FEMS Microbiology Ecology, 96, 2020, fiaa166

doi: 10.1093/femsec/fiaa166
Advance Access Publication Date: 20 August 2020
Research Article

RESEARCH ARTICLE

Drivers of richness and community composition of fungal endophytes of tree seeds

Iva Franić[1,2,3,*], René Eschen[1], Eric Allan[2], Martin Hartmann[4,†], Salome Schneider[3] and Simone Prospero[3]

[1]CABI, CH-2800 Delémont, Switzerland, [2]Institute of Plant Sciences, University of Bern, CH-3013 Bern, Switzerland, [3]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, CH-8903 Birmensdorf, Switzerland and [4]Institute of Agricultural Sciences, ETH Zürich, CH-8092 Zürich, Switzerland

*Corresponding author: CABI, Rue des Grillons 1, CH-2800 Delémont, Switzerland. Tel: +41 (0) 32 421 4887; Fax: +41 (0) 32 421 4871; E-mail: i.franic@cabi.org

One sentence summary: Host identity is the key driver of richness and community composition of fungal endophytes of tree seeds.

Editor: Petr Baldrian

†Martin Hartmann, http://orcid.org/0000-0001-8069-5284

ABSTRACT

Recent studies revealed a high diversity of fungal endophytes in traded tree seeds, including potential plant pathogens. The factors determining richness and composition of seed mycobiomes are poorly understood, but might be an important determinant for tree health. We assessed the relative impact of host identity, site, several site-specific environmental factors, and whether the host was sampled in its native or non-native distribution range, on the richness and composition of fungal seed endophytes of nine tree species across 15 sites in Europe and North America. Our results show that fungal richness was affected by host identity, but not by environmental variables or host distribution range. Fungal community composition was primarily driven by host identity, and to a lesser extent by environment. Around 25% of the 2147 amplicon sequence variants (ASVs) were generalists appearing on both continents and in both gymnosperms and angiosperms. Around 63% of the ASVs appeared in only gymnosperms or angiosperms, and 33% of the ASVs were associated with a single host species, while none were found in all tree species. Our results suggest that although seed trade might facilitate movements of fungi, their establishment and spread in the new environment might be limited by host availability.

Keywords: plant pathogens; seed-associated fungi; host affinity; site affinity; distribution range; climatic factors

INTRODUCTION

Endophytic fungi are a taxonomically and functionally diverse group occurring within living plant tissues showing no visible signs of presence (Carroll 1986; Petrini 1991; Arnold 2007; Sun and Guo 2012), and they can be involved in various interactions with their hosts (Stone, Polishook and White 2004; Sieber 2007). For example, endophytes of the family Clavicipitaceae are known to increase the fitness of their grass hosts (Poaceae) by providing protection against herbivores and pathogens (Saikkonen *et al.* 1998). On the other hand, endophytes of the family Botryosphaeriaceae can be latent pathogens, and cause diseases in stressed hosts (Slippers and Wingfield 2007). Since fungal endophytes are present in various tissues of probably all living plants, and in all environments, the endophytic mycobiome is expected to be highly diverse (Petrini 1991), but at the same time is poorly described. Previous studies on endophytic tree mycobiomes have mostly focused on foliar endophytes, which are considered more frequent and diverse than seed-associated fungal endophytes (Ganley and Newcombe 2006). However, recent studies revealed that tree seeds do in fact harbor a rich and diverse community of fungal endophytes, including potential plant pathogens (Cleary *et al.* 2019; Franić *et al.* 2019). This is of

Received: 16 April 2020; **Accepted:** 12 August 2020

particular concern because seeds are considered phytosanitary safe and almost no restrictions are applied to their exchange (Anon. 2016). The factors shaping the endophytic mycobiome of tree seeds are still largely unknown, but this knowledge would be beneficial for better understanding the risks associated with the exchange of forest tree seeds.

Endophytic communities associated with individual trees consist of a subset of the fungal species that occur in the surrounding environment, which acts as a main source of fungal inoculum (Arnold 2007). Their formation is mediated by a series of filtering processes imposed by the abiotic and biotic environment (e.g. climatic factors and tree itself, respectively). For example, abiotic factors (like climate) may determine the surrounding vegetation that serves as a source of fungal inoculum, and select fungal phenotypes that tolerate the abiotic conditions in the environment. On the other hand, plant traits (like biochemical and structural defenses) may act as biotic filters that further shape endophytic communities of trees. It is thus to be expected that both abiotic and biotic factors will affect (i) the number of fungal taxa in a given plant (quantitative aspect of fungal diversity, i.e. endophyte richness) and (ii) identity of fungal taxa in a given plant (qualitative aspect of fungal diversity, i.e. community composition). The relative importance of different abiotic and biotic factors in shaping endophyte richness and community composition of fungi in forest tree seeds is, however, yet to be described.

Since climate (i.e. temperature and precipitation) is known to have a significant impact on fungal germination, growth and reproduction (Agrios 2005), it may potentially affect endophyte richness and community composition. Some of the previous studies suggested that moderate temperatures (Vacher *et al.* 2008) and high moisture (Agrios 2005) might benefit fungal pathogens and thus increase pathogen species richness. Other studies found no relationship between mean annual temperature, and precipitation, and fungal richness in leaves or roots (U'Ren *et al.* 2012; Zimmerman and Vitousek 2012; Coince *et al.* 2014). Although abiotic environmental factors showed inconsistent influence on endophyte richness, a significant impact of mean annual temperature (Coince *et al.* 2014), and mean annual temperature and rainfall (Zimmerman and Vitousek 2012) on community composition of foliar tree endophytes was shown. The contradictory effect of climatic factors on endophyte richness and community composition might be due to other relatively more important factors that could determine endophyte richness, such as biotic factors.

Trees of different species differ in their traits related to tree structures and physiological processes, and these traits seem to be evolutionarily conserved (Kraft *et al.* 2007). Host identity can thus strongly affect endophyte communities, frequently in combination with site factors. For example, Higgins *et al.* (2007) suggested high levels of both host and site specificity in fungal communities of photosynthetic tissues of *Picea mariana*, *Dryas integrifolia* and *Huperzia selago* from two boreal and one arctic site. Similarly, Schlegel, Queloz and Sieber (2018) found that communities of foliar fungal endophytes of European ash (*Fraxinus excelsior*) and sycamore maple (*Acer pseudoplatanus*) in Switzerland are characterized by a few abundant and host-specific species. Since endophytes are more likely to colonize closely related than unrelated hosts probably due to a higher fraction of shared traits (Sieber 2007; Gilbert *et al.* 2012), endophytic mycobiomes may be affected by not only the characteristics of their host but also neighboring tree species. For instance, hosts surrounded by closely related species may accumulate more fungal species than hosts that are distantly related to their neighbors. Previous studies have shown positive relationship between plant diversity and diversity of associated organisms (Rottstock *et al.* 2014; Liebhold *et al.* 2018). However, the relationship between endophyte richness and the number of congeneric tree species at a site has rarely been assessed, but it would be useful for better understanding factors that determine the endophytic inoculum of seeds.

