Network modules and hubs in plant-root fungal biomes

Hirokazu Toju¹, Satoshi Yamamoto², Akifumi S. Tanabe³, Takashi Hayakawa⁴,⁵ and Hiroshi S. Ishii⁶

¹Graduate School of Human and Environmental Studies, Kyoto University, Saky, Kyoto 606-8501, Japan
²Graduate School of Human Development and Environment, Kobe University, 3-11 Tsunakabuto, Nada-ku, Kobe 657-8501, Japan
³National Research Institute of Fisheries Science, Fisheries Research Agency, 2-12-4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 226-8648, Japan
⁴Department of Wildlife Science (Nagoya Railroad Co., Ltd.), Primate Research Institute, Kyoto University, Inuyama, Aichi 484-8506, Japan
⁵Japan Monkey Centre, Inuyama, Aichi 484-0081, Japan
⁶Department of Environmental Biology and Chemistry, Graduate School of Science and Engineering, University of Toyama, 3190 Gofuku, Toyama 930-8555, Japan

Terrestrial plants host phylogenetically and functionally diverse groups of below-ground microbes, whose community structure controls plant growth/survival in both natural and agricultural ecosystems. Therefore, understanding the processes by which whole root-associated microbiomes are organized is one of the major challenges in ecology and plant science. We here report that diverse root-associated fungi can form highly compartmentalized networks of coexistence within host roots and that the structure of the fungal symbiont communities can be partitioned into semi-discrete types even within a single host plant population. Illumina sequencing of root-associated fungi in a monodominant south beech forest revealed that the network representing symbiont–symbiont co-occurrence patterns was compartmentalized into clear modules, which consisted of diverse functional groups of mycorrhizal and endophytic fungi. Consequently, terminal roots of the plant were colonized by either of the two largest fungal species sets (represented by *Oidiodendron* or *Cenococcum*). Thus, species-rich root microbiomes can have alternative community structures, as recently shown in the relationships between human gut microbiome type (i.e. ‘enterotype’) and host individual health. This study also shows an analytical framework for pinpointing network hubs in symbiont–symbiont networks, leading to the working hypothesis that a small number of microbial species organize the overall root–microbiome dynamics.

1. Introduction

Since their colonization to terrestrial biosphere 470 Ma, land plants have coevolved with diverse mutualistic and pathogenic microbes in soil [1–4]. Mycorrhizal fungi and various lineages of rhizosphere bacteria, for instance, enhance plant nutritional states and/or protect hosts from pathogenic soil microbes [2,5,6]. As plant growth and health is highly dependent on those root-associated microbes, understanding factors determining the structure of plant-root microbiomes is one of the major challenges in ecology and plant science [2,5]. However, the diversity of below-ground fungi and bacteria is enormous [7–9], making it difficult to reveal the key ecological processes that control the entire community structure of root-associated microbes.

Although uncovering the determinants of microbiome structure is difficult not only in interactions between plants and below-ground microbes but also in other host–symbiont systems, recent findings in human gut microbiome studies have revolutionized our views on the formation of microbiomes within/on host organisms [10–12]. Those studies have shown that human individuals are grouped into some major clusters defined by gut bacterial community structure and that...
such ‘enterotypes’ may be organized by facilitative and competitive interactions among microbial symbionts within hosts [10,13,14]. Moreover, an increasing number of studies have revealed close relationships between enterotypes and human health [12], illuminating the importance of symbiont–symbiont interactions in the performance of host individuals [11]. These analytical and conceptual frameworks developed in human enterotype studies are expected to make substantial contributions to plant science. Nonetheless, the existence of classifiable ‘rhizotypes’ [5] of plant-root microbiomes remains to be explored despite its potential importance in the diagnostics and control of root-associated microbial communities.

Here, we show a network depicting symbiont–symbiont co-occurrence patterns in hosts and examine whether discrete sets of symbiont community structures actually exist even within a single population of a single plant species. Among the major groups of below-ground plant–fungus interactions, we focus on ectomycorrhizal symbiosis [2]. Ectomycorrhizal fungi on the same host plant species potentially compete with each other for space and resources, and several pairs of them are known to show segregated (mutually exclusive) distribution patterns across host individuals as expected by competitive exclusion processes [15–17]. On the contrary, pairs of fungi in facilitative interactions, especially those showing functional complementarity, may coexist within the same terminal root tissue, displaying more aggregated patterns than expected by chance [18]. In addition, fungi adapting to the same soil or host physiological environments are expected to show correlated distribution patterns [19]. Therefore, we predicted that such segregated and aggregated patterns were indicative of potential symbiont–symbiont direct interactions and/or correlated environmental adaptation within host root systems and conducted high-throughput DNA barcoding analysis [4,20] to reveal how the network of symbiont–symbiont co-occurrence patterns [10,13] was structured throughout a plant population. Furthermore, to uncover how multiple phylogenetic and functional groups of fungi constitute the entire network, we also took into account fungi belonging to non-ectomycorrhizal lineages. Endophytic fungi, in particular, are conspicuous in below-ground microbiomes have been poorly understood [4,21,22]. Although taxonomic information may help to infer potential ecological roles of each fungus, it is important to acknowledge that fungi detected through high-throughput sequencing can be not only mutualistic, but also commensalistic or antagonistic to their host plants [4].

