Morriën et al. (24) observed that connectance increased with successional age..  
109 Comparisons of mycorrhizal networks between young and old forests in Estonia have also shown  
110 that networks are more connected in old-growth forests, and increase in their number of  
111 specialists (25), indicating that niche partitioning is greater at later successional stages (26).  
112 Together these studies indicate that increasing network complexity and specialization are  
113 properties that should be targeted when restoring interaction networks, but these ideas remain to  
114 be explicitly tested.

115 Habitat reforestation, the practice of out-planting native woody plants into disturbed areas  
116 with the goal of spurring secondary succession, is a common practice in restoration ecology. We  
117 contend that focusing on individual species or taxonomic groups, while disregarding their  
118 ecological interactions, can put restoration efforts at risk for stall or failure. To assess this  
119 concept, we examined plant symbiotic mycorrhizal fungal communities and their interactions  
120 with their host plants within a large-scale reforestation project on Hawaii Island that appears to  
121 be stalled in its development. There, despite substantial attempts at re-establishing a native  
122 subtropical montane forest, seedling recruitment is sparse to non-existent and plant diversity is  
123 limited to out-planted native species of a single age-class (~22 years old). We determined the  
124 network architecture of mycorrhizal communities between remnant forest patches that represent  
125 the benchmark for restoration, to those in adjacent areas that have been previously disturbed and  
126 have since undergone reforestation. We predicted that the mycorrhizal networks of remnant  
127 forests would be denser, more connected, and more specialized relative to restored forest  
128 networks. We expected that differing land-use histories between these habitats would result in  
129 disparate mycorrhizal fungal communities, coinciding with loss of keystone fungal taxa from  
130 restored forests. Additionally, we predicted that remnant forests would harbour a higher relative  
131 abundance fungal taxa with late successional life history strategies (*sensu* 27).

## 132 **Methods**

### 133 *Study Area*

134 Sampling for this study was conducted within the 13,240 ha Hakalau Forest National  
135 Wildlife Refuge (herein Hakalau) located on the eastern slope of Mauna Kea on the Island of  
136 Hawai'i (19°51'N; 155°18'W). During the 1800's large swaths of land below Mauna Kea were  
137 converted to pastureland for livestock grazing, which included activities such as large-scale  
138 removal of native vegetation by means of logging, bulldozing and fires, and planting of exotic  
139 grasses for pasture (28). In 1987, a large-scale reforestation project was initiated in Hakalau to  
140 re-establish habitat for rare and endangered native Hawaiian forest birds (28). Initially, over

141 390,000 seedlings of the native canopy tree *Acacia koa* (koa) were planted into open pastureland  
142 areas over a 2-3 year period (29). Additional out-plantings of other native woody species began  
143 approximately 10 years later. However, these costly efforts have not led to the reestablishment of  
144 native forests or birds (30, 31).

145 *Sampling*

146 Sampling was undertaken in July and August 2017. At the time of sampling, a gradient of  
147 habitat types existed within Hakalau (SI Appendix, Fig. S1A) which included; 1) un-restored  
148 open-pasture consisting of exotic pasture grasses, 2) restoration corridors of out-planted *A. koa*  
149 and an understory of grasses, 3) restoration areas of *A. koa* with additional out-planted native  
150 woody species and smaller remnant patches of grass, and 4) remnant forest patches that are co-  
151 dominated by the native canopy trees *A. koa*, and *Metrosideros polymorpha*, and include  
152 subcanopy native trees such as *Cheirodendron trigynum* ('ōlapa), *Myrsine lessertiana* (kōlea),  
153 and *Coprosma rhynchocarpa* (pilo); and shrubs such as *Rubus hawaiensis* ('ākala),  
154 *Leptecophylla tameiameia* (pukiawe), and *Vaccinium calycinum* ('ōhelō). These species were  
155 chosen for two reasons: 1) they are representative of the plant community composition within  
156 remnant forest patches, and 2) they provide important resources for native birds at risk for  
157 extinction (31). For the purposes of our study we focused on just two of the habitat types within  
158 Hakalau: restoration areas (herein restored forest) and remnant forest patches (herein remnant  
159 forest) as our reference sites.

160 We sampled the roots of the same host plant species in both remnant and restored forest  
161 plots. Sampling plots were chosen along two established parallel transects, both of which  
162 transitioned from pasture into each of the habitat types listed above (SI Appendix, Fig. S1A).  
163 Along each of the two transects, six plots were established, divided equally between remnant and  
164 restored forest habitat types, resulting in a grand total of 12 plots. Within each plot we sampled  
165 roots from seven hosts to arbuscular mycorrhizal (AM) fungi, including six native host species;  
166 *M. polymorpha*, *A. koa*, *C. trigynum*, *M. lessertiana*, *C. rhynchocarpa*, and *R. hawaiensis*, and  
167 non-native grasses (most of which were *P. clandestinum* but were not identified to species). Plots  
168 were ~12 m in diameter and their perimeters were separated by ≥ 20 m (SI Appendix, Fig. S1B).  
169 Within each plot, we sampled roots from underneath up to eight individuals of each target  
170 species (SI Appendix, Appendix I). Generating a grand total of over 625 root samples (SI  
171 Appendix, Table S1). For each individual host, we sampled roots by tracing fine roots back to  
172 larger branching roots of the host. In the field, roots samples were bagged and stored on ice until  
173 they could be transferred to a 1°C cold room where they were kept until they could be processed  
174 for DNA extraction.

175 *Molecular analysis*

176 Cleaned and dried root samples ≥ 0.5 g were cut into 1 cm fragments using sterilized  
177 scissors, and DNA was extracted from a 0.25 g subsample using the MP Bio FastDNA® spin kit

178 for plant and animal tissue (MP Biomedicals, LLC, Santa Ana, California, USA), following the  
179 manufacturer's instructions. Then, we carried out a two-step PCR reaction to first specifically  
180 amplify the small subunit (SSU) of arbuscular mycorrhizal fungi's ribosomal RNA (rRNA) and  
181 then a second reaction adhered Illumina barcodes and adaptors to our amplicons (SI Appendix,  
182 Appendix II; 32).

183 *Soil chemical analysis*

184 For a subset of soil samples from within each plot (ranging 20-28), soil chemical analyses  
185 for organic matter (OM), estimated total nitrogen (N), readily available phosphorus (P),  
186 extractable cations (potassium (K), magnesium (Mg), calcium (Ca), sodium (Na)), hydrogen (H),  
187 Sulfate-S (S), pH, and cation exchange capacity (CEC). Analyses were performed by A & L  
188 Western Agricultural Laboratories, Inc. (Modesto, CA, USA).

189 *Bioinformatics*

190 Bioinformatic processing was conducted using the open-source platform quantitative  
191 insights into microbial ecology (QIIME version 2; <https://qiime2.org>). For details on  
192 demultiplexing, quality control methods, and taxonomic identification of sequences see SI  
193 Appendix, Appendix III.

194 *Data Analysis*

195 All analyses were conducted in R (version 3.5.2; 33), using R studio (34).

196 *Plant-AM fungal networks*

197 Bipartite plant-AM fungal networks were built using the bipartite package (35). For each  
198 plot, bipartite networks were assembled using an aggregated species-level matrix, where mean  
199 relative abundances of each AM fungal species were calculated for each plant species. Networks  
200 were assembled for each plot individually, and each network was considered an independent  
201 unit. Meaning, for each habitat type we had six independently assembled plant-AM fungal  
202 networks.

203 We examined multiple network-level characteristics including: connectance, nestedness,  
204 modularity, linkage density, and specialization of the entire network ( $H^2'$ ) (see SI Appendix,  
205 Table S2 for descriptions of each characteristic). For all network-level characteristics, we  
206 calculated mean values for both remnant and restored forest habitats by averaging the observed  
207 metrics across plots for each habitat type. We then used Welch unequal variance t-tests (36) to  
208 compare the metrics between the two habitat types.

209 Network characteristics of each host plant were examined by quantifying host symbiont  
210 range (an indicator of the capacity of a host to interact with different AM fungi) and host  
211 specialization on particular AM fungi ( $d'$ ) (SI Appendix, Table S2). For both metrics, differences  
212 between habitat types and among plant hosts were evaluated using a generalized linear model  
213 (GLM) with a quasi-Poisson distribution and a log link function, using the *glm* function available

214 in the R stats package (SI Appendix, Appendix IVa). Pairwise comparisons for every  
215 combination of host and habitat type were examined using the *emmeans* function in the *emmeans*  
216 package (37), followed by the *pairs* function in the R graphics package (v3.5.2).

217 To test whether observed network characteristics differed from random, we used a null  
218 model approach. We used two separate null model algorithms available in the bipartite package;  
219 *swap.web* and *r2dtable*. *Swap.web* was chosen because it provides a more constrained null  
220 analysis and is therefore more realistic. We concurrently used the *r2dtable* null model because  
221 similar to *swap.web*, *r2dtable* maintains marginal totals, but randomizes network connectance.  
222 Null bipartite networks generated using *r2dtable* were used to determine whether observed  
223 connectance differed from random, while bipartite generated using *swap.web* were used for the  
224 other network characteristics. For both null algorithms, observed matrices were randomized 1000  
225 times using the *nullmodel* function, followed by network characteristics calculated on null  
226 matrices. For each network index, we compared observed network characteristics to the null  
227 calculations using Welch unequal variance t-tests (36).

228 *AM fungal co-occurrence networks*

229 AM fungal networks were assembled using the Sparse Inverse Covariance Estimation for  
230 Ecological Association Inference package in R (SPIEC-EASI; 38). We assembled AM fungal  
231 networks using the *spiec.easi* function with 9999 iterations (SI Appendix, Appendix IVb).  
232 Inferred networks were then converted to igraph objects using the *adj2igraph* function, and  
233 analyzed networks using the igraph package (39). Overall AM fungal network characteristics  
234 were examined by evaluating connectedness (degree) and centrality (betweenness) between  
235 habitat types (Table S2). Each metric was calculated on fungal-fungal networks that were  
236 assembled individually for each plot. Overall network topologies were compared between the  
237 two habitat types using Welch unequal variance t-tests (36).

238 To evaluate potentially important AM fungal taxa within the networks for each habitat  
239 type, we identified candidate keystone and indicator species (SI appendix, Appendix IVc). A  
240 keystone species has traditionally been described as a species that has a disproportionately large  
241 effect on its environment/community relative to its abundance (40). In our case, we branded  
242 keystone species within our networks as taxa that had a disproportionately large effect on the  
243 network relative to their prevalence (41). Candidate keystone species were identified on  
244 composite networks generated for each habitat type, and were denoted as nodes that were  
245 maximal in both centrality and degree, while also presenting a low prevalence score ( $\leq 0.005$ , SI  
246 Appendix, Appendix IVb).

247 *AM fungal diversity*

248 To examine the influence of habitat type, host species, and their interaction on AM fungal  
249 richness, we used a GLM with a Poisson distribution and a log link function. Pairwise  
250 comparisons for every combination of host and habitat type were examined using the *emmeans*  
251 function, followed by the *pairs* function. To help interpret patterns of AM richness between

252 habitats we compared the species abundance distributions (SADs) using the *sads* package (42).  
253 Observed SADs were assembled using the *octav* command, and then multiple species probability  
254 distribution models were fit to the observed distributions using the *fitsad* command. The best  
255 species distribution model was chosen using Akaike's information criterion (AIC; 43) using the  
256 *AICtab* function in the *bbmle* package (44). The best species distribution model was chosen  
257 based on which model had the lowest AIC.

258 To examine patterns of AM fungal community nestedness within hosts and habitats, we  
259 compared observed nestedness "temperature" to 9999 permuted null model communities  
260 generated non-sequentially preserving sample (row) frequencies using the *oecosimu* function  
261 available in the *vegan* package (45), using the 'r0' algorithm and a seed of 96822.

262 Changes in AM fungal community composition among hosts and habitats were  
263 determined using the Bray-Curtis dissimilarity index (46). Variation in AM fungal community  
264 composition was visualized using nonmetric multidimensional scaling (NMDS) using the  
265 *metaMDS* function available in the *vegan* package (45). Soil chemical properties were fitted to  
266 the NMDS ordination using the *envfit* command in *vegan*. The effect of habitat type and host  
267 species on AM fungal community composition were determined by permutational multivariate  
268 analysis of variance using distance matrices (PERMANOVA; 47) using the *adonis* function and  
269 9999 permutation in the *vegan* package (45). Pairwise community compositional differences  
270 among all combinations of habitat type by host species were then done using the  
271 *pairwise.perm.manova* function from the *RVAideMemoire* package in R (48).

272 To help further interpret diversity patterns, we examined changes in relative abundance  
273 of the major AM fungal families between the two habitat types. Differences in relative  
274 abundance of each family were determined by binning sequence reads by family and then  
275 comparing mean relative abundance for each family between the two habitat types using Welch  
276 unequal variance t-tests (36).

## 277 Results

### 278 Plant-AM fungal networks

279 As predicted, overall network specialization ( $H^2'$ ) was significantly higher in remnant  
280 than restored forests (Fig. 1A; Welch unequal variance t-test;  $t = -4.42$ ,  $df = 6.64$ ,  $p = 0.003$ ).  
281 Null model analyses revealed that remnant forest networks were more specialized than expected  
282 by chance, while restored forest networks did not differ from null expectations (SI Appendix,  
283 Table S3). Similarly, host specialization on AM fungi ( $d'$ ) was significantly higher in remnant  
284 than restored forests, and varied among species (Fig. 2A, SI Appendix, Table S4). Null model  
285 analyses revealed that within remnant forests three out of the seven hosts were more specialized  
286 than expected by chance, while in restored forest plots, only one of these seven hosts was  
287 significantly more specialized than randomly assembled networks (SI Appendix, Table S5).  
288 While no significant differences in host symbiont range were observed between habitat types  
289 (Fig. 2B), when we compared observed symbiont range to that of randomly assembled networks,

290 we found that within restored forests *M. polymorpha* had a smaller symbiont range than expected  
291 by chance, while grasses had larger symbiont range than expected (SI, Appendix, Table S5).

292 Contrary to our expectations, overall density (the proportion of realized connections per  
293 species) of the plant-AM fungal networks was greater in restored versus remnant forests (Fig.  
294 1B; Welch unequal variance t-test;  $t = 2.99$ ,  $df = 5.76$ ,  $p = 0.03$ ) and our null models revealed  
295 that network density did not significantly differ from null expectations in restored forests and  
296 was lower than expected by chance in remnant forests (SI Appendix, Table S3). Related to  
297 network density, the mean number of AM fungal links per plant species was also significantly  
298 greater in restored versus remnant forests, indicating less host specificity (Fig. 1C; Welch  
299 unequal variance t-test;  $t = 3.03$ ,  $df = 9.98$ ,  $p = 0.01$ ). Again, null model analyses revealed that  
300 plant species within remnant forests had fewer links than expected by chance, with no significant  
301 deviations from null expectations with restored plots (SI Appendix, Table S3). For both habitat  
302 types network connectance was lower -, while both nestedness and modularity were higher than  
303 expected by chance indicating interactions between plants and AM fungi are formed through  
304 non-random processes (SI Appendix, Table S3; 49).