The richness and community composition of fungal endophytes associated with a tree species may also depend on whether the specific tree species grows in its native or non-native range. When trees of a given species are moved outside their native range, they may carry only a fraction of the native endophytic mycobiome with them (Mitchell *et al.* 2010). Recruitment of new endophytes in the non-native range might be limited to generalist fungi and might only occur from closely related hosts (Keane and Crawley 2002). Previous studies, mostly based on literature data, suggested lower pathogen richness in the non-native than native range of plant species and identified some of the factors that increase pathogen recruitment in the non-native area, such as the time since the introduction or host geographic range size (Mitchell and Power 2003; Mitchell *et al.* 2010). Previous studies have also shown the differences in fungal community composition between the native and non-native ranges of tree species (Fisher, Petrini and Sutton 1993; Fisher *et al.* 1994). However, these studies focused on a small number of tree species and sites. Thus, comparative information on mycobiomes of multiple species in their native and non-native ranges is still lacking, but would reveal general patterns across habitats and species.

Disentangling the relative importance of host identity, site-specific environmental factors and host distribution range (native vs non-native; later referred to as range) in shaping endophyte communities is challenging because in natural ecosystems these factors often interact. 'Reciprocal transplant experiments' that involve swapping organisms between environments can help to overcome this limitation. In particular, botanical gardens, where native and non-native tree species grow together and are exposed to the same environment (Morales-Rodríguez *et al.* 2019), offer a unique opportunity to study the diversity of plant-associated microorganisms.

In this study, we aimed to assess the relative importance of different factors in determining the richness and community composition of seed-associated fungal endophytes. We looked at host identity, site, including site-specific environmental factors (i.e. mean annual temperature, mean annual precipitation, altitude and the number of congeneric tree species at a site), and native and non-native ranges of host species as predictors, and we analyzed seed-associated fungal endophytes from nine native and non-native tree species belonging to angiosperms and gymnosperms from Europe and North America. We hypothesized that (i) there is a significant positive relationship between endophyte richness and congeneric tree species richness at a site, (ii) endophyte richness is higher in the native than the non-native range of a tree species and (iii) endophyte community composition significantly differs among different hosts, sites, including site-specific environmental factors, and native and non-native ranges.

MATERIALS AND METHODS

Study design

Seeds of nine tree species belonging to three families and six genera of angiosperms and gymnosperms were collected in botanical gardens in Europe and North America in autumn 2016 (Table 1). While seeds of the Asian species *Acer palmatum*

Table 1. Natural distribution range of the nine tree species analyzed in this study and number of seed lots (i.e. 100 seeds of a tree species at a site) per tree species collected in Europe and North America. Number of seed lots per continent of host natural distribution range is bold. Host distribution range for collected seed lots (N and NN for native and non-native, respectively) is also indicated.

	Taxonomy		Continent of seed collection	
Natural distribution range	Family	Group	Europe	North America
Asia			**12**	**4**
Acer palmatum Thunb.	Sapindaceae	Angiosperm	5 NN	4 NN
Larix gmelinii (Rupr.) Kuzen.	Pinaceae	Gymnosperm	7 NN	0 NN
Europe			**34**	**8**
Acer pseudoplatanus L.	Sapindaceae	Angiosperm	10 N	2 NN
Fagus sylvatica L.	Fagaceae	Angiosperm	10 N	2 NN
Picea abies (L.) H. Karst.	Pinaceae	Gymnosperm	6 N	2 NN
Pinus sylvestris L.	Pinaceae	Gymnosperm	8 N	2 NN
North America			**12**	**8**
Acer macrophyllum Pursh	Sapindaceae	Angiosperm	5 NN	3 N
Pinus ponderosa Douglas ex Hook	Pinaceae	Gymnosperm	3 NN	3 N
Tsuga heterophylla (Raf.) Sarg.	Pinaceae	Gymnosperm	4 NN	2 N
Total			58	20

and *Larix gmelinii* were sampled in the non-native range only, seeds of European and North American species were collected in both their native and non-native ranges. *Acer palmatum* and *L. gmelinii* were included in the analyses because their inclusion adds to revealing the differences in fungal communities among tree species and sites (see the 'Results' section). Seeds of angiosperms and cones of gymnosperms were collected from one to five trees per tree species, directly from the branches, by collaborators of the botanical gardens. Angiosperm seeds were cleaned from detritus if necessary. *Fagus* nuts (later referred to as seeds) were extracted from cupules, and *Acer* seeds were detached from the samara's wings. Gymnosperm seeds were extracted from cones. Mechanically cleaned seeds were kept in the freezer at −20°C until further analysis.

Fungal community assessment by high-throughput sequencing (HTS)

A total of 100 seeds per tree species at each site (later referred to as a seed lot; SL) were used for fungal assessment. To eliminate environmental contamination, seeds were first surface sterilized in 0.5% sodium hypochlorite (NaClO) for 5 min, followed by rinsing in sterile water two times for 5 min (Gamboa, Laureano and Bayman 2003) and air drying in a laminar flow cabinet. The surface sterilized seeds were then ground using a batch mill Tube Mill (IKA-Werke GmbH & Co. KG, Staufen, Germany) or with a mortar and pestle, under liquid nitrogen, if seeds were small. Genomic DNA was extracted from 50 mg of ground seed tissue using the DNeasy PowerPlant Pro Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions, including the use of 40 μl of phenolic separation solution to improve DNA yield from phenolic rich samples. One negative control for each batch of DNA extraction was included to ensure the absence of laboratory contamination. DNA concentrations were quantified using the Qubit dsDNA BR Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) on a Qubit 3.0 Fluorometer (Thermo Fisher Scientific) and DNA was diluted to 10 ng/μl. Samples that yielded <10 ng/μl were not diluted. The internal transcribed spacer region 2 (ITS2) of the ribosomal operon was amplified with primers 5.8S-Fung and ITS4-Fung (Taylor *et al.* 2016) following the protocol described in Franić *et al.* (2019). Each sample was amplified in technical triplicates and successful PCR amplification confirmed by agarose gel electrophoresis before and after pooling the triplicates. Both positive and negative PCR controls were included to exclude false positives or negatives. Purification of PCR products, library preparation and sequencing on the Illumina MiSeq v3 platform (Illumina Inc., San Diego, CA, USA) were performed by the Génome Québec Innovation Center at McGill University (Montréal, Canada). Sequence quality filtering, clustering into amplicon sequence variants (ASVs; Callahan, McMurdie and Holmes 2017) and taxonomic assignments were done on a larger dataset consisting of seed samples obtained from trade and botanical gardens (total of 142 samples). Sequence quality filtering and clustering into ASVs were done with a customized pipeline largely based on UPARSE (Edgar 2013) implemented in USEARCH v.8 (Edgar 2010). After quality control, 6 551 214 high quality reads out of 7 860 151 raw reads remained, corresponding to 83% of the original dataset. Taxonomic classification of ASVs was performed using the naïve Bayesian classifier (Wang *et al.* 2007) implemented in MOTHUR (Schloss *et al.* 2009). These steps are described in Franić *et al.* (2019) where fungal diversity in traded tree seeds was analyzed. For the purposes of this manuscript, only 78 samples (seed lots) and corresponding ASV abundances obtained from seed material collected in botanical gardens were used. Raw HTS sequences have been deposited in the Sequence Read Archive (bioproject PRJNA550270, accession numbers: SRX8676992-SRX8677069).