2. Material and methods

2.1. Sampling

Sampling was conducted in a temperate forest of Fuscoporia cliffortioides (Hook.f.) Heenan & Smissen (Nothofagaceae) [23] in the Queenstown Lakes District, New Zealand (44°26′00″ S, 169°15′40″ E), from 16 to 20 January 2014. As the Fuscoporia species was the only tree species that reached the canopy of the forest, it provided an ideal research system for inferring how symbiont–symbiont interactions were structured in a wild host plant population. Along a 687 m mountain trail, we collected 2 cm segments of terminal root samples at 3 cm below the soil surface at 1 m horizontal intervals. The altitudes of the sampling points varied from 862 m (sample no. 1) to 710 m (sample no. 688). The collected 688 samples were carefully washed to remove adhering soil and immediately dried with ample silica gel.

As DNA-barcoding-based analysis per se does not provide any information on the nature of symbioses between plants and their root-associated fungi, we use the word ‘symbionts’ to refer to observed fungi irrespective of their potential effects on host plants (i.e. ‘symbionts’ in a broad sense) [24]. Although taxonomic information may help to infer potential ecological roles of each fungus, it is important to acknowledge that fungi detected through high-throughput sequencing can be not only mutualistic, but also commensalistic or antagonistic to their host plants [4].

2.2. Molecular analysis

Each of the 688 samples was pulverized with 4 mm zirconium balls using a Tissuelyser II (Qiagen) [22], and host plant and fungal symbiont DNA were simultaneously extracted with the cetyltrimethylammonium bromide method [25]. For the molecular identification of fungal symbionts, the nuclear internal transcribed spacer 1 (ITS1) region of fungi was amplified by polymerase chain reaction (PCR). In the PCR amplification of the fungal ITS region, we used the forward primer ITS1 [26] fused with 6-mer Ns (for improved ‘chastity’ in Illumina sequencing) [27] and the forward Illumina sequencing primer (5'-TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG [sequencing primer]-NNNNNNN [6-mer Ns]-GGA AGT AAA AGT CGT AAC AAG G [ITS5]-3') and the reverse primer ITS2_KYO2 [28] fused with 6-mer Ns and reverse sequencing primer (5'-GTC TCG TGG GCC CGG AGA TGT GGA TAA GAG ACA G AGA C [sequencing primer]-NNNNNNN [6-mer Ns]-TGY RCT RCG TTC TTC ATC [ITS2_KYO2]-3'). The PCR reaction was conducted using the buffer and DNA polymerase system of KOD FX Neo (TOYOBO), which has proofreading ability, with a temperature profile of 94°C for 2 min, followed by 35 cycles at 98°C for 10 s, 50°C for 30 s, 68°C for 50 s and a final extension at 68°C for 5 min. Illumina sequencing adapters were added in the subsequent PCR process using a forward fusion primer consisting of a P5 Illumina adaptor, 8-mer index tags for sample identification [29], the 5'-end of the sequencing adapter (5'-AAT GAT ACG GCC ACC ACC GAT GAC TAC AC [P5 adaptor]-AAAAA-XXXXXXX [8-mer tag]-TCG TCG GCA GCG TC [sequencing primer]-3') and a reverse fusion primer (5'-CAG GAA GAC GGG GTA CAG GAT [P7 adaptor]-AAAAA-XXXXXXX [8-mer tag]-GTC TCG TGG GCC CG [sequencing primer]-3'). The additional PCR was conducted using the KOD FX Neo system with a temperature profile of 94°C for 2 min, followed by eight cycles at 98°C for 10 s, 50°C for 30 s, 68°C for 50 s and a final extension at 68°C for 5 min. We also PCR-amplified plant chloroplast rbcL and trnH-psbA regions to confirm that the sampled roots were those of *F. cliffortioides*. In the first PCR step for the amplification of the two chloroplast regions, we performed a multiplex PCR by mixing equal concentrations of *rbcL* (rbcL_F3 [30] and rbcL_R4 [30]) and *trnH-psbA* (psbA31 [31] and *trnH* [32]) primers. The multiplex PCR products were then subjected to the second PCR step for adding the index and Illumina adaptor regions. For each step, the buffer/polymerase system and thermal-cycle protocols detailed above were applied.