### 305 *AM fungal networks*

306 Keystone species analysis revealed two keystone AM fungi within remnant forest  
307 networks; *Glomus VTX00290* (Degree = 8, Betweenness Centrality = 1578.5, Prevalence =  
308  $1.05E-05$ ) and *Acaulospora VTX00227* (Degree = 7, Betweenness Centrality = 1822, Prevalence =  
309 = 0.0045) (Fig. 3A). For restored forest AM fungal networks, we identified *Archaeospora*  
310 *trapeii*, which was not detected in within the remnant forest as a candidate keystone species (Fig.  
311 3B; Degree = 50, Betweenness Centrality = 9467, Prevalence =  $1.04E-09$ ). Between the two  
312 habitats, *Gl. VTX00290* was detected in a similar percentage of samples (remnant forest = 14%,  
313 restored forest = 15%) and at a similar relative abundance (remnant forest =  $7.31E-05$  relative  
314 read abundance, restored forest =  $7.89 E-05$ ), but was not classified as a keystone species as it  
315 was not a hub within the network (Degree = 4, Betweenness Centrality = 27.5). The remnant  
316 forest keystone species *Ac. VTX00227* was not detected in restored forest plots.

317 Indicator species analysis detected 17 AM fungal taxa significantly associated with  
318 remnant forests, and 10 with restored forests (SI Appendix, Table S6). We also found that  
319 Claroideoglomeraceae ( $p = 8.37E-04$ ), in addition to Ambisporaceae ( $p = 2.13E-02$ ),  
320 Archaeosporaceae ( $p = 1.13E-03$ ), and Paraglomeraceae ( $p = 3.75E-07$ ) all had significantly  
321 higher relative abundances in restored forest plots relative to remnant forest (SI Appendix, Fig  
322 S4). Conversely, we observed Diversisporaceae ( $p = 1.72E-10$ ) and Gigasporaceae ( $p = 5.74E-$   
323 03) to both have higher abundances in remnant forests (SI Appendix, Fig. S4; Table S12).

### 324 *AM fungal diversity*

325 AM fungal richness was similar between habitat types, but differed significantly among  
326 hosts (SI Appendix, Fig. S5; Table S8). In total, 183 AM fungal taxa were detected in restored

327 forest samples (sample mean and SD:  $19.1 \pm 5.5$ ), and 184 taxa were detected in remnant forest  
328 samples ( $18.4 \pm 6.1$ ). Overall, we detected 212 AM fungal taxa from 9 families and 13 genera.  
329 Ninety-two of our taxa were matched VT in the MaarJAM or NCBI databases, and the remaining  
330 120 were undetermined beyond the subphyla level of Glomeromycotina. Post-hoc comparisons  
331 revealed that *M. lessertiana* within remnant forests harboured the most AM fungal taxa, while *M.*  
332 *polymorpha* within remnant forests harboured the fewest (SI Appendix, Table S9). AM fungi  
333 from both remnant (Nested temperature = 15.00; P = 1e-04) and restored (Nested temperature =  
334 7.58, P = 0.001) forest habitat types had significant nestedness patterns among plant hosts when  
335 compared to null assembled communities (Fig. 4). Meaning that plants from the same habitat  
336 type hosted subsets of a species pool of AM fungi, rather than entirely new AM fungal species  
337 assemblages per host. Within remnant forests, AM fungal communities for all plant hosts were  
338 nested within *M. lessertiana* fungal communities (Fig. 4A), while in restored forests they were  
339 nested within grass fungal communities (Fig. 4B).

340 Observed SADs indicated that AM fungal communities from remnant forests had more  
341 rare species and a more symmetric SAD than restored forest communities (SI Appendix, Fig.  
342 S6). The observed SAD for restored forests was bimodal, having two peaks in species  
343 abundances at octave 8 and octave 11. Maximum likelihood estimation revealed that both  
344 remnant (Fig. 5A; P = 2.2e-16) and restored (Fig. 5B; P = 2.2e-16) AM fungal communities  
345 fitted significantly to a Poisson log-normal type of SAD.

#### 346 *AM fungal community membership*

347 AM fungal community composition was significantly influenced by habitat type, host,  
348 and their interaction (SI Appendix, Table S10). Pairwise comparisons revealed that AM fungal  
349 communities differed by habitat type (Pairwise Permutation MANOVA; p = 0.0001,  
350 permutations = 9999, Fig. 5), and soil pH ( $r^2 = 0.42$ , p = 0.001) was the best abiotic predictor of  
351 community composition (Table S11). All soil chemical variables, except Ca, Na, and P were  
352 significantly correlated with the ordination obtained (Fig. 5; SI Appendix, Table S11).  
353 Additionally, all soil chemical variables, except Ca, Mg, and P differed between habitat types (SI  
354 Appendix, Table S12). Within both habitats, AM fungal communities differed significantly  
355 among hosts (SI Appendix, Table S13). Congruent with greater host specificity, NMDS analysis  
356 revealed that AM fungal communities from the same hosts within remnant versus restored  
357 forests clustered closer together (SI Appendix, Fig. S7).

## 358 **Discussion**

359 We investigated whether the common restoration practice of out-planting native plant  
360 species into previously disturbed areas is successful in re-establishing critical belowground  
361 symbiotic networks. By comparing traditional measures of diversity, and contemporary measures  
362 of biotic interactions, we found that even after 20 years post reforestation these areas do not  
363 mimic more intact remnant forests and have not recovered to their pre-disturbance state (Fig. 5).

364 In the current study, the best indicators for remnant forests were high overall network and  
365 individual host species specialization (Figs 1A & 2A). The degrees of specialization within  
366 remnant forests were significantly higher than expected by chance, while being essentially  
367 indistinguishable from random within restored sites (SI Appendix, Table S3). We propose three  
368 possible non-exclusive mechanisms that could be contributing to these patterns. First, due to the  
369 age differences of the forests, their environmental conditions are distinct. For instance, higher  
370 root densities and lower micronutrient pools in the older remnant forests, could lead to greater  
371 competition for soil resources than in the younger restored forest (SI Appendix, Table S12). This  
372 in turn, may lead to more niche partitioning belowground and hosts partnering with the specific  
373 fungi that provide them with the most benefits for the lowest costs (50). Second, fungal taxa that  
374 found in the remnant forests are not establishing within restored sites, which could be hindering  
375 native plant recruitment (51). This idea is supported by our findings that: a) restored forests were  
376 missing keystone AM fungal taxa present in the remnant forests, b) each habitat type had distinct  
377 indicator taxa, and c) overall fungal community composition differed significantly between the  
378 two habitats. Furthermore, analyses of the species abundance distributions (SADs) of AM fungi  
379 in the two habitat types indicated a drop-off of rare taxa between remnant and restored sites (SI  
380 Appendix, Fig. S6), and ordination analysis revealed greater intermixing of AM fungi among  
381 hosts in restored sites (SI Appendix, Fig. S7), both of which are indicative of a more uniform  
382 AM fungal community.

383 A third potential source leading to decreased specialization within restored sites is their  
384 historical occupation by invasive plant hosts, namely grasses. Previous work has shown that  
385 specialist AM fungal taxa reluctantly colonize novel plant hosts (52), meaning that generalist  
386 fungi will favourably associate with invasive hosts novel to an ecosystem. Subsequently, positive  
387 feedbacks between invasive hosts and generalist fungi would increase the abundance of these  
388 fungi creating a soil legacy affect even after the removal of these hosts (53, such as in our  
389 restored sites). The species interaction and nestedness patterns we observed support this concept.  
390 Within restored sites we observed grasses to associate with a broader range of AM fungal taxa  
391 than expected by chance and all other out-planted native hosts' fungal communities nested  
392 within grasses (Fig. 4). This suggests that within restored forests, grass AM fungi are the "donor"  
393 pool from which all other hosts accumulate their mycorrhizal symbionts, and these fungi may be  
394 less beneficial to native hosts thereby hindering their re-establishment. For example, within  
395 restored sites *M. polymorpha*, a dominant species in native Hawaiian forests (54), harboured  
396 symbionts that were distinct from its specialized partners in remnant forests (Fig. 4, SI  
397 Appendix, Table S5). In the remnant forests, *M. lessertiana* harboured basically all of the AM  
398 fungal diversity found in all other hosts (Fig. 4). Therefore isolating and pre-inoculating native  
399 hosts with these fungi prior to outplanting may aid in the overall reestablishment of the target  
400 ecosystem.

401 The consequences of specialization loss within the restored forests is that native hosts are  
402 not re-establishing their specific mycorrhizal partnerships, but are instead associating with a

403 greater diversity of low-quality symbionts. This has led an interaction network with more links  
404 and greater density making it less susceptible to perturbations than randomly assembled, or the  
405 remnant forest networks (Fig. 1). At first, the creation of a more robust ecological network may  
406 seem like an unintended positive effect of reforestation. However, we would argue that this is not  
407 necessarily the case for the following reasons: 1) despite a lack of seed limitation from now  
408 mature out-planted trees, reforestation efforts have not led to natural regeneration indicating that  
409 specific belowground interactions may be needed for hosts to reproduce successfully and, 2) now  
410 that these robust networks are established, they will be harder to manipulate and may represent  
411 an alternative stable state to late-stage forest succession.

412 Two keystone AM fungal taxa from remnant forest networks were either lost, or did not  
413 perform as such within the networks of restored forests (Fig. 3). Specifically, we found that a  
414 previously unidentified *Acaulospora* species was absent within restored forest, while playing a  
415 major role within remnant forest networks (Fig. 3). This is a notable observation, as a loss of  
416 keystone taxa within interaction networks has been shown to lead to cascading effects (55). For  
417 example, the loss of keystone taxa can potentially proliferate throughout networks resulting in  
418 the loss, or decreasing abundance of other taxa with similar traits. For plant communities, a  
419 potential consequence of losing fundamental mycorrhizal symbionts is the alteration of  
420 community composition and/or ecosystem functioning. This pattern has been observed  
421 previously in other studies by Duhamel et al. and Banerjee et al. (56, 57). Notably, they found  
422 that the effects of disturbance altered mycorrhizal community composition, and especially  
423 influenced by a subset of rare taxa or keystone taxa which in turn affected plant establishment  
424 and biogeochemical cycling. In our system we suspect a similar dynamic exists, where the loss  
425 or alteration of rare and keystone taxa could be a possible explanation as to why native  
426 subtropical montane forests have not successfully established 20 years post-restoration.  
427 However, further work examining the specific roles of these keystone taxa, especially in terms of  
428 their function and physiological effect on hosts will be vital to understanding their potential  
429 importance in re-establishing forests (57).

430 Changes in AM fungal community composition between remnant and restored forest sites  
431 coincided with a decrease in relative abundance of taxa with late successional life history  
432 strategies (SI Appendix, Fig. S4; 17, 27). Specifically, we observed a decrease in  
433 Diversisporaceae and Gigasporaceae taxa between remnant and restored forests (SI Appendix,  
434 Fig. S4). The latter family was also an indicator of remnant forests. Two potential interacting  
435 mechanisms could be contributing to their observed decline. First, these families are  
436 characterized by high investment into extraradical hyphae in the soil (58, 59), which would  
437 enable hosts to access sparse soil nutrient pools where competition is high (60). By enabling  
438 hosts to access these pools, fungal taxa with such traits would theoretically receive higher  
439 rewards from their hosts, thus securing their place in the community (60). Second, families such  
440 as Gigasporaceae have slow recovery rates after mechanical disruptions in the soil (61).  
441 Therefore, such taxa would be less likely to re-establish post disturbance (such as our restored

442 sites). Coincidentally, we observed relative increase in the abundance of Archaeosporaceae in  
443 restored habitats. A recent meta-analysis by van der Heyde et al. (61) identified this family as  
444 highly tolerant to disturbance. A potential consequence of these shifts in AM fungal community  
445 composition that coincide with the alteration of specific functional traits is stalled of native plant  
446 recruitment and ecosystem succession like we have observed at Hakalau (56). To help mitigate  
447 these losses, and to help increase restoration success, AM fungal taxa from groups with late  
448 successional life history strategies should be targeted for future reforestation efforts (9, 20).

## 449 **Conclusions**

450 We proposed that by focusing on individual guilds of organisms such as plants, while  
451 disregarding their interactions with symbionts, puts restoration efforts are at risk for stall or  
452 failure. Our examination of the symbiotic fungal communities in an area that has undergone  
453 large-scale restoration, revealed that the ‘restored’ mycorrhizal communities were substantially  
454 different from remnant forest sites, and lacked many of their properties. Furthermore, we found  
455 that the interaction patterns between plants and mycorrhizal fungi were not restored, especially  
456 the specialized interactions of foundational native plant species. To help mitigate this issue, we  
457 suggest adopting a holistic approach to restoration, which focuses on restoring not only  
458 individuals and species, but also their ecological interactions. While this requires a thorough  
459 understanding of the ecology of the target ecosystem, the pay-off is substantial, as these  
460 connections are essential to long-term ecosystem functioning and resilience.

## 461 **Acknowledgments**

462 We would like to thank Corbin Amend, Dave Bouck, Ken Davidson, Sarah Schoepflin, and  
463 Stephanie Yelenik for help with sampling, and the Hakalau Forest National Wildlife Refuge for  
464 site access. Additionally, we would like to thank James Downey, Kacie Kajihara, Terrence  
465 McDermott, and Danyel Yogi for their dedication and long-hours processing samples for  
466 molecular and chemical analyses. We also wish to thank Anthony Amend, Carla D’Antonio,  
467 Pierre-Luc Chagnon, Miranda Hart, Vincent Merckx, Evan Rehm, Laura Tipton, and Stephanie  
468 Yelenik for discussion of ideas and/or reading early versions of this manuscript. This work was  
469 funded by the National Science Foundation (#1556856).

## 470 **Author’s Contributions**

471 All authors contributed to the conception of sampling design, sampling in the field, and  
472 conception of ideas for analysis. C.P.E. and S.O.I.S. conducted molecular analyses and prepared  
473 amplicon libraries for sequencing. C.P.E. analysed the data and led the writing of the manuscript.  
474 All authors contributed critically to the drafts and gave final approval for publication.