Statistical analysis

All statistical analyses were performed using R (R Core Team 2018) and tests were considered significant at $P < 0.05$.

Endophyte richness

To assess the effects of host identity, site-specific environmental variables and range on fungal richness (number of ASVs per seed lot), generalized linear mixed effect models (glmmTMB function from the glmmTMB package; Brooks *et al.* 2017) were used. Since Poisson models were overdispersed, we assumed a negative binomial distribution for the errors. Group (angiosperms and gymnosperms), mean annual temperature, mean annual precipitation, altitude, number of congeneric tree species in a botanical garden and range (whether the species was sampled

in its native or non-native range) were included in the model as fixed factors, and tree species and botanical garden were included as random effects. This analysis was additionally done for the dataset containing no Asian species to account for possible biases introduced by unbalanced sampling design. The analysis was also done for a filtered dataset containing no rare ASVs. To remove ASVs with low abundance within and among samples (i.e. rare ASVs), all ASVs with <5 reads and in <5 samples were excluded. With this filtering, we eliminated 85% of the ASVs (1829 out of 2147 ASVs) in the dataset.

The number of congeneric tree species in a garden was obtained from Botanic Gardens Conservation International (BGCI; www.bgci.org) for their member gardens. For non-member gardens, we obtained the information directly from the botanical gardens. Climatic factors such as annual mean temperature and annual mean precipitation were extracted from the WorldClim database (www.worldclim.org) at a resolution of 2.5 min. Altitude for each site was obtained from Google earth Pro (www.google.com/earth/).

Endophyte community composition

Differences in community composition among host species, sites and the native and non-native ranges of a species were assessed with permutational multivariate analysis of variance [PERMANOVA (Anderson 2001); adonis2 function from the vegan package (Oksanen *et al*. 2018)]. Significance of each variable was analyzed using marginal tests available in adonis2, which assess the unique impact of each variable when added after all others. Corresponding R^2 values that indicate the size of the effect of each variable were also calculated. Furthermore, we ran an additional PERMANOVA to assess the differences in endophyte community composition caused by host species, site-specific environmental variables and range by decomposing the site into the site-specific environmental variables (i.e. mean annual temperature, mean annual precipitation, altitude, the number of congeneric tree species at a site). PERMANOVAs were run on Sørensen dissimilarity matrix calculated from the whole dataset and from the dataset containing no rare ASVs, as for endophyte richness analysis. As mentioned for endophyte richness, PERMANOVAs were also run for the dataset containing no Asian species.

To visualize the differences in fungal community composition among hosts and sites (continents) we produced a heatmap showing the presence of ASVs in each sample. Furthermore, we performed cluster analyses on samples, and on ASVs to visualize clusters of samples that have similar community composition, and clusters of ASVs that are more likely to occur together. The heatmap was plotted using the function heatmap.plus from the package heatmap.plus (Day 2012), together with the results of cluster analyses. Prior to plotting, data were filtered to exclude ASVs with low abundance within and among samples, as for previous analyses. Additionally, all ASVs appearing in >50% of the samples (eight ASVs in total) were excluded to eliminate ubiquitous ASVs that will not contribute to the variation between samples. The data were then transformed to presence/absence data and used to produce a heatmap. Cluster analysis was performed using the function hclust from the package stats (R Core Team 2018), using the Ward method applied to a Sørensen dissimilarity matrix.

The strength of association between fungal genera and host genera was assessed by indicator species analysis using correlation indices (De Cáceres and Legendre 2009). Correlation indices allowed the assessment of the preference of the fungal genera for different host genera and their combinations, and were calculated with the function multipatt from the indicspecies package (De Cáceres and Legendre 2009). Bipartite graphs were used to visualize interactions between host and fungal genera identified in the study. The bipartite graph was plotted with the function plotweb from the package bipartite (Dormann, Gruber and Fruend 2008) based on an average number of ASVs of each fungal genus per seed lot of each of the host genera. Results of the indicator species analysis (35 fungal genera characteristic for one of the host genera) were incorporated in the visualization. Additionally, FUNGuild (Nguyen *et al*. 2016) was used to assess which fungal genera are known as plant pathogens and this information was incorporated in the same visualization.

RESULTS

Endophyte taxonomy

Fungal sequences were detected in all 78 analyzed seed lots, forming a total of 2147 ASVs. A large proportion of ASVs was rare, with 38% of ASVs occurring in one or two seed lots (513 and 312 ASVs, respectively). A total of 42% of ASVs appeared in >2 and <10 seed lots (902 ASVs). Around 20% of ASVs (420 ASVs) were found in >10 seed lots with only eight ASVs appearing in >40 seed lots (i.e. roughly 50% of seed lots). Most ASVs belonged to the phylum Ascomycota (71%), followed by Basidiomycota (22%), Mucoromycota (1%), Chytridiomycota (<1%) and Olpidiomycota (<1%). About 6% of ASVs could not be assigned to a phylum. The ASVs belonged to a total of 22 classes, with the numerically most important being the ascomycete classes Dothideomycetes (29% of ASVs), Sordariomycetes (16%), Eurotiomycetes (9%) and Leotiomycetes (9%), and the basidiomycete class Tremellomycetes (10%). A total of 1217 ASVs were assigned to 235 genera and 623 ASVs of those to 268 species. Around 43% and 71% of ASVs were not identified to genus or species, respectively.

Endophyte richness

Fungal richness differed between seed lots of angiosperms and gymnosperms with angiosperms having a significantly higher number of ASVs per seed lot than gymnosperms [239 ± 30 vs 123 ± 15 (mean ± se), respectively; Table 2]. Site-specific environmental factors (i.e. mean annual temperature, mean annual precipitation, altitude and the number of congeneric tree species at a site) and range had no significant influence on the number of ASVs per seed lot [181 ± 17 (overall mean ± se); Table 2]. When Asian tree species or rare ASVs were excluded from the analysis, qualitatively similar results were obtained (Table S1, Supporting Information, and Table 2, respectively).

Endophyte community composition

Differences in endophyte community composition were primarily driven by host identity, which explained around two times more variation than site (33–35% vs 17%, respectively; Table S2, Supporting Information; Fig. 1). Less than 10% of the variation was explained by site-specific environmental variables (i.e. mean annual temperature, mean annual precipitation, altitude and the number of congeneric tree species at a site), all of which except the number of congeneric host species at a site have significantly influenced community composition. Host range was not a significant predictor of community composition and it explained only around 1% of the variation (Table S2, Supporting Information; Fig. 1). Qualitatively similar results were

Table 2. Results of generalized linear mixed effect models for the differences in the number of ASVs per seed lot depending on the taxonomic group (angiosperms vs gymnosperms), the number of congeneric tree species at a site, mean annual temperature, mean annual precipitation, altitude at a site, and native and non-native ranges of a host. Shown are degrees of freedom (df), chi square (χ^2) values and *P* values. Significant *P* values ($P < 0.05$) appear in bold. Results of the analyses considering whole dataset and filtered dataset (no rare species contained) are presented.