The indexed PCR products of the 688 samples were pooled into a single library after purification with an AMPure XP Kit (Beckman Coulter). The ratio of sample volume to AMPure volume was set to 1 : 0.6 [27] to remove the remaining PCR primers. In the library, the ratio of ITS1 products to rbcL/[trnH-psbA] products was set to 4 : 1. The pooled library was then subjected to an Illumina MiSeq run (run centre: Graduate School of Human and Environmental Studies, Kyoto University (KYOTO-HE)) with the 2 × 300 cycle sequencing kit (20% PhiX spike-in).

2.3. Bioinformatics

The raw MiSeq data were converted into FASTQ files using the BCL2FASTQ program provided by Illumina. The FASTQ files were then demultiplexed using the program CLAIDENT v. 0.2.2015.03.11 [33,34]. To avoid possible errors resulting from low-quality index
sequences, the sequencing reads whose 8-mer index positions included nucleotides with low (less than 30) quality scores were discarded in this process. The forward and reverse sequencing reads were then fused with each other using the program PEAR v. 0.9.6 with a stringent criterion for merging ($p = 0.0001$).

Among the 11 948 484 reads obtained for the ITS1 region, 121 609 were excluded from the subsequent process because their sequences were less than 150 bp or because 10% or more of their nucleotides had low (less than 30) quality values. We also discarded potentially chimeric reads using the programs UCHIME v. 4.2 (de novo mode) [35]. In addition, noisy reads were removed using the approach of Li et al. [36] with CLADIENT, leaving 10 366 999 reads. The remaining reads were clustered with a cut-off sequence similarity of 97% based on a parallelized process of the genome assembler MINIMUS [37], which also enabled highly accurate clustering of PCR-amplified marker regions, as implemented in CLADIENT. The obtained consensus sequences were then used as operational taxonomic units (OTUs) in the subsequent community ecological analyses. In this clustering process, reads of each sample were clustered beforehand with a cut-off sequence similarity of 98%: the clustered-read membership of the within-sample clustering was used as guide information in order only to accelerate the 97% clustering process. Among the OTUs obtained, we excluded those whose sequencing reads were less than 10 [38] in all samples because their sequences were likely to contain PCR/sequencing errors. After this process, the number of remaining OTUs was 2886. For each of the obtained OTUs, taxonomic identification was conducted based on the query-centric auto-k-nearest-neighbour (QCAUTO) method [34] and subsequent taxonomic assignment with the lowest common ancestor (LCA) algorithm [39] using CLADIENT. A benchmark analysis has shown that the combination of the QCAUTO and LCA algorithms returns the most accurate taxonomic identification results among the existing methods of automated DNA barcoding [34]. Also, importantly, the QCAUTO method is applicable to the DNA barcoding of not only eukaryotic fungi but also diverse clades of endophytic fungi [22]. The QCAUTO taxonomic assignment was applied to our OTU dataset using the databases obtained by filtering out unreliable sequence entries from the National Center for Biotechnology Information (NCBI) ‘nt’ database (downloaded from ftp://ftp.ncbi.nih.gov/ on 27 January 2015) [34]. Among the filtered databases bundled with CLADIENT, we used the ‘semiall_genus’ database, from which Caenorhabditis, Drosophila and vertebrate sequences as well as sequences lacking genus-level taxonomic information were discarded [33]. The QCAUTO query search results with the database were then subjected to the LCA taxonomic assignment (LCA/genus). The default LCA process is very stringent and conservative in that it assigns taxonomic information at a given rank only when the information from all neighbourhood sequences is consistent with each other. Therefore, an additional taxonomic assignment was performed by tolerating 5% mismatches among neighbourhood sequences (relaxed-LCA/genus) [22]. To facilitate order-level taxonomic identification, we also conducted a QCAUTO search based on the ‘semiall_order’ filtered database, from which sequences lacking order-level taxonomic information were excluded, and we then applied the relaxed-LCA assignment to the search results (relaxed-LCA/order). The overall taxonomic assignment results were obtained by merging the LCA/genus, relaxed-LCA/genus and relaxed-LCA/order results in this priority order, i.e. results with less stringent settings were not used if they contradicted those with stringent settings [22]. To confirm the results with the QCAUTO–LCA process, we also performed taxonomic assignment with the UCLUST approach [40] using the UNITE v. 7 dynamic database [41] as implemented in QIME [42].

Based on the QCAUTO–LCA taxonomic assignment results, 965 non-fungal OTUs were excluded from the dataset. We then obtained a sample (row) × fungal OTU (column) data matrix, in which a cell entry indicated the number of the reads of each OTU in each sample. In the matrix, cell entries whose reads were less than 1% of the total read count of each sample were excluded. The final filtered electronic supplementary material, figure S1) because those rare entries could represent contamination from soil or among-sample contamination due to ‘mis-tagging’ [43]. The data matrix (‘592-OTU.matrix’ in the electronic supplementary material, data S1) was then rarefied to 1000 reads per sample (‘Rarefied.matrix’ in the electronic supplementary material, data S1) using the VEGAN v. 2.2-1 package of R v. 3.2.0 (electronic supplementary material, figure S1). Eight hundred and twelve and 24 rare OTUs were discarded in the filtering and rarefaction processes, respectively.