## 475 **Data Accessibility**

476 Data will be made available on FigShare (<https://figshare.com/>) following acceptance for  
477 publication.

478 **References**

- 479 1. R.L.B. Hooke, J.F. Martín-Duque, J. Pedraza, Land transformation by humans: a review.  
480 *GSA today*, **22**(12), 4-10 (2012). doi: 10.1130/GSAT151A.1.
- 481 2. International Union for the Conservation of Nature, Forest landscape restoration (2015).  
482 [https://www.iucn.org/about/work/programmes/forest/fp\\_our\\_work/fp\\_our\\_work\\_thematic/fp\\_our\\_work\\_flr/](https://www.iucn.org/about/work/programmes/forest/fp_our_work/fp_our_work_thematic/fp_our_work_flr/) (accessed September 2018).
- 484 3. H.P. Jones et al., Restoration and repair of Earth's damaged ecosystems. *Proceedings of the Royal Society B*, **285**(1873), pii: 20172577 (2018). doi: 10.1098/rspb.2017.2577.
- 486 4. M.A. Palmer, R.F. Ambrose, N. LeRoy Poff, Ecological theory and community restoration ecology. *Restoration Ecology*, **5**, 291-300 (1997). doi: 10.1046/j.1526-100X.1997.00543.x
- 488 5. C.E. Wainwright et al., Links between community ecology theory and ecological restoration are on the rise. *Journal of Applied Ecology*, **55**, 570-581 (2018). doi: 10.1111/1365-2664.12975
- 491 6. M.G.A. van der Heijden, R.D. Bardgett, N.M. van Straalen, The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters*, **11**, 296-310 (2008). doi: 10.1111/j.1461-0248.2007.01139.x
- 494 7. J.D. Bever, et al., Rooting theories of plant community ecology in microbial interactions. *Trends in Ecology and Evolution*, **25**, 468-478 (2010). doi: 10.1016/j.tree.2010.05.004
- 496 8. R.D. Bardgett, W.H. van der Putten, Belowground biodiversity and ecosystem functioning. *Nature*, **515**, 505-511 (2014). doi: 10.1038/nature13855
- 498 9. L. Koziol, J.D. Bever, Mycorrhizal feedbacks generate positive frequency dependence accelerating grassland succession. *Journal of Ecology*, **107**(2), 622-632 (2019). doi: 10.1111/1365-2745.13063
- 501 10. P.H. Thrall et al., Seed inoculation with effective root-nodule bacteria enhances revegetation success. *Journal of Applied Ecology*, **42**(4), 740-751 (2005). doi: 10.1111/j.1365-2664.2005.01058.x
- 504 11. P. Kardol, D.A. Wardle, How understanding aboveground-belowground linkages can assist restoration ecology. *Trends in Ecology and Evolution*, **25**(11), 670-679 (2010). doi: 10.1016/j.tree.2010.09.001
- 507 12. E.R.J. Wubs, W.H. van der Putten, M. Bosch, T.M. Bezemer, Soil inoculation steers restoration of terrestrial ecosystems. *Nature Plants*, **2**, 16107 (2016). doi: 10.1038/NPLANTS.2016.107
- 510 13. J. Harris, Soil microbial communities and restoration ecology: facilitators or followers? *Science*, **325**, 573-574 (2009). doi: 10.1126/science.1172975
- 512 14. M.M. Alguacil et al., The impact of tillage practices on arbuscular mycorrhizal fungal diversity in subtropical crops. *Ecological Applications*, **18**(2), 527-536 (2008). doi: 10.1890/07-0521.1
- 515 15. S.R. Holden, K.K. Treseder, A meta-analysis of soil microbial biomass response to forest disturbances. *Frontiers in Microbiology*, **4**, 163 (2013). doi.org/10.3389/fmicb.2013.00163
- 517 16. T.K. Schnoor, Y. Lekberg, S. Rosendahl, P. Axel Olsson, Mechanical soil disturbance as a determinant of arbuscular mycorrhizal fungal communities in a semi-natural grassland. *Mycorrhiza*, **21**, 211-220 (2011). doi: 10.1007/s00572-010-0325-3

- 520 17. M.M. Hart, P.D. Zaitsoff, M. van der Heyde, J. Pither, Testing life history and trait-based  
521 predictions of AM fungal community assembly. *Pedobiologia*, **59**, 203-213 (2016). doi:  
522 10.1016/j.pedobi.2016.06.001
- 523 18. S.E. Smith, D.J. Read, Mycorrhizal symbiosis, 3<sup>rd</sup> edn. (Academic Press 2008).
- 524 19. E.L. Middleton et al., Locally adapted arbuscular mycorrhizal fungi improve vigor and  
525 resistance to herbivory of native prairie plant species. *Ecosphere*, **6**(12), 276 (2015). doi:  
526 10.1890/ES15-00152.1
- 527 20. L. Koziol et al., The plant microbiome and native plant restoration: the example of native  
528 mycorrhizal fungi. *BioScience*, **68**(12), 996-1006 (2018). doi: 10.1093/biosci/biy125
- 529 21. L. Neuenkamp, S.M. Prober, J.N. Price, M. Zobel, R.J. Standish, Benefits of mycorrhizal  
530 inoculation to ecological restoration depend on plant functional type, restoration context and  
531 time. *Fungal Ecology*, **40**, 140-149 (2018). doi: 10.1016/j.funeco.2018.05.004
- 532 22. J. Bascompte, Disentangling the web of life. *Science*, **325**, 416-419 (2009). doi:  
533 10.1126/science.1170749
- 534 23. M.L. Forup, K.S.E. Henson, P.G. Craze, J. Memmott, The restoration of ecological  
535 interactions: plant-pollinator networks on ancient and restored heathlands. *Journal of  
536 Applied Ecology*, **45**, 742-752 (2008). doi: 10.1111/j.1365-2664.2007.01390.x
- 537 24. E. Morriën et al., Soil networks become more connected and take up more carbon as nature  
538 restoration progress. *Nature Communications*, **8**, 14349 (2017). doi: 10.1038/ncomms14349
- 539 25. A.E. Bennett et al., Arbuscular mycorrhizal fungal networks vary throughout the growing  
540 season and between successional stages. *PLoS ONE*, **8**(12), e83241 (2013).  
541 doi.org/10.1371/journal.pone.0083241
- 542 26. N. Blüthgen, F. Menzel, N. Blüthgen, Measuring specialization in species interaction  
543 networks. *BMC Ecology*, **6**, 9 (2006). doi.org/10.1186/1472-6785-6-9.
- 544 27. P.L. Chagnon, R.L. Bradley, H. Maherli, J.N. Klironomos, A trait-based framework to  
545 understand life history of mycorrhizal fungi. *Trends in Plant Science*, **18**(9), 484-491 (2013).  
546 doi: 10.1016/j.tplants.2013.05.001
- 547 28. United States Fish and Wildlife Service, Hakalau forest national wildlife refuge  
548 comprehensive conservation plan (2010).  
549 [https://www.fws.gov/pacific/planning/main/docs/HI-  
550 PI/Hakalau%20Forest%20NWR%20Draft%20CCP-EA.pdf](https://www.fws.gov/pacific/planning/main/docs/HI-PI/Hakalau%20Forest%20NWR%20Draft%20CCP-EA.pdf) (accessed September  
551 2018)
- 552 29. J. Jeffrey, B. Horiuchi, Tree planting at Hakalau forest national wildlife refuge - the right  
553 tool for the right stock type. *Native Plants*, **4**(1), 30-31 (2003). doi: 10.3368/npj.4.1.30
- 554 30. S.G. Yelenik, Linking dominant Hawaiian tree species to understory development in  
555 recovering pastures via impacts on soil and litter. *Restoration Ecology*, **25**(1), 42-52 (2017).  
556 doi: 10.1111/rec.12377
- 557 31. E.H. Paxton et al., Rapid colonization of a Hawaiian restoration forest by a diverse avian  
558 community. *Restoration Ecology*, **26**(1), 165-173 (2018). doi: 10.1111/rec.12540
- 559 32. CP Egan, D.W. Li, J.N. Klironomos, Detection of arbuscular mycorrhizal fungal spores in  
560 the air across different biomes and ecoregions. *Fungal Ecology*, **12**, 26-31 (2014). doi:  
561 10.1016/j.funeco.2014.06.004

- 562 33. R Core Team, R: A language and environment for statistical computing. R Foundation for  
563 Statistical Computing, Vienna, Austria (2019). URL <http://www.R-project.org/>.
- 564 34. R Studio Team, RStudio: Integrated Development for R. RStudio, Inc., Boston, MA URL  
565 (<http://www.rstudio.com/>).
- 566 35. C.F. Dormann, B. Gruber, J. Fruend, Introducing the bipartite package: analysing ecological  
567 networks. *R news*, **8**(2), 8-11 (2008).
- 568 36. B.L. Welch, The generalization of “student”s’ problem when several different population  
569 variances are involved. *Biometrika*, **34**(12), 28–35 (1947). doi: 10.1093/biomet/34.1-2.28
- 570 37. R. Lenth, emmeans: Estimated Marginal Means, aka Least-Squares Means. R package  
571 version 1.3.3. (2019).
- 572 38. Z. Kurtz, C. Mueller, E. Miraldi, R. Bonneau, SpiecEasi: Sparse Inverse Covariance for  
573 Ecological Statistical Inference. R package version 1.0.5 (2019).
- 574 39. G. Csardi, T. Nepusz, The igraph software package for complex network research, *Inter  
575 Journal, Complex Systems* 1695 (2006). <http://igraph.org>
- 576 40. R.T. Paine, A note on trophic complexity and community stability. *The American Naturalist*,  
577 **103**(929), 91–93 (1969). doi:10.1086/282586.
- 578 41. D. Berry, S. Widder, Deciphering microbial interactions and detecting keystone species with  
579 co-occurrence networks. *Frontiers in Microbiology*, **5**, 219 (2014). doi:  
580 10.3389/fmicb.2014.00219
- 581 42. P.I. Prado, M.D. Miranda, A. Chalom, sads: Maximum Likelihood Models for Species  
582 Abundance Distributions. R package version 0.4.2. (2018). <https://CRAN.R-project.org/package=sads>
- 583 43. H. Akaike, Information theory and an extension of maximum likelihood principle. In: Petrov  
584 BN, Csaki F, Proceedings of the Second International Symposium of Information Theory.  
585 (Akademiai Kiado 1973), pp 267-281.
- 586 44. B. Bolker, bbmle: Tools for General Maximum Likelihood Estimation. R package version  
587 1.0.20 (2017). <https://CRAN.R-project.org/package=bbmle>
- 588 45. J. Oksanen et al., vegan: Community Ecology Package. R package version 2.5.4 (2019).  
589 <https://CRAN.R-project.org/package=vegan>
- 590 46. J.R. Bray, J.T. Curtis, An ordination of the upland forest communities of southern  
591 Wisconsin. *Ecological Monographs*, **27**(4), 325-349 (1957). doi: 10.2307/1942268
- 592 47. M.J. Anderson, A new method for non-parametric multivariate analysis of variance. *Austral  
593 Ecology*, **26**, 32-46 (2001). doi: 10.1111/j.1442-9993.2001.01070.pp.x
- 594 48. M. Hervé, RVAideMemoire: Testing and Plotting Procedures for Biostatistics. R package  
595 version 0.9.72. (2019). <https://CRAN.R-project.org/package=RVAideMemoire>
- 596 49. S.K. Sepp et al., Non-random association patterns in a plant-mycorrhizal fungal network  
597 reveal host-symbiont specificity. *Molecular Ecology*, **28**(2), 365-378 (2019). doi:  
598 10.1111/mec.14924
- 599 50. E.T. Kiers et al., Reciprocal rewards stabilize cooperation in the mycorrhizal symbiosis.  
600 *Science*, **333**(6044), 880-882 (2011). doi: 10.1126/science.1208473
- 601 51. L. Koziol, J.D. Bever, Mycorrhizal response trades off with plant growth rate and increases  
602 with plant successional status. *Ecology*, **96**(7), 1768-1774 (2015). doi: [doi.org/10.1890/14-2208.1](https://doi.org/10.1890/14-2208.1)

- 605 52. M. Moora et al., Alien plants associate with widespread generalist arbuscular mycorrhizal  
606 fungal taxa: evidence from a continental-scale study using 454 sequencing. *Journal of*  
607 *Biogeography*, **38**, 1305-1317 (2011). doi:10.1111/j.1365-2699.2011.02478.x
- 608 53. Q. Zhang et al., Positive feedback between mycorrhizal fungi and plants influences plant  
609 invasion success and resistance to invasion. *PLoS ONE*, **5**(8), e12380 (2010). doi:  
610 10.1371/journal.pone.0012380
- 611 54. D. Mueller-Dombois, Forest dynamics in Hawaii. *Trends in Ecology & Evolution*, **2**(7), 216-  
612 220 (1987). doi: 10.1016/0169-5347(87)90024-3
- 613 55. P.R. Guimarães Jr, P. Jordano, J.N. Thompson, Evolution and coevolution in mutualistic  
614 networks. *Ecology Letters*, **14**(9), 877-885 (2011). doi: 10.1111/j.1461-0248.2011.01649.x
- 615 56. M. Duhamel et al., Plant selection initiates alternative successional trajectories in the soil  
616 microbial community after disturbance. *Ecological Monographs*, In Press, e01367 (2019).  
617 doi: 10.1002/ecm.1367
- 618 57. S. Banerjee, K. Schlaeppi, M.G.A. van der Heijden, Keystone taxa as drivers of microbiome  
619 structure and functioning. *Nature Reviews Microbiology*, **16**, 567-576 (2018). doi:  
620 10.1038/s41579-018-0024-1
- 621 58. M.M. Hart, R.J. Reader, Host plant benefit from association with arbuscular mycorrhizal  
622 fungi: variation due to size of mycelium. *Biology and Fertility of Soils*, **36**(5), 357-366  
623 (2002). doi: 10.1007/s00374-002-0539-4
- 624 59. H. Maherli, J.N. Klironomos, Phylogenetic trait-based assembly of arbuscular mycorrhizal  
625 fungal communities. *PLoS ONE*, **7**(5), e36695 (2012). doi: 10.1371/journal.pone.0036695
- 626 60. P.L. Chagnon, R.L. Bradley, J.N. Klironomos, Trait-based partner selection drives  
627 mycorrhizal network assembly. *Oikos*, **124**, 1609-1616 (2015). doi: 10.1111/oik.01987
- 628 61. M. van der Heyde, B. Ohsowski, L.K. Abbot, M.M. Hart, Arbuscular mycorrhizal fungus  
629 responses to disturbance are context-dependent. *Mycorrhiza*, **27**(5), 431-440 (2017). doi:  
630 10.1007/s00572-016-0759-3
- 631
- 632
- 633
- 634
- 635
- 636
- 637
- 638
- 639

640 **Figure Legends**

641

642 **Fig. 1. Dot plots displaying bipartite arbuscular mycorrhizal fungal and plant host**  
643 **network metrics between remnant (green) and restored (blue) habitats within the Hakalau**  
644 **Forest National Wildlife Refuge.** Only bipartite metrics that differed significantly between  
645 habitat types are shown, and include: overall network specialization ( $H^2$ ; A), linkage density  
646 (B), and mean number of links per species in the network (C). Dots represent group means and  
647 whiskers represent group standard deviation around the mean. Differences in bipartite metric  
648 means between habitat types were determined using a Welch unequal variance t-test. Asterisks  
649 indicate statistical differences between habitat types where; ns =  $p > 0.05$ , \* =  $p \leq 0.05$ , \*\* =  $p \leq$   
650 0.01, \*\*\* =  $p \leq 0.001$ , \*\*\*\* =  $p \leq 0.0001$ .

651

652 **Fig. 2. Boxplots showing host specialization (A) and hosts' symbiont range (B) on**  
653 **arbuscular mycorrhizal (AM) fungi in restored forest (left-side) and remnant forest (right-**  
654 **side) habitat types within the Hakalau Forest National Wildlife Refuge.** The bottom and the  
655 top of the boxes represent the first and third quartiles, the dark band inside boxes represents the  
656 median, the whiskers contain the upper and lower 1.5 interquartile range, and the dots represent  
657 outliers. Boxplots are coloured by host species. Symbiont range was significantly influenced by  
658 host species ( $p = 1.67 \times 10^{-10}$ , with a marginal effect of habitat type ( $p = 0.099$ ). Host  
659 specialization was significantly influenced by both habitat type ( $p = 7.67 \times 10^{-10}$ ) and host species  
660 ( $p = 0.016$ ). For both metrics, pairwise comparisons were determined by estimated marginal  
661 means, where statistical differences are signified by boxes without shared letters.

