New Phytologist

Research

Core arbuscular mycorrhizal fungi are predicted by their high abundance–occupancy relationship while host-specific taxa are rare and geographically structured

Kacie T. Kajihara[1]*, Cameron P. Egan[1,2]*, Sean O. I. Swift[1], Christopher B. Wall[1,3], Christopher D. Muir[4] and Nicole A. Hynson[1]

[1]Pacific Biosciences Research Center, University of Hawai'i at Mānoa, 1993 East-West Road, Honolulu, HI 96822, USA; [2]Department of Biology, Okanagan College, 1000 KLO Road, Kelowna, BC VIY 4X8, Canada; [3]Biological Sciences, University of California San Diego, 9500 Gilman Dr, La Jolla, CA 92093, USA; [4]School of Life Sciences, University of Hawai'i at Mānoa, 2538 McCarthy Mall, Honolulu, HI 96822, USA

Author for correspondence:
Kacie Kajihara
Email: kaciekaj@hawaii.edu

Received: *6 January 2022*
Accepted: *16 February 2022*

New Phytologist (2022)
doi: 10.1111/nph.18058

Key words: community ecology, deforestation, ecological restoration, microbial ecology, mycorrhizal symbiosis, plant–soil feedback, tropical ecology.

Summary

- Habitat restoration may depend on the recovery of plant microbial symbionts such as arbuscular mycorrhizal (AM) fungi, but this requires a better understanding of the rules that govern their community assembly.
- We examined the interactions of soil and host-associated AM fungal communities between remnant and restored patches of subtropical montane forests.
- While AM fungal richness did not differ between habitat types, community membership did and was influenced by geography, habitat and host. These differences were largely driven by rare host-specific AM fungi that displayed near-complete turnover between forest types, while core AM fungal taxa were highly abundant and ubiquitous. The bipartite networks in the remnant forest were more specialized and hosts more specific than in the restored forest. Host-associated AM fungal communities nested within soil communities in both habitats, but only significantly so in the restored forest.
- Our results provide evidence that restored and remnant forests harbour the same core fungal symbionts, while rare host-specific taxa differ, and that geography, host identity and taxonomic resolution strongly affect the observed distribution patterns of these fungi. We suggest that host-specific interactions with AM fungi, as well as spatial processes, should be explicitly considered to effectively re-establish target host and symbiont communities.

Introduction

Anthropogenic disturbance has left habitats increasingly fragmented, with over 70% of remaining forests lying within 1 km of a forest edge (Haddad *et al.*, 2015). Coinciding losses in biodiversity impair ecosystem function world-wide (Pereira *et al.*, 2012), emphasizing the importance of forest restoration (Bastin *et al.*, 2019). Current approaches to terrestrial habitat and ecosystem restoration often fall short of restoring target ecosystem functions (Godefroid *et al.*, 2011), likely because they oversimplify complex systems (Palmer *et al.*, 1997) and omit key ecosystem components or processes (Wainwright *et al.*, 2018). Notably, plant symbiotic microorganisms such as arbuscular mycorrhizal (AM) fungi are often overlooked during restoration (Perring *et al.*, 2015), which is a conspicuous oversight considering their importance in host plant fitness (Smith & Read, 2008; Brundrett & Tedersoo, 2018), and influences on plant community diversity, composition and productivity (van der Heijden *et al.*, 1998a,b, 2008). Thus, manipulating AM fungal communities represents a promising tool to restore degraded habitats (Wubs *et al.*, 2016), but this requires a better understanding of the factors affecting their community assembly. While AM fungal communities are commonly shown to segregate by habitat (Öpik *et al.*, 2006; Kivlin *et al.*, 2011) or vegetation type (Li *et al.*, 2010; Martínez-García *et al.*, 2015), individual taxa can further diverge into those that are generalists or specialists to habitats or hosts, which may abide by different assembly rules (Barnes *et al.*, 2016).

Where a host or symbiont lands on the generalist–specialist spectrum is often linked to their disturbance response (Aizen *et al.*, 2012). Previous work has shown that generalist AM fungi are favoured in disturbed habitats compared to those found with specific hosts (Helgason *et al.*, 2007; Bennett *et al.*, 2013), but that host-specific AM fungi may be necessary to restore late-successional plant communities in degraded ecosystems (Koziol & Bever, 2017). However, the presence of these host-specific fungi is often diminished by disturbance and the introduction of novel invasive plants (Moora *et al.*, 2011) and may be especially difficult to maintain if restoration sites are distant in time or

*These authors contributed equally to this work.

space from their source pools (Wall *et al.*, 2020). Biogeography is a particularly important consideration for AM fungal community dynamics as many host-specific taxa are also generally considered poor dispersers (Egan *et al.*, 2014), while a few generalists are globally ubiquitous (Davison *et al.*, 2015). Together, these findings suggest that host-specific AM fungi may respond differently to disturbances such as land conversion and restoration techniques such as reforestation, but this has yet to be explicitly tested in a natural setting.

Ecological networks are an especially powerful tool to assess the degree of specificity, both of entire symbiotic communities and of individual hosts (Bascompte, 2009). Network patterns also provide valuable information on the stability of biotic interactions in the face of perturbations. For example, some studies suggest that increasing network complexity and specialization are properties that should be targeted when restoring ecological networks, but these ideas remain to be explicitly tested (Forup *et al.*, 2008; Morriën *et al.*, 2017). Regardless of the implications of network architecture for ecological restoration, because they measure complex interactions, network approaches provide a framework for assessing the system-wide status of restored areas relative to a reference, such as remnant primary forests.

While the presence of specialized network interactions between hosts and symbionts may be necessary to support ecosystem biodiversity, core symbiont community members or those that persist with a host even in the face of disturbance have also been shown to be indicators of host health (Risely, 2020). Despite advances in the last decade in enumerating diverse microbial communities, identifying core microbial symbiont communities remains challenging (Koskella *et al.*, 2017). We propose that ecological patterns such as abundance–occupancy relationships, which reflect the tendency for abundant species to be widespread and rare species to be limited in their distributions (Gaston *et al.*, 2000), can be used to infer core or host-specific community members. Under this framework, core community members should be derived from a resident community and have a high abundance–occupancy relationship within a host across habitats, while specialists would only be found in a specific host and exhibit a low abundance–occupancy relationship (Shade & Stopnisek, 2019).

In the context of AM fungi, the soil can be considered the resident community from which hosts are colonized, but not all host-associated AM fungi are always observed simultaneously residing in the soil (Hempel *et al.*, 2007; Varela-Cervero *et al.*, 2015). This discrepancy between soil and host-associated AM fungal communities is possibly due to methodological biases or temporal variation (Varela-Cervero *et al.*, 2015, 2016). However, including both soil and root-inhabiting fractions of the AM fungal community can provide important information on core community membership and host filtering. The current study examines the interactions between spatial patterns, habitat and host filtering (specificity) and the abundance–occupancy relationships for two main groups of AM fungi: (1) fungi found to co-occur in the soil and individual hosts; and (2) fungi found actively colonizing the roots of specific host plant species, but not detected in the soil, the latter of which we refer to as 'host-specific', while recognizing that all AM fungi at some stage of their life cycle inhabit both the soil and root niches.

Numerous barriers exist to understanding how AM fungal communities assemble in nature and how restoration practices such as the removal of invasive plant species and outplanting of native ones influence these communities. First, for example, AM fungal communities can be highly diverse with long-tail exponential distribution curves consisting of a few abundant and many rare taxa (Hughes *et al.*, 2001). Even with the advent of high-throughput sequencing, many studies have only been able to capture a fraction of the predicted diversity (e.g. Sepp *et al.*, 2018), making it difficult to reliably parse ecological patterns from sampling bias. Second, isolating habitat or environmental effects on host-associated AM fungal communities requires the same plant species to be sampled in each habitat type of interest, but because habitats and hosts typically co-vary, identical plant representation is rarely found (e.g. Davison *et al.*, 2015). Additionally, it has been shown that clustering sequences into operational taxonomic units or virtual taxa at 97% sequence similarity can obscure important measures of biogeography, diversity and species abundance for fungi (Tipton *et al.*, 2021). Here, we have overcome these barriers as we have: (1) sampled all host and soil AM fungal communities to completion, indicating that patterns arising from our data are biological rather than artefacts of sampling bias; (2) controlled for the conflation of habitat and host on AM fungal community assembly by sampling the same hosts in two distinct habitats (remnant primary forest and restored forest); and (3) employed the more granular taxonomic unit of amplicon sequence variants (ASVs), which are resolved down to the level of single-nucleotide differences, allowing host-, soil- or habitat-specific AM fungi and their overarching ecological patterns to be identified more precisely, at a taxonomic resolution that is comparable across studies (Callahan *et al.*, 2017).

We contrasted AM fungal community membership between habitats with differing land-use histories in the Hakalau Forest National Wildlife Refuge (Hakalau) of Hawai'i Island, USA, one of the largest reforestation projects in the State of Hawai'i. We predicted that hosts and habitats would be significant contributors to AM fungal community composition. We then considered the contribution of spatial proximity to AM fungal community composition for all fungi found across hosts and habitats, as well as for host-specific taxa only. In both cases, we predicted that host and habitat more than spatial proximity would determine AM fungal taxa turnover. To assess whether there are core AM fungi, we determined how soil pools of AM fungi contribute to host-associated communities in remnant primary and restored forests. We predicted that restored forest host-associated AM fungal communities would be more nested in soil communities than remnant forest communities due to their having less time to establish specialist interactions (Bennett *et al.*, 2013). Using a network approach, we also compared the specificity of AM fungal networks in remnant and restored forest patches and compared changes in host specificity between the two habitat types. We hypothesized that remnant forest networks would be more specialized than restored forest networks, as would their hosts. Finally, we tested whether host-specific and core AM fungi

clustered distinctly by ASV abundance and sample occupancy to assess the viability of the abundance–occupancy distribution as a predictor of core vs host-specific AM fungi. We predicted that AM fungal taxa with high abundance and occupancy would also be core community members, while the opposite would be true for host-specific taxa.

