

Spatial structure of fungal DNA assemblages revealed with eDNA metabarcoding in a forest river network in western Japan

Shunsuke Matsuoka¹, Yoriko Sugiyama², Hirotooshi Sato², Izumi Katano^{3,4}, Ken Harada⁵, Hideyuki Doi¹

1 Graduate School of Simulation Studies, University of Hyogo, 7-1-28 Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan.

2 Graduate School of Human and Environmental Studies, Kyoto University, Yoshida-nihonmatsu-cho, Sakyo-ku, Kyoto 606-8501, Japan.

3 Graduate School of Humanities and Sciences, Nara Women's University, Kitanoya-nishimachi, Nara 630-8506, Japan.

4 KYOSEI Science Center for Life and Nature, Nara Women's University, Kitanoya-nishimachi, Nara 630-8506, Japan.

5 School of Human Science and Environment, University of Hyogo, Shinzaike-Honcho, Himeji 670-0092, Japan.

Corresponding author: *Shunsuke Matsuoka* (code_matuoka@hotmail.com)

Academic editor: *Sofia Duarte* | Received 20 May 2019 | Accepted 21 June 2019 | Published 10 June 2019

Abstract

Growing evidence has revealed high diversity and spatial heterogeneity of fungal communities in local habitats of terrestrial ecosystems. Recently, the analysis of environmental DNA has been undertaken to study the biodiversity of organisms, such as animals and plants, in both aquatic and terrestrial habitats. In the present study, we investigated fungal DNA assemblages and their spatial structure using environmental DNA metabarcoding targeting the internal transcribed spacer 1 (ITS1) region of the rRNA gene cluster in habitats across different branches of rivers in forest landscapes. A total of 1,956 operational taxonomic units (OTUs) were detected. Of these, 770 were assigned as Ascomycota, 177 as Basidiomycota, and 38 as Chytridiomycota. The river water was found to contain functionally diverse OTUs of both aquatic and terrestrial fungi, such as plant decomposers and mycorrhizal fungi. These fungal DNA assemblages were more similar within, rather than between, river branches. In addition, the assemblages were more similar between spatially closer branches. This spatial structuring was significantly associated with geographic distances but not with vegetation of the catchment area and the elevation at the sampling points. Our results imply that information on the terrestrial and aquatic fungal compositions of watersheds, and therefore their spatial structure, can be obtained by investigating the fungal DNA assemblages in river water.

Key Words

diversity, environmental DNA, forest river, fresh water, fungi, ITS, metabarcoding, spatial structure

Introduction

Understanding patterns of biodiversity is fundamental for the maintenance of ecological processes during this era of environmental change (Margules and Pressey 2000; Pecl et al. 2017). The kingdom Fungi encompasses high species diversity and performs fundamental ecological roles in both terrestrial and aquatic ecosystems via the decomposition of organic substrates and symbiotic or parasitic interactions with other organisms, such as animals and plants

(Peay et al. 2016; Hawksworth and Lücking 2017; Grosart et al. 2019). Recently, with the growing use of DNA metabarcoding techniques, an increasing number of studies are investigating fungal diversity on a variety of substrates in terrestrial and aquatic habitats (Duarte et al. 2015; Peay et al. 2016; Nilsson et al. 2019). However, as fungi show high local diversity and spatial heterogeneity (Bahram et al. 2015, 2016), the investigation of fungal diversity over large spatial scales, such as landscape (tens of km) and regional scales (hundreds to thousands of km) requires a significant amount of sampling effort, time for analysis, and costs.

Recently, the availability of environmental DNA in river water for biodiversity exploration has been discussed in several studies (Deiner et al. 2016; Nakagawa et al. 2018). Deiner et al. (2016) suggest that river water can integrate DNA from the catchment area (~tens of km) and work as conveyor belt of diversity information for various eukaryotes in both aquatic and terrestrial habitats. Aquatic fungi live in river water. They have diverse lineages and functional groups, including saprophytes which degrade plant substrates (e.g., fallen leaves and wood) and fungi that are parasitic toward planktons (Voronin 2014; Raja et al. 2018; Grossart et al. 2019). Aquatic fungi reach substrates by the movement of propagules like spores along the river flow. In addition to these aquatic fungi, it is expected that the spores and mycelia of terrestrial fungi enter the rivers allowing for the detection of the terrestrial fungi (Voronin 2014). Indeed, recent studies have detected the DNA of supposedly not only aquatic fungi but also terrestrial fungi from river water (Deiner et al. 2016; LeBrun et al. 2018), indicating that the diversity information of fungal diversity in a catchment area can be accessed by examining fungal DNA assemblages in river water. In this scenario, even investigators not familiar with fungi could infer the fungal diversity of a catchment area from river water samples. By using this method, the fungal diversity information of any habitat type can be obtained in less time and at a lower cost than through traditional methods.

Rivers typically consist of a network of multiple branches that originate from the main stream. Both aquatic and terrestrial fungal communities have spatial structures wherein geographically closer sites share similar fungal communities which reflects the surrounding environment and dispersal ability of each species (Talbot et al. 2014; Matsuoka et al. 2016b; Peay et al. 2016; Duarte et al. 2017). Assuming that fungal DNA assemblages in river water reflect the fungal community composition of surrounding land, the assemblages should also indicate the spatial structure. Rivers have a directional water flow, from upstream to downstream. Therefore, the DNA assemblages of river water could also show river-specific fungal assemblages. Samples from the same branch may be more similar to one another than those from different branches. Information regarding fungal DNA assemblages in river water and their spatial structure is essential for investigations into the broad spatial diversity of fungi. However, no previous studies have explored fungal DNA assemblages in river water focusing on the spatial structure.

In the present study, we aim to characterize fungal DNA assemblages in river water across branches over forest landscapes and clarify their spatial structure. In particular, we analyze (1) whether fungal DNA assemblages are more similar within, rather than between, river branches, and (2) whether fungal DNA assemblages show a spatial structure that reflects geographical distance or environmental factors.

Methods

Study sites and sampling

In the present study, the same water samples as in Katano et al. (2017) were used. Study sites are located in the headwater tributaries of the Ibo River in Hyogo, Japan (Fig. 1). Seven sites are located in the Akazai River (A1–A7) and 12 sites in the tributaries of the Ibo River (I1–I12). The orders of the streams and width of the tributaries are 1–2 and 3–5 m, respectively. A1–A7, I7 and I9, and I8 and I9 are located on the same tributary. At each sampling site, longitude, latitude, and altitude were recorded using a GPS device (Garmin etrex 30×).

A 1 L sample was collected from the surface of the water at the central part of each stream in the study sites using a DNA-free polypropylene bottle. Samples were stored in a cooler box with a blank. The blank contained 1 L distilled water which was brought to the field and treated identically to the other water sampling bottles, except that it was not opened at the field sites. The sampling was conducted in late September, 2015. The collected water samples and blank were vacuum-filtered through 47 mm GF/F glass filters (pore size 0.7 µm, GE Healthcare, Little Chalfont, UK). The filter was then wrapped in commercial aluminum foil and stored at -20 °C before DNA extraction.