		Whole dataset		Filtered dataset	
Factor	Df	χ^2	P	χ^2	P
Group	1	13.92	**<0.001**	16.72	**<0.001**
Congeneric tree species richness	1	0.76	0.383	0.08	0.777
Mean annual temperature	1	0.75	0.385	0.30	0.585
Mean annual precipitation	1	2.49	0.114	0.74	0.389
Altitude	1	3.40	0.065	0.20	0.656
Range	1	0.55	0.458	0.81	0.368

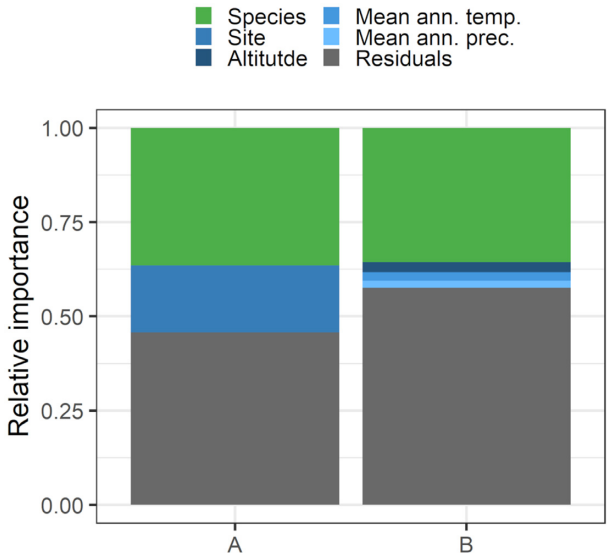

Figure 1. Relative importance of tree species, site and site-specific environmental variables (i.e. altitude, mean annual temperature, mean annual precipitation) in explaining differences in endophyte community composition (R^2; PERMANOVA). The data were analyzed considering **(A)** host species, sites and ranges, and **(B)** host species, site-specific environmental variables (i.e. altitude, mean annual temperature, mean annual precipitation and congeneric tree species richness at a site) and ranges. The effects of the range and the number of congeneric tree species at a site were not shown since they were not significant. The proportion of unexplained variance (residuals) is also indicated.

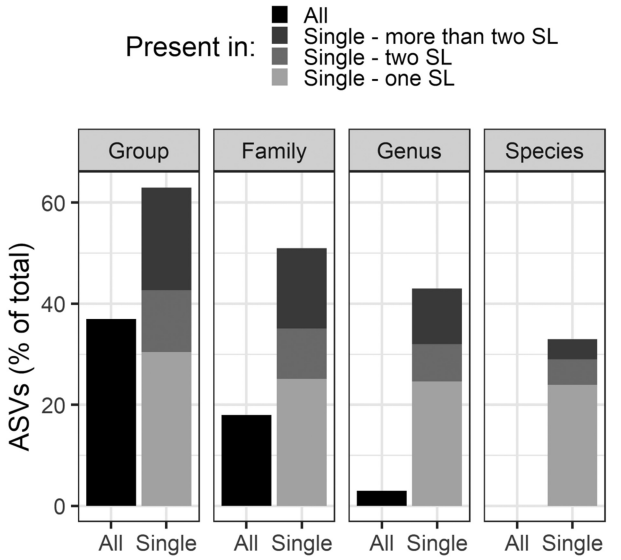

Figure 2. Proportion of the total number of ASVs found in this study occurring in a single or all host groups (angiosperms or gymnosperms), families, genera and species. For ASVs occurring in a single host group, family, genus and species we indicate the number of seed lots (SL) in which ASVs were found. The percentages add to 100% only for the groups since ASVs can only be present in one of the two or both groups. In all other cases, the percentages do not add to 100% since there are more than two families, genera and species.

obtained when rare ASVs or Asian tree species were excluded from the analysis (Tables S2 and S3, Supporting Information, respectively). However, when rare ASVs were excluded from the analysis, host identity explained roughly 10% more variation than when the whole dataset was considered (Table S2, Supporting Information).

Of the 2147 ASVs recorded in this study, around 25% were shared among continents and host groups. Most ASVs (63–33%) were found in a single group, family, genus or species and no ASVs were shared among all tree species. The proportion of fungi associated with a single host group, family, genus or species was greater for higher than lower taxonomic levels (i.e. it decreased from group to species). Furthermore, most of ASVs associated with a single host group, family, genus and species were rare and appeared in one or two seed lots (Fig. 2). Around 58% of all ASVs were recorded either in Europe or North America. European samples contained more than three times as many unique ASVs as North American samples (55% and 15%, respectively). About 36% of the ASVs were found in both the native and non-native ranges of European and North American species.

Host taxonomy was the key driver of differences in fungal community composition (Fig. 3). We filtered rare or ubiquitous ASVs and analyzed the occurrence of the 310 remaining ASVs using a clustering analysis. The first split separated angiosperms from gymnosperms, and the second split separated Fagaceae (*Fagus sylvatica*) from Sapindaceae (*Acer* spp.) within angiosperms and *Larix gmelinii* from other gymnosperms. Only the third split separated samples among continents of collection. The majority of *Acer* spp. samples collected in North America appeared in cluster A together with *Acer palmatum* from Europe, while cluster B consisted almost exclusively of *Acer* spp. samples collected in Europe. All samples of Pinaceae spp. collected in North America clustered within cluster D, while cluster E included *Pinus sylvestris* samples collected in Europe. Lower levels of clustering within cluster D separated *Tsuga heterophylla* (c) from *Picea abies* samples in Europe (b) and *P. abies* and *Pinus* spp. samples collected in North America (a).

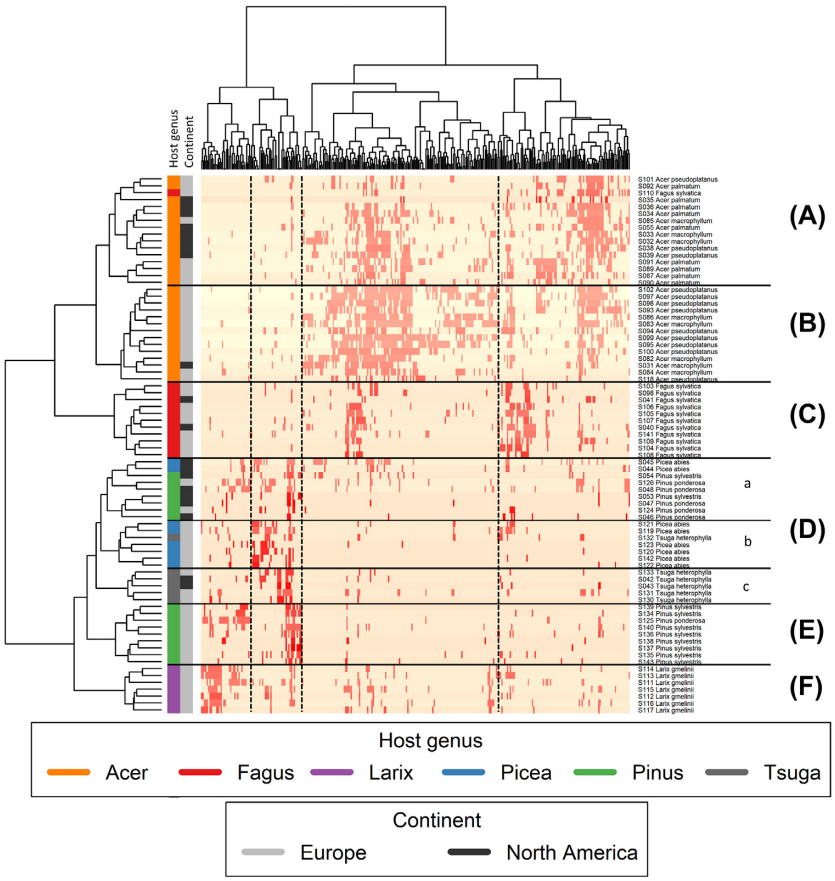

Figure 3. Separation of fungal communities between hosts and origins. The figure shows the distribution of the 310 most characteristic ASVs among samples. Horizontal lines separate samples into clusters of similar ASV community composition and vertical lines separate ASVs into clusters. Capital letters on the right indicate main clusters as a result of third-level branching. Small letters indicate further clustering.

