Environmental factors and host genetic variation shape the fungal endophyte communities within needles of Scots pine (Pinus sylvestris)

Fungal Ecology 57–58 (2022) 101162

Contents lists available at ScienceDirect

Fungal Ecology

journal homepage: www.elsevier.com/locate/funeco

Research Paper

Environmental factors and host genetic variation shape the fungal endophyte communities within needles of Scots pine (*Pinus sylvestris*)

K. Schönrogge [a,*], M. Gibbs [a], A. Oliver [a], S. Cavers [b], H.S. Gweon [a,c], R.A. Ennos [d], J. Cottrell [e], G.R. Iason [f], J. Taylor [g]

[a] UK Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Wallingford, OX10 8BB, UK
[b] UK Centre for Ecology & Hydrology, Bush Estate, Penicuik, Midlothian, EH26 0QB, UK
[c] Biological Sciences, University of Reading, Reading, UK
[d] Institute of Evolutionary Biology, University of Edinburgh, Ashworth Building, King's Buildings, Edinburgh, EH9 3JT, UK
[e] Forest Research, Northern Research Station, Roslin, Midlothian, EH25 9SY, UK
[f] James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK
[g] Royal Botanic Garden Edinburgh, 20A Inverleith Row, Edinburgh, EH3 5LR, UK

ARTICLE INFO

Corresponding Editor: James White

Keywords:
Endophytic fungi
Provenance-progeny trial
Genetic component
Community diversity
Lophodermium
Culture-based sequencing
Culture-free metabarcoding
Core-satellite taxa

ABSTRACT

To determine the role of environmental and host genetic factors in shaping fungal endophyte communities we used culturing and metabarcoding techniques to quantify fungal taxa within healthy Scots pine (*Pinus sylvestris*) needles in a 7-y old provenance-progeny trial replicated at three sites. Both methods revealed a community of ascomycete and basidiomycete taxa dominated by the needle pathogen *Lophodermium seditiosum*. Differences in fungal endophyte taxon composition and diversity indices were highly significant among trial sites. Within two sites, fungal endophyte communities varied significantly among provenances. Furthermore, the communities differed significantly among maternal families within provenances in 11/15 and 7/15 comparisons involving culture and metabarcoding data respectively. We conclude that both environmental and host genetic variation shape the fungal endophyte community of *P. sylvestris* needles.

1. Introduction

The tissues of plants are naturally inhabited by distinct communities of endophytic fungi (Rodriguez et al., 2009; Partida-Martinez and Heil, 2011). Within leaves and needles the foliar endophyte community, dominated by ascomycete and basidiomycete taxa, becomes established following spore-based transmission (Porras-Alfaro and Bayman, 2011). Experimental evidence demonstrates that fungal endophytes can alter the expression of many genes within the tissues they inhabit (Mejía et al., 2014). Consequently, variation in endophyte communities may influence a range of existing host phenotypic traits (Hawkes et al., 2021), including resistance to herbivory and pathogen attack, (Arnold et al., 2003; Mejía et al., 2008; Telford et al., 2014; Ridout and Newcombe, 2015; Christian et al., 2017), and plant - water relations (Rodriguez et al., 2009; Albrectsen and Witzell, 2012). Given the ecological importance of the foliar endophyte community to the host plant, it is vital that we understand the factors that influence its composition.