662

663 **Fig. 3. Scatterplots to identify potential keystone species. Betweenness centrality plotted**  
664 **against node degree for all arbuscular mycorrhizal (AM) fungal species within the**  
665 **networks from remnant (A) and restored forests (B) habitat types within the Hakalau**  
666 **Forest National Wildlife Refuge.** AM fungal species with high betweenness centrality represent  
667 potentially key connector species within networks, while AM fungal species with high degree  
668 represent hubs in the network. Points representing AM fungal species are sized by prevalence.  
669 Candidate keystone AM fungal species that had high betweenness centrality and degree values,  
670 while also exhibiting a prevalence score  $\geq 0.005$ , are highlighted in both plots.

671

672 **Fig. 4. Matrices depicting the nested temperature of arbuscular mycorrhizal (AM) fungal**  
673 **communities in remnant (A) and restored (B) forest habitat types within Hakalau.** Host  
674 plant species are rows and AM fungal taxa are columns. Presences of an AM fungal taxon within  
675 hosts are indicated by colored boxes.

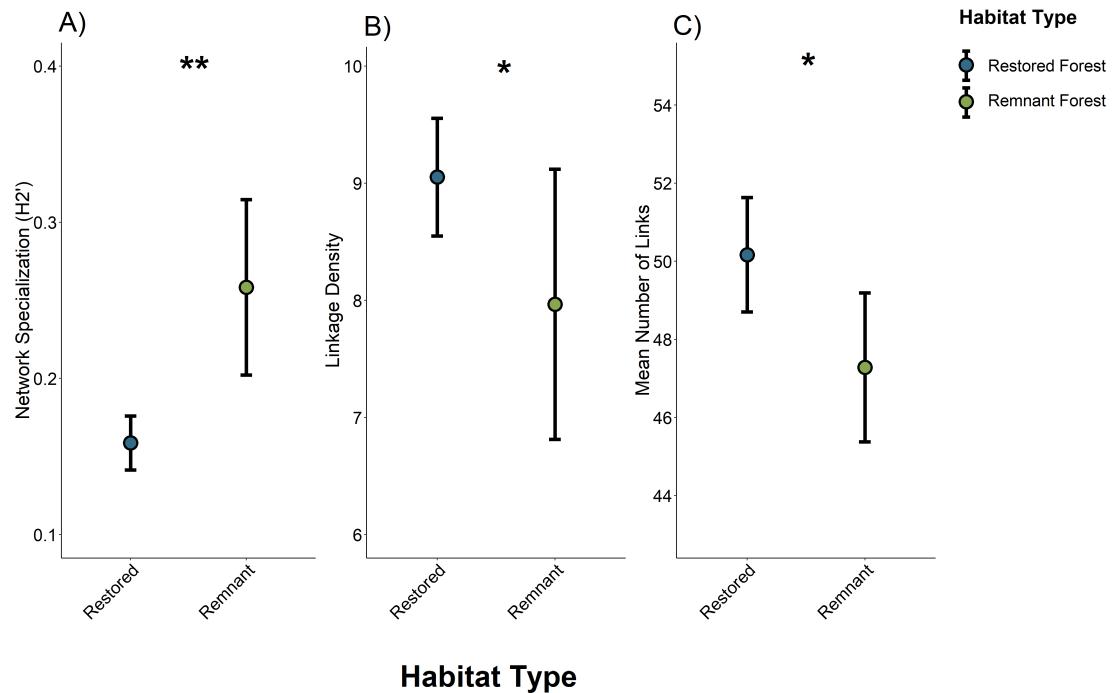
676

677 **Fig. 5. Non-metric dimensional scaling (NMDS) ordination plot of the community**  
 678 **composition of arbuscular mycorrhizal (AM) fungal communities collected from Remnant**  
 679 **(green circles) and Restored forest (blue diamonds) habitat types within Hakalau.**  
 680 Compositional differences are based on Bray-Curtis dissimilarity among samples. Communities  
 681 are colored by habitat type. Ellipses represent the 95% confidence region based on the centroid  
 682 for each host. Environmental variables with significant correlation with the ordination are shown.  
 683 Arrowhead size is proportional to the strength of correlation (SI Appendix, Table S10).

684

## 685 Figures

### 686 Fig. 1.



687

688

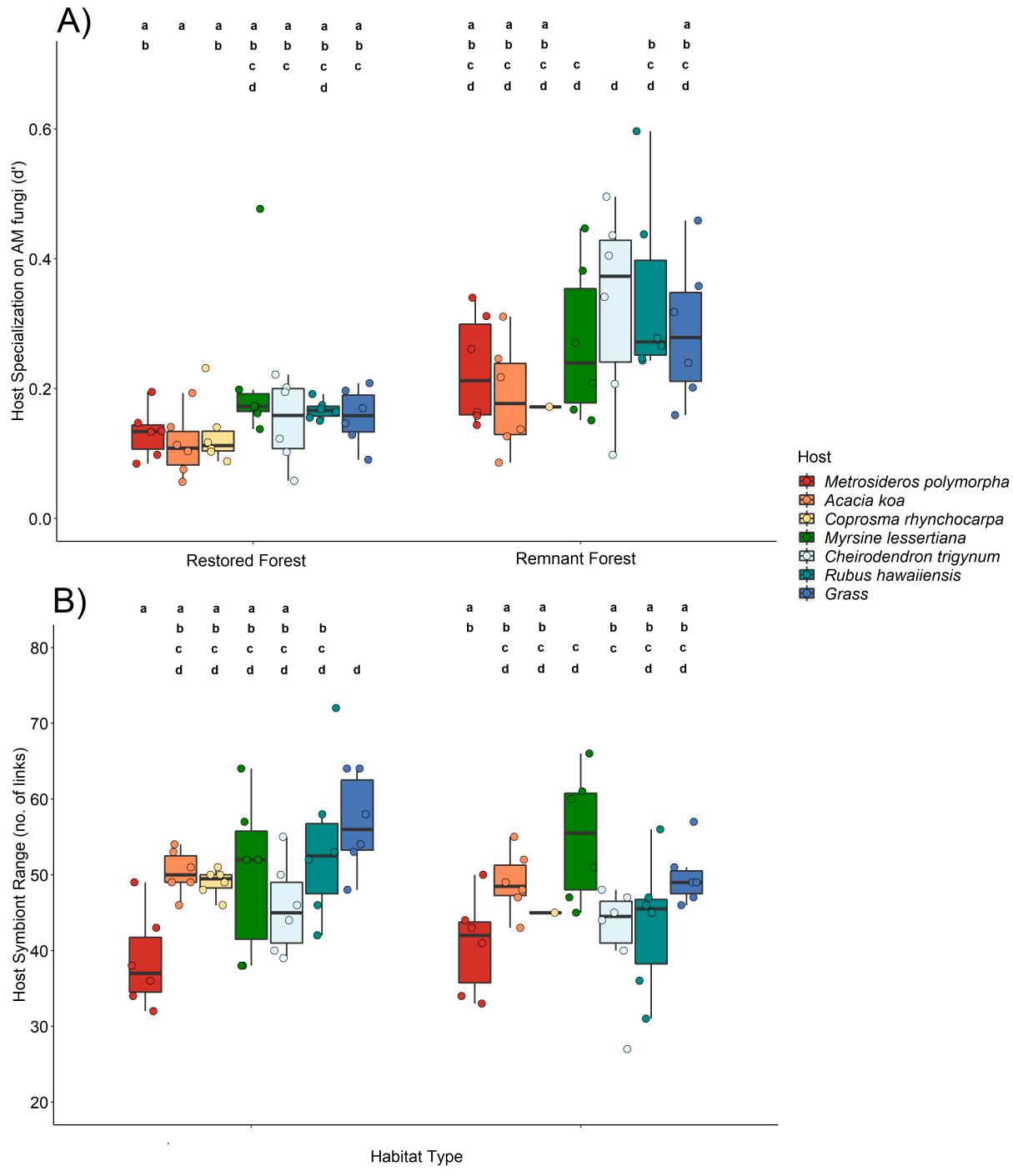
689

690

691

692

693

694 **Fig. 2.**

695

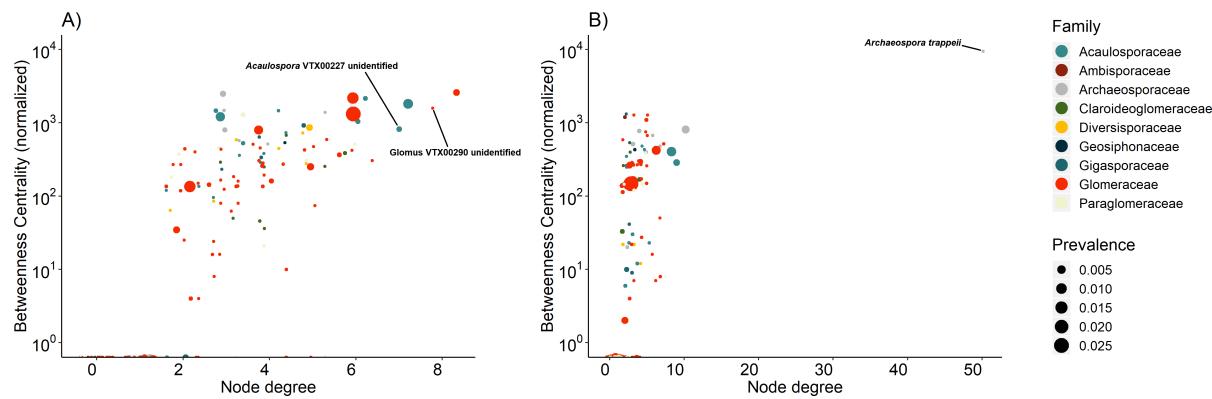
696

697

698

699

700 **Fig. 3.**

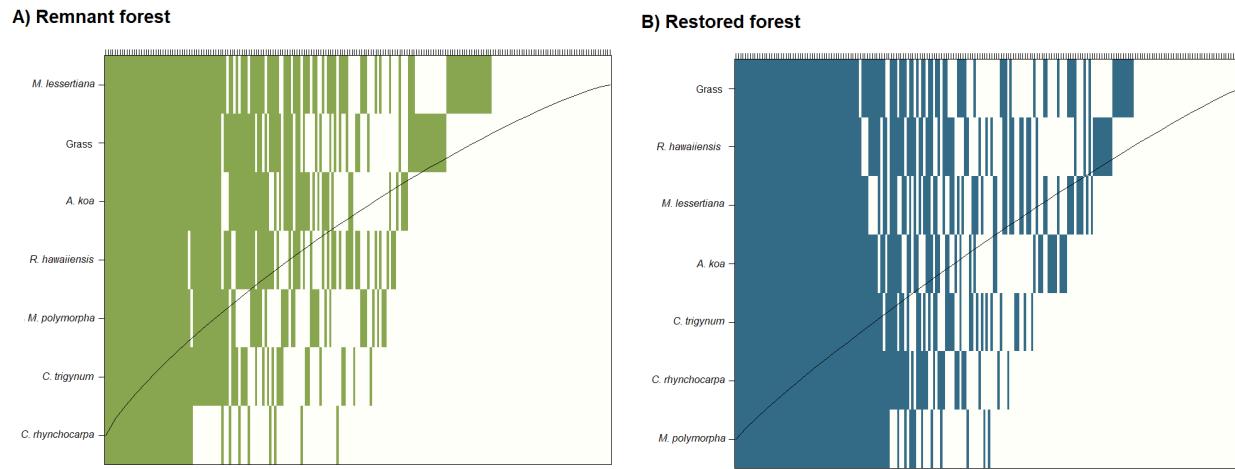


701

702

703

704 **Fig. 4.**



705

706

707

708

709

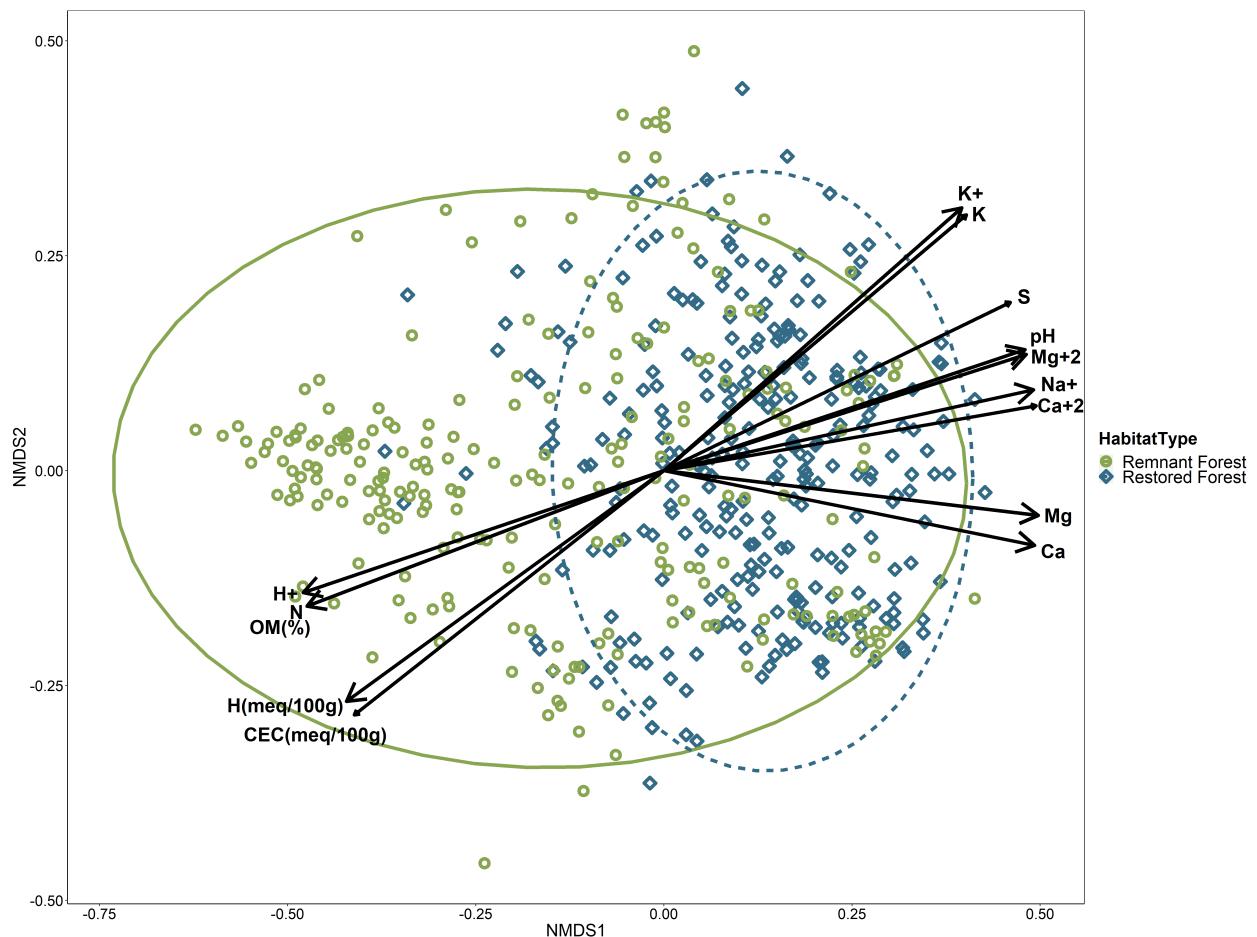
710

711

712

713

714 Fig. 5.



715

## **Supplementary Information**

### **The interaction networks of plants and symbiotic fungi are not restored following reforestation of a degraded subtropical ecosystem**

Cameron P. Egan, Sean O.I. Swift, Nicole A. Hynson

## **Contents**

- I. **Appendix I: Sampling - description of spatial sampling within plots**
- II. **Appendix II: Molecular Methods**
- III. **Appendix III: Bioinformatics**
- IV. **Appendix IV: Data Analyses**
- V. **Supplementary Tables**
- VI. **Supplementary Figures**
- VII. **Additional References**

### **I. Appendix I: Sampling - description of spatial sampling within plots**

Because plant communities within our study system are spatially heterogeneous, in some instances, we were not able to detect eight individuals for each species within the 12m radius, in which case we sampled as many individuals as were detected, increasing our sampling intensity to have the same number of replicate root samples per host and per plot (SI Appendix, Table S1). We ensured that samples were taken greater than 2 m apart, as it has been shown that AM fungal communities are spatially structured at scales of <1 m, and spatial autocorrelation ceases at distances >1m (1, 2).