Host taxonomy was equally reflected in ASV clustering with the first split separating ASVs typical for angiosperms and gymnosperms. The second split separated ASVs of Fagaceae from Sapindaceae within angiosperms, and fungi typical for *Larix gmelinii* from those in other gymnosperms.

Similarly, the bipartite plot revealed differences in associated fungal communities between host genera that roughly matched the results obtained from clustering (Fig. 4). The 310 ASVs were assigned to 72 fungal genera. Almost half of the genera (34 out of 72) were characteristic for a single host genus as revealed by indicator species analysis. Additional 13 fungal genera were characteristic for combinations of two fungal genera, and only five fungal genera were characteristic for three or more host genera (Table S4, Supporting Information). *Acer* had at least twice as many associated identified fungal genera and ASVs belonging to these genera per seed lot than any other host genus (Table 3, Fig. 4). *Fagus* and *Larix* had ~50%, while the remaining host genera contained 20–30% of the total number of fungal genera associated with *Acer* (Table 3, Fig. 4). Among the 60 fungal genera found in *Acer* seeds, nine appeared in *Acer* only (Table 3, Fig. 4). Indicator species analysis revealed 24 fungal genera being characteristic for *Acer* (Table 3, Fig. 4; Table S4, Supporting Information). Furthermore, 16 fungal genera found in *Acer* were classified as plant pathogens, and six of them were characteristic for this genus (e.g. *Colletotrichum*, *Septoria* etc.; Table 3, Fig. 4; Table S4, Supporting Information). Seeds belonging to *Fagus* contained only one unique fungal genus among a total of 38 fungal genera (i.e. *Neocladophialophora*; Table 3, Fig. 4). However, four fungal genera were found to be characteristic for *Fagus* species, and two of them were classified as plant pathogens (i.e. *Parastaganospora* and *Muriphaeosphaeria*; Table 3, Fig. 4; Table S4, Supporting Information). Additional nine plant pathogens were found in *Fagus* seeds (Table 3, Fig. 4). *Larix* seeds contained a diverse fungal community including 33 fungal genera, five of which were characteristic for *Larix*, although no unique fungal genera were detected (Table 3, Fig. 4; Table S4, Supporting Information). A total of eight fungal genera were classified as plant pathogens, three of which were found to be characteristic for *Larix* (i.e. *Rhodotorula*, *Celosporium* and *Didymella*; Table 3, Fig. 4; Table S4, Supporting Information). *Pinus*, *Picea* and *Tsuga* had no unique fungal genera among 43, 37 and 18 fungal genera associated with them, respectively (Table 3, Fig. 4). One fungal genus, however, was characteristic for *Pinus* (i.e. *Kabatina*), and one genus for *Picea* (i.e. *Strobilobius*; Table 3, Fig. 4; Table S4, Supporting Information). No characteristic fungal genera were found for *Tsuga* (Table 3, Fig. 4; Table S4, Supporting Information). None of eight, nine and two plant pathogenic fungal genera found in *Pinus*, *Picea* and *Tsuga*, respectively, were characteristic for one of these genera (Table 3, Fig. 4; Table S4, Supporting Information).

DISCUSSION

Endophyte richness

Host identity was the only significant driver of endophyte richness. Our results show significant differences among

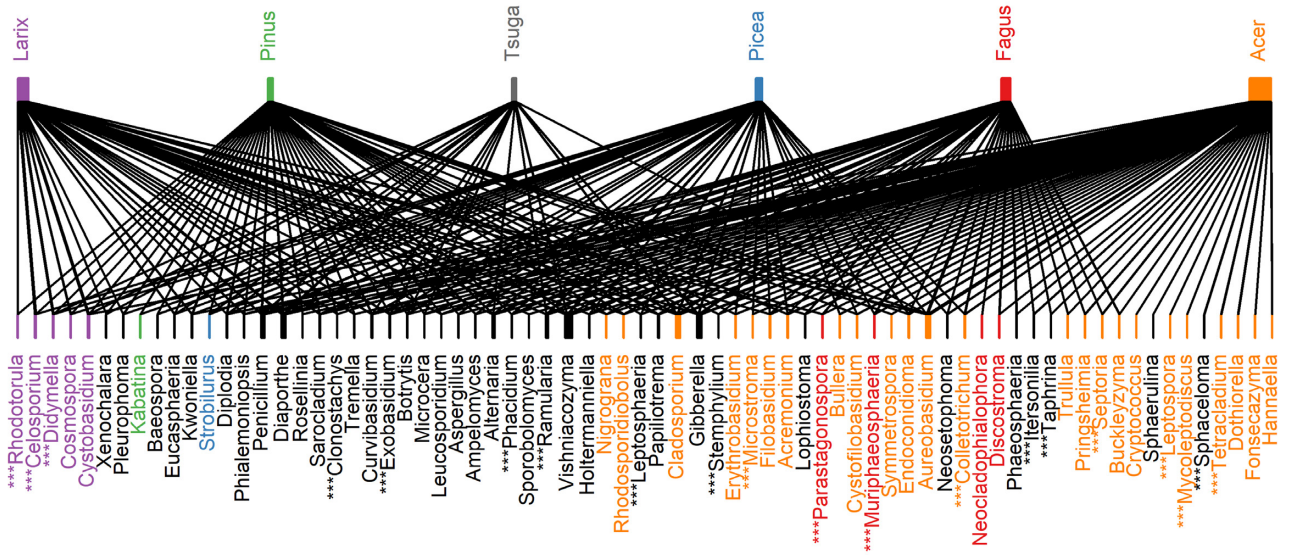

Figure 4. Bipartite association web summarizing the presence of fungal ASVs identified to genus in seed lots of the six host genera. The size of the boxes at the top indicates the number of ASVs associated with each host genus per seed lot. The size of the boxes at the bottom indicates the number of ASVs identified to each genus in association with seed lots of study host genera. The width of lines connecting host genera and fungal genera corresponds to the number of ASVs of each fungal genus per seed lot of host genus. The colors of the boxes at the bottom indicate which fungal genus is characteristic of which host genus, as revealed by indicator species analysis. Fungal genera revealed as containing ASVs that are potential plant pathogens are marked with '***' before the genus name.