Materials and Methods

Study area

Samples were collected within Hakalau, which is located on the eastern slope of Mauna Kea on Hawai'i Island (19°51′N; 155°18′W). Before the refuge's establishment, upper-elevation forests withstood over 150 yr of habitat destruction through commercial ranching activities such as conversion to grassland, livestock grazing and logging. Reforestation of the now-abandoned 13 240 hectares of pasturelands began in 1987 with mass outplantings of the native nitrogen-fixing canopy tree *Acacia koa* (koa), followed by additional outplantings of native woody species *c.* 10 yr later (Jeffrey & Horiuchi, 2003). The goals of this restoration effort included the reduction in non-native grass cover and provision of mixed-canopy habitat for rare and endangered native Hawaiian bird species. But thirty years later, monodominant stands of *A. koa* and the continued dominance of exotic grasses in restored forest understories have created unamenable conditions for the recruitment of native seedlings and avian groups (Yelenik, 2017; Paxton *et al.*, 2018; Rehm *et al.*, 2019).

Sampling

Sampling was conducted in July and August 2017 from two main habitat types within Hakalau: (1) remnant forest areas with intact native over- and understory, co-dominated by the canopy trees *A. koa* and *Metrosideros polymorpha* ('ōhi'a lehua); and (2) restoration corridors of *A. koa* with large patches of pasture grass, and outplanted native woody species such as *M. polymorpha*, *Cheirodendron trigynum* ('ōlapa), *Myrsine lessertiana* (kōlea), *Coprosma rhynchocarpa* (pilo), *Rubus hawaiensis* ('ākala) and *Vaccinium calycinum* ('ōhelo). Hereafter, the two habitats will be referred to as either remnant or restored forests.

Sampling occurred along two parallel transects, each spanning from the restored to the remnant forest, which have similar climates, topography, elevations and soil parent materials (Supporting Information Fig. S1). Restored plots were chosen in areas with plant communities most resembling those in the remnant forest. Within each 12-m-diameter plot, roots and nearby soil were sampled from five native plant species known to host AM fungi (*A. koa, M. polymorpha, C. trigynum, M. lessertiana* and *R. hawaiensis*), as well as non-native grasses that were dominantly *Pennisetum clandestinum,* but not all grasses were identified to species. Roots were sampled from up to eight individuals of each host species by tracing from large to fine roots for a total of 569 root samples across 12 plots. Soil was sampled by taking the top 10 cm of mineral soil directly adjacent to where roots were sampled for a total of 583 soil samples including samples taken for the soil chemistry analyses in Wall *et al.* (2020). If eight individuals of each host species were not present in a given plot, we increased sampling intensity so that each plot yielded the same number of root and soil samples per host. Root and soil samples were bagged, stored on ice and kept in a 1°C cold room for up to 1 wk until DNA extractions could be conducted.

DNA extraction and sequencing

DNA was extracted from 0.25 g subsamples of prepared soil and root material (Notes S1) in separate batches for each sample type using the MPBio FastDNA® Spin Kit for soil and the MPBio FastDNA® Spin Kit for Plant and Animal Tissue (MP Biomedicals LLC, Santa Ana, CA, USA) for soil and roots, respectively, following the manufacturer's instructions. We then conducted a two-step PCR to amplify the small subunit (SSU) of AM fungi's ribosomal RNA (rRNA) and adhere Illumina barcodes and adaptors to our amplicons. The first reaction amplified 550 bp of AM fungal SSU rRNA using the universal eukaryotic primer WANDA (Dumbrell *et al.*, 2011) with the Fluidigm tag CS1 on the 5′ end (Fluidigm Inc., South San Francisco, CA, USA) and the Glomeromycotinan-specific primer AML2 (Lee *et al.*, 2008) with the Fluidigm tag CS2 on the 3′ end. PCRs were performed in 25 µl volumes using 2.5 µl of DNA template, 7.8 µl of nuclease-free H_2O, 12.5 µl of 2× Kapa Plant PCR Buffer (containing 1.5 mM $MgCl_2$ (1×) and 0.2 mM of each dNTP), 1.5 µl of 25 mM $MgCl_2$, 0.5 µl of 50 µM AML2/CS2, 0.5 µl of 50 µM WANDA/CS1 and 0.2 µl 2.5 U μl^{-1} Kapa3G DNA Polymerase (Kapa Biosystems, Wilmington, MA, USA). An Eppendorf Mastercycler Nexus Thermal Cycler (Eppendorf North America, Hauppauge, NY, USA) was used under the following conditions: 95°C for 2 min; 30 cycles of 95°C for 30 s, 54°C for 40 s and 72°C for 1 min; and 72°C for 10 min. Extracted soil DNA was diluted to 1 : 10 concentrations, and this diluted DNA was used as template for the PCR1 product.

The second PCR used the same Fluidigm tags as above (CS1 or CS2) along with an 8-bp Illumina Nextera barcode and an Illumina adapter (Egan *et al.*, 2018). Amplicons from the first PCR were first diluted 15-fold, and then, PCRs were performed in 20 µl volumes using 1 µl of diluted PCR1 product, 6.89 µl of nuclease-free H_2O, 10 µl of 2× Kapa Plant PCR Buffer (containing 1.5 mM $MgCl_2$ (1×) and 0.2 mM of each dNTP), 1.2 µl of 25 mM $MgCl_2$, 0.75 µl of 2 µM Illumina adaptor/barcode primer complex and 0.16 µl 2.5 U μl^{-1} Kapa3G DNA Polymerase. Thermocycler conditions were as follows: 95°C for 1 min; 12 cycles of 95°C for 30 s, 60°C for 30 s and 68°C for 1 min; and 68°C for 5 min.

Amplicons were purified and normalized to 25 ng using the Just-a-Plate™ PCR Purification and Normalization Plate (Charm Biotech, San Diego, CA, USA) following the manufacturer's instructions. Root and soil amplicon pools made by pooling 10 µl of eluted 25 ng DNA from each sample were purified again and size-selected for amplicons *c.* 650 bp in length using AMPure XP beads (Beckman Coulter, Indianapolis, IN, USA). Two Illumina MiSeq runs were conducted, one for each

sample type (soil and roots), using 2 × 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/).

Bioinformatics

Demultiplexing, quality control and taxonomic assignment were performed with QIIME2 v.2020.8 (Bolyen *et al.*, 2019). Forward and reverse reads were demultiplexed using the *q2-demux* plugin, resulting in 4435 672 paired soil sequences and 9464 322 paired root sequences being assigned sample barcodes. Low-quality sequences and chimeras were removed with DADA2 (Callahan *et al.*, 2016) via *q2-dada2*, using a trim-left-*f* value of 20 and trim-left-*r* value of 22 to remove conserved primer regions. Sequences passing quality checks were assembled into amplicon sequence variants (ASVs) for a total of 2300 ASVs representing 2290 893 soil sequences and 3704 ASVs representing 2853 318 root sequences. The two ASV tables were merged using *q2-feature-table*, resulting in a total of 5196 ASVs representing 5144 211 sequences.

Taxonomy was assigned to ASVs using MAARJAM reference sequences (Öpik *et al.*, 2010) via *q2-feature-classifier* (Bokulich *et al.*, 2018), which uses the BLAST+ algorithm (Camacho *et al.*, 2009). We aimed to identify ASVs with the subphylum level of Glomeromycotina, which encompasses only arbuscular mycorrhizal (AM) fungi. Searches were conducted in order of decreasing sequence similarity, while sequence alignment was kept at ≥ 90% and BLAST *e*-value at < 1e-50 for all. At ≥ 95% sequence similarity, 944 ASVs (25% of ASVs) were identified, representing 2494 818 sequences (48% of quality-checked sequences). Unassigned sequences were searched again at ≥ 90% sequence similarity, resulting in 198 ASVs (5% of ASVs) representing 32 063 sequences (*c.* 0.01% of quality-checked sequences). A second search of unassigned sequences was done at ≥ 80% sequence similarity (QIIME2 default value). From this search, 665 more ASVs were identified (18% of ASVs) representing 108 608 sequences (2% of quality-checked sequences). A final search of unassigned sequences was done against 18S Glomeromycotina reference sequences downloaded from the NCBI Nucleotide database (accessed September 2020) at ≥ 80% sequence similarity, resulting in 277 ASVs (7% of ASVs) representing 535 340 sequences (10% of quality-checked sequences).

Glomeromycotinan ASVs from all four taxonomy searches were merged via *q2-feature-table* after removing unassigned sequences and sequences from negative PCR controls via *q2-taxa*. The filtered ASV table and taxonomy classification file were exported from QIIME2 for downstream analyses in R v.4.0.2 (R Core Team, 2020), and the QIIME2R package (Bisanz, 2018) was used to import qza files into RSTUDIO (RStudio Team, 2020). Singleton sequences were discarded, generating a complete data set of 1108 samples representing 1932 Glomeromycotinan ASVs. A reduced version of the data set was generated for analyses based on distance matrices of combined soil and root AM fungal communities (NMDS ordination and beta dispersion tests), in which samples with < 1000 sequences were discarded. The reduced data set contained 800 samples representing 1766 ASVs.

Data analysis

All analyses were conducted in R v.4.0.2 (R Core Team, 2020) using RSTUDIO (RStudio Team, 2020). Accumulation curves of ASV richness generated in the iNEXT package (Hsieh *et al.*, 2016) reached asymptotes for all host and soil communities in both habitats (Fig. S2). Differences in AM fungal richness by host between habitats were assessed with Welch's two-sample *t*-tests, and *P*-value corrections were made using the Benjamini and Hochberg control of the false discovery rate (Benjamini & Hochberg, 1995). Richness among hosts within habitats was evaluated via one-way analysis of variance (ANOVA) and Tukey's *post hoc* tests. Overlap in community membership was visualized with the area-proportional Euler diagrams (Larsson & Gustafsson, 2018).