Molecular analyses

The protocols described by Uchii et al. (2016) were followed to extract DNA from the filters. In brief, each filter was incubated in 400 µl of Buffer AL (Qiagen, Hilden, Germany) and 40 µl of Proteinase K (Qiagen, Hilden, Germany). Then, 220 µl of TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH: 8.0) was added to the filter and centrifuged. The dissolved DNA in the eluted solution was purified using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol.

For MiSeq sequencing, the fungal internal transcribed spacer 1 (ITS1) region of rDNA was amplified (Schoch et al. 2012). The methods of DNA analysis followed those described by Ushio et al. (2017) with some modifications, such as number of PCR cycles. The first-round PCR (first PCR) amplified the ITS1 region using the ITS1-F-KYO2 (5'- TAG AGG AAG TAA AAG TCG TAA -3') and ITS2-KYO2 (5'- TAG AGG AAG TAA AAG TCG TAA -3') primer set (Toju et al. 2012) fused with an Illumina sequencing primer and six random bases (N). The 1st PCR was performed in a 12 µl volume with the buffer system of KODFX NEO (TOYOBO, Osaka, Japan), which contained 2.0 µl of template DNA, 0.2 µl of KOD FX NEO, 6.0 µl of 2× buffer, 2.4 µl of dNTP, and 0.7 µl each of the two primers (5 µM). The PCR conditions were as follows; an initial incubation for 2 min at 94 °C followed by 5 cycles of 10 s at 98 °C, 30 s at 68 °C for annealing and 30 s at 68 °C, 5 cycles of 10 s at 98 °C, 30 s at 65 °C and 30 s at 68 °C; 5 cycles of 10 s at 98 °C, 30 s at

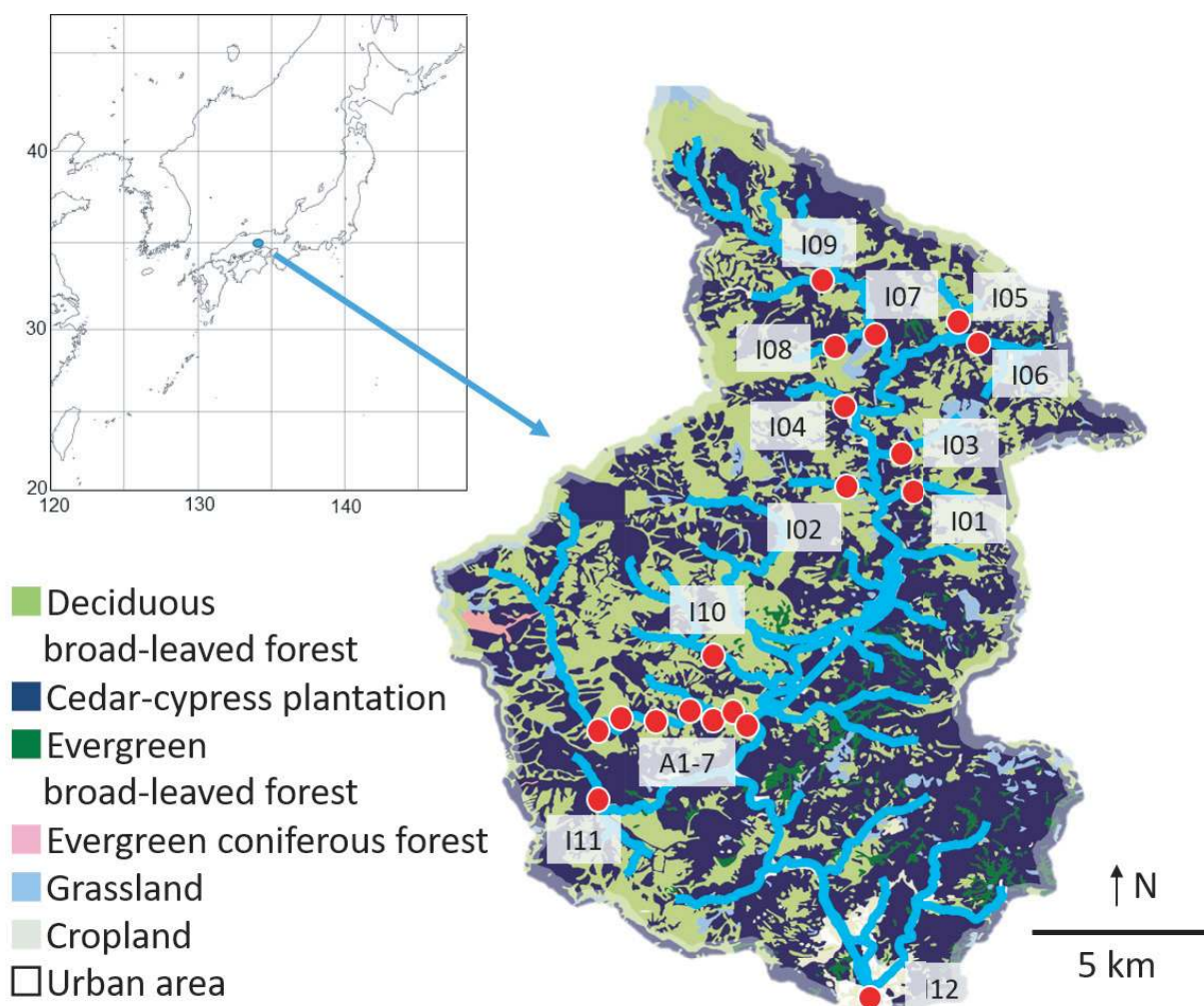


Figure 1. Sampling sites and land use of the catchment area.

62 °C and 30 s at 68 °C; 25 cycles of 10 s at 98 °C, 30 s at 59 °C and 30 s at 68 °C, and a final extension of 5 min at 68 °C. Eight replicate first-PCRs (per sample) were performed to mitigate the reaction-level PCR bias. Then, the duplicated first PCR amplicons (per sample) were combined, resulting in a template per sample for the second PCR. The PCR templates were purified using Agencourt AMPure XP (Beckman Coulter, Brea, California, USA) before the second PCR. The second PCR amplified the first PCR amplicons to add the MiSeqP5/P7 adapter and the 8X bases index sequences (Hamady et al. 2008) using the sequence primers. The second PCR was carried out with 12 µl reaction volume containing 1.0 µl of template, 6 µl of 2× KAPA HiFi HotStart ReadyMix (KAPA Biosystems, Wilmington, Washington, USA), 1.4 µl of each primer (2.5 µM), and 2.2 µl of sterilized distilled water. The PCR conditions were as follows; an initial incubation for 3 min at 95 °C followed by 12 cycles of 20 s at 98 °C, 15 s at 72 °C for annealing and extension, and a final extension of 5 min at 72 °C.

The second PCR amplicons were pooled and purified using AMPure XP. A target-sized DNA of the purified library (approximately 380–510 base pairs [bp]) was then

excised using E-Gel SizeSelect (ThermoFisher Scientific, Waltham, MA, USA). The DNA sample was applied to the Illumina MiSeq platform at Ryukoku University, Japan. The sequence data were deposited in the Sequence Read Archive of the DNA Data Bank of Japan (accession number: DRA007786). See Appendix 1 for details of molecular analyses.