Table 3. Fungal diversity associated with six study genera. Number of seed lots (SL), mean number (and standard error) of fungal genera, mean number (and standard errors) of fungal ASVs, total number of fungal genera, number of unique fungal genera, number of characteristic fungal genera as revealed by indicator species analysis, total number of plant pathogenic fungal genera and the number of characteristic plant pathogenic fungal genera as revealed by indicator species analysis per each host genus is indicated.

		Fungal genera		Fungal ASVs			Unique	Characteristic	Plant pathogenic	Characteristic plant
Host genus	SL	Mean	SE	Mean	SE	Total genera	genera	genera	genera	pathogenic genera
Acer	29	21.9	1.3	54.6	3.9	60	9	24	16	6
Fagus	12	11.5	1.0	22.7	2.2	38	1	4	11	2
Larix	7	13.4	0.8	27.4	2.3	33	0	5	8	3
Picea	8	9.0	2.2	16.3	4.4	43	0	1	8	0
Pinus	16	9.2	1.2	12.6	1.9	37	0	1	9	0
Tsuga	6	5.7	1.1	10.7	2.1	18	0	0	2	0

angiosperms and gymnosperms in the richness of seed-associated fungal endophytes, which confirms our previous results (Franić *et al.* 2019). The higher fungal richness in angiosperm seeds is to a large extent driven by the high diversity of *Acer* endophytes. Unlike gymnosperm seeds, *Acer* seeds have no protective structures that may physically hinder the recruitment of fungi from the environment. In the present study, however, one single genus (*Acer*) accounted for the higher diversity of endophytes in angiosperms. Thus, to confirm that angiosperm endophyte communities are more diverse, further investigations including additional angiosperm genera would be required.

For assessing fungal communities associated with seeds of the selected tree species, we used well-established methods. However, since seeds of the different tree species show anatomical and chemical differences (e.g. size, mass, surface texture, chemical compounds), it may be possible that certain biases were introduced at different steps of fungal assessment. For example, surface sterilization and DNA extraction might not have been equally efficient for seeds of all study species. Although such biases are difficult to avoid, using a standard method is widely used approach when comparing mycobiomes of different tree species (Pellitier, Zak and Salley 2019; U'Ren *et al.* 2019).

We found that environmental variables (i.e. mean annual temperature, mean annual precipitation, altitude) do not significantly affect fungal richness in tree seeds. Previous studies also found no effects of environmental variables on fungal richness in other plant tissues (U'Ren *et al.* 2012; Coince *et al.* 2014), suggesting that endophyte richness may be generally only weakly influenced by climate. However, broader environmental gradients than those in our (temperature: 8.1–11.3°C, precipitation: 560–1311 mm, altitude a.s.l.: 19–961 m) or in previous studies (U'Ren *et al.* 2012; Coince *et al.* 2014) might show larger effects of environmental variables on tree endophyte richness. Nevertheless, our results indicate that the host tree environment might be relatively stable and consistent despite different growing conditions (Terhonen *et al.* 2019). Therefore, a given host tree may have a largely constant endophyte richness even under different climatic conditions.

We expected to find a significant positive correlation between the number of congeneric tree species in a botanical garden and fungal richness because pathogen spillovers (host jumps) are most likely to occur between closely related hosts (Gilbert

et al. 2012). However, in our study a higher number of congeneric species at a site did not translate into higher endophyte richness in the seeds of the target host. This could suggest that not only the number of congeneric tree species plays a role but also other characteristics of the surrounding plant community, such as the total plant diversity, age of the trees or length of co-existence. Additional studies looking at the effects of surrounding plant communities on fungal richness in trees are needed for better understanding the factors that contribute to environmental fungal inoculum.

Endophyte community composition

Overall, our results show that the host is the key driver of fungal endophyte communities associated with tree seeds, which is consistent with previous reports showing high host affinity of fungal endophytes associated with photosynthetic tree tissues (Higgins *et al.* 2007; Koyama, Maherali and Antunes 2019; U'Ren *et al.* 2019) and tree bark (Pellitier, Zak and Salley 2019). Our study reveals that seeds of angiosperms and gymnosperms are characterized by different endophytic mycobiomes. The clear separation of the two groups may reflect the early divergence of angiosperms and gymnosperms that resulted in the evolution of different traits characteristic for each group, and consequently, the formation of fungal communities adapted to these traits. Interestingly, the two orders that dominate the endophytic communities in photosynthetic organs of gymnosperms and angiosperms (Helotiales and Diaporthales, respectively; Sieber 2007) are not the most frequent in the metabarcoding-based seed mycobiomes. This seems to confirm previous results by Ganley and Newcombe (2006) showing that the dominant foliar endophytes in a pine species were absent from its seeds. Thus, in addition to strong differences between host species, host tissues may also support different endophyte communities that might be driven by differences in substrate composition and quality (i.e. complex carbohydrates in seeds and simple carbohydrates in photosynthetic tissues).

We also found host effects on fungal communities at lower taxonomic levels of the host. The further separation of fungal communities followed the taxonomic separation of host families, but became less obvious among genera of the same family, as suggested by the absence of unique genera, and low number of characteristic genera in *Picea*, *Tsuga* and *Pinus* within family Pinaceae. However, to further explore the differences in fungal communities among and within lower taxonomic levels of a host, samples from more genera within the same families, and species within the same genera would be needed. Furthermore, it is striking that fungal communities associated with *Larix* were so different from other Pinaceae given the fact that *Larix* is more closely related to *Picea* and *Pinus* than *Tsuga*. Possible reason for this is that *Larix gmelinii* is evolutionarily much younger than other Pinaceae (Wang, Tank and Sang 2000), which might have resulted in evolution of different traits and thus formation of distinct fungal communities. Furthermore, *L. gmelinii* was the only gymnosperm species in this study with Asian origin, and the only species sampled only in non-native range (i.e. Europe). It is thus possible that its fungal community was partially adopted from closely related species not included in our study (e.g. other *Larix* and *Abies* species). This is, however, unconfirmed, and larger study sampling more species at the same site, and same species at different sites would be needed to determine the sources of fungal endophytes for a given tree species. Characterizing fungal communities associated with environmental substrates (e.g. air, herbaceous vegetation and soil) would also help to reveal the origin of fungal tree endophytes.

Fungal community composition differed significantly between sites, which explained around three times less variation in fungal communities than host species (17% vs 33–35% of variance explained). The significant effect of climate and altitude on fungal community composition found in this study is concordant with previous findings (Zimmerman and Vitousek 2012; Coince *et al.* 2014) and suggests that climate and other environmental factors might contribute to determining fungal inoculum at a site. When only site-specific, environmental variables were considered, they explained roughly one-third of the variation as site alone, which suggests that other site-specific factors not included in this study, such as plant community composition or geographic proximity of two sites (Peay, Garbelotto and Bruns 2010; U'Ren *et al.* 2012), might be important drivers of the differences in endophyte communities among sites. Furthermore, fungal traits related to survival and dispersal (e.g. ability to sporulate) might also affect the formation of environmental fungal pools, and thus endophytic communities of trees.