To examine the effect of host plant identity, habitat type (remnant or restored forest) and spatial proximity (as principal coordinates of neighbour matrices (PCNMs); see description below) on overall AM fungal community composition (soil and roots), we constructed a nonmetric multidimensional scaling (NMDS) ordination from the Bray–Curtis distances of the reduced data set using the *metaMDS* function in the VEGAN package (Oksanen *et al.*, 2013). The *envfit* function (Oksanen *et al.*, 2013) was used to fit each variable to the NMDS and test their effects on overall AM fungal community composition. Multivariate homogeneity of group dispersion was tested using *betadisper* (Oksanen *et al.*, 2013).

To compare how compositionally nested AM fungal root communities from the two habitats were within their respective soil communities, nestedness was evaluated via two metrics: (1) matrix temperature (Atmar & Patterson, 1993; Rodríguez-Gironés & Santamaría, 2006) and (2) the nestedness metric based on the overlap and decreasing fill (NODF; Almeida-Neto *et al.*, 2008), calculated with the *nestedtemp* and *nestednodf* functions from the VEGAN package using the presence–absence data (Notes S2).

The bipartite package (Dormann *et al.*, 2008) was used to assess whole network ($H2'$) and host specificity (d') by aggregating AM fungal ASVs by host plant species (Blüthgen *et al.*, 2006). Networks were assembled independently for each plot using observed abundance, which generated six plant–AM fungal networks per habitat type. The between-habitat differences were assessed using Welch's two-sample *t*-tests. *P*-values were corrected (Benjamini & Hochberg, 1995) for both sets of comparisons (network and host level).

To further examine how hosts and habitats differed in terms of AM fungal community composition, the relative abundance of AM fungal families was calculated for three groups of ASVs: (1) ASVs shared between individual hosts and soil; (2) individual host- and soil-specific ASVs; and (3) ASVs shared between all hosts and soil, which formed a subset of the first group. Differences in relative abundance within hosts between habitats were analysed using Welch's two-sample *t*-tests, and *P*-values were corrected (Benjamini & Hochberg, 1995). To compare the overall

contribution of each group to AM fungal diversity, within each habitat we calculated the relative abundance of ASVs by family and then averaged across sample type (soil or roots).

Grouped ASVs were then further categorized by habitat specificity. To assess the contribution of spatial proximity and habitat type to the high rate of host-specific AM fungal ASV turnover (see the 'Results' section), we reanalysed our Bray–Curtis matrix including only host-specific ASVs and generated PCNM (Borcard & Legendre, 2002; Oksanen *et al.*, 2013) vectors using our georeferenced plot locations. This reduced the dimensions of our geographic data and created spatial predictors ('Geo' vectors) that could be incorporated into a principal coordinates analysis (PCoA) that also included our categorical variable of habitat type. The first six vectors were used to represent the contribution of geography to community composition and were fit to the matrix using the *envfit* function (Dray *et al.*, 2006; Oksanen *et al.*, 2013; Darcy *et al.*, 2020).

To visualize the abundance–occupancy relationship of core and host-specific AM fungi from our study, ASV abundance was regressed against the number of samples in which ASVs were found (occupancy). To estimate whether core and host-specific AM fungi differed in their abundance and occupancy relationships, we used a Bayesian multiresponse Poisson regression model, modelling abundance and occupancy as the Poisson-distributed response variables. The likelihood of observed counts by latent parameters is modelled using a log link from a bivariate normal distribution; hence, the effect of category (core or host-specific) is modelled through coefficients on the Poisson rate parameters. We fit the model in STAN v.2.26.1 (Stan Development Team, 2021) with a CMDSTANR v.0.4.0 backend (Gabry & Češnova, 2021). It ran on four parallel chains for 1000 warm-up iterations and 1000 sampling iterations. All parameters converged ($\hat{R} \approx 1$), and the effective sample size from the posterior exceeded 1000 (Vehtari *et al.*, 2021). We used the posterior median for point estimates and calculated uncertainty with the 95% quantile interval from the posterior distribution.

Results

In total, 1932 AM fungal ASVs were detected in remnant and restored forest samples (1165 ASVs in the remnant forest and 1188 ASVs in the restored; Fig. S3a). ASV accumulation curves were saturated for soil and every host in both habitats, indicating that our sampling and sequencing captured the vast majority of AM fungal taxa from our study system (Fig. S2). There was no evidence that the average AM fungal richness associated with each host differed between habitats (Table S1), but there was moderate-to-strong evidence that richness differed between some hosts within habitats (Fig. S4; Table S2). For soil AM fungi, there was moderate evidence of higher richness in the restored forest relative to the remnant ($P = 0.013$, $t = -3.139$, Fig. S4; Table S1) and strong evidence that richness was lower in each habitat relative to all hosts except for *C. trigynum* and *M. polymorpha* (Table S2). ASVs detected in soil samples accounted for 37% of all ASVs (724 ASVs) and were divided evenly between the two habitats (Fig. S3b). The remaining 63% of ASVs (1208 ASVs) were detected only in hosts (Fig. S3b), but the degree to which these ASVs were shared among hosts differed between habitat types.

Relative to the null model, AM fungal communities showed very strong evidence of nesting within soil communities and by host in the restored forest (Fig. 1b; $T = 36.9$, $P < 0.001$; NODF = 30.7, $P < 0.001$). Remnant forest communities, despite having similar nestedness values, were more modular and showed no evidence of differing from the null (Fig. 1a; $T = 34.7$, $P = 0.545$; NODF = 32.0, $P = 0.121$). In both habitats, grass hosted much of the AM fungal diversity represented in other species (Fig. 1). Except for *Myrsine lessertiana* and grass, the ordering of hosts in the nestedness hierarchy differed by forest type. In the remnant forest, the foundational native tree species *A. koa* and *M. polymorpha* were less nested (more modular) than *R. hawaiensis* and *C. trigynum*. Conversely in the restored forest, *A. koa* and *M. polymorpha* were more nested within the other host AM fungal communities.

Bipartite network analyses revealed moderate evidence that remnant forest networks were more specialized (higher H2′) ($P = 0.010$, $t = -3.736$, Table S3; Fig. S5). Host specificity (d') was on average higher in the remnant forest, moderately so for *M. polymorpha* ($P = 0.040$, $t = 3.799$) and marginally so in all remaining hosts except for *A. koa* and *M. lessertiana* whose d' values showed no evidence of differing between habitat types (Fig. 2; Table S4).

Geographic distance, habitat type and host plant identity showed very strong evidence of predicting AM fungal community composition ($P < 0.001$), explaining 30.1%, 15.2% and 6.0% of the variation, respectively (Fig. S6; Table 1). Community dispersion also showed very strong evidence of differing between hosts ($P < 0.001$, $F = 5.573$) and habitats ($P < 0.001$, $F = 14.787$).

ASVs shared between individual hosts and soil made up 23.2% of the total observed in the remnant forest and 21.5% in the restored, but only half of these ASVs were considered core or detected in the same host and soil in both habitats (53.7% in the remnant forest and 56.8% in the restored). Core taxa were highly abundant in both habitats, contributing an average of 93.6% (SD ± 1.9) of the total relative abundance for a given host community in the remnant forest and 93.8% (SD ± 2.9) in the restored (Fig. 3; Table S5). Overall, there was no evidence that the relative abundance of AM fungal families shared between hosts and soil differed across habitat except among a select few hosts and two fungal families (Table S5). However, the spotty presence of Diversisporaceae made statistical comparisons of relative abundance for this family intractable (Table S5).

Half of the total ASVs in each habitat were defined as 'host-specific' because they were found to only colonize the roots of specific host plant species and were not detected in the soil (49.2% in the remnant forest and 52.5% in the restored). The average relative abundance of these host-specific ASVs was low relative to core taxa at 2.2% (SD ± 2.6) in the remnant forest and 1.7% (SD ± 2.1) in the restored (Fig. 4; Table S6). There was no evidence of differences in the relative abundance of AM fungal families in the host- and soil-specific ASVs between habitats (Table S6). At the ASV level, however, we observed a near-complete turnover in the identity of host-specific AM fungi

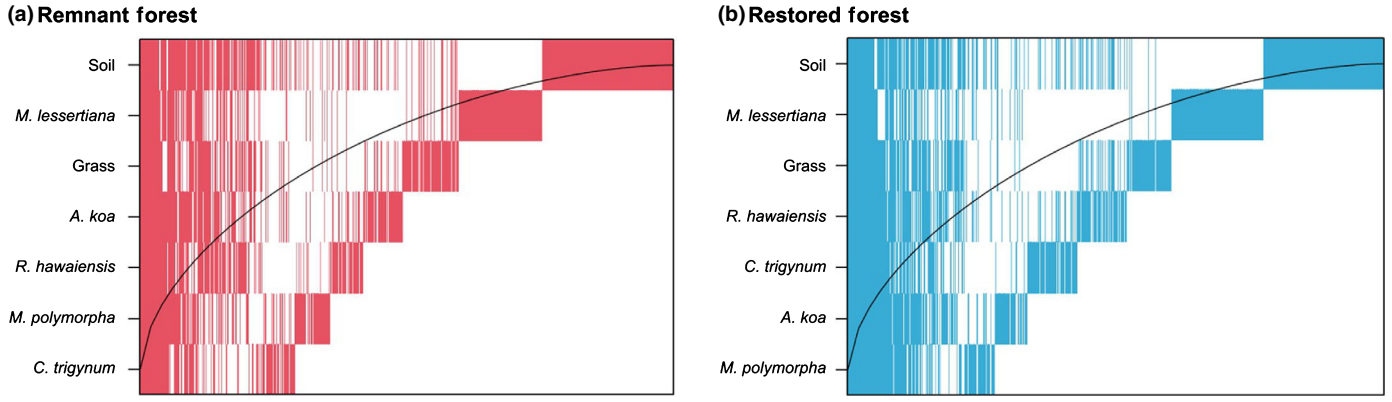

Fig. 1 Nestedness plots of arbuscular mycorrhizal (AM) fungal amplicon sequence variants (ASVs) aggregated by host or soil community membership in remnant (a; red) and restored (b; blue) forests within the Hakalau forest national wildlife refuge. Restored forest communities showed strong evidence of nestedness relative to a null model ($T = 36.9$, $P < 0.001$; nestedness metric based on the overlap and decreasing fill (NODF) = 30.7, $P < 0.001$), while remnant forest communities did not ($T = 34.7$, $P = 0.545$; NODF = 32.0, $P = 0.121$). Plant hosts included *Acacia koa, Cheirodendron trigynum, Metrosideros polymorpha, Myrsine lessertiana, Rubus hawaiensis* and grasses that were dominantly *Pennisetum clandestinum,* but not all grasses were identified to species. Each vertical line represents the presence of an AM fungal ASV within host or soil communities. Perfect nestedness ($T = 0$ or NODF = 100) would be achieved if all ASVs occurred in the left of the curved black line.