Bioinformatics

The procedures used for bioinformatics and data analyses followed those described previously (Matsuoka et al. 2016a; Ushio et al. 2017). The raw MiSeq data were converted into FASTQ files using the bcl2fastq program provided by Illumina. The FASTQ files were then demultiplexed using the commands implemented in Claident pipeline (Tanabe and Toju 2016; software available online: <https://www.claident.org/>). The forward and reverse sequences were then merged with each other. The total 611,092 reads ($32,162 \pm 10,784$ reads per sample, mean \pm SD, $n = 19$) were assembled using Claident v0.2.2016.07.05. First, short (<150 bp) reads were removed, then potentially chimeric sequenc-

es and sequencing errors were removed using UCHIME v4.2.40 (Edgar et al. 2011) and algorithms in CD-HIT-OTU (Li et al. 2012), respectively. The remaining sequences were assembled at a threshold similarity of 97% (Osono 2014), and the resulting consensus sequences represented molecular operational taxonomic units (OTUs). Then, singletons were removed. Through these procedures, 2,549 OTUs (602,304 reads) in total were obtained. An OTU table (i.e., matrix of OTUs and samples with sequence reads in each cell entry) was generated after this process. Then, coverage-based rarefaction was conducted to standardize the effect of sequencing depth (Chao and Jost 2012, rarefaction curves for each sample are shown in Suppl. material 1: Fig. S1). For each of the obtained OTUs, taxonomic identification was conducted based on the query-centric auto-k-nearest-neighbor (QCauto) method (Tanabe and Toju 2016) with NCBI database and subsequent taxonomic assignment with the lowest common ancestor algorithm (Huson et al. 2007) using Claident. The functional guild of each fungal OTU was estimated based on the FUNGuild database (Nguyen et al. 2016). Thirty-nine OTUs (718 reads) that were identified as non-fungal organisms were discarded and the remaining 1,956 OTUs (175,101 reads) were used in further analyses (Suppl. material 2: Table S1).

Vegetation information of catchment area

To test the effect of vegetation of the catchment area on fungal DNA assemblages, the gross area and vegetation of the catchment area of each sampling site was calculated by using geographic information system software (QGIS v.2.14.8). First, the catchment area of each site was calculated based on the GPS coordinates of the site. Then, by consulting the vegetation map made by the Ministry of the Environment of Japan, vegetation of the catchment area was divided into seven categories (deciduous broad-leaved forest, cedar-cypress plantation, ever-green broad-leaved forest, ever-green coniferous forest, grassland, cropland, and urban area) and the area of each category was calculated (Fig. 1).

Statistical analyses

The presence or absence of the OTUs was recorded for each sample regardless of the number of reads and used for all statistical analyses as binary data. All analyses were performed using R v.3.4.3 (R Core Team 2017). Note that in the present study, analyses described below were conducted on the datasets with unidentified OTUs (those that could not be identified even at the kingdom level). However, to investigate the effect of inclusion of unidentified OTUs, and the use of presence/absence data instead of read number, additional analyses were conducted by using datasets for only fungal OTUs (1,031 OTUs, see Suppl. material 2: Table S1), and by considering read counts of each OTU.

The variation of community composition of OTUs among sampling sites and the community dissimilarity of the OTUs among plots were visualized using non-metric multidimensional scaling (NMDS, Minchin 1987). Presence/absence data of OTUs for each site were converted into a dissimilarity matrix using the standardized effect size (SES) of the Jaccard dissimilarity index to show the community dissimilarity without dependence on richness gradient among sites, which is an issue of long-standing importance (Chase and Myers 2011). The SES was defined as: $(\beta_{\text{obs}} - \beta_{\text{null}}) / \beta_{\text{sd}}$, where β_{obs} is the observed β diversity, β_{null} is the mean of the null distribution of β diversity, and β_{sd} is the standard deviation of the null distribution. The null distribution was calculated based on 999 randomizations preserving both the site occurrence and the OTU richness. The randomized community data were generated using “randomizeMatrix” function in “picante” package. The SES value indicates the magnitude of the deviation from the expectation of a random assembly process, and positive and negative values indicate more and less β -diversity, respectively, than expected by chance. I12 is the only site located in an urban area, and the NMDS ordination indicates that the community in I12 is largely different from the others (Suppl. material 1: Fig. S2). Therefore, we excluded this site from further analyses. The correlation of NMDS structure with geographic coordinates (latitude and longitude) was tested through permutation tests (“envfit” command in the vegan package, 9999 permutations) to show the spatial structure of OTU composition among sites. Further, only the results for presence/absence data are shown. Then, to test whether the OTU compositions are more similar within a tributary than between tributaries, SES values were compared calculating 99% confidence interval (CI).

To test the effects of environmental variables (i.e., elevation, catchment area, and vegetation) and spatial distance on the OTU composition, a multivariate permutational analysis of variance (PERMANOVA) was implemented (“adonis” command in the vegan package, 9999 permutations). In these analyses, only one sample was used for each branch as samples from the same branch (A1-7 and I7-8) share most of their catchment. Therefore, eleven samples (A4, I1-6, and I8-11) were used for the following analyses and others were excluded. Preliminary analyses confirmed that the choice of sample inside the branch does not affect the results (data were used in Suppl. material 2: Table S1). Presence/absence data of ECM fungal OTUs for each plot were converted into a dissimilarity matrix using the modified Raup–Crick dissimilarity index (Chase et al. 2011). The Raup–Crick dissimilarity index is calculated based on the null model approach, which is akin to the SES value described above. Then, vegetation and spatial variables were calculated. Vegetation variables were extracted based on principal coordinate analysis (PCA). First, the vegetation data of the catchment area of each sampling site were transferred into the Euclidian distance matrix. Then, PCA vectors were extracted. The first two PCA axes explained the variations in the vegetation among sites of 86.9% and 11.2%, respectively, (Suppl.

material 1: Fig. S3) and were used for the analysis. Spatial variables were extracted based on principal components of neighbor matrices (PCNM, Borcard et al. 2004). The PCNM analysis produced a set of orthogonal variables derived from the geographical coordinates of the locations of sampling sites. Two PCNM variables that best accounted for autocorrelation were used for the analysis.

Results

Fungal diversity

After quality filtering and denoising, 602,304 reads of the original 611,092 reads were retained. After rarefaction analysis, 175,101 reads from 19 samples were grouped into 1,956 OTUs with 97% sequence similarity (Suppl. material 2: Table S1). In total, 770 OTUs were assigned as Ascomycota (38.4 % of the total number of fungal OTUs), 177 OTUs as Basidiomycota (8.8 %), and 38 OTUs as Chytridiomycota (1.9 %). The remaining 941 OTUs were not assigned to any phylum. The proportions of OTU numbers of each phylum were similar between the sampling sites except for I12 (Suppl. material 1: Fig. S4). At the order level, 552 Ascomycota OTUs belonged to 31 orders and 122 Basidiomycota OTUs belonged to 21 orders. The common ascomycete orders were Capnodiales (216 OTUs), Helotiales (88 OTUs), and Diaporthales (47 OTUs). The common basidiomycete orders were Tremellales (31 OTUs), Polyporales (15 OTUs), and Filobasidiales (12 OTUs). Thirty-six of 38 chytrid OTUs belonged to Rhizophydiales. At the genus level, 224 Ascomycota OTUs were assigned to 114 genera and 88 Basidiomycota OTUs to 55 genera. The OTU rich genera were *Ramularia* (36 OTUs) and *Mycosphaerella* (16 OTUs). The detailed taxonomic assignments are shown in Suppl. material 2: Table S1.