The composition of naturally occurring foliar endophyte communities is strongly dependent upon the host plant species (Arnold and Lutzoni, 2007; U'Ren et al., 2019; Christian et al., 2020; Romeralo et al., 2022). However, within a host taxon, foliar endophyte communities can vary both among populations, and among individuals within populations (Helander et al., 2007; Peršoh, 2013; Reignoux et al., 2014; Millberg et al., 2015; Bowman and Arnold, 2018). Endophyte community composition within host taxa can potentially be affected both by biotic and abiotic environmental factors, and by host genetic factors. Environmental factors include the local rain of endophyte spores, and climatic variables that interact with both the host and endophyte species to affect the probability of infection of plant tissue (Kraft et al., 2015; Cadotte and Tucker, 2017; Seabloom et al., 2019). Genetic variation within the host plant can also potentially affect the probability of infection and persistence of endophyte species, and hence the composition of the endophyte community (Cordier et al., 2012; Balint et al.,

* Corresponding author.
E-mail address: ksc@ceh.ac.uk (K. Schönrogge).

https://doi.org/10.1016/j.funeco.2022.101162
Received 19 January 2022; Received in revised form 6 May 2022; Accepted 9 May 2022
Available online 21 May 2022

2013, 2015; Rajala et al., 2013; Qian et al., 2018) shaping it as the extended phenotype of the host tree (Whitham et al., 2003; Partida-Martinez and Heil, 2011; Fort et al., 2021). This genetic variation may be present among populations of the host plant species, and/or among individuals within these populations.

To disentangle and compare the influence of environmental and host genetic effects on endophyte community composition, a common garden experimental approach is needed (Elamo et al., 1999). Here host plants from the same array of populations, and families within populations, are grown at multiple sites and exposed to natural infection by endophytes. Comparisons among sites measure the environmental influence over endophyte communities. Comparisons among host populations, or among families within a host population, allow the influence of host genetic variation to be assessed (Balint et al., 2013, 2015; Lamit et al., 2014; Albrectsen et al., 2018).

A major technical issue with implementing these studies is the development of suitable techniques for quantifying, with sufficient depth and accuracy, the fungal endophyte communities residing within large numbers of host individuals. Traditional studies of foliar endophyte communities have been based on culturing of fungi isolated from surface sterilised leaves and needles followed by identification of cultures on the basis of morphology alone, or with additional information from sequencing of selected loci e.g. ITS (Ko et al., 2011; Sun and Guo, 2012; Koljalg et al., 2013). There are several inherent limitations and biases associated with this approach. Unculturable fungi will not be identified, and taxa that grow faster on artificial media are more likely to be recorded. In addition, the technique is very labour intensive and it may be impractical to conduct the large number of isolations required to detect low frequency taxa. Culturing provides a limited window on the foliar endophyte community and may therefore be subject to significant bias.

In recent years the study of fungal endophyte communities has been revolutionised with the development of metabarcoding based on amplification of ITS sequences from DNA extracted from individual samples, including in pines (Peršoh, 2013; Bullington and Larkin, 2015; Millberg et al., 2015; Taudiere et al., 2018, Lazarević and Menkis, 2020). This technique has the ability to detect unculturable and slow-growing taxa, can signal the presence of taxa at very low frequency (Thomas et al., 2016; Zinger et al., 2019), and with suitable caution (Amend et al., 2010), can be used to estimate the relative frequencies of taxa from counts of amplicon number. Rigorous screening of amplified DNA sequence data to remove artefacts is a prerequisite for any such analysis. So too is a comprehensive ITS sequence database of the fungal species that are likely to be present in the study system (Nilsson et al., 2019a) ideally integrated into software designed to link ITS sequence information to fungal taxa (Gweon et al., 2015). If these prerequisites are met, metabarcoding can provide an extremely powerful method for quantifying fungal endophyte communities in large numbers of samples.

In this study we jointly apply culturing and metabarcoding techniques to the same set of samples to study the degree to which environmental and genetic factors affect the fungal endophyte community present in the needles of Scots pine, *Pinus sylvestris*. This study system has been chosen because there is already a wealth of information on the fungal endophytes inhabiting *P. sylvestris* needles (Minter, 1981; Kowalski, 1982; Helander et al., 1994; Millberg et al., 2015; Oono et al., 2015; Taylor et al., 2019; Rim et al., 2021). Furthermore, a comprehensive database of ITS sequences has been compiled for these taxa, UNITE (https://unite.ut.ee/), and software has been developed to link ITS sequences generated in metabarcoding to species identities (Gweon et al., 2015).