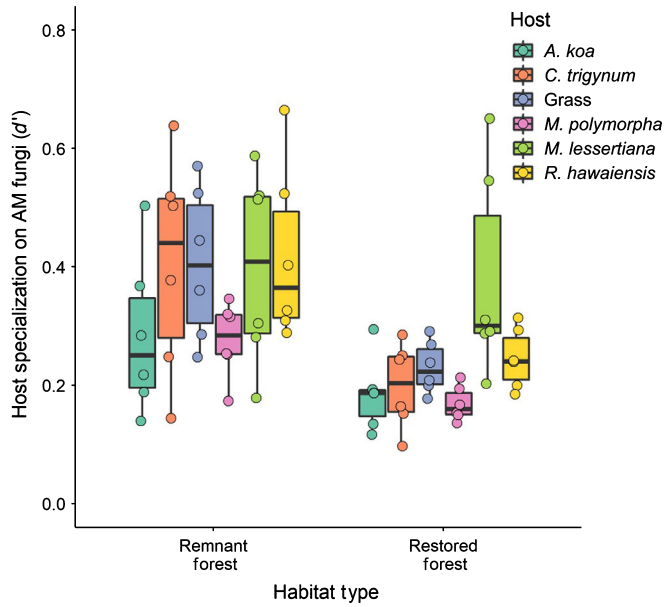

Fig. 2 Boxplots displaying host specialization (d') on arbuscular mycorrhizal (AM) fungi in the remnant (left) and restored (right) forests. The bottom and top edges of the boxes represent the first and third quartiles, the line inside the boxes represents the median, the whiskers contain the lower and upper 1.5 interquartile range, and the dots represent a host's d' value for each plot in a given habitat. Boxes are coloured by host identity. Plant hosts included *Acacia koa, Cheirodendron trigynum, Metrosideros polymorpha, Myrsine lessertiana, Rubus hawaiensis* and grasses that were dominantly *Pennisetum clandestinum,* but not all grasses were identified to species.

Table 1 Contributions of geographic distance, habitat type (remnant or restored forest) and host plant identity to arbuscular mycorrhizal (AM) fungal beta diversity within the Hakalau forest national wildlife refuge.

Variable	Vector	R^2	P
Geography	Geo1	0.025	< 0.001
	Geo2	0.022	< 0.001
	Geo3	0.138	< 0.001
	Geo4	0.055	< 0.001
	Geo6	0.061	< 0.001
Habitat		0.152	< 0.001
Host		0.060	< 0.001

Calculations were performed using the *envfit* function in the VEGAN R package (Oksanen *et al*., 2013) on a Bray–Curtis dissimilarity matrix. Principal coordinates of neighbour matrices (PCNMs) were used to generate 'Geo' vectors, using georeferenced plot locations for every sample, to reduce the dimensions of geographic data and allow for comparisons with categorical variables. The first six vectors were used to represent the contribution of geography to AM fungal community composition. R^2 values represent the proportion of variance explained by each variable or vector. All variables shown had very strong evidence of contributing to beta diversity ($P < 0.001$).

between habitats (Fig. 4). Nearly, all host-specific ASVs (98.8% in the remnant forest and 98.9% in the restored) also appeared to be habitat-specific (Fig. 4).

Within the pool of core ASVs was a small group of generalists shared between all hosts and soil (3.3% of ASVs in the remnant forest and 4.7% of ASVs in the restored). Depending on the host or soil community, there was moderate-to-strong evidence of differences in relative abundance within sample type between habitats for these generalist taxa in AM fungal families Acaulosporaceae, Archaeosporaceae, Gigasporaceae, Glomeraceae and those that could not be classified beyond subphylum Glomeromycotina (Table S7). A subset of the generalist ASVs (2.5% of ASVs in the remnant forest and 2.4% of ASVs in the restored) were 'universal' or shared across all hosts and soil in both habitats. Within the remnant forest, universal ASVs were most abundant in grass, soil and *R. hawaiensis* and least abundant in *C. trigynum* and *M. polymorpha*, and within the restored forest there were no significant differences among hosts or soil (Fig. S7; Table S7).

To assess the contribution of habitat type and geographic distance to the high turnover of AM fungal host-specific ASVs, we used a PCoA on the Bray–Curtis community dissimilarity matrix of just these taxa, and asked whether habitat type and our

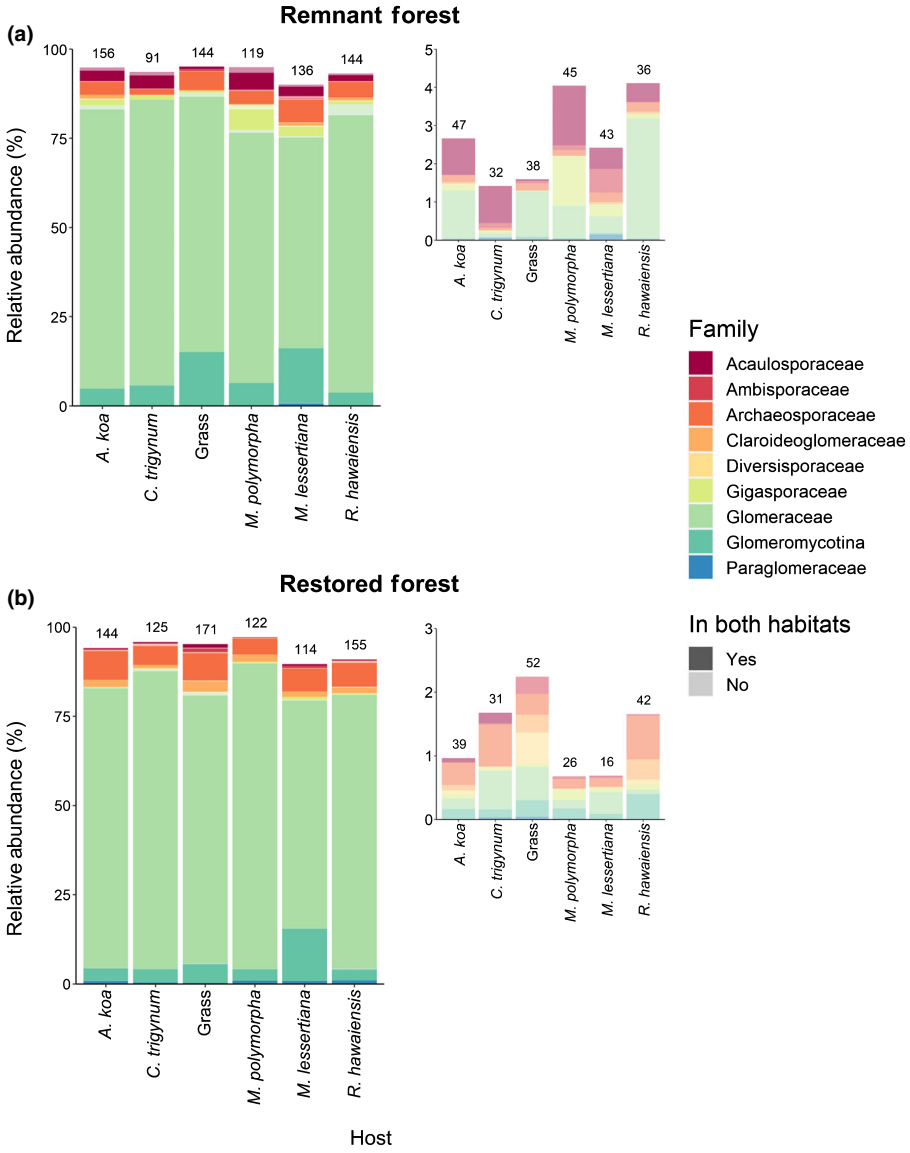

Fig. 3 Relative abundance of arbuscular mycorrhizal (AM) fungal families detected in amplicon sequence variants (ASVs) shared between individual hosts and soil in remnant (a) and restored (b) forest habitats. Bar labels represent ASV counts. ASVs present in the same host in both habitats ('core' ASVs) are shown in solid colours, while habitat-specific ASVs are shown in fainter colours. Smaller inset bar plots depict habitat-specific ASVs only; note the different scales for relative abundance on the *y*-axis. Plant hosts included *Acacia koa, Cheirodendron trigynum, Metrosideros polymorpha, Myrsine lessertiana, Rubus hawaiensis* and grasses that were dominantly *Pennisetum clandestinum*, but not all grasses were identified to species. ASVs that could not be assigned at the family level are represented as Glomeromycotina, the subphylum that encompasses all AM fungi.

PCNMs (representative of geographic distance) were strong explanatory factors of community composition. Both habitat and geography showed moderate-to-very strong evidence of predicting host-specific AM fungal community composition, where geographic distance explained 16.1% of the variation, and habitat type explained 6.3% (Table 2; Fig. 5).