To exclude non-Fuscospora root samples from the dataset, the plant rbcL and trnH-psbA read data were, respectively, clustered with a cut-off sequence similarity of 99.8%. Ten root samples, which turned out to be the roots of non-Fuscospora plants, were then excluded from the dataset. Overall, we obtained a data matrix including 620 root samples and 592 fungal OTUs (electronic supplementary material, data S1 and S2): 58 samples from which the number of sequencing reads was less than 1000 were discarded in the above-mentioned processes. Hereafter, we use the word ‘species’ instead of ‘OTUs’ for simplicity, paying careful attention to the fact that OTUs defined with a fixed sequence similarity value do not necessarily represent fungal species. On average, each root sample was colonized by 11.1 fungal species (s.d. = 3.7; electronic supplementary material, figure S1).

### 2.4. Symbiont–symbiont network

To reveal the structure of the symbiont–symbiont co-occurrence network, we evaluated the extent of the aggregation of fungal symbionts within plant-root samples. For each pair of fungal species, we first calculated the togetherness score ($T$-score) [44], which was defined as follows:

$$T = \frac{N + S - R_i - R_j}{\sqrt{N(S - 1)}}$$

where $N$ is the total number of root samples examined, $R_i$ and $R_j$ are the total number of the occurrences (root sample counts) of species $i$ and $j$, and $S$ is the number of co-occurrences of species $i$ and $j$. By using the togetherness score, we performed a randomization test to evaluate the extent of aggregation for each pair of fungal species. In the randomization analysis for each pair of fungal species, the entry of one species was randomized across root samples (100 000 permutations). To evaluate how the observed togetherness deviated from randomized ones, we calculated standardized togetherness as follows:

$$\text{standardized togetherness} = \frac{T_{\text{observed}} - \text{mean}(T_{\text{randomized}})}{\text{s.d.}(T_{\text{randomized}})}$$

where $T_{\text{observed}}$ is the togetherness of the original data, and mean($T_{\text{randomized}}$) and s.d.($T_{\text{randomized}}$) were the mean and standard deviation of the togetherness scores of randomized data, respectively. In the togetherness analysis, we used the data of the 52 fungal species that occurred in 30 or more root samples (electronic supplementary material, data S3). The results for 1326 fungal species pairs were subjected to multiple comparison analysis based on the false discovery rate (FDR) [45]. We then drew a symbiont–symbiont co-occurrence network by compiling linking between pairs of fungal species that displayed statistically significant (FDR < 0.05) signs of aggregation (togetherness) (electronic supplementary material, data S3). Fungal species within the network were placed using the FORCEATLAS2 algorithm [46].

We also evaluated how pairs of fungal species showed mutually segregated distribution across root samples using the checkerboard score (C-score) [44], which was calculated as follows:

$$C = (R_i - S) \times (R_j - S).$$
For each of the 1326 fungal species pairs, a randomization analysis of checkerboard scores was conducted (100 000 permutations). Pairs of fungal species with statistically significant (FDR, 0.05) signs of segregation were then indicated on the above-mentioned co-occurrence network.

In addition to the togetherness and checkerboard score analyses for the presence/absence dataset format, we also performed analyses of possible symbiont–symbiont associations based on two methods using sequencing-read count information. One used the information of compositional correlations between pairs of species (the sparse correlations for compositional data (SparCC) method [47]) and the other was based on the concept of ‘conditional independence’ between pairs of species (the sparse inverse covariance estimation for ecological association inference (SPIEC-EASI) method [48]). In the SparCC analysis, the threshold of absolute correlation coefficients was set to 0.3 as in a benchmark study comparing the SparCC and SPIEC-EASI approaches [48]. In the SPIEC-EASI analysis, the Meinshausen and Bühlmann (MB) algorithm [49] was applied. As these composition-based methods are usually applied to data matrices without rare species [47,48], the 52 fungal species analysed in the togetherness/checkerboard tests were screened from the original data matrix (‘592-OTU.matrix’ in the electronic supplementary material, data S1). We also screened samples with sufficient compositional (read-count) information by removing those with less than 5000 sequencing reads. As a result, the input data matrix for the SparCC and SPIEC-EASI analyses consisted of 277 samples and the 52 fungal species (‘SparCC. SPIEC-EASI.-matrix’ in the electronic supplementary material, data S1). In general, analyses based on sequencing-read counts deserve utmost care because they can be more vulnerable to biases resulting from interspecific variation in the number of ribosomal DNA