The samples analysed here are derived from an experiment in which the same set of Scots pine populations (provenances) and families within these provenances have been planted at three sites within Scotland and allowed naturally to develop needle endophyte communities. Our study design enables us to investigate three important questions. The first is the extent to which the environment occupied by the host affects the community of fungal endophytes that colonise its needles. The second is whether genetic differences among provenances, and/or among families within these provenances, affect endophyte community composition. The third is the degree to which the endophyte communities revealed by culturing and metabarcoding are congruent. Although there are previous studies that have compared endophyte communities assessed by culturing and metabarcoding (Johnston et al., 2017; Dissanayake et al., 2018; U'Ren et al., 2019; Hoyer and Hodkinson 2021; Oita et al., 2021), no such studies have been carried out for *Pinus*.

2. Materials and methods

2.1. Scots pine trial

A common garden provenance-progeny trial comprising material of Scots pine (*Pinus sylvestris*) sampled from across its natural range in Scotland was replicated at three contrasting planting sites (Fig. 1). These sites were:

1. Glensaugh in northeast Scotland with a relatively cool/dry climate. The planting site is former agricultural land adjacent to a plantation of mature *P. sylvestris*.
2. Inverewe in northwest Scotland with a relatively warm/wet climate. The site was previously occupied by a clearfelled plantation of Sitka spruce (*Picea sitchensis*) and *Pinus contorta*.
3. Yair in southeast Scotland with a relatively warm/dry climate. The site was previously an *Abies* plantation and is now surrounded by plantations of *Picea sitchensis* and *P. contorta*

In 2007, to establish the trial, open-pollinated seed was collected from 8 mother trees in each of 21 native Caledonian pinewood populations (provenances). Of these, five were sampled for needle endophytes in this research: Beinn Eighe (−5.349°E, 57.630°N), Glen Affric (−4.921°E, 57.273°N), Glen Loy (−5.130°E, 56.908°N), Glen Tanar (−2.862°E, 57.048°N), and Rhidorroch (−4.979°E, 57.894°N) (Fig. 1). The rationale for choosing these five provenances was that they cover the full climatic range occupied by *P. sylvestris* within Scotland and

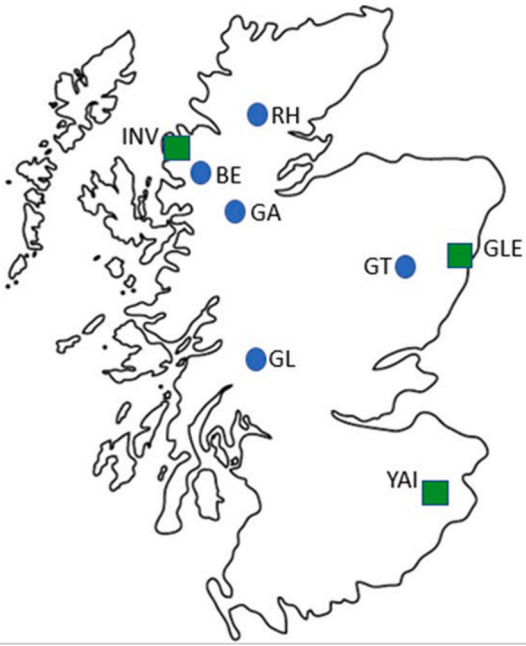

Fig. 1. Geographic locations of the five provenances of *Pinus sylvestris* sampled from the provenance/progeny trial (circles) GA, Glen Affric; GL, Glen Loy; BE, Beinn Eighe; RH, Rhidorroch; GT, Glen Tanar and the three trial sites (squares) INV, Inverewe; GLE, Glensaugh; YAI, Yair.

are located in five of the seven biochemical Scots pine seed zones (Salmela et al., 2010). Seeds were sown in 2007 and grown on at the James Hutton Institute, Aberdeen, UK. Saplings were planted out in 2012 in a balanced, randomised block design at each of the three sites, with a single sapling from every family in every block, and three (Inverewe) or four (Glensaugh, Yair) blocks per site.