### **II. Appendix II: Molecular Methods**

For our first PCR reactions, a 550 bp of AM fungal SSU rRNA was targeted using the universal eukaryotic primer WANDA (SI from 3) with the Fluidigm tag CS1 attached to the 5' end (Fluidigm Inc., South San Francisco, CA, USA), and the Glomeromycotina specific primer AML2 (4) with the Fluidigm tag CS2 attached to the 5' end. Reactions were carried out in 25 µL volumes using 2.5 µL of DNA template, 7.8 µL of nuclease-free H<sub>2</sub>O, 12.5 µL of 2x Kapa plant

PCR buffer (containing 1.5 mM MgCl<sub>2</sub> (1X) and 0.2 mM of each dNTP), 1.5 µL of 25 mM MgCl<sub>2</sub>, 0.5 µL of 50 µM AML2/CS2, 0.5 µL of 50 µM WANDA/CS1, and 0.2 µL 2.5 U/ µL Kapa3G DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA). Thermocycler conditions were: 95°C for 2 min; 30 cycles of 95°C for 30 seconds, 54°C for 40 seconds, 72°C for 1 min; and then 72°C for 10 min, on an Eppendorf Mastercycler Nexus Thermal Cycler (Eppendorf North America, Hauppauge, NY, USA). Preliminary test PCR reactions indicated that adherence of primers to AM fungal DNA in soil samples was being inhibited, therefore we diluted extracted soil DNA to 1:10 concentrations to be used as template for our first PCR reactions.

Second PCR primer complexes consisted of the same Fluidigm tags (CS1 or CS2) as our PCR1 primers, an 8 bp Illumina Nextera barcode, and an Illumina adapter (5). For PCR2 reactions, products from the PCR1 reactions were first diluted 15-fold. Reactions were then carried out in 20 µL volume using 1 µL of diluted PCR1 product, 6.89 µL of nuclease-free H<sub>2</sub>O, 10 µL of 2x Kapa plant PCR buffer (containing 1.5 mM MgCl<sub>2</sub> (1X) and 0.2 mM of each dNTP), 1.2 µL of 25 mM MgCl<sub>2</sub>, 0.75 µL of 2 µM Illumina adaptor/barcode primer complex, and 0.16 µL 2.5 U/ µL Kapa3G DNA Polymerase. Thermocycler conditions were: 95°C for 1 min, followed by 12 cycles of 95°C for 30 sec, 60°C for 30 sec, 68°C for 1 min; and then 68°C for 5 min, on an Eppendorf Mastercycler Nexus Thermal Cycler.

Samples were then purified and normalized to 25 ng using the Just-a-Plate™ PCR purification and normalization plate (Charm Biotech, San Diego, CA, USA) following the manufacturers protocol. Separate amplicon pools for roots and soil samples were then created by pooling 10 µL of eluted 25 ng DNA from each sample. Pooled amplicon libraries were purified again, and size selected for amplicons ~650 bp in length using AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA). Root libraries were sequenced on a single Illumina MiSeq run, using 2 x 300 paired-end (PE) sequencing and 600 cycles (Illumina Inc., San Diego, CA, USA) by the Genomics High Throughput Facility at the University of California, Irvine (<https://ghtf.biochem.uci.edu/>).

### III. Appendix III: Bioinformatics

The sequencing run generated four *fastq* files: forward barcode reads, forward target reads, reverse barcode reads, and reverse target reads, along with a total of 9,464,322 target reads passed the filter.

Forward and reverse barcode reads were first extracted for each sample, and then re-assembled into a new *fastq* file representing each sample's unique barcode combination using the *combine\_fastq\_barcodes.py* command (<https://gist.github.com/walterst/6284164>). Next, forward and reverse reads were demultiplexed using the *qiime demux emp-paired* command. Sequences were quality checked, chimeras were removed, and sequences were dereplicated and assembled into exact sequence variants (ESV) using DADA2 (6) and the *dada2 denoise-paired* command

using the default parameters in QIIME. Of those initial reads, 2,884,613 sequences were kept for analysis after quality checking, chimera detection.

Quality filtered sequences were identified to virtual taxa (VT) in MaarjAM (7) using the BLAST+ algorithm (8), a sequence similarity of  $\geq 97\%$ , sequence alignment  $\geq 97\%$ , and a BLAST e-value  $<1e-50$ . Using these parameters, 991 ESVs were identified to VT (20% of ESVs, containing 2,540,905 sequences,  $\sim 49\%$  of sequences that passed initial quality filters). Reads that did not get a hit with these criteria were subjected to a further BLAST+ search against the NCBI nucleotide database (accessed February 2019) using a sequence similarity of  $\geq 90\%$ , and a BLAST e-value of  $<1e-50$ . A further 281 ESVs (5% of total ESVs, containing 24,705 sequences,  $\sim 0.5\%$  of sequences that passed initial quality filters) were identified during this second round of taxonomic assignment. Because our study system is relatively understudied in AM fungal research (9), we expected to detect a high incidence of Glomeromycotinan sequences not previously identified and deposited into a curated database (10). Therefore, we identified remaining ESVs using the MaarjAM database, and a relaxed sequence similarity of  $\geq 80\%$  (default in QIIME2) while maintaining a BLAST e-value of  $<1e-50$ . We used this method as previously unidentified Glomeromycotinan species may be similar to taxa already present in databases but will not be identified using strict search criteria. Using our relaxed criteria, remaining sequences were identified to the closest VT in MaarjAM. Newly identified ESVs were denoted by appending ‘\_unidentified’ after the VT name (e.g. ESVs identified to VTX00242 were identified as VTX00242\_unidentified). Using this method, an additional 2,548 SVs ( $\sim 50\%$  of total ESVs, containing 1,872,756 sequences,  $\sim 36\%$  of sequences that passed initial quality filters) were taxonomically identified during this final round of taxonomic assignment. Remaining unidentified sequences or sequences identified as non-Glomeromycotinan were then filtered from down stream analyses using the *qiime taxa filter-table* command. To control for differences in read abundance generated for each sample, we rarefied to 1,500 sequences (SI Appendix, Fig. S2). ESVs were then binned by species using the *qiime taxa collapse* command. We exported our rarefied species table, taxonomy identifier table, and associated sample metadata from QIIME for downstream analyses. in R (version 3.5.2; R Core Team, 2018).

## IV. Appendix IV: Data Analyses

### a) Plant-AM fungal Networks

Comparisons of network metrics between habitat types and among hosts were conducted using a generalized linear model (GLM) with a quasi-Poisson distribution. Quasi-Poisson distributions were chosen as this best fit the data when we examined model assumptions (normally distributed residuals, homogeneity of variance, and the linear relation between the response and the linear predictor). GLMs were built considering host species, habitat type, and their interaction, and model significance was assessed using an analysis of deviance.

### **b) AM fungal co-occurrence networks**

To infer network structure, we used the neighborhood selection approach in SPIEC-EASI, which applies the Meinshausen and Bühlmann (MB) algorithm (11) that reconstructs the network on a node-by-node basis, where, for each node (in our case an AM fungal species) co-occurrence correlations with other nodes in the network are normalized using a penalized regression problem (12). SPIEC-EASI estimates this penalty parameter using a form of bootstrapping called Stability Approach to Regularization Selection (StARS; 13). StARS consists in repeating the inference on random sub-sets of the data and selecting the penalty parameter such that the resulting network is least variable across repetitions.

Using a method from Tipton et al. (14), we first assembled composite networks for root and soil fungal communities within restored and remnant forest plots (Fig. S3). We then evaluated important nodes (AM fungal species) within networks by calculating centrality and degree for each node. High centrality scores are indicative of key connector species within networks, as these nodes act as bridges between other nodes within a network (15). High degree scores are indicative of nodes that are highly connected within a network. To estimate abundance within the community, we calculated the prevalence of each node. Prevalence was calculated by multiplying the percent of samples that a node was detected in and multiplying that value by that node's relative read abundance. Higher prevalence scores were indicative of higher abundances within a community.

### **c) Indicator species analysis**

To determine whether any identified keystone AM fungal taxa could also be potential indicators of host species or habitat types within Hakalau we determined their indicator value (IndVal; 16), which measures the strength of the association between an AM fungal taxon and a defined group (i.e. different host species and either remnant or restored forest). We identified AM fungal taxa significantly associated with a particular habitat type or host using the *multipatt* function from the indispecies package (17), using 9999 permutations.

## **V. Supplementary Tables**

**Table S1 Metadata associated with samples taken from the Hakalau Forest National Wildlife Refuge.**

**Table S2 Definitions of the network metrics quantified for plant-AM fungal bipartite and AM fungal co-occurrence networks.**

**Table S3 Results of Welch unequal variance t-tests on observed versus null network-level bipartite metrics within Remnant forest and Restored forest.**

Metric	Habitat Type	Observed Mean	Null Mean	t	df	P
Connectance	Restored forest	0.4899245	0.8782528	-34.082	5.0038	4.05E-07
	Remnant forest	0.5004124	0.8999253	-20.834	5.002	4.70E-06
Nestedness	Restored forest	62.07922	52.86312	10.821	5.022	1.14E-04
	Remnant forest	60.64218	52.4509	6.2659	5.0188	1.50E-03
Modularity	Restored forest	0.4542609	0.2338583	20.963	5.183	3.28E-06
	Remnant forest	0.4393174	0.2451128	13.589	5.0885	3.41E-05
Linkage density	Restored forest	8.550043	8.664144	-0.87848	5.0414	0.4196
	Remnant forest	7.097882	8.353661	-2.6802	5.0032	0.04378
Links per plant species	Restored forest	50.14216	51.82874	-2.2527	5.0153	0.07387
	Remnant forest	46.99849	49.409	-3.3482	5.0215	2.02E-02
Host robustness	Restored forest	0.7332503	0.7908414	-18.1684	5.0057	5.45E-04
	Remnant forest	0.7233384	0.774355	-6.1177	5.0092	1.68E-03
Network specialization ( $H_2'$ )	Restored forest	0.1635254	0.1441939	1.9398	5.0126	0.11
	Remnant forest	0.2794375	0.1503079	5.3246	5.0041	3.12E-03

**Table S4 Results of Generalized Linear Model (GLM) analysis on observed host interaction specialization ( $d'$ ) on AM fungi between remnant and restored forests (Habitat Type) and among plant host species.**

	Df	Deviance	Residual	Residual	Pr(>Chi)
			Df	Deviance	
Null			78	4.3315	
Habitat Type	1	1.34702	77	2.9845	7.67E-10
Host	6	0.55778	71	2.4267	0.016
Habitat Type x Host	6	0.15572	65	2.271	0.626

**Table S5 Results of Generalized Linear Model (GLM) analysis on observed host symbiont range between habitat types (remnant vs restored forest) and among plant host species.**

	Df	Deviance	Residual	Residual	Pr(>Chi)
			Df	Deviance	
Null			78	114.741	
Host	6	35.517	72	79.224	4.5530E-06
Habitat Type	1	2.757	71	76.467	0.09987
Habitat Type x Host	6	9.687	65	66.779	0.147

**Table S6 Comparisons of observed plant species-level symbiont range and specialization ( $d'$ ) on arbuscular mycorrhizal (AM) fungi to random chance for bipartite networks between plant and AM fungi within the Hakalau Forest National Wildlife Refuge.** Null networks were assembled using the *swap.web* null model available in the bipartite package in R (18) using 1000 permutations. Comparisons between observed and null network metrics for each host were done using Welch unequal variance t-tests.

Network Metric	Habitat Type	Host	Observed Mean	Null mean	t	df	P
<b>Host Symbiont Range</b>	Restored Forest	<i>Metrosideros polymorpha</i>	39	49	-4.084	5.0244	0.009
		<i>Acacia koa</i>	50	49	0.929	5.1155	0.395
		<i>Coprosma rhynchocarpa</i>	49	49	-0.292	5.3231	0.782
		<i>Myrsine lessertiana</i>	50	49	0.220	5.0092	0.835
		<i>Cheirodendron trigynum</i>	46	49	-1.421	5.0274	0.214
		<i>Rubus hawaiiensis</i>	54	49	1.076	5.0092	0.331
		Grass	57	49	2.912	5.0238	0.033
	Remnant Forest	<i>Metrosideros polymorpha</i>	41	47	-2.208	5.0227	0.0781
		<i>Acacia koa</i>	49	47	1.386	5.0535	0.224
		<i>Coprosma rhynchocarpa</i> <sup>†</sup>	46	45	1.911	999	5.63E-02
		<i>Myrsine lessertiana</i>	55	47	2.403	5.0134	0.061
		<i>Cheirodendron trigynum</i>	42	47	-1.508	5.0157	0.192
		<i>Rubus hawaiiensis</i>	44	47	-0.875	5.0123	0.422
		Grass	50	47	1.988	5.0608	0.103
<b>Host specialization (<math>d'</math>)</b>	Restored Forest	<i>Metrosideros polymorpha</i>	0.1321	0.1402	-0.512	5.0337	0.630
		<i>Acacia koa</i>	0.1138	0.1395	-1.287	5.0212	0.254
		<i>Coprosma rhynchocarpa</i>	0.1312	0.1398	-0.402	5.0183	0.704
		<i>Myrsine lessertiana</i>	0.2201	0.1395	1.551	5.003	0.182
		<i>Cheirodendron trigynum</i>	0.1502	0.1415	0.329	5.0124	0.756
		<i>Rubus hawaiiensis</i>	0.1673	0.1410	4.346	5.2425	0.007
		Grass	0.1568	0.1406	0.898	5.0258	0.410
	Remnant Forest	<i>Metrosideros polymorpha</i>	0.2298	0.1569	2.087	5.0098	0.091
		<i>Acacia koa</i>	0.1873	0.1586	0.828	5.01	0.445
		<i>Coprosma rhynchocarpa</i> <sup>†</sup>	0.1789	0.1719	1.789	999	0.074
		<i>Myrsine lessertiana</i>	0.2711	0.1582	2.304	5.005	0.069
		<i>Cheirodendron trigynum</i>	0.3305	0.1572	2.820	5.0031	3.71E-02
		<i>Rubus hawaiiensis</i>	0.3447	0.1584	3.186	5.0035	2.44E-02
		Grass	0.2892	0.1554	2.956	5.0054	0.032

**Table S7 Indicator arbuscular mycorrhizal (AM) fungal taxa for the two different habitat types sampled within Hakalau.** Values are shown for: the positive predictive power of the taxon being an indicator of that habitat type (A), the probability of finding the taxon within that habitat type (B), the strength of the association (IndVal), and the statistical significance of this relationship based on 9999 permutations (P-value). Values of A indicate the degree to which the taxon is found only within that habitat type (i.e., A=1.000 means the taxon is only found within that habitat type). Values of B indicate how often that taxon was detected within plots from that habitat type (i.e., B=1.000 indicates it is found in all plots sampled from that habitat type). IndVal values indicate the strength of the association between that taxon and the habitat type, where higher values indicate a stronger indicator of that host).