FUNGuild assigned 539 OTUs to functional guilds, 147 OTUs of which were saprotrophs and the others included pathogens (117 OTUs) and symbionts (19 OTUs) (Suppl. material 2: Table S1). The remaining 256 OTUs are those that were assigned to more than one functional guild, or whose function can change depending on circumstances. The detected OTUs included aquatic hyphomycetes, e.g., *Flagellospora* sp. (OTU_0413) and *Tetracladium marchalianum* (OTU_1491), that were dominantly detected from 19 and 16 out of 19 samples, respectively. On the other hand, OTUs belonging to terrestrial taxonomic or functional groups were also detected; plant decomposer (e.g., *Mycena* sp., OTU_0020); plant pathogen (e.g., *Ciboria shiraiana*, OTU_1523); and mycorrhizal fungi (e.g., Glomeromycetes sp. (OTU_1449) and *Tuber* sp. (OTU_0010)).

Spatial structures of fungal OTU composition

The Mantel test and NMDS ordination revealed biogeographic changes in the fungal OTU compositions. The

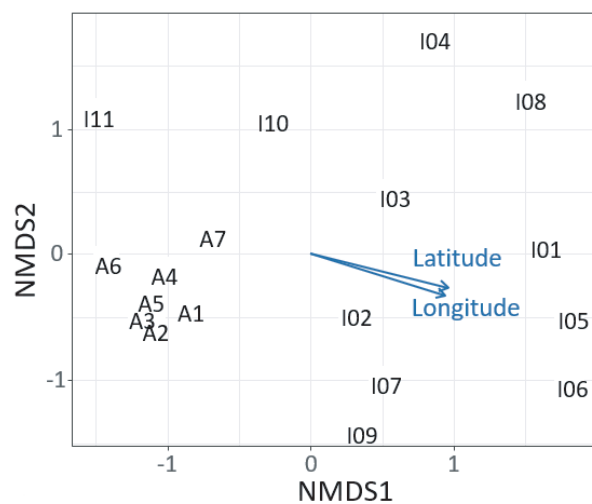


Figure 2. Dissimilarity of the DNA assemblages among sites as revealed by nonmetric multidimensional scaling (NMDS) ordination (stress value = 0.154). Numbers are identical with site numbers in Fig. 1. Fungal DNA assemblages were similar within the same branch (A1-7) and changed with latitude and longitude across the study site.

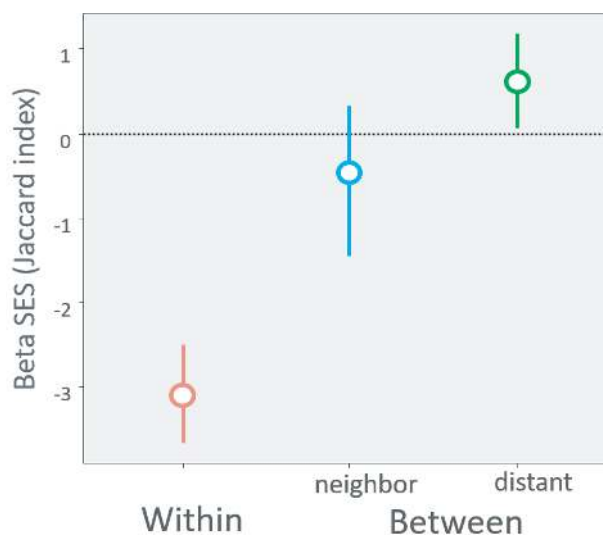


Figure 3. Standard effect size (SES) of the Jaccard index for comparison of the compositional dissimilarity within a same branch and between the adjacent and distant branches. Error bars indicate 99% confidence interval. The horizontal dotted line represents SES = 0, indicating a non-significant effect.

dissimilarity of OTU composition among samples was significantly correlated with latitudinal and longitudinal distances (Mantel test, latitude, $r = 0.585$, $P < 0.001$; longitude, $r = 0.510$, $P < 0.001$). The NMDS ordination showed the geographical change of OTU composition among sites (Fig 2, stress value = 0.154). The ordination was significantly correlated with the latitude and longitude ('envfit' function; latitude, $r^2 = 0.737$, $P < 0.001$; longitude, $r^2 = 0.781$, $P < 0.001$). The dissimilarity indices of

Table 1. PERMANOVA results for the composition of DNA assemblages.

	Degree of freedom	Sums of Squares	F Model	R ²	P value
null	10	1.801		1	
Elevation	1	0.392	2.502	0.218	0.169
Catchment area	1	0.248	1.434	0.137	0.356
Vegetation PC1	1	0.150	0.819	0.083	0.509
Vegetation PC2	1	0.001	0.003	0.001	0.838
PCNM1	1	0.778	6.838	0.432	0.009
PCNM2	1	-0.125	-0.584	-0.069	0.942

OTU composition (SES value of Jaccard index) were significantly lower than 0 when compared between the sites in the same branch, higher when compared between distant branches, and intermediate when compared between adjacent branches (Fig. 3, based on the 99% confidence interval). This indicates that OTU compositions are significantly more similar inside the branch than expected by chance and more different between spatially distinct different branches. PERMANOVA showed that the differences in OTU composition among the branches were only related to the spatial variable (PCNM1) (F-model = 6.938, R² = 0.432, P = 0.009) and did not significantly correlate with the other environmental variables including elevation and vegetation PCA axis (Table 1). In addition, using datasets composed only by fungal OTUs and taking into account reads abundance of each out, it did not affect the results. The results for these additional analyses are provided in Appendix 2.

Discussion

In the present study, we demonstrated the applicability of eDNA metabarcoding for the investigation of fungal DNA assemblages and their spatial structure in river water. Our results show that the fungal DNA assemblages are more similar within branches than between. The results also show that the fungal DNA assemblages are more similar between spatially closer branches. Therefore, this study for the first time shows the spatial structure, especially compositional similarity, of fungal DNA assemblages across river branches.

Fungal diversity in river water

River water contained both phylogenetically and functionally diverse fungal DNA. In the present study, phylogenetically diverse fungi, primarily ascomycetes and basidiomycetes, were detected in the river, as shown in the recent studies using molecular methods (Duarte et al. 2015, Wurzbacher et al. 2016; LeBrun et al. 2018). In river water, aquatic fungi have diverse lineages and functional groups, including saprophytes, which decompose plant substrates, and parasites of plants and animals. In the present study, some genera known as aquatic hyphomycetes (e.g., *Flagellospora* and *Tetracladium*; Raja

et al. 2018; Grossart et al. 2019) were detected. The OTUs belonging to these taxa are thought to degrade plant substrates that have entered the river from land (fallen leaves and wood). In particular, OTUs such as OTU_0413 (*Flagellospora* sp.) and OTU_1491 (*Tetracladium marchalianum*) were detected in almost all samples, suggesting that they are widely distributed in the surveyed landscape. These aquatic fungi can be detected based on the eDNA obtained from the river water owing to the presence of spores or mycelial fragments in the water.