2.2. Needle sampling and processing

Sampling of needles from the three common garden trials was undertaken over a 14-d period in 2014 (Glensaugh 7th September, Inverewe 20th September, Yair 15th September) when trees were 7 y old. At each trial site, we aimed to sample needles from three trees from each of the eight families in the five provenances, making a total of 360 samples. However, due to early tree mortality (deer damage) we were only able to collect samples from 316 trees (Glensaugh n = 115, Inverewe n = 94, Yair = 107).

From each sampled tree, nine pristine, apparently healthy needles of the previous years growth flush (spring 2013, 16 months prior) were randomly collected across all aspects and heights of each tree and kept refrigerated for a maximum of 5 d prior to processing. We deployed a comprehensive surface sterilization and treatment protocol to ensure that we assayed only true endophytes within needle tissue. Glassware used in this process was cleaned in a Miele laboratory dishwasher with Neodisher detergent using sterile distilled water throughout. Whole needles were surface sterilised by immersion for 1 min in 70% ethanol followed by 5 min in 3.5% sodium hypochlorite, 30 sec in 70% ethanol, and finally, 3 min in 0.5 ml/L Tween20 solution in an ultrasonicator. After surface sterilization the nine needles from each tree were dissected into 5 mm length fragments.

2.3. Endophyte culturing and identification with Sanger sequencing

A single 5 mm fragment from each of the nine surface sterilised needles per plant was plated onto malt extract agar (MEA, Fluka, Sigma Aldrich, Germany) supplemented with 0.3 g/L streptomycin sulphate (Sigma Aldrich, Germany) to culture endophytes. Fungal isolates growing from sterilised needle fragments were sub-cultured, with the plates checked regularly for further fungal development over a period of 3 months.

Sub-cultured endophyte colonies were designated as morphotypes, and the number of each of the different morphotypes isolated from the nine needle pieces plated for each tree was recorded. DNA was extracted from individual morphotypes for identification by Sanger sequencing of the ITS region using primers ITS1F (Gardes and Bruns, 1993) and ITS4A (Larena et al., 1999). Morphotypes were assigned to taxa using information from the UNITE database (Nilsson et al., 2019a), and taxonomic names for all ranks are also aligned to this database.

2.4. DNA extraction for metabarcoding

A further 5 mm fragment was taken from each of the nine surface sterilised needles per plant, ensuring that three tip-, base- and mid-sections were included, to create a pooled sample for each individual tree. The pooled needle fragments were placed in 2 ml Eppendorf tubes and stored at −80 °C before DNA extraction. All samples from a trial site were processed within 4 days.

Prior to DNA extraction samples were freeze dried (Christ Alpha 1–4 LD plus, Osterode im Harz, Germany) and finely powdered using a Qiagen Tissue Lyser (LT, Hilden, Germany) at 50Hz for 2 min. DNA extraction was undertaken using DNeasy Plant Mini Kits (Qiagen, Hilden, Germany) as per the manufacturer's instructions, with the following modifications. After addition of AP buffer, samples were placed for 1 h on a Thermomixer C (Eppendorf, Germany) at 800 rpm, and as the final step, two volumes of 50 μl TE were eluted through the filter provided in the kit.