Regression analysis showed very strong evidence of a correlation between ASV abundance and sample occupancy (Fig. S8, $R^2 = 0.763$, $P < 0.001$). There was very strong evidence that core and host-specific ASVs had different abundance–occupancy relationships. Based on 4000 sampling iterations, the difference in abundance between the two groups was always > 0, where the core taxa had high abundance and occupancy and host-specific the opposite (Fig. 6).

Discussion

Here, we examined the relationship between soil and AM fungal communities and how those interactions varied with respect to land-use history. Similar to many areas in the tropics, at our study site, land-use changes include the conversion of forest to pasture over 150 yr ago, and then, a massive reforestation effort initiated over three decades ago. These types of disturbances can lead to significant alterations in edaphic factors such as soil chemistry (Wubs *et al.*, 2016). In a previous study, Wall *et al.* (2020) found that despite the similar climate, slope, aspect, elevation and soil parent material, the soil chemical properties of reforested (restored) pasture sites differed from remnant forest patches and that these differences affect soil AM fungal community membership. In the present study, our aim was to examine the interactions between soil pools of AM fungi, geography and habitat type on host-associated communities. Our results reaffirmed the significance of geography, habitat and host filtering as drivers of community assembly (Öpik *et al.*, 2006; Kivlin *et al.*, 2011; Egan *et al.*, 2014), but revealed that the taxonomic resolution at which observations are made and the abundance–occupancy distribution are important considerations and predictors of community membership.

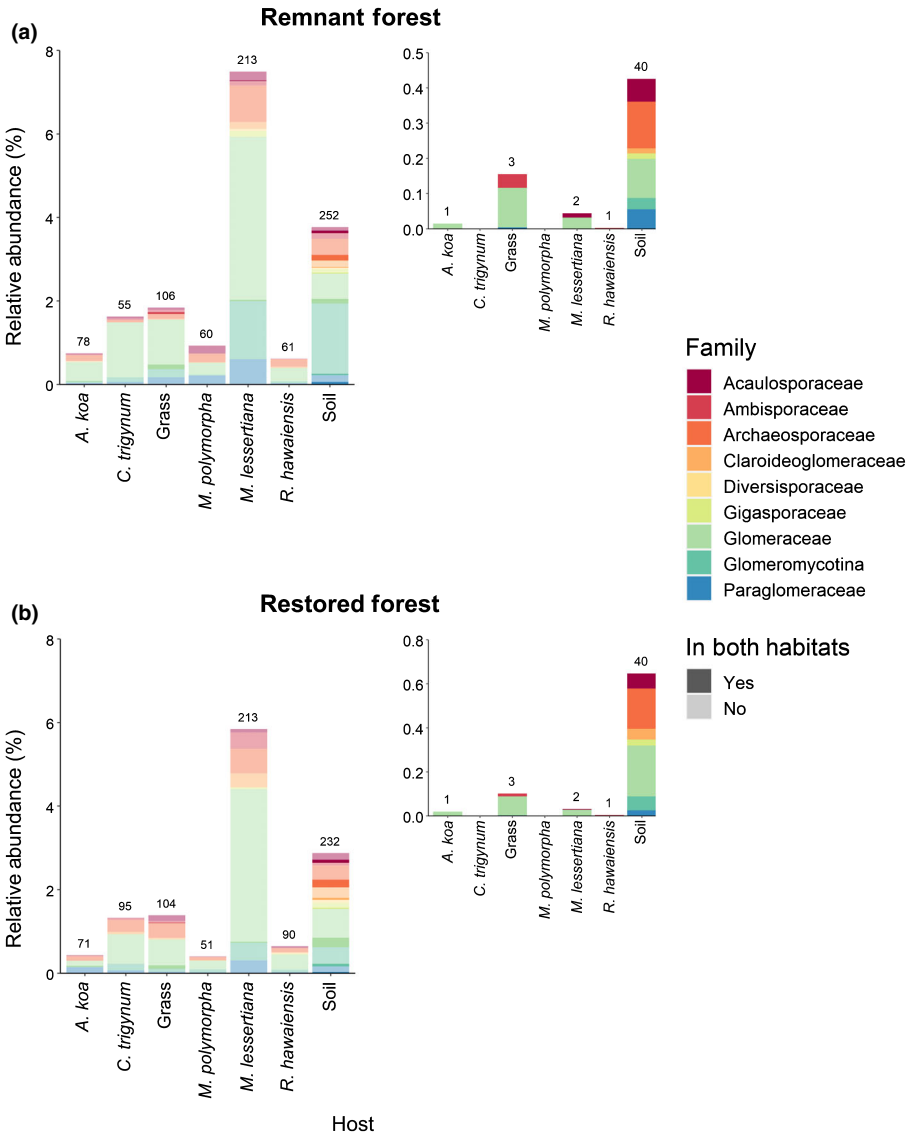

Fig. 4 Relative abundance of arbuscular mycorrhizal (AM) fungal families detected in host- and soil-specific amplicon sequence variants (ASVs) in remnant (a) and restored (b) forest habitats. Bar labels represent ASV counts. ASVs present in the same host or soil community in both habitats are shown in solid colours, while habitat-specific ASVs are shown in fainter colours. Smaller inset bar plots depict ASVs present in both habitats only; note the different scales for relative abundance on the y-axis. Plant hosts included *Acacia koa, Cheirodendron trigynum, Metrosideros polymorpha, Myrsine lessertiana, Rubus hawaiensis* and grasses that were dominantly *Pennisetum clandestinum,* but not all grasses were identified to species. ASVs that could not be assigned at the family level are represented as Glomeromycotina, the subphylum that encompasses all AM fungi.

There was little evidence that AM fungal ASV richness changed between the remnant and restored forests (Fig. S3a; Table S1), indicating that unlike prior studies, habitat alteration does not always lead to a decrease in AM fungal richness (House & Bever, 2018; Hart *et al.*, 2016, but see Lekberg *et al.*, 2013). Instead, we found differences in community composition between habitats, especially for host-specific ASVs. AM fungal communities in the restored forest were more homogeneous with a greater percentage of generalist and universally shared ASVs (Fig. S7). Host communities in the restored forest also showed strong evidence of nesting within the soil community, implying a lack of specialization (Fig. 1; Bennett *et al.*, 2013). Further, bipartite network analyses revealed moderate evidence that restored forest networks were less specialized than remnant forest networks (Fig. S6). Together, these results indicate a more uniform AM fungal community across hosts in the restored forest made up of more generalist taxa than in the remnant forest, which could be impeding restoration by negatively impacting native seedling recruitment in the restored forest, while favouring generalist hosts such as non-native grass.

A more uniform AM fungal community in the restored forest could be due to differences in forest age. For instance, higher root densities in the older remnant forest could lead to greater competition for soil resources than in the younger restored forest resulting in more niche partitioning by taxonomically distinct fungi (Chagnon *et al.*, 2013). Another possible explanation for the diminished specialization in the restored forest is the persistence of AM fungi from the historic pasture grass community. In both habitats, we observed grass to act as a generalist partnering with more AM fungal symbionts than any host besides the outlier *M. lessertiana* (Fig. 1). The grass AM fungal community may be acting as a donor from which native hosts accumulate mycorrhizal partners, but in the restored forest, this community may be dominated by legacy pasture fungi that are of lesser mutualistic quality to native host plants (Chagnon *et al.*, 2013). This legacy effect might also explain why the AM fungal communities

Table 2 Contributions of geographic distance and habitat type (remnant or restored forest) to host-specific arbuscular mycorrhizal (AM) fungal amplicon sequence variant (ASV) turnover.

Variable	Vector	R^2	P
Geography	Geo3	0.056	< 0.001
	Geo4	0.014	0.036
	Geo6	0.091	< 0.001
Habitat		0.063	< 0.001

Calculations were performed using the *envfit* function in the VEGAN R package (Oksanen *et al*., 2013) on the Bray–Curtis dissimilarity matrix for host-specific ASVs only. Principal coordinates of neighbour matrices (PCNMs) were used to generate 'Geo' vectors, using georeferenced plot locations for every sample, to reduce the dimensions of geographic data and allow for comparisons with categorical variables. The first six vectors were used to represent the contribution of geography to AM fungal community composition. R^2 values represent the proportion of variance explained by each variable or vector. All variables shown had moderate-to-very strong evidence of contributing to beta diversity ($P \leq 0.036$).

associated with the foundational native tree species *A. koa* and *M. polymorpha* were less nested in the remnant than in the restored forest (Fig. 1), which suggests a shift in their functional roles from symbiont donors to recipients. In accordance with Moora *et al.* (2011), the generalist alien grasses may have facilitated the proliferation of generalist AM fungi, which in turn can suppress the diversity of local and host-specific AM fungi and may be curtailing the community assembly processes and function of the restored forest.

With few exceptions, AM fungal ASVs contributed from the soil pool to individual hosts were also core community members, in that they were recruited by hosts in both habitat types independent of land-use history (Fig. 3). In general, core ASVs were most taxon-rich and abundant within host communities in both the remnant and restored forests (Table S5; Figs 3, 6). This, combined with our result that host-associated AM fungal communities nest within the soil community, suggests that the soil is a significant source pool for hosts, contributing an average 51.4% of a host's ASVs (SD ± 9.8) regardless of habitat. We also found that hosts' access to various AM fungal families did not tend to differ between habitat types (Table S5), indicating that these core AM fungal communities at the levels of both families and ASVs have been maintained despite intensive changes in land use. Core microbiomes have been studied in many systems (Turnbaugh *et al.*, 2009; Lundberg *et al.*, 2012; Ainsworth *et al.*, 2015) and are thought to define the 'healthy' state of a host organism (Zaura *et al.*, 2009). Core community members have been assumed to persist across disturbance gradients due to their importance to host function (Shade & Stopnisek, 2019; Risely, 2020), but despite robust core communities observed within hosts between habitat types at Hakalau, restored forest recovery has been stalled, and native seedling recruitment is low (Yelenik, 2017), suggesting that additional AM fungal taxa, such as host specialists, may be necessary to restore the target ecosystem state.