OTUs of terrestrial fungi such as the plant decomposer *Mycena* and the mycorrhizal fungi *Tuber* and Glomeromycota were detected, as well as the OTUs of aquatic fungi. DNA of these fungi may be released from the propagules, which flowed into the rivers from land. For example, Basidiomycota acting as plant decomposers often produce epigeous sporocarps and disperse their spores into air. Such airborne spores have the potential to enter the river system. On the other hand, species in the genus *Tuber* produce hypogeous sporocarps and do not disperse their spores into air. This indicates that the propagules and mycelial fragments in soil and plant materials (e.g., decomposing leaves) flow into the river via the surface current or with the direct addition of these substrates (Voronin 2014). Recent studies that investigated the environmental fungal DNA of lakes and rivers also detected a significant amount of DNA for terrestrial fungi, as well as for aquatic fungi (Deiner et al. 2016, Khomich et al. 2017; LeBrun et al. 2018). Our results are consistent with the results of these previous studies.

The spatial structure of fungal DNA assemblages

The fungal OTU compositions in the studied landscape were more similar within the same branch than between branches. This may be because some fungal DNA, which entered upstream flowed downstream and was therefore detected both upstream and downstream in the branches. Indeed, recent evidence suggests that rivers work as a dispersal pathway of terrestrial fungi (Deiner et al. 2016; LeBrun et al. 2018). Other possibilities are that the same aquatic fungal species live across the branches, or that the same terrestrial fungal DNA enters the river at several different sites. For example, Tsui et al., (2001) investigated aquatic fungi in upstream and downstream of the same river and reported the presence of common species in both areas of the stream. However, there is currently no information in

support of either of these possibilities. Further investigations into the composition and distribution of aquatic fungi along the river, the composition of fungal DNA flowing downstream, as well as the air and surface currents may help to clarify the mechanisms behind these results.

Fungal DNA assemblages show geographical structures at a scale of around 15 km in the forest site. This means that fungal DNA assemblages are more similar between spatially closer branches than between distant ones. Such geographical structures have also been observed by Khomich et al. (2017); fungal DNA assemblages along latitudes and longitudes were investigated in 77 lakes in the Scandinavia Peninsula. In the lakes, the surrounding rivers brought and accumulated the DNA of terrestrial fungi from the surrounding areas (Khomich et al. 2017). Our study shows that such geographical structures can also be observed in streams, where water has a directional flow and is constantly replaced. The geographical structure of fungal DNA may be generated by both aquatic and terrestrial fungal DNA. For aquatic fungi, environmental differences (e.g. water chemical properties or current velocity/discharge) across branches or dispersal limitations can affect the fungal communities (Heino et al. 2014; Duarte et al. 2017), causing geographically closer branches to share more fungal species. Similarly, environmental factors such as vegetation type, soil chemical properties, and dispersal limitation contribute to the formation of similar terrestrial fungal communities in the same vicinity (Toljander et al. 2006; Kadowaki et al. 2014). The propagules of these fungi flow into the nearby river, causing the DNA assemblages of geographically closer branches to resemble one another.

In the present study, geographic variations in the fungal DNA assemblage were explained only by spatial variables, and no significant relationship between fungal DNA assemblage and the vegetation of the catchment area was detected. These results suggest that in the study landscape, the spatial structure of the fungal DNA assemblages may be influenced by spatial factors such as dispersion limitation (Talbot et al. 2014; Peay et al. 2016) or unmeasured environmental factors. Previous studies have shown that the community composition of aquatic fungi is more strongly influenced by spatial variations in the environment, such as water quality (e.g., pH), compared to spatial distance among sites (Heino et al. 2014; Duarte et al. 2017). For example, LeBrun et al. (2018) investigated fungal DNA assemblage in river water with P enrichment problems and showed that the fungal DNA assemblage changes with P concentration. In our study sites, no such nutrient enrichment problems have been reported to date (Ministry of Land, Infrastructure, Transport and Tourism 2007) and the range of variation in the environment is considered to be smaller than that of previous studies (Heino et al. 2014; Duarte et al. 2017), which dealt with community changes at regional and continental scales. However, it is highly likely that unmeasured water qualities influence DNA assemblages in our study sites. In addition, the discharge of the surveyed river may affect the DNA compositions via changing the size of particles carried in the water col-

umn. In the present study, the brown particles, such as the fine particle organic matter (FPOM), were attached to the filter and fungal DNA attached to the FPOM might have been detected. Wurzbacher et al. (2016) showed that the detected fungal DNA can vary depending on the size and amount of FPOM. These unmeasured factors and relative importance of environmental variables and spatial distance should be considered in future studies. In a forest ecosystem, the types of vegetation present have a stronger effect on fungal communities compared to other environmental variables (Peay et al. 2013). Therefore, if the proportion of terrestrial fungal DNA is high, the fungal DNA assemblage may be explained by the vegetation present. The lack of a relationship between the vegetative and fungal DNA assemblages can be partly attributable to the fact that the catchment area and distribution of fungal individuals of potential DNA origin did not correspond. For example, at a landscape scale with tens of kilometers, similar to our study site, fungal DNA would possibly flow or be blown from surrounding forests, including outside the catchment area. Otherwise, fungal DNA can be derived from areas closer to the river catchment area. There is limited information available on the dispersal ability of fungi. However, a previous study has reported that for terrestrial fungi, most spores fall within several meters of the sporocarps (Li et al. 2005; Galante et al. 2011). This indicates that spores of fungal species and individuals colonizing more than several meters apart from the river should have little chance to fall into the river. Further studies are required to specify the range from which most fungal DNA is derived.

Methodological limitations

In the present study, the ITS region was used as a genetic marker, but many OTUs remained unidentified. Compared to other terrestrial habitats, fewer genomic studies have been conducted in aquatic habitats. As a result, fewer ITS sequence records are available and new phylogenetic groups are expected to be found (Grossart et al. 2019; Khomich et al. 2018). Furthermore, for basal fungi living in an aquatic habitat, such as chytrids, more conservative regions such as large subunit (LSU) and small subunit (SSU) regions of rDNA are often used (Nilsson et al. 2019). However, few studies have investigated whether a difference in genetic markers affects the analytical results (Nilsson et al. 2019). Therefore, this may be an interesting topic for future studies to investigate. This database deficiency and handling of resulting unidentified OTUs as well as the selection of appropriate marker genes and primer limitation/specificity towards a particular group can be a key problem in fungal DNA metabarcoding.

Conclusions

This is the first study to indicate that fungal DNA assemblages in river water show a spatial structure in the catch-

ment area. Our results also demonstrated that river water in a forest contains phylogenetically and functionally diverse fungal DNA that is derived from both land and aquatic regions. The results imply that by investigating fungal DNA assemblages in river water, information on both the land and aquatic fungal compositions of the catchment area can be accessed. Further studies are required to clarify to what extent these DNA assemblages reflect spatial structure of the fungal communities, and what generates such a spatial structure of DNA assemblages. Importantly, fungal communities are known to vary with time in both aquatic and terrestrial habitats (Nikolcheva and Bärlocher 2005; Matsuoka et al. 2016b; Song et al. 2018). The influence of temporal changes in fungal communities on the spatial structure of fungal DNA assemblages should be investigated in the future.