Figure 1. Symbiont–symbiont co-occurrence network. (a) Scores representing the extent of aggregation of fungal symbionts within host root samples. For each pair of fungal species, a togetherness score was examined in a randomization analysis to evaluate aggregated distribution (100 000 permutations). Multiple comparisons were performed based on FDR. (b) Scores representing the extent of segregation of fungal symbionts within host root samples. For each pair of fungal species, a checkerboard score was examined in a randomization analysis to evaluate segregated distribution (100 000 permutations). (c) Network of aggregated and segregated patterns. Fungal species are linked by the lines indicating statistically significant (FDR < 0.05) aggregation (blue) and segregation (red). The thickness of links is proportional to the standardized togetherness or checkerboard scores. The circles representing fungal species (yellow, ectomycorrhizal fungi; grey, fungi with unknown functions) are placed based on the aggregation patterns with the FORCEATLAS2 algorithm. The outer parts of the circles represent fungal taxonomy (brown, Ascomycota; green, Basidiomycota; white, unidentified).
tandem repeats and compositional biases introduced in PCR-amplification processes than analyses based on presence/absence information [50]. However, because relative abundance information was used across samples as detailed previously [48], our SparCC and SPIEC-EASI results were unlikely to be affected greatly by those potential read-count biases.

2.5. Symbiont modules
We examined how the symbiont–symbiont co-occurrence network was partitioned into the modules of frequently coexisting fungal species. Modules were detected based on a ‘data-compression-based’ approach using the INFOMAP algorithm [51], which was known to find network modules the most accurately among available methods [52]. The fungal species composition of each module was inferred based on consensus [53] over 1000 INFOMAP runs with the default setting.

By focusing on pairs of fungal species belonging to different modules, we evaluated the relationships among the detected modules. Specifically, the ratio of significant aggregation links to possible symbiont–symbiont combinations was calculated as follows:

$$\text{ratio of among-module aggregation} = \frac{\text{S}_{ij}}{N_i \times N_j},$$

where $S_{ij}$ denotes the number of statistically significant (FDR < 0.05) aggregations (togetherness scores) between fungal species in modules $i$ and $j$, and $N_i$ and $N_j$ represent the number of fungal species in modules $i$ and $j$, respectively. The ratio of among-module segregation was also calculated in the same way based on the analysis of checkerboard scores.

2.6. Clustering analysis of root sample
In light of the statistical method used in the ‘enterotyping’ of the human gut microbiome [10], we conducted the clustering of fungal species compositions of the root samples. For each pair of the root samples, the Bray–Curtis β-diversity of the fungal species composition was calculated (electronic supplementary material, method S1). Plant root samples were then partitioned into clusters in terms of their fungal species compositions based on the partitioning around medoids algorithm of clustering for a given number of clusters [10]. Based on the results with various a priori cluster numbers, the optimal number of clusters was estimated with the Calinski–Harabasz index [54]. Non-metric multidimensional scaling (NMDS) was then performed to visualize the inferred clusters. In the clustering and NMDS visualization, the VEGAN, CLUSTER v. 2.0.1 and CLUSTERSIM v. 0.44.2 packages of R were used.

2.7. Network hubs
To evaluate the topological properties of each fungal species within the symbiont–symbiont co-occurrence network, we calculated betweenness [55, 56] centrality. Fungal species with high betweenness are expected to play important ‘topological roles’ in interconnecting pairs of other fungal species in the symbiont–symbiont co-occurrence network [55, 56]. The obtained betweenness values were z-standardized (zero-mean; unit-variance). In addition to the betweenness analysis, topological roles in interconnecting species in different modules (partition coefficient [55, 57]) and the number of links with species in the same module (within-module degree) were calculated. The former can vary from 0 (species linked only with species in the same modules) to 1 (species interacting indiscriminately with species in all modules), while the latter was z-standardized.

2.8. Spatial scales of sampling
Because the roots analysed were collected randomly at 1 m intervals within the forest, our samples as a whole may have included those from the same Fuscospora individuals. Thus, we conducted an additional analysis in which each root sample was expected to represent a plant individual. As sampling was conducted in a mature forest with a closed canopy, roots collected at 5 m intervals were possibly those of different host plant individuals. Therefore, we divided the 1 m interval full data into five partial datasets, each of which consisted of the root samples collected at 5 m intervals (electronic supplementary material, data S4). For each of the five partial datasets, the randomization analysis of the togetherness and checkerboard scores was performed for each pair of fungal species. Fungal species that occurred in 10 or more root samples in each partial dataset were subjected to the analysis.

3. Results
3.1. Architecture of the symbiont–symbiont co-occurrence network
The symbiont–symbiont co-occurrence network in the Fuscospora forest displayed a highly organized structure in terms of the sets of fungal species that frequently coexisted.
Table 1. Major fungal species in the two module groups. For each module group A (module 1) and B (modules 2–5) (figure 2), the top 10 fungal species with the highest sample counts (the number of root samples) are shown. The information for the lowest taxonomic rank assigned by the UCLUST algorithm with the UNITE v. 7 dynamic database is also shown with the results from the QC un-LCA approach.