Geographic separation of endophyte communities suggests that further seed movements could potentially lead to introductions of new species. Around one-third of the fungal genera (20 out of 72) identified from a filtered dataset (no rare and ubiquitous ASVs) were classified as potentially plant pathogenic. This supports previous studies indicating that 20–30% of seed-associated fungi were potential plant pathogens (Cleary *et al.* 2019; Franić *et al.* 2019). Hence, our results highlight one more time that potential pathogens could be introduced to new areas via seeds. Given the high host affinity of seed endophytes, their establishment might, however, highly depend on the availability of suitable hosts. If only a small number of trees are cultivated from non-native imported seeds (e.g. botanical gardens, urban parks or private gardens), potential pathogens might establish only if closely related tree species are present in the surrounding. On the other hand, if whole plantations are established from imported seeds, potential pathogens might easily infect other trees in the plantation and form a local population. Moreover, chances for seed endophytes to become established may also strongly depend on the frequency of vertical transmission to the seedling, which is still largely unknown. Pathogens that are not vertically transmitted could, however, survive in the soil, become a part of the environmental inoculum and infect trees at the site.

All seeds analyzed in this study were collected from trees in botanical gardens, which were probably established from native seeds. Reciprocal sampling of the same species in their native and non-native ranges revealed similar fungal richness and community composition between the same hosts in their native and non-native ranges, and between native and non-native species at a site. This suggests that, while hosts may indeed have been introduced with a subset of the fungal community from the native range, they also recruited new fungi in the non-native environment, perhaps from congeneric species.

CONCLUSIONS

Our study clearly shows that host identity is the key driver of fungal richness and community composition of seed-associated endophytes of trees. Although environmental factors had no impact on endophyte richness, endophyte community composition differed between sites, which suggests that the

environment plays a certain role in determining fungal inoculum. Since tree species still show specific distribution patterns (i.e. everything is not everywhere), seed movement may contribute to a certain homogenization of the endophytic tree mycobiomes. However, host jumps could probably happen only between closely related species. Predicting the consequences for tree health of moving seeds is challenging, mostly because we do not know which seed endophytes can be successfully transmitted to the seedlings. Prediction is further complicated by the fact that lifestyle (mutualism, commensalism, parasitism) of most tree endophytes, including the seed-associated ones, is still unknown. Detailed investigations on the extent to which seed endophytes are successfully transmitted and become established in mature trees would be extremely helpful to assess and mitigate the risk of seed movements.

ACKNOWLEDGMENTS

Iva Franić, Eric Allan, René Eschen and Simone Prospero designed the study. Iva Franić supervised seed collections in botanical gardens and arboreta. Salome Schneider optimized molecular laboratory protocols. Iva Franić completed laboratory work. Martin Hartmann supervised amplicon sequencing approach and performed the bioinformatic analyses of the HTS sequence data. Iva Franić, René Eschen and Simone Prospero analyzed the data and wrote the manuscript with an input from all co-authors.

We gratefully acknowledge following arboreta and botanical gardens for providing seeds for this study, as well as all participants that helped with seed collections: Hoyt Arboretum (Portland, OR, USA; Martin Nicholson), Washington Park Arboretum (Seattle, Washington, USA; Raymond J. Larson), Idaho Botanical Garden (Boise, Idaho, USA; Daniel Murphy and Michele Lesica), Leach Botanical Garden (Portland, OR, USA; Courtney Vengarick), UC Botanical Garden (Berkeley, California, USA; Holly Forbes), Arboretum Aubonne (Aubonne, CH; Jean-Paul Degletagne), Botanical Garden of the University of Bern (Bern, CH; Steffen Boch), Forest Botanical Garden Eberswalde (Eberswalde, DE; Andreas Linde, Harald Schill and Bernhard Götz), Royal Botanical Garden Edinburgh (Edinburgh, UK; William Hinchliffe, Martyn Dickson and Katherine Hayden), Arboretum Hørsholm (Hørsholm, DK; Hans Peter Ravn), Meise Botanic Garden (Meise, BE; Anne Ronse), Jardin des Plantes de Paris et Arboretum de Chevreloup (Paris, FR; Frédéric Achille), Arboreti di Vallombrosa (Vallombrosa, IT; Alberto Santini and Giovanni Galipò) and WSL (Birmensdorf, CH). We also thank Dinka Matošević for collecting seed in Zagreb and Jastrebarsko (HR). We acknowledge the contribution of the BGCI (Richmond, UK) in providing contacts and information about botanical garden's plant diversity. We thank Quirin Kupper, Hélène Blauenstein and Stephanie Pfister (WSL, Birmensdorf, CH; Gabor Reiss) for their assistance with laboratory work conducted in Switzerland, and Niklaus J. Grünwald (USDA Horticultural Crops Research Unit, Corvallis, OR, USA), Richard Sniezko (Dorena Genetic Resource Center, USDA Forest Service, Cottage Grove, OR, USA) and Wyatt Williams (Private Forests Division, Oregon Department of Forestry, Salem, OR, USA) and their lab members for their help with laboratory work conducted in the USA. We acknowledge the Genetic Diversity Centre (GDC) at ETH Zurich for providing computational infrastructure and acknowledge the contribution of McGill University and Génome Québec Innovation Center (Montréal, Canada) for pair-end sequencing on Illumina MiSeq. We are grateful to the three anonymous reviewers for the helpful comments on the manuscript.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

FUNDING

This work was supported by the Swiss National Science Foundation [Project C15.0081] and the Swiss Federal Office for the Environment [Grant 00.0418.PZ/P193–1077]. Contributions of Iva Franić and René Eschen were supported by CABI with core financial support from its member countries (see http://www.cabi.org/about-cabi/who-we-work-with/keydonors/).

Conflicts of interest. None declared.

REFERENCES

Agrios G. *Plant Pathology*. 5th edn. Amsterdam: Elsevier Academic Press, 2005.

Anderson MJ. A new method for non-parametric multivariate analysis of variance. *Austral Ecol* 2001;**26**:32–46.

Anon. REGULATION (EU) 2016/2031 OF THE EUROPEAN PARLIAMENT OF THE COUNCIL of 26 October 2016 on protective measures against pests of plants, amending Regulations (EU) No 228/2013, (EU) No 652/2014 and (EU) No 1143/2014 of the European Parliament and of the Council and repealing Council Directives 69/464/EEC, 74/647/EEC, 93/85/EEC, 98/57/EC, 2000/29/EC, 2006/91/EC and 2007/33/EC. 2016.https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R2031 (Accessed March 10, 2020).

Arnold AE. Understanding the diversity of foliar endophytic fungi: progress, challenges, and frontiers. *Fungal Biol Rev* 2007;**21**:51–66.

Brooks ME, Kristensen K, van Benthem KJ *et al*. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J* 2017;**9**:378–400.

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* 2017;**11**:2639–43.

Carroll G. The biology of endophytism in plants with particular reference to woody plants. In: Fokkema N, van den Heuvel J (eds). *Microbiology of the Phyllosphere*. Cambridge: Cambridge University Press, 1986, 205–22.

Cleary M, Oskay F, Doğmuş HT *et al*. Cryptic risks to forest biosecurity associated with the global movement of commercial seed. *Forests* 2019;**10**:459.