Habitat Type	Taxa	Family	A	B	IndVal	P
Remnant Forest	<i>Glomus_VTX00345</i>	Glomeraceae	0.951	0.734	0.835	0.0001
	<i>Diversispora_VTX00347_unidentified</i>	Diversisporaceae	0.863	0.667	0.759	0.0001
	<i>Glomus_VTX00099</i>	Glomeraceae	0.795	0.633	0.709	0.0001
	<i>Scutellospora_VTX00261</i>	Gigasporaceae	0.812	0.544	0.665	0.0001
	<i>Acaulospora_VTX00272</i>	Acaulopsporaceae	0.916	0.300	0.524	0.0001
	<i>Acaulospora_laevis</i>	Acaulopsporaceae	0.970	0.253	0.496	0.0001
	<i>Acaulospora_VTX00014</i>	Acaulopsporaceae	0.759	0.245	0.431	0.0001
	<i>Acaulospora_VTX00024</i>	Acaulopsporaceae	0.996	0.152	0.389	0.0001
	<i>Glomus_VTX00084</i>	Glomeraceae	0.777	0.152	0.343	0.0001
	<i>Scutellospora_VTX00260</i>	Gigasporaceae	0.941	0.114	0.327	0.0001
	<i>Acaulospora_VTX00017_unidentified</i>	Acaulopsporaceae	0.969	0.076	0.271	0.0003
	<i>Scutellospora_VTX00049</i>	Gigasporaceae	0.927	0.059	0.234	0.0041
	<i>Acaulospora_VTX00012</i>	Acaulopsporaceae	1.000	0.046	0.215	0.0003
	<i>Glomus_VTX00292</i>	Glomeraceae	0.986	0.042	0.204	0.0013
	<i>Paraglomus_occultum</i>	Paraglomeraceae	0.782	0.051	0.199	0.0412
Restored Forest	<i>Acaulospora_VTX00272_unidentified</i>	Acaulopsporaceae	1.000	0.021	0.145	0.0194
	<i>Glomus_VTX00206_unidentified</i>	Glomeraceae	1.000	0.017	0.130	0.0426
	<i>Glomus_VTX00126</i>	Glomeraceae	0.886	0.139	0.351	0.0002
	<i>Glomus_VTX00223</i>	Glomeraceae	0.822	0.125	0.320	0.0064
	<i>Glomus_VTX00327</i>	Glomeraceae	0.814	0.089	0.269	0.0064
	<i>Glomeraceae uncultured</i>	Glomeraceae	0.926	0.075	0.263	0.0002
	<i>Glomus_VTX00253</i>	Glomeraceae	0.918	0.075	0.262	0.0015
	<i>Claroideoglomus_VTX00402</i>	Claroideoglomeraceae	0.806	0.082	0.257	0.0048
	<i>Glomus_VTX00200_unidentified</i>	Glomeraceae	0.830	0.061	0.224	0.0332
	<i>Diversispora_VTX00263</i>	Gigasporaceae	0.882	0.053	0.217	0.0214

**Table S8 Indicator arbuscular mycorrhizal (AM) fungal taxa for hosts within Hakalau.**

Values are shown for: the positive predictive power of the taxon being an indicator of the host (A), the probability of finding the taxon associating with the host (B), the strength of the association (IndVal), and the statistical significance of this relationship based on 9999 permutations (P-value). Values of A indicate the degree to which the taxon is found only associating with that host (i.e., A=1.000 means the taxon only associates with that host). Values of B indicate how often that taxon is detected associating with that host (i.e., B=1.000 indicates it is found in all host individuals sampled). IndVal values indicate the strength of the association between that taxon and the host, where higher values indicate a stronger indicator of that host).

Host	Taxa	Family	A	B	IndVal	P
<i>M. lessertiana</i>	<i>Glomus_VTX00111_unidentified</i>	Glomeraceae	0.809	0.219	0.421	0.0001
	<i>Paraglomus_occultum</i>	Paraglomeraceae	0.596	0.172	0.320	0.0001
	<i>Acaulospora_VTX00231_unidentified</i>	Acaulosporacaea	0.830	0.078	0.255	0.0019
	<i>Glomus_VTX00191</i>	Glomeraceae	0.779	0.078	0.247	0.0042
	<i>Glomus_VTX00182_unidentified</i>	Glomeraceae	0.722	0.063	0.212	0.0052
	<i>Glomus_VTX00407_unidentified</i>	Glomeraceae	0.933	0.047	0.209	0.0173
	<i>Glomus_VTX00206_unidentified</i>	Glomeraceae	0.932	0.047	0.209	0.0078
	<i>Glomus_VTX00371</i>	Glomeraceae	0.923	0.047	0.208	0.0096
	<i>Glomus_VTX00080</i>	Glomeraceae	1.000	0.031	0.177	0.0259
	<i>Glomus_VTX00178_unidentified</i>	Glomeraceae	1.000	0.031	0.177	0.0257
	<i>Glomus_VTX00234_unidentified</i>	Glomeraceae	1.000	0.031	0.177	0.0267
	<i>Glomus_VTX00053</i>	Glomeraceae	0.925	0.031	0.170	0.0391
<i>R. hawaiiensis</i>	<i>Glomus_VTX00412_unidentified</i>	Glomeraceae	0.891	0.500	0.667	0.0001
	<i>Claroideoglomus_VTX00225</i>	Claroideoglomeraceae	0.725	0.057	0.203	0.0153
	<i>Glomus_VTX00180_unidentified</i>	Glomeraceae	0.889	0.034	0.174	0.0377
<i>C. trigynum</i>	<i>Glomus_VTX00322_unidentified</i>	Glomeraceae	0.842	0.514	0.658	1.00E-04
All native hosts	<i>Scutellospora_VTX00261</i>	Gigasporaceae	0.950	0.428	0.638	0.0004
	<i>Paraglomus_VTX00281</i>	Paraglomeraceae	0.940	0.340	0.565	0.0182
All natives except <i>A. koa</i>	<i>Glomus_VTX00068_unidentified</i>	Glomeraceae	0.937	0.369	0.588	0.0208
Grass	<i>Paraglomus_VTX00238_unidentified</i>	Paraglomeraceae	0.663	0.148	0.313	0.004
	<i>Acaulospora_VTX00328</i>	Acaulosporacaea	0.625	0.125	0.280	0.0065
	<i>Acaulospora_VTX00030</i>	Acaulosporacaea	0.635	0.102	0.255	0.0118

**Table S9 Differences in the relative abundance of AM fungal families detected in roots between Remnant and Restored forest habitat types within Hakalau.** Significant differences were determined using Welch unequal variance t-tests on relative read abundances.

Family	Mean Relative Abundance		t	df	P
	Restored Forest	Remnant Forest			
Acaulosporaceae	0.13681	0.16171	-1.802	450.49	7.22E-02
Ambisporaceae	0.00413	0.00230	2.310	493.52	2.13E-02
Archaeosporaceae	0.07720	0.05422	3.277	483.86	1.13E-03
Claroideoglomeraceae	0.01651	0.00751	3.361	482.66	8.37E-04
Diversisporaceae	0.00427	0.02015	-6.649	261.86	1.72E-10
Geosiphonaceae	0.00053	0.00044	0.437	470.55	6.62E-01
Gigasporaceae	0.01415	0.02521	-2.777	400.60	5.74E-03
Glomeraceae	0.70900	0.71199	-0.169	475.25	8.66E-01
Paraglomeraceae	0.03739	0.01754	5.159	496.22	3.75E-07

**Table S10 Generalized linear model (GLM) results of the effect of host, habitat type (remnant vs restored forest) and their interaction on AM fungal species richness.**

	Deviance					
	Df	(Chi-sq)	Resid.Df	Resid.Dev	Pr(>Chi)	Significance
NULL			993	1544.4		
Host	6	75.394	987	1469	3.19E-14	***
Habitat Type	1	2.854	986	1466.2	9.12E-02	.
Host x Habitat Type	6	30.393	980	1435.8	3.31E-05	***

**Table S11 Arbuscular mycorrhizal (AM) fungal taxa richness for hosts within restored and remnant forest patches within the Hakalau Forest National Wildlife Refuge.**

Habitat Type	Host	Mean Richness	sd
Restored forest	<i>Acacia koa</i>	18.4	4.8
	<i>Cheirodendron trigynum</i>	18.9	5.6
	<i>Coprosma rhynchocarpa</i>	19.2	4.4
	Grass	21.0	4.9
	<i>Metrosideros polymorpha</i>	17.0	4.7
	<i>Myrsine lessertiana</i>	20.5	5.6
	<i>Rubus hawaiensis</i>	18.4	6.0
Remnant forest	<i>Acacia koa</i>	19.2	5.7
	<i>Cheirodendron trigynum</i>	16.1	5.3
	<i>Coprosma rhynchocarpa</i>	19.3	4.7
	Grass	18.7	4.7
	<i>Metrosideros polymorpha</i>	18.1	4.3
	<i>Myrsine lessertiana</i>	21.6	5.9
	<i>Rubus hawaiensis</i>	17.5	5.8

**Table S12 Correlation of measured soil chemical factors with non-metric dimensional scaling (NMDS) plot (Figure 6 and S6).** Both the correlation ( $R^2$ ) and significant of that correlation (P) are shown for each chemical measured.

Soil Chemical	$R^2$	P
pH	0.422976	0.001
$H^+$	0.350518	0.001
Total nitrogen (lbs/A)	0.279557	0.001
Organic matter (%)	0.27942	0.001
$Ca^{+2}$	0.266718	0.001
$Mg^{+2}$	0.214542	0.001
Sulfate-S (ppm)	0.204703	0.001
Hydrogen (meq/100g)	0.136019	0.001
$K^+$	0.102082	0.001
Potassium (ppm)	0.078229	0.001
Cation Exchange Capacity (meq/100g)	0.045774	0.001
Magnesium (ppm)	0.034723	0.007
$Na^+$	0.032783	0.001
Calcium (ppm)	0.026779	0.012
Readily available phosphorus (ppm)	0.012433	0.135
Sodium (ppm)	0.002668	0.655

**Table S13 Differences in soil chemistry between Remnant and Restored habitat types within Hakalau.** Significant differences were determined using Welch unequal variance t-tests. Soil chemical analyses were performed by A & L Western Agricultural Laboratories INC (Modesto,CA).

Soil Chemical	Mean in Habitat Type				
	Remnant	Restored	t	df	P
pH	4.04	4.80	-15.408	210.33	2.20E-16
Organic matter (%)	71.52	58.45	9.378	248.22	2.20E-16
Total nitrogen (lbs/A)	1480.53	1209.95	9.378	248.05	2.20E-16
Readily available phosphorus (ppm)	11.51	11.99	-0.622	328.27	5.35E-01
Potassium (ppm)	58.15	70.25	-4.362	314.14	1.75E-05
Magnesium (ppm)	52.53	58.36	-1.346	294.98	1.80E-01
Calcium (ppm)	205.04	244.97	-1.914	327.44	5.65E-02
Sodium (ppm)	27.28	30.63	-2.359	324.05	1.89E-02
Sulfate-S (ppm)	32.77	54.62	-11.405	326.33	2.20E-16
Hydrogen (meq/100g)	3.95	2.00	7.338	213.12	4.49E-12
Cation Exchange Capacity (meq/100g)	5.67	3.84	5.034	238.22	9.47E-07
K <sup>+</sup>	3.92	6.24	-6.536	320.76	2.49E-10
Mg <sup>+2</sup>	8.73	13.05	-8.999	249.95	2.20E-16
Ca <sup>+2</sup>	17.90	28.58	-11.686	327.16	2.20E-16
H <sup>+</sup>	66.30	47.32	14.673	270.56	2.20E-16
Na <sup>+</sup>	2.80	4.56	-7.505	303.28	6.88E-13

**Table S14 Variation in arbuscular mycorrhizal (AM) fungal community composition as determined by permutational multivariate analysis of variance (PERMANOVA) on Bray-Curtis dissimilarity among samples within the Hakalau Forest National Wildlife Refuge.** Changes in community composition were examined between habitat types (remnant forest versus restored forest) and host plant species, and their interactions. Significant contributions of each factor and their combination were determined using 9999 permutations with terms added sequentially from first to last. All factors and their combination were shown to significantly contribute to differences in Bray-Curtis dissimilarity among AM fungal communities.

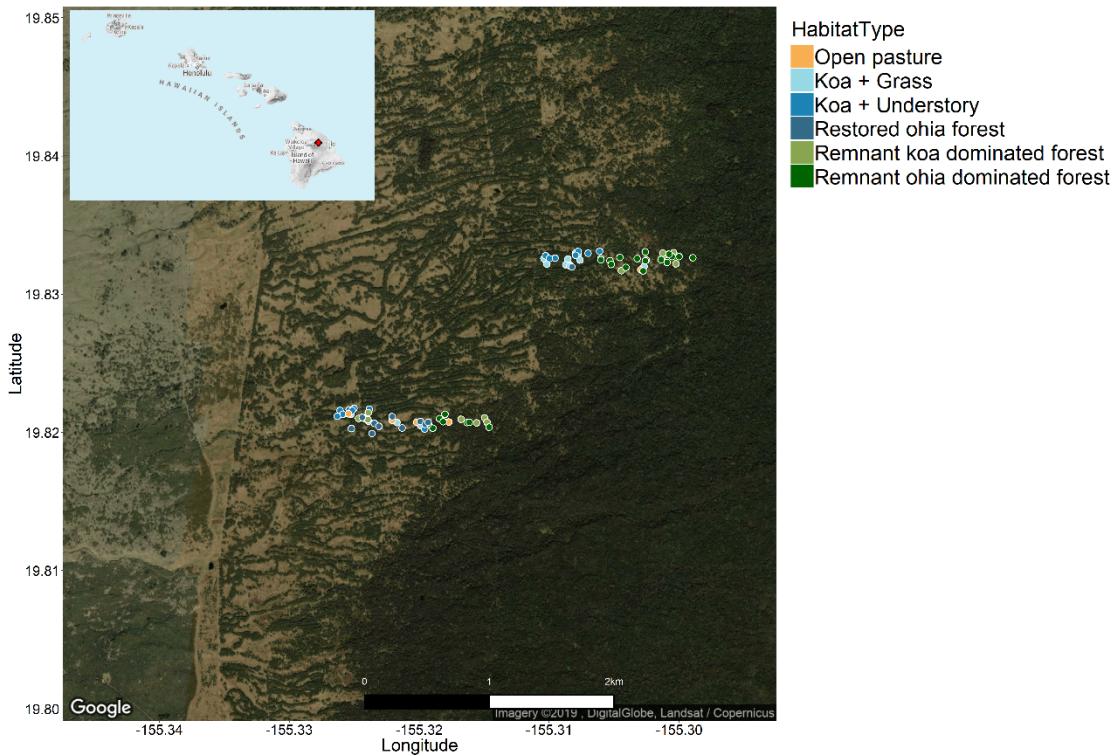
	Df	SS	MS	F Model	R2	Pr(>F)
Habitat Type	1	15.516	15.516	70.521	0.10465	1.00E-04
Host	6	14.514	2.419	10.994	0.09789	1.00E-04
Habitat Type • Host	6	7.351	1.2251	5.568	0.04958	1.00E-04
Residuals	504	110.89	0.22		0.74789	
Total	517	148.271			1	

**Table S15 Pairwise perMANOVA results showing differences in location of the AM fungal community group centroid for all host species in both habitat types.** Community dissimilarity was determined using Bray-Curtis dissimilarity. Pairwise comparisons were done using permutational MANOVAs on Bray-Curtis distances using 9999 permutations. P-values <0.05 indicate significant differences in the position of the group centroid.