<table>
<thead>
<tr>
<th>OTU</th>
<th>module</th>
<th>N samples</th>
<th>phylum</th>
<th>class</th>
<th>order</th>
<th>family</th>
<th>genus</th>
<th>functional group</th>
<th>UNITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>F3</td>
<td>A (1)</td>
<td>245</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>Dermateaceae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F4</td>
<td>A (1)</td>
<td>237</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Myxotrichaceae</td>
<td>Oidiodendron</td>
<td>unknown</td>
<td>Oidiodendron</td>
<td></td>
</tr>
<tr>
<td>F5</td>
<td>A (1)</td>
<td>235</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Myxotrichaceae</td>
<td>Oidiodendron</td>
<td>unknown</td>
<td>Oidiodendron</td>
<td></td>
</tr>
<tr>
<td>F6</td>
<td>A (1)</td>
<td>222</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Myxotrichaceae</td>
<td>Oidiodendron</td>
<td>unknown</td>
<td>Oidiodendron</td>
<td></td>
</tr>
<tr>
<td>F7</td>
<td>A (1)</td>
<td>210</td>
<td>Basidiomycota</td>
<td>Agaricomycetes</td>
<td>Agaricales</td>
<td>Cortinarius</td>
<td>unknown</td>
<td>ectomycorrhizal</td>
<td>Cortinarius</td>
</tr>
<tr>
<td>F8</td>
<td>A (1)</td>
<td>207</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>Mortierellales</td>
<td>Mortierella</td>
<td></td>
</tr>
<tr>
<td>F9</td>
<td>A (1)</td>
<td>161</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F10</td>
<td>B (2)</td>
<td>270</td>
<td>Ascomycota</td>
<td>Dothideomycetes</td>
<td>Glomiaceae</td>
<td>Genococum</td>
<td>unknown</td>
<td>ectomycorrhizal</td>
<td>Genococum</td>
</tr>
<tr>
<td>F11</td>
<td>B (3)</td>
<td>159</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F12</td>
<td>B (3)</td>
<td>117</td>
<td>Ascomycota</td>
<td>Eurotiomycetes</td>
<td>Chaetothyriales</td>
<td>Herpotrichiellaceae</td>
<td>Cladophialophora</td>
<td>unknown</td>
<td>Cladophialophora</td>
</tr>
<tr>
<td>F13</td>
<td>B (2)</td>
<td>114</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>Dermateaceae</td>
<td>unknown</td>
<td>unknown</td>
<td>Dermateaceae</td>
</tr>
<tr>
<td>F14</td>
<td>B (2)</td>
<td>98</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F15</td>
<td>B (1)</td>
<td>114</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F16</td>
<td>B (1)</td>
<td>87</td>
<td>Basidiomycota</td>
<td>Agaricomycetes</td>
<td>Agaricales</td>
<td>Cortinarius</td>
<td>unknown</td>
<td>unknown</td>
<td>Cortinarius</td>
</tr>
<tr>
<td>F17</td>
<td>B (5)</td>
<td>77</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F18</td>
<td>B (2)</td>
<td>64</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F19</td>
<td>B (2)</td>
<td>63</td>
<td>Ascomycota</td>
<td>Leotiomycetes</td>
<td>Helotiales</td>
<td>unknown</td>
<td>unknown</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
within the narrow space of host root systems (figure 1). The network representing statistically significant aggregation patterns was partitioned into five modules (excluding modules containing only one species) and each of the modules included fungi in phylogenetically diverse lineages (figures 1 and 2; see also electronic supplementary material, figure S2). A complementary network analysis based on checkerboard scores further indicated that fungi in different network modules often showed segregated patterns (figure 1c). In particular, fungi in module 1 (module group A) seldom co-occurred with those in modules 2–5 (module group B), while fungi in the latter three modules frequently coexisted within host root systems (figures 1 and 2).

There were some characteristics in the taxonomic compositions of the module groups (table 1; electronic supplementary material, table S1). First, both module groups included ectomycorrhizal fungi in Cortinariaceae as well as fungi in the ascomycete order Helotiales, which were known to include endophytic and ectomycorrhizal lineages [58] (table 1). Second, other than Helotiales fungi, the module group A was represented by fungi in the genus Oidiodendron, while the module group B was dominated by a fungus in the ectomycorrhizal genus Cenococcum (table 1). Third, whereas some Oidiodendron fungi were included not only in module group A but also in module group B, Cenococcum appeared only in module group B (electronic supplementary material, table S1).

Additional analyses based on sequencing-read count information (the SparCC and SPIEC-EASI analyses) further indicated the existence of those modules or module groups (figure 3). Meanwhile, the number of links connecting fungal species was fewer in the SparCC/SPIEC-EASI analyses than in the togetherness/checkerboard analyses (cf. figures 1 and 3). As a result, 14 of the 52 fungal species examined did not have links, and the module or module groups (figure 1) appeared as discrete clusters (figure 3a,b).