The AM fungal specificity of hosts was on average lower in the restored forest than in the remnant except for *A. koa* and *M. lessertiana*, the latter of which associated primarily with host-specific taxa regardless of habitat type (Figs 1, 2). Our findings that core AM fungal communities appear to be similar between habitats and that *A. koa* – the original outplant species for reforestation – has a similar degree of specificity regardless of habitat type potentially indicate the successful restoration or maintenance of AM fungal symbionts. However, while at the familial level the identities and relative abundance of host-specific AM fungal taxa showed no evidence of differing between habitat types (Table S6), at the more granular level of ASVs, we observed nearly complete turnover in these communities for all hosts between the remnant and restored forests (Fig. 4). This mass turnover cannot be attributed to undersampling because we have sequenced all host-associated AM fungal communities to completion (Fig. S2). This result, along with our finding that *c.* 6% of the host-specific AM fungal community composition was owed

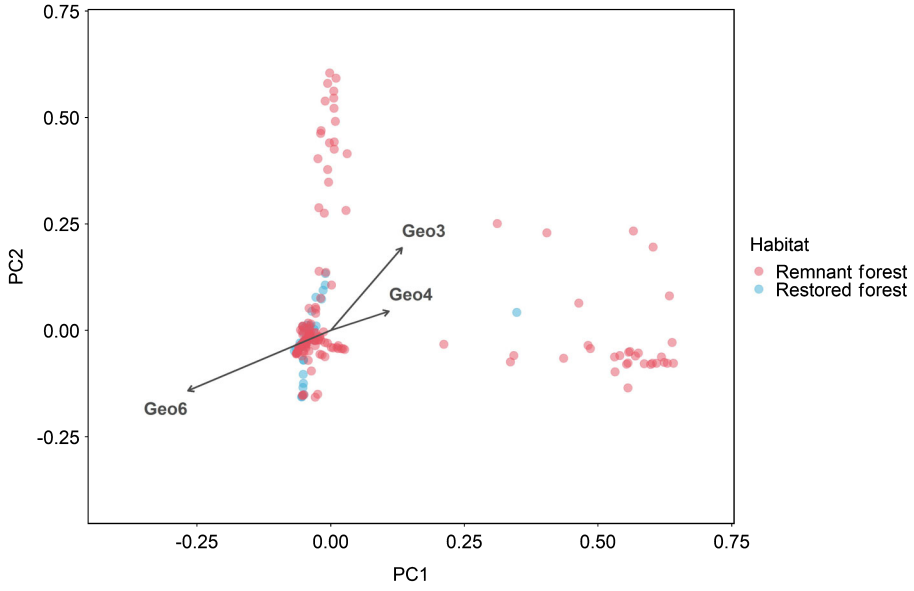

Fig. 5 Principal coordinates analysis (PCoA) ordination plot of host-specific arbuscular mycorrhizal (AM) fungal community composition between remnant (red) and restored (blue) forests. Distances are based on the Bray–Curtis dissimilarity matrix. Arrows correspond to principal coordinates of neighbour matrix (PCNM) vectors representing moderate-to-very strong evidence of geographic distance contributing to beta diversity ($P \leq 0.036$).

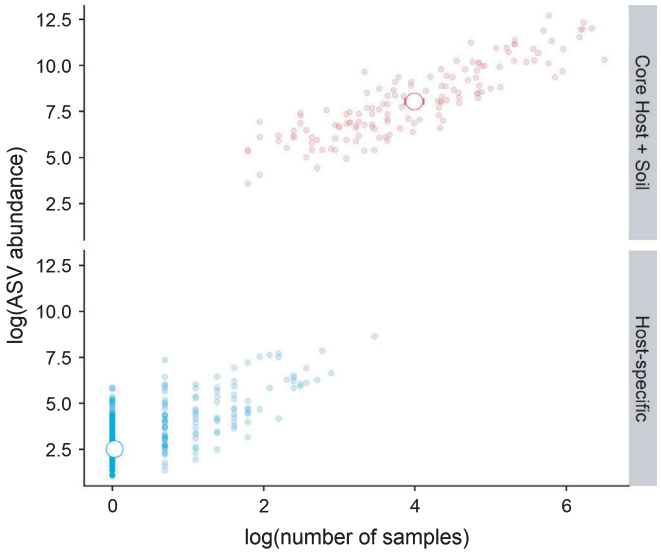

Fig. 6 Log–log scatter plots of amplicon sequence variant (ASV) abundance against a number of samples by core ASVs and host-specific ASVs. Red points (upper) represent ASVs shared between individual hosts, soil and habitats (core community members). Blue points (lower) represent all host-specific ASVs. Means are represented by large white circles. The 95% confidence interval of each mean is represented by coloured lines extending outward from a white circle; this interval is too narrow to be visible for the host-specific ASVs (lower). The difference in abundance between groups was > 0 in all 4000 sampling iterations.

to habitat (Table 2), indicates that these taxa may be important considerations for restoration, with the caveat that spatial processes also need to be taken into account (Vályi *et al.*, 2016). We found that when spatial proximity of our plots along with habitat type were taken into consideration, geography explained more than twice the variation in host-specific AM fungal community turnover than habitat (Table 2). Thus, the make-up of host-specific communities, while affected by differences in biotic and abiotic conditions between habitats, is likely also owed to processes such as dispersal limitation (Vályi *et al.*, 2016).

While recently receiving less attention than core microbiomes, host-specific symbionts have shown to be crucial mediators of host health in mutualisms throughout nature and are often required to ensure the fitness of one or both partners (Poulsen *et al.*, 2005; Wang *et al.*, 2012). Though the AM symbiosis has traditionally been viewed as generalist (Smith & Read, 2008), many studies have since demonstrated that specific AM fungi affect plant growth, seedling establishment and plant–soil feedbacks differently by host species (Kiers *et al.*, 2000; Klironomos, 2000; Bruns *et al.*, 2002; Mangan *et al.*, 2010), which in turn influence plant community structure and productivity for the whole ecosystem (van der Heijden *et al.*, 1998a,b). In the case of restoration, it may be important to consider the proper pairing of hosts with specific symbionts that will provide the most benefit in the particular environmental conditions of a restoration site, especially since based on our results, these fungi do not appear to have persistent soil reservoirs.

We found that the abundance–occupancy relationship of AM fungi is an indicator of their role in the community as either core or host-specific taxa. Core taxa were more abundant and ubiquitous than host-specific, which were rare (Fig. 6). While this result is somewhat intuitive, it provides a new and useful framework for understanding the ecology of cryptic soil microorganisms such as AM fungi. Furthermore, the fact that host-specific AM fungi were also rare implies that they are at higher risk of extinction (Harnik *et al.*, 2012), an additional and important consideration in the context of ecosystem restoration.

In summary, habitat restoration has successfully restored or maintained core AM fungal communities from the reference ecosystem, in this case, montane tropical native forests. However, differences in network and host specificity, as well as the high turnover in the rare host-specific AM fungi between the restored and reference ecosystems, could be impairing host plant fitness and leading to an incomplete state of recovery in the restored forest, but this concept needs to be tested through additional manipulative studies. Furthermore, we found that spatial proximity is an important determinant of AM fungal community composition, indicating that the closer a restoration site is to a reference ecosystem, the higher the likelihood of finding similar AM fungal communities. Also, family-level comparisons may be too coarse to detect important differences in community composition that may affect host health; thus, a better understanding of the level at which phylogenetic conservatism informs AM fungal function is desperately needed (Chagnon *et al.*, 2013). We suggest that future restoration efforts consider legacy effects and pre-inoculating hosts with reference ecosystem AM fungal communities or those that provide the greatest benefit to hosts in the specific environmental conditions of the restoration site (Maltz & Treseder, 2015; Neuenkamp *et al.*, 2019). Future work should examine the functional roles of core and host specialist communities to better understand their roles in host health and potential importance in restoring ecosystems.

Acknowledgements

We thank Corbin Amend, Dave Bouck, Ken Davidson, Sarah Schoepflin and Stephanie Yelenik for help with sampling, the Hakalau Forest National Wildlife Refuge for site access, and James Downey, Terrance McDermott and Danyel Yogi for their dedication and long hours processing samples for molecular and chemical analyses. Additionally, we thank Anthony Amend, Carla D'Antonio, Miranda Hart, Vincent Merckx, Evan Rehm, Laura Tipton and Stephanie Yelenik for stimulating discussions of ideas. We would also like to acknowledge the helpful feedback from three expert anonymous reviewers and our Editor, Maarja Öpik. This work was funded by the National Science Foundation (no.1556856) and the W. M. Keck Foundation.

Author contributions

NAH, SOIS and CPE designed the experiment. CPE, SOIS, KTK and NAH collected the data. CPE and SOIS performed laboratory analyses. CPE, CBW and KTK performed bioinformatics and analysed the data. CDM performed abundance–occupancy analyses. KTK and NAH wrote the manuscript. KTK and CPE contributed equally to this work.

ORCID

Cameron P. Egan https://orcid.org/0000-0001-6863-071X
Nicole A. Hynson https://orcid.org/0000-0001-7098-9641
Kacie T. Kajihara https://orcid.org/0000-0002-1741-5400
Christopher D. Muir https://orcid.org/0000-0003-2555-3878
Sean O. I. Swift https://orcid.org/0000-0002-1170-4711
Christopher B. Wall https://orcid.org/0000-0002-7164-3201

Data availability

All data for analyses are available at https://github.com/kkajihara/hakalau. DNA sequences and project metadata are archived in the NCBI Sequence Read Archive (SRA) under BioProject accession no. PRJNA644447. Code for abundance–occupancy analyses associated with Fig. 6 is available at https://github.com/cdmuir/abundance-occupancy. All other code is available at https://github.com/kkajihara/hakalau.