2.5. Amplicon library construction and sequencing for metabarcoding

For each of the 304 successfully extracted pooled DNA samples a fungal ITS region 2 (ITS2) gene library was constructed. We employed the dual indexing strategy of Kozich et al. (2013) utilising the fITS7 (forward; GTGARTCATCGAATCTTTG) and ITS4 (reverse: TCCTCCG CTTATTGATATGC) primers described in Ihrmark et al. (2012), which anneal to the 5.8S and LSU rRNA genes flanking the ITS2 region. Briefly, each primer consisted of the appropriate Illumina adapter, an 8-nt index sequence, a 10-nt pad sequence, a 2-nt linker and the gene-specific primer. Triplicate amplicons were generated using a high-fidelity DNA polymerase (Q5 Taq; New England Biolabs) and pooled. PCR was conducted on 10 ng of template DNA employing an initial denaturation of 30 sec at 95 °C, followed by 30 cycles of 30 sec at 95 °C, 30 sec at 52 °C and 2 min at 72 °C. A final extension of 10 min at 72 °C was also included to complete the reaction.

Amplicons were quantified using the Agilent 2200 Tape Station bioanalyser, and an equimolar pool (library) was prepared prior to purification by gel extraction (QIAEX II; Qiagen). The final concentration of the library was calculated using a SYBR green quantitative PCR (qPCR) assay with primers specific to the Illumina adapters (Kappa, Anachem).

The ITS2 library was sequenced at a concentration of 5.4 pM with a 0.6 pM addition of an Illumina generated PhiX control library. Sequencing runs, generating 2 x 300 bp reads were performed on an Illumina MiSeq using V3 chemistry. The read 1 (R1), read 2 (R2) and index sequencing primers used were also ITS specific: R1 = sequence of the combined pad, linker and fITS7; R2 = sequence of the combined pad, linker and ITS4; I = reverse complement of the R2 primer (See Fig. S1 in Gweon et al., 2015). Sequencing generated a total of 24,469,237 paired-end sequences across the 304 samples. All data used for this study are available at https://www.ebi.ac.uk/ena/data/view/PRJEB21684. Two samples, INVE271 & INVE231, produced less than 1000 reads and were not included in any further analyses.

2.6. Converting sequence data to OTUs

The raw sequence data were processed using the pipeline PIPITS (https://github.com/hsgweon/pipits) described in detail by Gweon et al. (2015). All processing steps were performed with PIPITS default settings. Briefly, all raw read pairs were joined at the overlapping region and then quality filtered, chimera filtered, singleton filtered, contaminant filtered, merged, and clustered into OTUs defined at 97% sequence similarity. The taxonomic assignment of OTUs was performed using RDP Classifier (Wang et al., 2007) against the UNITE 7.0 fungal reference data (Koljalg et al., 2013) at an RDP bootstrap threshold of 85%. The resulting table consisted of a total of 7610 OTUs.

2.7. Rarefaction of OTU data

Using this dataset we inspected the relationship between the number of ITS sequences generated per tree sample (a measure of the intensity of sampling for endophytes) and OTU richness by generating rarefaction curves using the specaccum function with a random method in the Vegan package 2.3–5 (Oksanen et al., 2016) in R (v. 3.1.3) (R Core Team, 2015). The number of OTUs was uneven between samples and thus the dataset was rarefied to 1138, the smallest number of reads in any sample using rrarefy in Vegan v.2.3–5 (Oksanen et al., 2016). Rarefying resulted in a reduction to 2079 OTUs.

2.8. Assigning taxa from OTU data

Within the rarefied dataset some of the 2079 OTUs have unresolved taxonomic assignments, i.e. matches at lower taxonomic levels that fall below the confidence threshold applied in PIPITs (RDP bootstrap threshold ≥ 85%). In some cases, multiple OTUs were assigned to the

same species hypotheses (SH) within the reference database UNITE. For these, our strategy was to pool the OTUs referring to the same SH and remove the ones not assigned to species level; 821 taxa were retained of which 600 (73%) have matches to specific, named UNITE species hypotheses. The remaining 27% were matched to a sequence in UNITE but the sequence was of an unknown species. Where a taxon at a frequency of >1% fell into this latter category, further DNA database searches were made in GenBank to determine whether a species identity had very recently been assigned to the sequence. Subsequent analyses were conducted on the rarefied dataset that recorded the number of reads for each of the 821 taxa for each of the trees sampled. Taxonomic names at all ranks are aligned with the UNITE database.