3.2. Clustering of fungal symbiont communities
The characteristic structure of the symbiont–symbiont co-occurrence network was reflected in the formation of fungal community type in the Fuscospora host plant. That is, the fungal symbiont composition of terminal root samples in the forest was partitioned into two semi-discrete statistical clusters (figure 4a,b; electronic supplementary material, figure S3). The two clusters corresponded to the compartmentalized pattern of the symbiont–symbiont network, i.e. one cluster consisted of root samples frequently colonized...
by fungi in module group A, while the other represented samples harbouring fungi in module group B at high frequency (figure 4c). Although a small fraction of samples hosted both module groups of fungi at comparative proportions, the fungal composition of most root samples was biased towards colonization by either of the fungal module groups (figure 4d). An additional analysis showed that there was a spatially auto-correlated pattern in the distribution of fungal community clusters within the forest (figure 4e; see also electronic supplementary material, figure S4).

3.3. Network hubs within the symbiont–symbiont co-occurrence network

We then focused on how each fungal species was embedded within the symbiont–symbiont co-occurrence network and found that several fungal species in the community were placed at the core of the network (figure 5). Some of those ‘network hub’ [55,59] species interconnected fungi in different modules within the symbiont–symbiont co-occurrence network (e.g. an endophytic fungus in Herpotrichiellaceae (‘23_Herpotrichiella’)), while others interconnected most fungal species within each module (e.g. an ectomycorrhizal fungus in the genus Cenococcum (‘2_Cenococcum’)) (figures 1 and 5a,b). Although generalist fungi that occurred in most samples could be the former type of network hubs (hereafter, ‘inter-module hubs’), the most frequently observed fungi (fungi observed from more than 200 samples) within the dataset (figure 5c) had the latter type of topological characteristics (hereafter, ‘within-module hubs’) (figure 5i). When the sample counts for each fungal species (i.e. the number of root samples from which each species was detected; figure 5c) was controlled, inter-module hubs were distinguished from within-module hubs as well as peripheral (rarer) species in the network (figure 5d).

3.4. Spatial scales of sampling

In the analysis based on the 5 m interval partial datasets, the number of fungal pairs that displayed statistically significant (FDR < 0.05) signs of aggregation/segregation was inevitably reduced due to the decreased sample size in the partial datasets (electronic supplementary material, figure S5 and data S4). However, many of the core symbiont–symbiont aggregation/segregation patterns found in the full-data analysis (figure 1) were reproduced in the additional analysis (electronic supplementary material, figure S5 and data S4), although care should be paid to the possibility that the 5 m interval partial datasets could still include some samples from the same host plant individuals.

4. Discussion

There are some potential mechanisms that can generate the observed differentiation of fungal symbiont compositions among host plant samples. For example, fungi in each module group (figure 2a) may share ecological niches [19], adapting to the same fine-scale environments in soil [38]. The spatial autocorrelation observed in the distribution of fungal community clusters within the forest (figure 4e) might reflect the suspected effects of such environmental factors.

Another important possibility, albeit not mutually exclusive with the former one, is that the observed semi-discrete community structures are organized mainly by direct symbiont–symbiont interactions. There has been clear experimental evidence that ectomycorrhizal fungal species compete for space within host root systems and that they strongly prevent the colonization of late comers through ‘priority effects’ [15,16,60]. Such competitive exclusion mechanisms have been reported not only between ectomycorrhizal
fungi but also between arbuscular mycorrhizal fungi [60]. In contrast to those negative interactions between fungal symbionts, pairs of fungi in facilitative interactions, especially those showing functional complementarity, are expected to coexist within the same terminal root tissue, displaying more aggregated patterns than expected by chance [18]. In this respect, the result that each module group included both ectomycorrhizal and endophytic fungi (figures 1c and 2) is interesting. This study was designed to screen for the signs of potential interactions between symbionts and revealed how diverse phylogenetic and functional groups of fungi constitute modules in a symbiont–symbiont co-occurrence network. Although the relative contributions of soil–environmental niche partitioning/sharing and direct interspecific interactions to the observed community patterns should be examined in future experimental studies, the analytical framework shown here provides a basis for understanding the mechanisms by which (semi-)discrete symbiont community structures are organized at the network level.