References

Ainsworth DT, Krause L, Bridge T, Torda G, Raina J-B, Zakrzewski M, Gates RD, Padilla-Gamiño JL, Spalding HL, Smith C *et al.* 2015. The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. *ISME Journal* **9**: 2261–2274.

Aizen MA, Sabatino M, Tylianakis JM. 2012. Specialization and rarity predict nonrandom loss of interactions from mutualist networks. *Science* **335**: 1486–1489.

Almeida-Neto M, Guimarães P, Guimarães PR, Loyola RD, Ulrich W. 2008. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**: 1227–1239.

Atmar W, Patterson BD. 1993. The measure of order and disorder in the distribution of species in fragmented habitat. *Oecologia* **96**: 373–382.

Barnes CJ, Burns CA, van der Gast CJ, McNamara NP, Bending GD. 2016. Spatio-temporal variation of core and satellite arbuscular mycorrhizal fungus communities in *Miscanthus giganteus*. *Frontiers in Microbiology* **7**: 1278.

Bascompte J. 2009. Disentangling the web of life. *Science* **325**: 416–419.

Bastin J-F, Finegold Y, Garcia C, Mollicone D, Rezende M, Routh D, Zohner CM, Crowther TW. 2019. The global tree restoration potential. *Science* **365**: 76–79.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B: Methodological* **57**: 289–300.

Bennett AE, Daniell TJ, Öpik M, Davison J, Moora M, Zobel M, Selosse M-A, Evans D. 2013. Arbuscular mycorrhizal fungal networks vary throughout the growing season and between successional stages. *PLoS ONE* **8**: e83241.

Bisanz JE. 2018. *qiime2R: importing QIIME2 artifacts and associated data into R sessions*. R package v.20. [WWW document] URL https://github.com/jbisanz/qiime2R [accessed 16 September 2020].

Blüthgen N, Menzel F, Blüthgen N. 2006. Measuring specialization in species interaction networks. *BMC Ecology* **6**: 1–12.

Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**: 1–17.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F *et al.* 2019. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nature Biotechnology* **37**: 852–857.

Borcard D, Legendre P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling* **153**: 51–68.

Brundrett MC, Tedersoo L. 2018. Evolutionary history of mycorrhizal symbioses and global host plant diversity. *New Phytologist* **220**: 1108–1115.

Bruns TD, Bidartondo MI, Taylor DL. 2002. Host specificity in ectomycorrhizal communities: what do the exceptions tell us? *Integrative and Comparative Biology* **42**: 352–359.

Callahan BJ, McMurdie PJ, Holmes SP. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal* **11**: 2639–2643.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581–583.

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Chagnon P-L, Bradley RL, Maherali H, Klironomos JN. 2013. A trait-based framework to understand life history of mycorrhizal fungi. *Trends in Plant Science* **18**: 484–491.

Darcy JL, Swift SO, Cobian GM, Zahn GL, Perry BA, Amend AS. 2020. Fungal communities living within leaves of native Hawaiian dicots are structured by landscape-scale variables as well as by host plants. *Molecular Ecology* **29**: 3102–3115.

Davison J, Moora M, Öpik M, Adholeya A, Ainsaar L, Bâ A, Burla S, Diedhiou AG, Hiiesalu I, Jairus T *et al.* 2015. Global assessment of arbuscular mycorrhizal fungus diversity reveals very low endemism. *Science* **349**: 970–973.

Dormann CF, Gruber B, Fründ J. 2008. Introducing the bipartite package: analysing ecological networks. *R News* **8**(2): 8–11.

Dray S, Legendre P, Peres-Neto PR. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling* **196**: 483–493.

Dumbrell AJ, Ashton PD, Aziz N, Feng G, Nelson M, Dytham C, Fitter AH, Helgason T. 2011. Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. *The New Phytologist* **190**: 794–804.

Egan C, Li D-W, Klironomos J. 2014. Detection of arbuscular mycorrhizal fungal spores in the air across different biomes and ecoregions. *Fungal Ecology* **12**: 26–31.

Egan CP, Rummel A, Kokkoris V, Klironomos J, Lekberg Y, Hart M. 2018. Using mock communities of arbuscular mycorrhizal fungi to evaluate fidelity associated with Illumina sequencing. *Fungal Ecology* **33**: 52–64.

Forup ML, Henson KS, Craze PG, Memmott J. 2008. The restoration of ecological interactions: plant–pollinator networks on ancient and restored heathlands. *Journal of Applied Ecology* **45**: 742–752.

Gabry J, Češnova R. 2021. *CmdStanR: R Interface to 'CmdStan'. v.0.4.0.* [WWW document] URL https://mc-stan.org/cmdstanr/reference/cmdstanr-package.html#:~:text=CmdStanR%20(cmdstanr%20package)%20is%20a,users%2Finterfaces%2Fcmdstan [accessed 23 March 2021].

Gaston KJ, Blackburn TM, Greenwood JJD, Gregory RD, Quinn RM, Lawton JH. 2000. Abundance–occupancy relationships. *Journal of Applied Ecology* **37**: 39–59.

Godefroid S, Piazza C, Rossi G, Buord S, Stevens A-D, Aguraiuja R, Cowell C, Weekley CW, Vogg G, Iriondo JM *et al.* 2011. How successful are plant species reintroductions? *Biological Conservation* **144**: 672–682.

Haddad NM, Brudvig LA, Clobert J, Davies KF, Gonzalez A, Holt RD, Lovejoy TE, Sexton JO, Austin MP, Collins CD *et al.* 2015. Habitat fragmentation and its lasting impact on Earth's ecosystems. *Science Advances* **1**: e1500052.

Harnik PG, Simpson C, Payne JL. 2012. Long-term differences in extinction risk among the seven forms of rarity. *Proceedings of the Royal Society B: Biological Sciences* **279**: 4969–4976.

Hart M, Zaitsoff D, van der Heyde M, Pither J. 2016. Testing life history and trait-based predictions of AM fungal community assembly. *Pedobiologia* **59**: 203–213.

van der Heijden MGA, Bardgett RD, van Straalen NM. 2008. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecology Letters* **11**: 296–310.

van der Heijden MGA, Boller T, Wiemken A, Sanders IR. 1998a. Different arbuscular mycorrhizal fungal species are potential determinants of plant community structure. *Ecology* **79**: 2082–2091.

van der Heijden MGA, Klironomos JN, Ursic M, Moutoglis P, Streitwolf-Engel R, Boller T, Wiemken A, Sanders IR. 1998b. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. *Nature* **396**: 69–72.

Helgason T, Merryweather JW, Young JPW, Fitter AH. 2007. Specificity and resilience in the arbuscular mycorrhizal fungi of a natural woodland community. *Journal of Ecology* **95**: 623–630.

Hempel S, Renker C, Buscot F. 2007. Differences in the species composition of arbuscular mycorrhizal fungi in spore, root and soil communities in a grassland ecosystem. *Environmental Microbiology* **9**: 1930–1938.

House GL, Bever JD. 2018. Disturbance reduces the differentiation of mycorrhizal fungal communities in grasslands along a precipitation gradient. *Ecological Applications* **28**: 736–748.

Hsieh TC, Ma KH, Chao A. 2016. iNEXT: an R package for rarefaction and extrapolation of species diversity (Hill numbers). *Methods in Ecology and Evolution* **7**: 1451–1456.

Hughes JB, Hellmann JJ, Ricketts TH, Bohannan BJM. 2001. Counting the uncountable: statistical approaches to estimating microbial diversity. *Applied and Environmental Microbiology* **67**: 4399–4406.

Jeffrey J, Horiuchi B. 2003. Tree planting at Hakalau forest national wildlife refuge—the right tool for the right stock type. *Native Plants Journal* **4**: 30–31.

Kiers ET, Lovelock CE, Krueger EL, Herre EA. 2000. Differential effects of tropical arbuscular mycorrhizal fungal inocula on root colonization and tree seedling growth: implications for tropical forest diversity. *Ecology Letters* **3**: 106–113.

Kivlin SN, Hawkes CV, Treseder KK. 2011. Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry* **43**: 2294–2303.

Klironomos JN. 2000. Host-specificity and functional diversity among arbuscular mycorrhizal fungi. *Microbial Biosystems: New Frontiers* **1**: 845–851.

Koskella B, Hall LJ, Metcalf CJE. 2017. The microbiome beyond the horizon of ecological and evolutionary theory. *Nature Ecology & Evolution* **1**: 1606–1615.

Koziol L, Bever JD. 2017. The missing link in grassland restoration: arbuscular mycorrhizal fungi inoculation increases plant diversity and accelerates succession. *Journal of Applied Ecology* **54**: 1301–1309.

Larsson J, Gustafsson P. 2018. A case study in fitting area-proportional Euler diagrams with ellipses using eulerr. In: *SetVR 2018*, 84–91. [WWW document] URL http://ceur-ws.org/Vol-2116/paper7.pdf [accessed 29 September 2020].

Lee J, Lee S, Young JPW. 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology* **65**: 339–349.

Lekberg Y, Gibbons SM, Rosendahl S, Ramsey PW. 2013. Severe plant invasions can increase mycorrhizal fungal abundance and diversity. *ISME Journal* **7**: 1424–1433.

Li L-F, Li T, Zhang Y, Zhao Z-W. 2010. Molecular diversity of arbuscular mycorrhizal fungi and their distribution patterns related to host-plants and habitats in a hot and arid ecosystem, southwest China. *FEMS Microbiology Ecology* **71**: 418–427.

Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, Tremblay J, Engelbrektson A, Kunin V, Del Rio TG *et al.* 2012. Defining the core *Arabidopsis thaliana* root microbiome. *Nature* **488**: 86–90.

Maltz MR, Treseder KK. 2015. Sources of inocula influence mycorrhizal colonization of plants in restoration projects: a meta-analysis. *Restoration Ecology* **23**: 625–634.