Acknowledgments

MiSeq sequencing was conducted in the Department of Environmental Solution Technology, Faculty of Science and Technology, Ryukoku University. We thank H. Yamana for supporting the MiSeq sequencing. This study received partial financial support from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (MEXT) to SM (Grant No. 17K15199) and IK (18K11678) and the Environment Research and Technology Development Fund of Environmental Restoration and Conservation Agency (4-1602).

References

- Bahram M, Kohout P, Anslan S, Harend H, Abarenkov K, Tedersoo L (2016) Stochastic distribution of small soil eukaryotes resulting from high dispersal and drift in a local environment. *ISME Journal* 10: 885–896. <https://doi.org/10.1038/ismej.2015.164>
- Bahram M, Peay KG, Tedersoo L (2015) Local-scale biogeography and spatiotemporal variability in communities of mycorrhizal fungi. *New Phytologist* 205: 1454–1463. <https://doi.org/10.1111/nph.13206>
- Borcard D, Legendre P, Avois-Jacquet C, Tuomisto H (2004) Dissecting the spatial structure of ecological data at multiple scales. *Ecology* 85: 1826–1832. <https://doi.org/10.1890/03-3111>
- Chao A, Jost L (2012) Coverage-based rarefaction and extrapolation: Standardizing samples by completeness rather than size. *Ecology* 93: 2533–2547. <https://doi.org/10.1890/11-1952.1>
- Chase JM, Myers JA (2011) Disentangling the importance of ecological niches from stochastic processes across scales. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 2351–2363. <https://doi.org/10.1098/rstb.2011.0063>
- Chase JM, Kraft NJB, Smith KG, Vellend M, Inouye BD (2011) Using null models to disentangle variation in community dissimilarity from variation in α -diversity. *Ecosphere* 2: art24. <https://doi.org/10.1890/ES10-00117.1>
- Deiner K, Fronhofer EA, Mächler E, Walser JC, Altermatt F (2016) Environmental DNA reveals that rivers are conveyor belts of biodiversity information. *Nature Communications* 7. <https://doi.org/10.1038/ncomms12544>
- Duarte S, Bärlocher F, Trabulo J, Cássio F, Pascoal C (2015) Stream-dwelling fungal decomposer communities along a gradient of eutrophication unraveled by 454 pyrosequencing. *Fungal Diversity* 70: 127–148. <https://doi.org/10.1007/s13225-014-0300-y>
- Duarte S, Cássio F, Pascoal C (2017) Environmental drivers are more important for structuring fungal decomposer communities than the geographic distance between streams. *Limnética* 36: 491–506.
- Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R (2011) UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27: 2194–200. <https://doi.org/10.1093/bioinformatics/btr381>
- Grossart H-P, Van den Wyngaert S, Kagami M, Wurzbacher C, Cunliffe M, Rojas-Jimenez K (2019) Fungi in aquatic ecosystems. *Nature Reviews Microbiology*. <https://doi.org/10.1038/s41579-019-0175-8>
- Hamady M, Walker JJ, Harris JK, Gold NJ, Knight R (2008) Error-correcting barcoded primers for pyrosequencing hundreds of samples in multiplex. *Nature Methods* 235–237. <https://doi.org/10.1038/nmeth.1184>
- Hawksworth DL, Lücking R (2017) Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiology Spectrum* 5. <https://doi.org/10.1128/microbiolspec.FUNK-0052-2016>
- Heino J, Tolkkinen M, Pirttilä AM, Aisala H, Mykrä H (2014) Microbial diversity and community – environment relationships in boreal streams. *Journal of Biogeography* 41: 2234–2244. <https://doi.org/10.1111/jbi.12369>
- Huson DH, Auch AF, Qi J, Schuster SC (2007) MEGAN analysis of metagenomics data. *Genome Research* 17: 377–386. <https://doi.org/10.1101/gr.5969107>
- Katano I, Harada K, Doi H, Souma R, Minamoto T (2017) Environmental DNA method for estimating salamander distribution in headwater streams, and a comparison of water sampling methods. *PLoS ONE* 12: 1–15. <https://doi.org/10.1371/journal.pone.0176541>
- Khomich M, Davey ML, Kauserud H, Rasconi S, Andersen T (2017) Fungal communities in Scandinavian lakes along a longitudinal gradient. *Fungal Ecology* 27: 36–46. <https://doi.org/10.1016/j.funeco.2017.01.008>
- Khomich M, Cox F, Andrew CJ, Andersen T, Kauserud H, Davey ML (2018) Coming up short: identifying substrate and geographic biases in fungal sequence databases. *Fungal Ecology* 36: 75–80. <https://doi.org/10.1016/j.funeco.2018.08.002>
- LeBrun ES, Taylor DL, King RS, Back JA, Kang S (2018) Rivers may constitute an overlooked avenue of dispersal for terrestrial fungi. *Fungal Ecology* 32: 72–79. <https://doi.org/10.1016/j.funeco.2017.12.003>
- Li W, Fu L, Niu B, Wooley J (2012) Ultrafast clustering algorithms for metagenomic sequence analysis. *Briefings in Bioinformatics* 13: 656–668. <https://doi.org/10.1093/bib/bbs035>
- Matsuoka S, Kawaguchi E, Osono T (2016a) Temporal distance decay of similarity of ectomycorrhizal fungal community composition in a subtropical evergreen forest in Japan. *FEMS Microbiology Ecology* 92: 1–11. <https://doi.org/10.1093/femsec/fiw061>
- Matsuoka S, Mori AS, Kawaguchi E, Hobara S, Osono T (2016b) Disentangling the relative importance of host tree community, abiotic environment and spatial factors on ectomycorrhizal fungal assemblages along an elevation gradient. *FEMS Microbiology Ecology* 92: 1–11. <https://doi.org/10.1093/femsec/fiw044>

- Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405: 243. <https://doi.org/10.1038/35012251>
- Minchin PR (1987) An evaluation of the relative robustness of techniques for ecological ordination. *Vegetatio* 69: 89–107. https://doi.org/10.1007/978-94-009-4061-1_9
- Ministry of Land, Infrastructure, Transport and Tourism (2007) Outline of the basin and river of the Ibo River system (in Japanese). https://www.mlit.go.jp/river/shinngikai_blog/shaseishin/kasenbunkakai/shouinikai/kihonhoushin/070119/pdf/refl-2.pdf [accessed on June 21, 2019]
- Nakagawa H, Yamamoto S, Sato Y, Sado T, Minamoto T, Miya M (2018) Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. *Freshwater Biology* 63: 569–580. <https://doi.org/10.1111/fwb.13094>
- Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, Schilling JS, Kennedy PG (2016) FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecology* 20: 241–248. <https://doi.org/10.1016/j.funeco.2015.06.006>
- Nikolcheva LG, Bärlocher F (2005) Seasonal and substrate preferences of fungi colonizing leaves in streams: traditional versus molecular evidence. *Environmental Microbiology* 7: 270–280. <https://doi.org/10.1111/j.1462-2920.2004.00709.x>
- Nilsson RH, Anslan S, Bahram M, Wurzbacher C, Baldrian P, Tedersoo L (2019) Mycobiome diversity: high-throughput sequencing and identification of fungi. *Nature Reviews Microbiology* 17: 95–109. <https://doi.org/10.1038/s41579-018-0116-y>
- Osono T (2014) Metagenomic approach yields insights into fungal diversity and functioning. In: Sota T, Kagata H, Ando Y, Utsumi S, Osono T (Eds) *Species Diversity and Community Structure*. Springer, Berlin, 1–23. https://doi.org/10.1007/978-4-431-54261-2_1
- Peay KG, Baraloto C, Fine PV (2013) Strong coupling of plant and fungal community structure across western Amazonian rainforests. *The ISME Journal* 7: 1852–1861. <https://doi.org/10.1038/ismej.2013.66>
- Peay KG, Kennedy PG, Talbot JM (2016) Dimensions of biodiversity in the Earth mycobiome. *Nature Reviews Microbiology* 14: 434–447. <https://doi.org/10.1038/nrmicro.2016.59>
- Pecl GT, Araújo MB, Bell JD, Blanchard J, Bonebrake TC, Chen IC, Clark TD, Colwell RK, Danielsen F, Evengård B, et al. (2017) Biodiversity redistribution under climate change: Impacts on ecosystems and human well-being. *Science* 355: 9214. <https://doi.org/10.1126/science.aai9214>
- R Core Team (2017) R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. <http://www.R-project.org/> [25 April 2019, date last accessed]
- Raja HA, Shearer CA, Tsui CKM (2018) *Freshwater Fungi*. John Wiley & Sons, Ltd: Chichester. <https://doi.org/10.1002/9780470015902.a0027210>
- Schoch CL, Seifert K a, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W (2012) Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* 109: 6241–6246. <https://doi.org/10.1073/pnas.1117018109>
- Song P, Tanabe S, Yi R, Kagami M, Liu X, Ban S (2018) Fungal community structure at pelagic and littoral sites in Lake Biwa determined with high-throughput sequencing. *Limnology* 19: 241–251. <https://doi.org/10.1007/s10201-017-0537-8>
- Talbot JM, Bruns TD, Taylor JW, Smith DP, Branco S, Glassman SI, Erlandson S, Vilgalys R, Liao H-L, Smith ME, Peay KG (2014) Endemism and functional convergence across the North American soil mycobiome. *Proceedings of the National Academy of Sciences of the United States of America* 111: 6341–6346. <https://doi.org/10.1073/pnas.1402584111>
- Tanabe AS, Toju H (2016) Correction: Two New Computational Methods for Universal DNA Barcoding: A Benchmark Using Barcode Sequences of Bacteria, Archaea, Animals, Fungi, and Land Plants. *PLoS ONE* 11: e0152242. <https://doi.org/10.1371/journal.pone.0152242>
- Toju H, Tanabe AS, Yamamoto S, Sato H (2012) High-coverage ITS primers for the DNA-based identification of ascomycetes and basidiomycetes in environmental samples. *PLoS ONE* 7: e40863. <https://doi.org/10.1371/journal.pone.0040863>
- Tsui CKM, Hyde KD, Hodgkiss IJ (2001) Colonization patterns of wood-inhabiting fungi on baits in Hong Kong rivers, with reference to the effects of organic pollution. *Antonie van Leeuwenhoek* 79: 33–38. <https://doi.org/10.1023/A:1010210631215>
- Uchii K, Doi H, Minamoto T (2016) A novel environmental DNA approach to quantify the cryptic invasion of non-native genotypes. *Molecular Ecology Resources* 16: 415–422. <https://doi.org/10.1111/1755-0998.12460>
- Ushio M, Aiba SI, Takeuchi Y, Iida Y, Matsuoka S, Repin R, Kitayama K (2017) Plant-soil feedbacks and the dominance of conifers in a tropical montane forest in Borneo. *Ecological Monographs* 87: 105–129. <https://doi.org/10.1002/ecm.1236>
- Voronin LV (2014) Terrigenous micromycetes in freshwater ecosystems (review). *Inland Water Biology* 7: 352–356. <https://doi.org/10.1134/S1995082914040191>
- Wurzbacher C, Grimm IJ, Bärlocher F (2016) Metabarcoding-based fungal diversity on coarse and fine particulate organic matter in a first-order stream in Nova Scotia, Canada. *F1000 Research* 4: 1378. <https://doi.org/10.12688/f1000research.7359.2>

Appendix 1

The detailed procedure of molecular analyses

Molecular analyses

The protocols described by Uchii et al. (2016) were followed to extract DNA from the filters. Each filter was incubated in 400 µl of Buffer AL (Qiagen, Hilden, Germany) and 40 µl of Proteinase K (Qiagen, Hilden, Germany), using a Salivette tube (Sarstedt, Nümbrecht, Germany) at 56 °C for 30 min. The Salivette tube with the filter was centrifuged at 3,500 × g for 5 min. Next, 220 µl of TE buffer (10 mM Tris-HCl and 1 mM EDTA, pH: 8.0) was added to the filter and centrifuged at 5,000 × g for 5 min. The dissolved DNA in the eluted solution was purified using a DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. The final volume of the extracted sample was eluted in 100 µl of Buffer AE of the DNeasy Blood & Tissue Kit.

For MiSeq sequencing, the fungal internal transcribed spacer 1 (ITS1) region of rDNA was amplified. The ITS region has been proposed as the formal fungal barcode (Schoch et al. 2012). The methods of DNA analysis followed those described by Ushio et al. (2017) with some modifications, such as number of PCR cycles. The first-round PCR (first PCR) amplified the ITS1 region using the ITS1-F-KYO2 and ITS2-KYO2 primer set, which is capable of amplifying the ITS1 region of most fungal groups (Toju et al. 2012). An Illumina sequencing primer and six random bases (N) were combined to produce each primer. Thus, the forward primer sequence was: 5'- *ACA CTC TTT CCC TAC ACG ACG CTC TTC CGA TCT* NN-NNNN TAG AGG AAG TAA AAG TCG TAA -3' and the reverse primer sequence was: 5'- *GTG ACT GGA GTT CAG ACG TGT GCT CTT CCG ATC T* NNNNNN TTY RCT RCG TTC TTC ATC- 3'. The italic and normal letters represent MiSeq sequencing primers and fungi-specific primers, respectively. The six random bases (N) were used to enhance cluster separation on the flowcells during initial base call calibrations on MiSeq.

The 1st PCR was performed in a 12 µl volume with the buffer system of KODFX NEO (TOYOBO, Osaka, Japan), which contained 2.0 µl of template DNA, 0.2 µl of KOD FX NEO, 6.0 µl of 2× buffer, 2.4 µl of dNTP, and 0.7 µl each of the two primers (5 µM). The PCR conditions were as follows; an initial incubation for 2 min at 94 °C followed by 5 cycles of 10 s at 98 °C, 30 s at 68 °C for annealing and 30 s at 68 °C, 5 cycles of 10 s at 98 °C, 30 s at 65 °C and 30 s at 68 °C; 5 cycles of 10 s at 98 °C, 30 s at 62 °C and 30 s at 68 °C; 25 cycles of 10 s at 98 °C, 30 s at 59 °C and 30 s at 68 °C, and a final extension of 5 min at 68 °C. Eight replicate first-PCRs (per sample) were performed to mitigate the reaction-level PCR bias. Then, the duplicated first PCR amplicons (per sample) were combined, resulting in a template per sample for the second PCR. The PCR templates were purified using Agencourt AMPure XP (PCR product: AMPure XP beads = 1:0.8; Beckman Coulter, Brea, California, USA) before the second PCR.