In general, the presence of alternative community compositions is represented by the term ‘alternative stable state’ [61,62]. Conceptually, there are two different contexts defining alternative stable states [63]. In one definition, shifts between alternative community structures occur in response to changes in state variables (e.g. population densities of respective species) [64,65], while, in the other definition, they occur as a consequence of changes in environmental parameters (e.g. host nutritional conditions) [66,67]. Although the former definition is frequently used in recent studies of community ecology [65], the latter definition would attract more attention in the context of applied microbiology, whose focus is on the possible relationships between microbiome structure and host physiological states [5,11,12,68]. As symbiont community compositions can not only be the indicators of hosts’ physiological states but also the determinants of hosts’ health [14,69,70], it is essential to investigate whether alternative structures of root-associated fungal communities are equal or different in their effects on plants’ physiology and performance.
The observed difference in taxonomic compositions between module groups A and B (table 1) is of particular interest for this point. Although both module groups included Cortinariaceae and Helotiaceae fungi as major components, module group A was represented by *Oidiodendron* fungi, which are known to be saprotrophic or ericoid mycorrhizal fungi [71]. In contrast, module group B was dominated by a fungus in a well-characterized ectomycorrhizal genus, *Cenococcum*, which surrounds host root-tips with heavily melanized mycelia [72] and produces antibiotics against pathogenic bacteria [73]. Given the ambiguous symbiotic status of *Oidiodendron* and the unique ectomycorrhizal feature of *Cenococcum* [74], the two fungal module groups observed in this study (figure 2 and table 1) may be playing distinct ecological roles in the *C. cliffortioides* population.

Another future research direction is illuminated by the working hypothesis that a small fraction of symbiont species can play essential roles in the assembly of plant-root microbiomes. The existence of topological hubs in symbiont–symbiont networks leads to the hypothesis that a small fraction of microbes play predominant roles in the organization of symbiont community structure (or rhizotype). Specifically, the presence of within-module hub species may facilitate the subsequent root colonization of other mycorrhizal, endophytic and pathogenic fungal species belonging to the same modules or module groups, while it may prevent the colonization of fungi in other module groups (see studies examining possible fungus-to-fungus interactions within host root systems [15,60,75,76]). Accordingly, the rhizotype of the plant root system might be determined, in large part, by which hub species first colonize the root tissue [60,77]. Given that potential within-module hubs had the highest sample counts (i.e. the number of samples from which they were observed) in our data (figure 5c), they may actually colonize host tissue earlier than others, organizing the microbiome structure within the hosts through priority effects. Meanwhile, inter-module hubs (figure 5d), albeit absent in the analyses based on sequencing-read count data (figure 3), may also play important roles in, for instance, the switching of alternative rhizotypes. However, our knowledge of such shifts among alternative symbiont community structures is still limited.

Although the observed patterns in the symbiont–symbiont co-occurrence network allow us to raise some intriguing hypotheses on microbiome assembly processes, our results are based on an analysis of only one monodominant forest, thereby providing limited chances for extrapolating the above discussion to other forest, grassland or agricultural ecosystems. In addition, the fully observational approach of our study precludes explicit testing of the existence of alternative stable states (or rhizotypes) and possible mechanisms underlying within-host dynamics of fungal symbiont communities. Also, importantly, the use of molecular OTUs as units of statistical analysis has been subject to continuing methodological challenges in microbiology [4,78]. Nonetheless, we herein have shown how to reconstruct the networks of potential symbiont-to-symbiont interactions based on field sampling and high-throughput sequencing. Moreover, the working hypothesis that a small number of ‘fixer’ species within a symbiont–symbiont co-occurrence network regulate within-host microbial communities deserves attention in both basic and applied ecology, providing a basis for future experimental and theoretical studies.

Overall, analysis of symbiont–symbiont networks is crucial in finding hub species, whose compatibility with plant genotypes and physiological conditions is likely to be the key to understanding the mechanisms that organize symbiont community structures. Thus, even the virtually complex dynamics of communities involving hundreds or more of root-associated microbial species may be reduced to the genetics or ecology of those hub species [79], if the architecture of symbiont–symbiont networks is properly estimated. Specifically, we may be able to manipulate plant-associated microbial communities by inoculating plant seedlings with hub microbial species or optimizing genetic compatibility between host plants and those hub microbes. More observational and experimental studies targeting other microbial groups (e.g. bacteria [80]) in various ecosystems are awaited to address the validity of such a reductionistic control of plant-associated microbiomes.

Data accessibility. The accession number of the DDBJ Sequence Read Archive: DRA003730. The data matrices supporting this article have been uploaded as part of the electronic supplementary material.

Authors’ contributions. H.T. designed the research. H.T. and H.S.I. performed the fieldwork. H.T., S.Y. and T.H. conducted the molecular experiments. H.T. and A.S.T. analysed the data. H.T. wrote the paper.

Competing interests. We declare that we have no competing interests.

Funding. This work was financially supported by a JSPS KAKENHI grant (no. 26711026) and the Funding Program for Next Generation World-Leading Researchers of Cabinet Office, the Government of Japan (GS014) to H.T. H.S.I. was also supported by a JSPS KAKENHI grant (no. 25409006).

Acknowledgements. We thank Yukuto Sato and Atsushi Nagano for their advice on next-generation sequencing and Minato Kodama for her support in molecular experiments. We are also grateful to three anonymous reviewers for their productive comments that improved the manuscript.

References


15. Kennedy PG, Peay KG, Burns TD. 2009 Root tip competition among ectomycorrhizal fungi: are priority effects a rule or an exception? Ecology 90, 2098–2107. (doi:10.1890/08-1291.1)