Mangan SA, Schnitzer SA, Herre EA, Mack KM, Valencia MC, Sanchez EI, Bever JD. 2010. Negative plant–soil feedback predicts tree-species relative abundance in a tropical forest. *Nature* **466**: 752–755.

Martínez-García LB, Richardson SJ, Tylianakis JM, Peltzer DA, Dickie IA. 2015. Host identity is a dominant driver of mycorrhizal fungal community composition during ecosystem development. *New Phytologist* **205**: 1565–1576.

Moora M, Berger S, Davison J, Öpik M, Bommarco R, Bruelheide H, Kühn I, Kunin WE, Metsis M, Rortais A *et al.* 2011. Alien plants associate with widespread generalist arbuscular mycorrhizal fungal taxa: evidence from a continental-scale study using massively parallel 454 sequencing. *Journal of Biogeography* **38**: 1305–1317.

Morriën E, Hannula SE, Snoek LB, Helmsing NR, Zweers H, de Hollander M, Soto RL, Bouffaud M-L, Buée M, Dimmers W *et al.* 2017. Soil networks become more connected and take up more carbon as nature restoration progresses. *Nature Communications* **8**: 1–10.

Neuenkamp L, Prober SM, Price JN, Zobel M, Standish RJ. 2019. Benefits of mycorrhizal inoculation to ecological restoration depend on plant functional type, restoration context and time. *Fungal Ecology* **40**: 140–149.

Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'hara RB, Simpson GL, Solymos P, Stevens MHH, Wagner H. 2013. *Package 'vegan'. Community ecology package* v.2.1-295 [WWW document] URL https://cran.r-project.org/web/packages/vegan/index.html.

Öpik M, Moora M, Liira J, Zobel M. 2006. Composition of root-colonizing arbuscular mycorrhizal fungal communities in different ecosystems around the globe. *Journal of Ecology* **94**: 778–790.

Öpik M, Vanatoa A, Vanatoa E, Moora M, Davison J, Kalwij JM, Reier Ü, Zobel M. 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist* **188**: 223–241.

Palmer MA, Ambrose RF, Poff NL. 1997. Ecological theory and community restoration ecology. *Restoration Ecology* **5**: 291–300.

Paxton EH, Yelenik SG, Borneman TE, Rose ET, Camp RJ, Kendall SJ. 2018. Rapid colonization of a Hawaiian restoration forest by a diverse avian community: bird colonization of tropical restoration forest. *Restoration Ecology* **26**: 165–173.

Pereira HM, Navarro LM, Martins IS. 2012. Global biodiversity change: the bad, the good, and the unknown. *Annual Review of Environment and Resources* **37**: 25–50.

Perring MP, Standish RJ, Price JN, Craig MD, Erickson TE, Ruthrof KX, Whiteley AS, Valentine LE, Hobbs RJ. 2015. Advances in restoration ecology: rising to the challenges of the coming decades. *Ecosphere* **6**: 1–25.

Poulsen M, Cafaro M, Boomsma JJ, Currie CR. 2005. Specificity of the mutualistic association between actinomycete bacteria and two sympatric species of *Acromyrmex* leaf-cutting ants. *Molecular Ecology* **14**: 3597–3604.

R Core Team. 2020. *A language and environment for statistical computing, v.4.0.2.* Vienna, Austria: R Foundation for Statistical Computing. [WWW document] URL http://www.R-project.org/.

R Studio Team. 2020. *RStudio: integrated development for R, v.4.1.0.* Boston, MA: RStudio, PBC. [WWW document] URL http://www.rstudio.com/ [accessed 1 September 2020].

Rehm E, Thomas M, Yelenik S, Bouck D, D'Antonio C. 2019. Bryophyte abundance, composition and importance to woody plant recruitment in natural and restoration forests. *Forest Ecology and Management* **444**: 405–413.

Risely A. 2020. Applying the core microbiome to understand host–microbe systems. *Journal of Animal Ecology* **89**: 1549–1558.

Rodríguez-Gironés MA, Santamaría L. 2006. A new algorithm to calculate the nestedness temperature of presence–absence matrices. *Journal of Biogeography* **33**: 924–935.

Sepp S-K, Jairus T, Vasar M, Zobel M, Öpik M. 2018. Effects of land use on arbuscular mycorrhizal fungal communities in Estonia. *Mycorrhiza* **28**: 259–268.

Shade A, Stopnisek N. 2019. Abundance-occupancy distributions to prioritize plant core microbiome membership. *Current Opinion in Microbiology* **49**: 50–58.

Smith SE, Read DJ. 2008. *Mycorrhizal symbiosis.* Amsterdam, the Netherlands & Heidelberg, Germany: Elsevier, AP.

Stan Development Team. 2021. *Stan modeling language users guide and reference manual.* [WWW document] URL https://mcstan.org [accessed 23 March 2021].

Tipton L, Zahn GL, Darcy JL, Amend AS, Hynson NA. 2021. Hawaiian fungal amplicon sequence variants reveal otherwise hidden biogeography. *Microbial Ecology* **1**: 10.

Turnbaugh PJ, Hamady M, Yatsunenko T, Cantarel BL, Duncan A, Ley RE, Sogin ML, Jones WJ, Roe BA, Affourtit JP *et al.* 2009. A core gut microbiome in obese and lean twins. *Nature* **457**: 480–484.

Vályi K, Mardhiah U, Rillig MC, Hempel S. 2016. Community assembly and coexistence in communities of arbuscular mycorrhizal fungi. *ISME Journal* **10**: 2341–2351.

Varela-Cervero S, López-García Á, Barea JM, Azcón-Aguilar C. 2016. Differences in the composition of arbuscular mycorrhizal fungal communities promoted by different propagule forms from a Mediterranean shrubland. *Mycorrhiza* **26**: 489–496.

Varela-Cervero S, Vasar M, Davison J, Barea JM, Öpik M, Azcón-Aguilar C. 2015. The composition of arbuscular mycorrhizal fungal communities differs among the roots, spores and extraradical mycelia associated with five Mediterranean plant species. *Environmental Microbiology* **17**: 2882–2895.

Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C. 2021. Rank-normalization, folding, and localization: an improved $\hat{R}$ for assessing convergence of MCMC (with Discussion). *Bayesian Analysis* **16**: 667–718.

Wainwright CE, Staples TL, Charles LS, Flanagan TC, Lai HR, Loy X, Reynolds VA, Mayfield MM. 2018. Links between community ecology theory and ecological restoration are on the rise. *Journal of Applied Ecology* **55**: 570–581.

Wall CB, Egan CP, Swift SIO, Hynson NA. 2020. Three decades post-reforestation has not led to the reassembly of arbuscular mycorrhizal fungal communities associated with remnant primary forests. *Molecular Ecology* **29**: 4234–4247.

Wang D, Yang S, Tang F, Zhu H. 2012. Symbiosis specificity in the legume – rhizobial mutualism. *Cellular Microbiology* **14**: 334–342.

Wubs ER, Van der Putten WH, Bosch M, Bezemer TM. 2016. Soil inoculation steers restoration of terrestrial ecosystems. *Nature Plants* **2**: 1–5.

Yelenik SG. 2017. Linking dominant Hawaiian tree species to understory development in recovering pastures via impacts on soils and litter. *Restoration Ecology* **25**: 4252.

Zaura E, Keijser BJ, Huse SM, Crielaard W. 2009. Defining the healthy 'core microbiome' of oral microbial communities. *BMC Microbiology* **9**: 259.

Supporting Information

Additional Supporting Information may be found online in the Supporting Information section at the end of the article.

Fig. S1 Map of the Hakalau Forest National Wildlife Refuge in the Hawaiian archipelago and sampling transects within Hakalau.

Fig. S2 Accumulation curves of amplicon sequence variant richness in raw arbuscular mycorrhizal fungal community data from remnant and restored forests.

Fig. S3 Venn diagrams of amplicon sequence variant membership by habitat and sample type.

Fig. S4 Boxplots displaying arbuscular mycorrhizal fungal community richness per sample within remnant and restored forests.

Fig. S5 Dot plots displaying the specialization of arbuscular mycorrhizal fungal and plant host networks between remnant and restored forests.

Fig. S6 Non-metric dimensional scaling ordination plot of arbuscular mycorrhizal fungal community composition between remnant and restored forests.

Fig. S7 Relative abundance of generalist arbuscular mycorrhizal fungal family amplicon sequence variants shared between all hosts and soil in remnant and restored forests.

Fig. S8 Log-log scatter plots of amplicon sequence variants (ASVs) abundance against number of samples for all arbuscular mycorrhizal fungal ASVs in remnant and restored forests.

Notes S1 Additional methodological information regarding sample preparation.

Notes S2 Additional methodological information regarding the nestedness analysis.

Table S1 Differences in the richness of arbuscular mycorrhizal fungal amplicon sequence variants detected in host and soil communities between remnant and restored forests within the Hakalau forest national wildlife refuge.

Table S2 Pairwise comparisons of the richness of arbuscular mycorrhizal fungal amplicon sequence variants between host and soil communities within remnant and restored forests.

Table S3 Differences in plant–arbuscular mycorrhizal fungal network specialization (H2′) between remnant and restored forests.

Table S4 Differences in host interaction specialization (d') on arbuscular mycorrhizal fungi between remnant and restored forests.

Tables S5 Differences in the relative abundance of arbuscular mycorrhizal fungal families shared between soil and individual hosts in remnant and restored forests within the Hakalau forest national wildlife refuge.

Table S6 Differences in the relative abundance of host- and soil-specific arbuscular mycorrhizal fungal families in remnant and restored forests within the Hakalau forest national wildlife refuge.

Table S7 Differences in the relative abundance of arbuscular mycorrhizal fungal families shared between all hosts and soil in remnant and restored forests within the Hakalau forest national wildlife refuge.

Please note: Wiley Blackwell are not responsible for the content or functionality of any Supporting Information supplied by the authors. Any queries (other than missing material) should be directed to the *New Phytologist* Central Office.