Coince A, Cordier T, Lengellé J *et al*. Leaf and root-associated fungal assemblages do not follow similar elevational diversity patterns. *PLoS One* 2014;**9**:e100668.

Day A. heatmap.plus: heatmap with more sensible behavior. *R Packag version 13* 2012. https://cran.r-project.org/web/packages/heatmap.plus/heatmap.plus.pdf (Accessed January 8, 2020).

De Cáceres M, Legendre P. Associations between species and groups of sites: indices and statistical inference. *Ecology* 2009;**90**:3566–74.

Dormann CF, Gruber B, Fruend J. Introducing the bipartite package: analysing ecological networks. *R News* 2008;**8**:8–11.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Edgar RC. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Methods* 2013;**10**:996–8.

Fisher PJ, Petrini O, Petrini LE *et al*. Fungal endophytes from the leaves and twigs of *Quercus ilex* L. from England, Majorca and Switzerland. *New Phytol* 1994;**127**:133–7.

Fisher PJ, Petrini O, Sutton BC. A comparative study of fungal endophytes in leaves, xylem and bark of *Eucalyptus* in Australia and England. *Sydowia* 1993;**45**:338–45.

Franić I, Prospero S, Hartmann M *et al*. Are traded forest tree seeds a potential source of nonnative pests? *Ecol Appl* 2019;**29**:e01971.

Gamboa MA, Laureano S, Bayman P. Measuring diversity of endophytic fungi in leaf fragments: Does size matter? *Mycopathologia* 2003;**156**:41–5.

Ganley RJ, Newcombe G. Fungal endophytes in seeds and needles of *Pinus monticola*. *Mycol Res* 2006;**110**:318–27.

Gilbert GS, Magarey R, Suiter K *et al*. Evolutionary tools for phytosanitary risk analysis: phylogenetic signal as a predictor of host range of plant pests and pathogens. *Evol Appl* 2012;**5**:869–78.

Higgins KL, Arnold AE, Miadlikowska J *et al*. Phylogenetic relationships, host affinity, and geographic structure of boreal and arctic endophytes from three major plant lineages. *Mol Phylogenet Evol* 2007;**42**:543–55.

Keane RM, Crawley MJ. Exotic plant invasions and the enemy release hypothesis. *Trends Ecol Evol* 2002;**17**:164–70.

Koyama A, Maherali H, Antunes PM. Plant geographic origin and phylogeny as potential drivers of community structure in root-inhabiting fungi. *J Ecol* 2019;**107**:1720–36.

Kraft NJB, Cornwell WK, Webb CO *et al*. Trait evolution, community assembly, and the phylogenetic structure of ecological communities. *Am Nat* 2007;**170**:271–83.

Liebhold AM, Yamanaka T, Roques A *et al*. Plant diversity drives global patterns of insect invasions. *Sci Rep* 2018;**8**:1–5.

Mitchell CE, Blumenthal D, Jarosik V *et al*. Controls on pathogen species richness in plants' introduced and native ranges: roles of residence time, range size and host traits. *Ecol Lett* 2010;**13**:1525–35.

Mitchell CE, Power AG. Release of invasive plants from fungal and viral pathogens. *Nature* 2003;**421**:625–7.

Morales-Rodríguez C, Anslan S, Auger-Rozenberg MA *et al*. Forewarned is forearmed: harmonized approaches for early detection of potentially invasive pests and pathogens in sentinel plantings. *NeoBiota* 2019;**123**:95–123.

Nguyen NH, Song Z, Bates ST *et al*. FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* 2016;**20**:241–8.

Oksanen J, Blanchet FG, Friendly M *et al*. vegan: Community Ecology Package. *R Packag version 25-3* 2018. https://cran.r-project.org/web/packages/vegan/vegan.pdf (Accessed January 8, 2020).

Peay KG, Garbelotto M, Bruns TD. Evidence of dispersal limitation in soil microorganisms: isolation reduces species richness on mycorrhizal tree islands. *Ecology* 2010;**91**:3631–40.

Pellitier PT, Zak DR, Salley SO. Environmental filtering structures fungal endophyte communities in tree bark. *Mol Ecol* 2019;**28**:5188–98.

Petrini O. Fungal endophytes of tree leaves. In: Andrews J, Hirano S (eds.). *Microbial Ecology of Leaves*. New York: Springer, 1991, 179–97.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2018.

Rottstock T, Joshi J, Kummer V *et al*. Higher plant diversity promotes higher diversity of fungal pathogens, while it decreases pathogen infection per plant. *Ecology* 2014;**95**:1907–17.

Saikkonen K, Faeth SH, Helander M *et al*. Fungal endophytes: a continuum of interactions with host plants. *Annu Rev Ecol Syst* 1998;**29**:319–43.

Schlegel M, Queloz V, Sieber TN. The endophytic mycobiome of European ash and Sycamore maple leaves – geographic patterns, host specificity and influence of ash dieback. *Front Microbiol* 2018;**9**:1–20.

Schloss PD, Westcott SL, Ryabin T *et al*. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol* 2009;**75**:7537–41.

Sieber TN. Endophytic fungi in forest trees: are they mutualists ? *Fungal Biol Rev* 2007;**21**:75–89.

Slippers B, Wingfield MJ. Botryosphaeriaceae as endophytes and latent pathogens of woody plants: diversity, ecology and impact. *Fungal Biol Rev* 2007;**21**:90–106.

Stone JK, Polishook JD, White JFJ. Endophytic fungi. In: Mueller G, Bills G, Foster M (eds). *Biodiversity of Fungi: Inventory and Monitoring Methods*. Boston: Elsevier Inc., 2004, 241–70.

Sun X, Guo LD. Endophytic fungal diversity: review of traditional and molecular techniques. *Mycology* 2012;**3**:65–76.

Taylor DL, Walters WA, Lennon NJ *et al*. Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. *Appl Environ Microbiol* 2016;**82**:7217–26.

Terhonen E, Blumenstein K, Kovalchuk A *et al*. Forest tree microbiomes and associated fungal endophytes: functional roles and impact on forest health. *Forests* 2019;**10**:1–32.

U'Ren JM, Lutzoni F, Miadlikowska J *et al*. Host and geographic structure of endophytic and endolichenic fungi at a continental scale. *Am J Bot* 2012;**99**:898–914.

U'Ren JM, Lutzoni F, Miadlikowska J *et al*. Host availability drives distributions of fungal endophytes in the imperilled boreal realm. *Nat Ecol Evol* 2019;**3**:1430–7.

Vacher C, Vile D, Helion E *et al*. Distribution of parasitic fungal species richness: influence of climate versus host species diversity. *Divers Distrib* 2008;**14**:786–98.

Wang Q, Garrity GM, Tiedje JM *et al*. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl Environ Microbiol* 2007;**73**:5261–7.

Wang XQ, Tank DC, Sang T. Phylogeny and divergence times in Pinaceae: evidence from three genomes. *Mol Biol Evol* 2000;**17**:773–81.

Zimmerman NB, Vitousek PM. Fungal endophyte communities reflect environmental structuring across a Hawaiian landscape. *Proc Natl Acad Sci USA* 2012;**109**:13022–7.