The second PCR amplified the first PCR amplicons using the primers (forward) 5'-*AAT GAT ACG GCG ACC ACC GAG ATC TAC AC* XXXXXXXX TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG-3' and (reverse) 5'-*CAA GCA GAA GAC GGC ATA CGA GAT* XXXXXXXX GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA G- 3'. The italic and normal letters represent the MiSeqP5/P7 adapter and sequencing primers, respectively. The 8X bases represent dual-index sequences inserted to identify different samples (Hamady et al. 2008). The second PCR was carried out with 12 µl reaction volume containing 1.0 µl of template, 6 µl of 2× KAPA HiFi HotStart ReadyMix (KAPA Biosystems, Wilmington, Washington, USA), 1.4 µl of each primer (2.5 µM), and 2.2 µl of sterilized distilled water. The PCR conditions were as follows; an initial incubation for 3 min at 95 °C followed by 12 cycles of 20 s at 98 °C, 15 s at 72 °C for annealing and extension, and a final extension of 5 min at 72 °C.

The indexed second PCR amplicons were pooled to make a library to be sequenced on MiSeq. The volume of each sample added to the library was adjusted to normalize the concentrations of each second PCR product. The pooled library was purified using Agencourt AMPure XP. A target-sized DNA of the purified library (approximately 380–510 base pairs [bp]) was then excised using E-Gel SizeSelect (ThermoFisher Scientific, Waltham, MA, USA). The double-stranded DNA concentration of the library was then adjusted to 4 nmol/L using Milli-Q water, and the DNA sample was applied to the Illumina MiSeq platform at Ryukoku University, Japan. The sequence data were deposited in the Sequence Read Archive of the DNA Data Bank of Japan (accession number: DRA007786).

Appendix 2

The results for additional analyses with different datasets

In the manuscript, results for the dataset employing presence/absence data for all OTUs, except for chimeric and

Table A1-1. PERMANOVA results for the composition of DNA assemblages.

	Degree of freedom	Sums of Squares	F Model	R ²	P value
(a) Only fungal OTUs included					
null	10	1.900			
Elevation	1	0.334	1.918	0.176	0.222
Catchment area	1	0.173	0.900	0.091	0.494
Vegetation PC1	1	0.167	0.866	0.088	0.485
Vegetation PC2	1	-0.106	-0.476	-0.056	0.954
PCNM1	1	0.794	6.459	0.418	0.007
PCNM2	1	0.017	0.080	0.009	0.781
(b) When read number considered					
null	10	1.194			
Elevation	1	0.452	2.931	0.246	0.126
Catchment area	1	0.269	1.540	0.146	0.335
Vegetation PC1	1	0.159	0.849	0.086	0.501
Vegetation PC2	1	-0.015	-0.075	-0.008	0.863
PCNM1	1	0.766	6.416	0.416	0.012
PCNM2	1	-0.088	-0.411	-0.048	0.905

non-fungal (but identified) OTUs, are shown. However, this dataset may include non-fungal OTUs. Therefore, the most conservative method of analysis is to include only those OTUs identified as fungus. In addition, some studies have treated a sequence read number as the proxy of abundance for each OTU, although there are known issues with the quantitative use of read numbers generated from amplicon sequencing (Amend et al. 2010; Elbrecht and Leese 2015). Therefore, to test whether these differences in datasets can affect the analytical results, we conducted

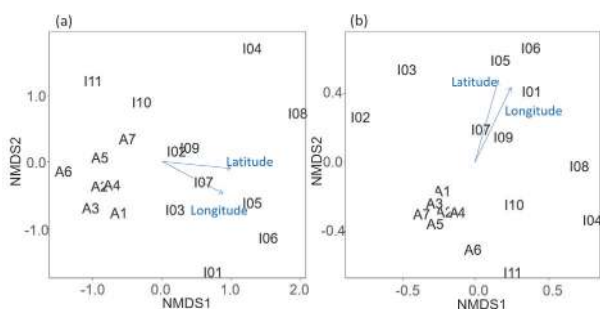


Figure A1-1. Dissimilarity of the DNA assemblages among sites as revealed by nonmetric multidimensional scaling (NMDS) ordination for (a) datasets with only fungal OTUs (stress value = 0.159) and (b) datasets using the read number (stress value = 0.153). Numbers are consistent with site numbers in Fig. 1 in the main text. The ordinations were significantly correlated with latitude and longitude for both datasets ('envfit' function; latitude, (a) $r^2 = 0.749$, $P = 0.0002$, (b) $r^2 = 0.684$, $P = 0.0006$; longitude, (a) $r^2 = 0.7931$, $P = 0.0002$, (b) $r^2 = 0.802$, $P = 0.0002$).



Figure A1-2. Standard effect size (SES) of the Jaccard index for comparison of compositional dissimilarity within a same branch and between the adjacent and distant branches. Error bars indicate 99% confidence interval. The horizontal dotted line represents SES = 0, indicating a non-significant effect.

the same sets of analyses for the dataset including only fungal OTUs (Suppl. material 2: Table S1) and using the sequence read number as a proxy of abundance for each OTU. The analytical procedure is the same as described in the manuscript except in the analyses considering the read number where the Bray-Curtis dissimilarity index was used instead of the Raup-Crick index. In addition, since the same randomization for presence/absence data cannot be conducted for non-binomial data, the pairwise comparison of SES values between and within branches (Fig. 3 in the main text) were not conducted on the dataset with the read number.

Both datasets yielded the same results as in the manuscript; the dissimilarity of OTU composition among sites correlated with the latitude and longitude (Fig. A1-1). The OTU composition was more similar within the same branches than between branches (Fig A1-1 and Fig A1-1) and only the PCNM1 vector significantly related to the difference in OTU composition among sites (Table A1-1). These results indicate that the result shown in the manuscript is not vulnerable to the dataset choice.

References

- Amend AS, Seifert KA, Bruns TD (2010) Quantifying microbial communities with 454 pyrosequencing: Does read abundance count? *Molecular Ecology* 19: 5555–5565. <https://doi.org/10.1111/j.1365-294X.2010.04898.x>
- Elbrecht V, Leese F (2015) Can DNA-based ecosystem assessments quantify species abundance? Testing primer bias and biomass-sequence relationships with an innovative metabarcoding protocol. *PLoS ONE* 10: e0130324. <https://doi.org/10.1371/journal.pone.0130324>

Supplementary material 1

Figures S1–S4

Authors: Shunsuke Matsuoka, Yoriko Sugiyama, Hiroto Sato, Izumi Katano, Ken Harada, Hideyuki Doi

Data type: multimedia

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited. Link: <https://doi.org/10.3897/mbmg.3.36335.suppl1>

Supplementary material 2

Table S1. List of all OTUs, their consensus sequence, and taxonomic and functional assignments.

Authors: Shunsuke Matsuoka, Yoriko Sugiyama, Hiroto Sato, Izumi Katano, Ken Harada, Hideyuki Doi

Data type: molecular data

Copyright notice: This dataset is made available under the Open Database License (<http://opendatacommons.org/licenses/odbl/1.0/>). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited. Link: <https://doi.org/10.3897/mbmg.3.36335.suppl2>