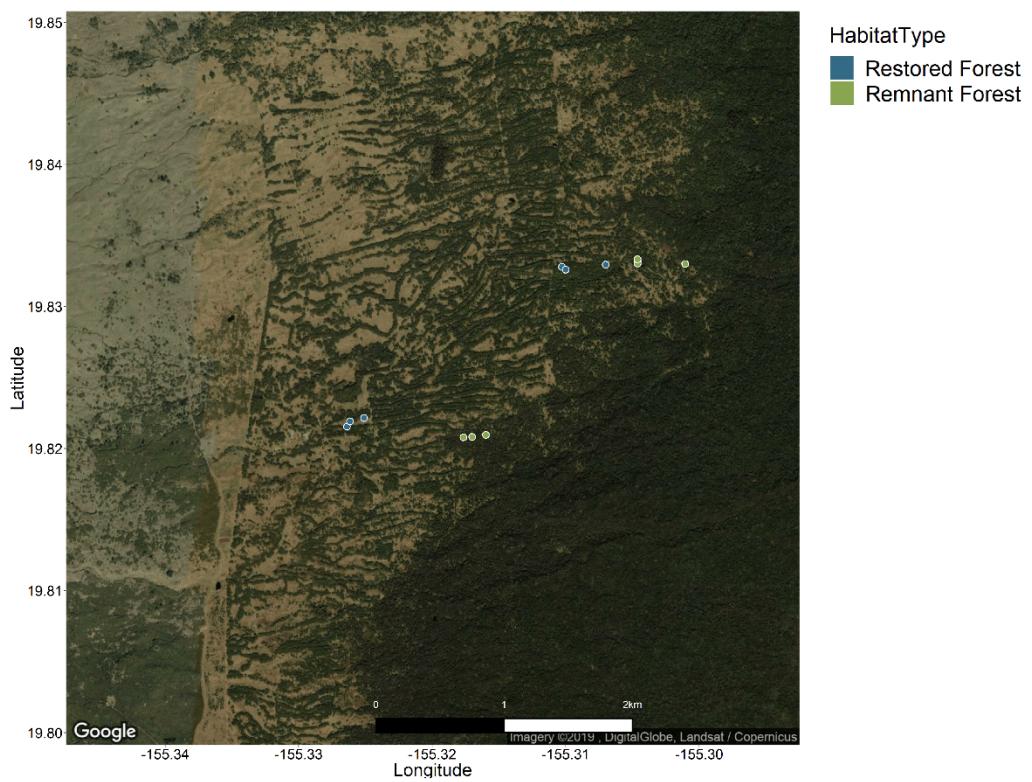
Host		<i>M. polymorpha</i>	<i>A. koa</i>		<i>C. rhyncocarpa</i>	<i>M. lessertiana</i>		<i>C. trigynum</i>		<i>R. hawaiiensis</i>	Grass
	Habitat Type	Remnant	Restored	Remnant	Restored	Remnant	Restored	Remnant	Restored	Remnant	Restored
<i>M. polymorpha</i>	Restored	0.00014	-	-	-	-	-	-	-	-	-
<i>A. koa</i>	Remnant	0.00799	0.00014	-	-	-	-	-	-	-	-
	Restored	0.00014	0.01592	0.00014	-	-	-	-	-	-	-
<i>C. rhyncocarpa</i>	Remnant	0.00014	0.00041	0.02437	0.00284	-	-	-	-	-	-
	Restored	0.00014	0.00014	0.00014	0.00014	0.00855	-	-	-	-	-
<i>M. lessertiana</i>	Remnant	0.00014	0.00014	0.00014	0.00014	0.00741	0.00014	-	-	-	-
	Restored	0.00014	0.00197	0.00014	0.00014	0.00303	0.00014	0.02434	-	-	-
<i>C. trigynum</i>	Remnant	0.1875	0.00014	0.0015	0.00014	0.00065	0.00014	0.00014	0.00014	-	-
	Restored	0.00014	0.12224	0.00014	0.05102	0.00065	0.00014	0.00014	0.0066	0.00014	-
<i>R. hawaiiensis</i>	Remnant	0.00014	0.00014	0.00014	0.00014	0.01809	0.00014	0.00014	0.00014	-	-
	Restored	0.00014	0.00014	0.00014	0.00028	0.00065	0.00014	0.00014	0.00014	0.00014	0.00576
Grass	Remnant	0.00014	0.00028	0.00014	0.00014	0.00239	0.00014	0.00014	0.00014	0.00014	0.00014
	Restored	0.00014	0.00206	0.00014	0.0303	0.00103	0.00014	0.00014	0.00014	0.00126	0.00028

## VI. Supplementary Figures

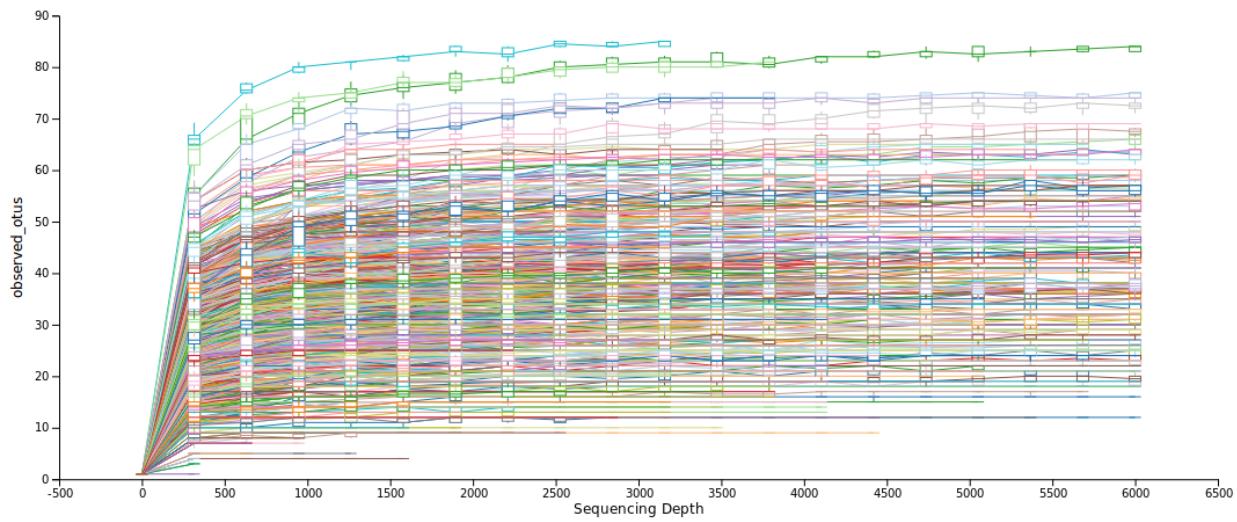
A.)



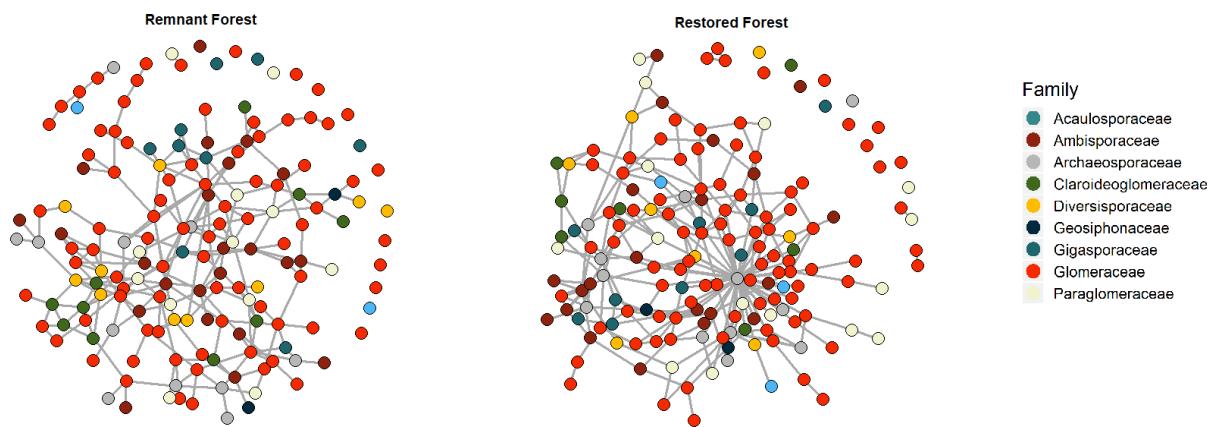
B.)



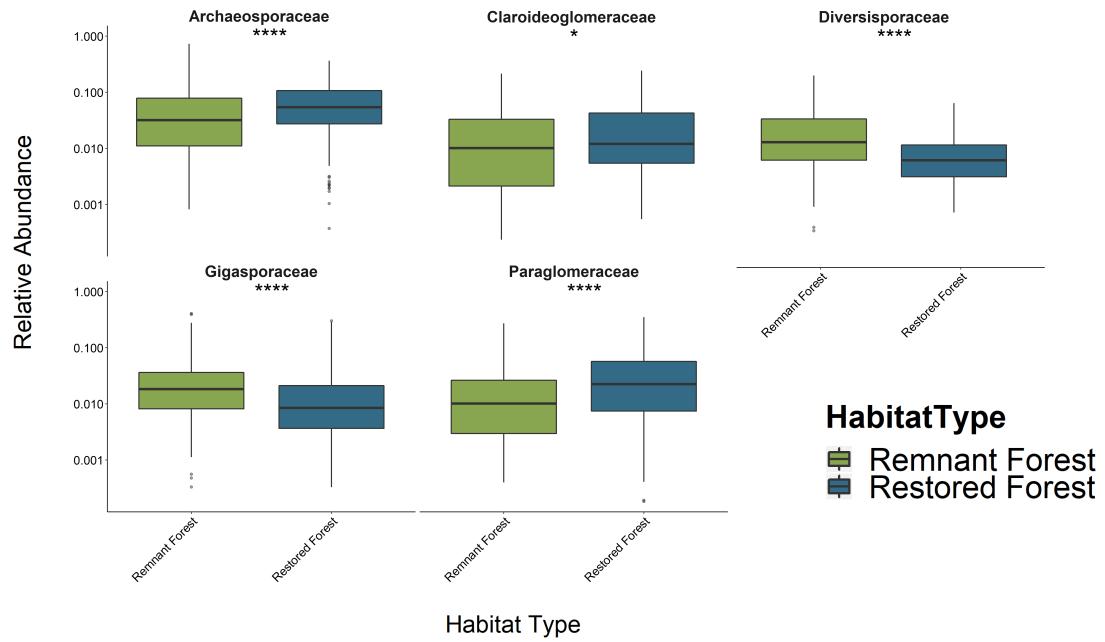
**Fig. S1 Sampling transects within the Hakalau National Forest Reserve.** Inset map showing the location of the Hakalau Forest National Wildlife Refuge (red point) in relation to the major Hawaiian islands (Fig. S1A). Sampling plots are coloured by restoration status/habitat type at each position. Sample plots established for the current study are shown in Fig S1B. Both Koa + Understory and Restored ohia forest habitat types were reclassified as being Restored Forest (blue) habitat for our study. Similarly, both Remnant koa dominated forest and Remnant forest dominated forest were re-classified as Remnant Forest (green) for our study.



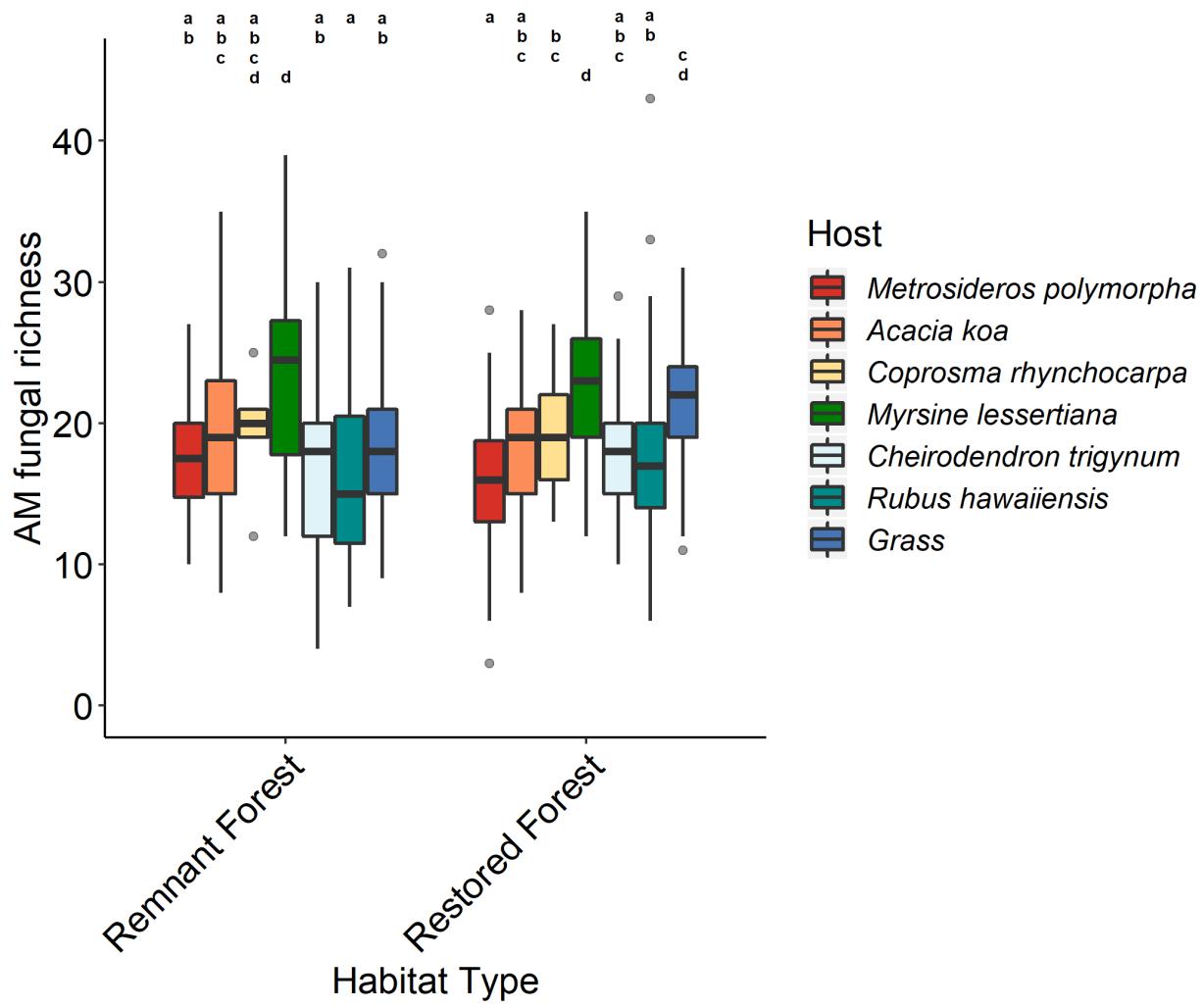
**Fig. S2** Accumulation curve of exact sequence variant (ESV) richness of Glomeromycotina sequences. Individual samples are denoted by different coloured lines.



**Fig. S3** Arbuscular mycorrhizal (AM) fungal co-occurrence networks for fungal communities in remnant forest (left) and restored forest (right) habitat types within the Hakalau Forest National Wildlife Refuge. Networks inferred using the Meinshausen and Bühlmann algorithm (11) with 9999 randomizations. Nodes represent AM fungal species and are coloured by AM fungal family. Edges represent non-random interactions between nodes.

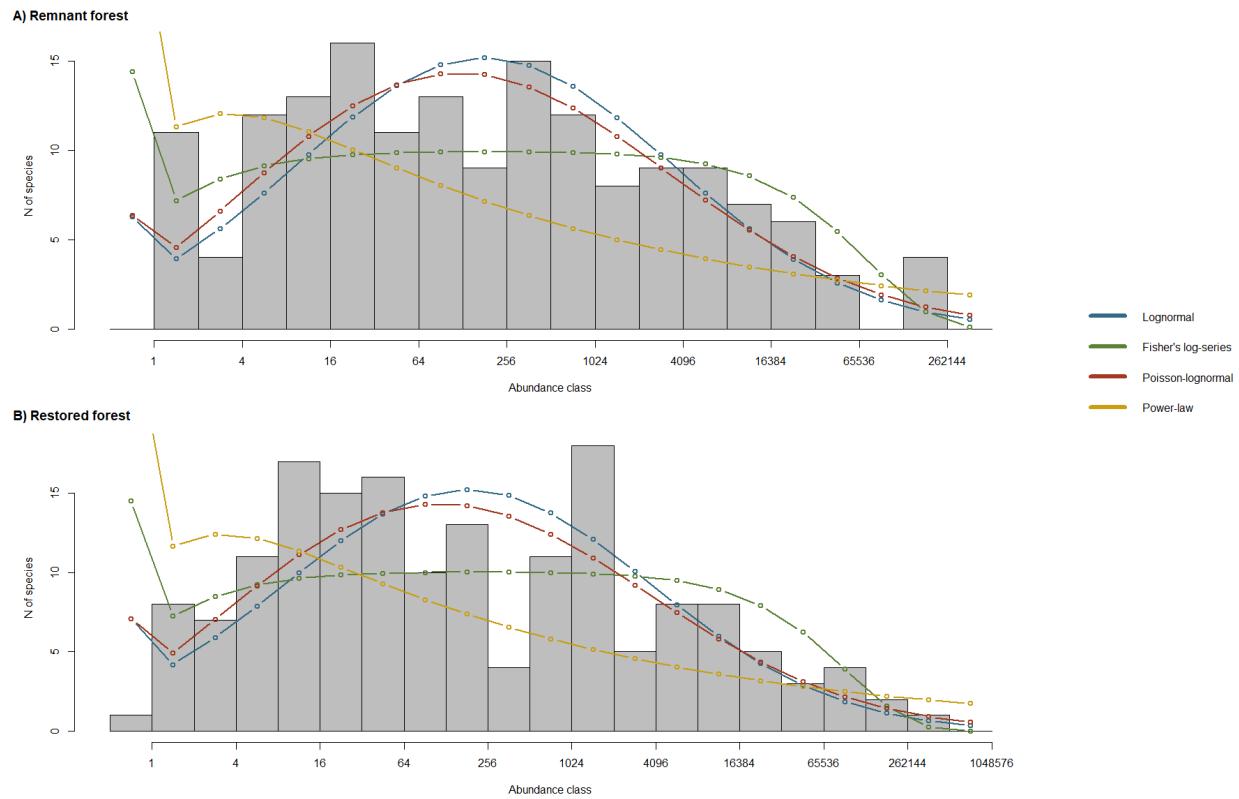


**Fig. S4 Relative sequence abundances of arbuscular mycorrhizal (AM) fungal families that differed significantly between the two habitat types sampled.** Boxplots displaying difference in read relative abundance for the major arbuscular mycorrhizal AM fungal families that differed significantly between Remnant (green) and Restored forest (blue) habitat types within the Hakalau Forest National Wildlife Refuge. The bottom and the top of the boxes represent the first and third quartiles, the dark band inside boxes represents the median, the whiskers contain the upper and lower 1.5 interquartile range, and the dots represent outliers. Mean read relative abundances within roots and soil were compared separately between habitat types using a Welch unequal variance t-test. Asterisks indicate statistical differences between habitat types where; ns =  $p > 0.05$ , \* =  $p \leq 0.05$ , \*\* =  $p \leq 0.01$ , \*\*\* =  $p \leq 0.001$ , \*\*\*\* =  $p \leq 0.0001$ . Log-scaled y-axis is for graphing purposes only, Welch unequal variance t-tests were done on non-transformed relative abundance data.

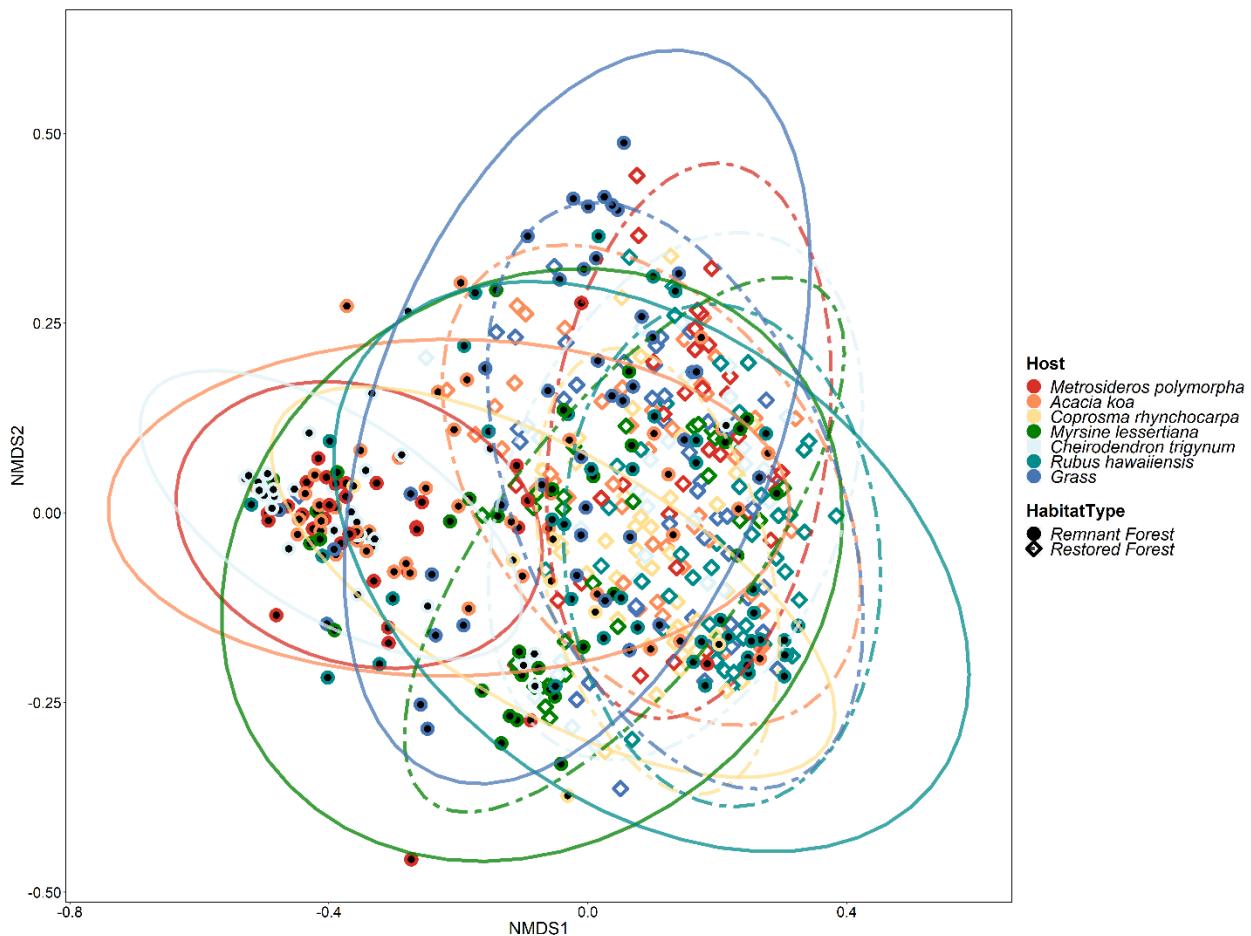


**Fig. S5 Arbuscular mycorrhizal (AM) fungal community richness within Hakalau.**

Boxplots displaying arbuscular mycorrhizal (AM) fungal richness per sample in relation to habitat type and plant host species. The bottom and the top of the boxes represent the first and third quartiles, the dark band inside boxes represents the median, the whiskers contain the upper and lower 1.5 interquartile range, and the dots represent outliers. Boxplots are coloured by host species. AM fungal richness varied significantly among hosts with habitat type having a marginal effect (Table S2). Boxes without shared letters indicate significant difference in AM fungal richness determined by estimated marginal means.



**Fig. S6 Observed species abundance distributions (SADs) for arbuscular mycorrhizal (AM) fungal communities sampled from remnant (A) and restored (B) forest habitat types within Hakalau.** Empirical SADs were fitted with multiple species distributions models (different coloured lines). Models were compared using Akaike's information criterion (AIC). SADs from both habitat types fit the Poisson-lognormal distribution model the best.



**Fig. S7 Community compositional differences between habitat types and among plant host species within Hakalau.** Non-metric dimensional scaling (NMDS) ordination plot of the community composition of arbuscular mycorrhizal (AM) fungi collected from the roots of different plant host species from both Remnant (filled circles) and Restored forest (empty diamonds) habitats within the Hakalau Forest National Wildlife Refuge. Compositional differences are based on Bray-Curtis dissimilarity among samples. Ellipses represent the 95% confidence region based on the centroid for each community, are coloured by host, with line type indicating whether the ellipse is for the Remnant (solid line) or Restored forest (dashed line) AM fungal community.

## VII. Additional References

1. D.L. Mumme, M.C. Rillig, Spatial characterization of arbuscular mycorrhizal fungal molecular diversity at the submetre scale in a temperate grassland. *FEMS Microbiology Ecology*, **64**(2), 260-270 (2008). doi: 10.1111/j.1574-6941.2008.00475.x
2. T.E. Cheeke et al., Spatial soil heterogeneity has a greater effect on symbiotic arbuscular mycorrhizal fungal communities and plant growth than genetic modification with *Bacillus thuringiensis* toxin genes. *Molecular Ecology*, **24**(10), 2580-2593 (2015). doi: 10.1111/mec.13178
3. A.J. Dumbrell, M. Nelson, T. Helgason, C. Dytham, A.H. Fitter, Idiosyncrasy and overdominance in the structure of natural communities of arbuscular mycorrhizal fungi: is there a role for stochastic processes? *Journal of Ecology*, **98**(2), 419-428 (2010). doi: 10.1111/j.1365-2745.2009.01622.x
4. J. Lee, S. Lee, P.W. Young, Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology*, **65**, 339-349 (2008). doi: 10.1111/j.1574-6941.2008.00531.x
5. CP Egan, D.W. Li, J.N. Klironomos, Detection of arbuscular mycorrhizal fungal spores in the air across different biomes and ecoregions. *Fungal Ecology*, **12**, 26-31 (2014). doi: 10.1016/j.funeco.2014.06.004
6. B.J. Callahan et al., DADA2: High - resolution sample inference from Illumina amplicon data. *Nature Methods*, **13**, 581-583 (2016). doi: 10.1038/nmeth.3869
7. M. Öpik et al., The online database MaarJAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist*, **188**, 223-241 (2010). doi: 10.1111/j.1469-8137.2010.03334.x.
8. C. Camacho et al., BLAST+: architecture and applications. *BMC Bioinformatics*, **10**, 421 (2009). doi: 10.1186/1471-2105-10-421.
9. M. Öpik, M. Moora, J. Liira, M. Zobel, Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. *Journal of Ecology*, **94**, 778-790 (2006). doi: 10.1111/j.1365-2745.2006.01136.x
10. B.M. Ohsowski, P.D. Zaitsoff, M. Öpik, M.M. Hart, Where the wild things are: looking for uncultured Glomeromycota. *New Phytologist*, **204**, 171-179 (2014). doi: 10.1111/nph.12894
11. N. Meinshausen, P. Bühlmann, High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, **34**(3), 1436-1462 (2006). doi: 10.1214/009053606000000281
12. Z. Kurtz, C. Mueller, E. Miraldi, R. Bonneau, SpiecEasi: Sparse Inverse Covariance for Ecological Statistical Inference. R package version 1.0.5 (2019).
13. H. Liu, K. Roeder, L. Wasserman, Stability Approach to Regularization Selection (StARS) for High Dimensional Graphical Models. *Advances in Neural Information Processing Systems*, **24**(2), 1432-1440 (2010). arXiv:1006.3316
14. L. Tipton et al., Fungi stabilize connectivity in the lung and skin microbial ecosystems. *Microbiome*, **6**(1), 12 (2018). doi: 10.1186/s40168-017-0393-0

15. L.C. Freeman, Centrality in social networks conceptual clarification. *Social Networks*, **1**(3), 215-239 (1979). doi: 10.1016/0378-8733(78)90021-7
16. M. Dufrêne, P. Legendre, Species assemblages and indicator species: the need for a flexible asymmetrical approach. *Ecological Monographs*, **67**(3), 345-366 (1997). doi: 10.1890/0012-9615(1997)067[0345:SAAIST]2.0.CO;2
17. M. De Cáceres, P. Legendre, Associations between species and groups of sites: indices and statistical inference. *Ecology*, **90**(12), 3566-3574 (2009). doi: 10.1890/08-1823.1
18. C.F. Dormann, B. Gruber, J. Fruend, Introducing the bipartite package: analysing ecological networks. *R news*, **8**(2), 8-11 (2008).
19. N. Blüthgen, F. Menzel, N. Blüthgen, Measuring specialization in species interaction networks. *BMC Ecology*, **6**, 9 (2006). doi.org/10.1186/1472-6785-6-9.
20. M. Almeida-Neto, P. Guimaraes, P.R. Guimaraes, R.D. Loyola, W. Ulrich, A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227-1239. (2008). doi: 10.1111/j.0030-1299.2008.16644.x
21. G. Strona, J.A. Veech, A new measure of ecological network structure based on node overlap and segregation. *Methods in Ecology and Evolution*, **6**(8), 907-915 (2015). doi: 10.1111/2041-210X.12395
22. L.F. Bersier, C. Banašek-Richter, M.-F. Cattin, Quantitative descriptors of food-web matrices. *Ecology*, **83**(9), 2394-2407 (2002). doi: 10.1890/0012-9658(2002)083[2394:QDOFWM]2.0.CO;2
23. S.K. Sepp et al., Non-random association patterns in a plant-mycorrhizal fungal network reveal host-symbiont specificity. *Molecular Ecology*, **28**(2), 365-378 (2019). doi: 10.1111/mec.14924
24. R.M. May, Will a large complex system be stable? *Nature*, **238**, 413-414 (1972). doi: 10.1038/238413a0