3. Data analysis

3.1. Data from cultures

3.1.1. Number of isolates

Analysis of variance was used to determine the effect of trial site, provenance, and family within provenance on the total number of endophyte isolates obtained from each tree sample after 3 months culturing.

3.1.2. Endophyte community composition and diversity

Alpha diversity at each site was measured as taxon richness and Shannon diversity index using data on total number of isolates of each taxon at each site.

3.1.3. Effect of site, provenance, and family within provenance on endophyte community composition

Contingency $\chi 2$ analysis was used to determine the significance of differences in the relative frequency of endophyte taxa among trial sites, among provenance within each site, and among families within provenance and site. For tests at the provenance level, the categories compared were each of the seven most common taxa together with a category containing all other taxa combined. At the provenance and family levels, where relatively small numbers of most taxa were present, difference in the relative frequencies of taxa were tested by collapsing the data into two categories: the most common taxon present and all remaining taxa.

3.2. Data from metabarcoding

3.2.1. Endophyte community composition and diversity

The rarefied read numbers for the dataset of 821 taxa were used to quantify taxon abundance for analyses of alpha diversity (Taylor et al., 2016). Taxon richness and Shannon diversity were calculated for each site using combined data from all the trees at that site. Taxon richness, chao1 estimates of taxon richness, and Shannon diversity were also calculated for individual trees to test for significant differences in alpha diversity of individual trees among sites.

3.2.2. Distinguishing core and satellite taxa

Core taxa, those that consistently occupy the environment of our Scots pine needle samples, were distinguished from opportunistic satellite taxa using the approach of Magurran and Henderson (2003) and van der Gast et al. (2011). This posits that the abundance distribution for core taxa follows a lognormal distribution, while that for satellite taxa matches a log-series or exponential distribution. We applied their method by jointly analysing the number of trees (samples) with a taxon present (occupancy), and the maximum abundance it reaches in any of the samples (both log-transformed). A piecewise linear model was employed where the break point was moved along the x-axis, (log (occurrence)), to identify the point where residual variance was minimised (Crawley, 2007). Species below that point were putative satellite species and those above it were putative core species. Abundance distributions in both groups were assessed using the fitdist and gofstat functions in the R package fitdistrplus (Delignette-Muller and Dutang, 2015).

3.2.3. Effect of site, host provenance, and host family within provenance on endophyte community composition

To assess the impact of site, host provenance and host family within provenance on variation in community composition, we employed multivariate general linear models using methods available in the package mvabund 3.12 (Wang, Eijkemans, et al., 2012; Wang, Naumann, et al., 2012). Here the abundance patterns for each taxon were analysed separately for the effect of the relevant factor, and the results summarised across the community (Wang, Naumann, et al., 2012). Consequently, unlike dissimilarity-matrix-based methods, multivariate GLMs do not confound location with dispersion effects, which can inflate type 1 and 2 error (Warton et al., 2012).

The models on which mvabund 3.12 is based do not allow for analysis of nested designs as used in our experiment. Therefore, multivariate general models were applied separately to analyse variation in community composition at three different levels; among the three sites, among host provenances within each site, and among families within each provenance at each site. The data used in the analyses were rarefied read counts for each of the 821 taxa recorded for each tree. Models were specified with negative-binomial errors and significance was assessed using permutation tests with Monte Carlo resampling based on log-likelihood ratios.

4. Results

4.1. Data from cultures

4.1.1. Taxa detected by culturing

Fungal endophytes were isolated from needle samples for 308 of the 316 trees sampled (Glensaugh n = 115, Inverewe n = 94, Yair = 107) yielding a total of 2436 isolates, grouped into 11 morphotypes. The vast majority of the taxa isolated were ascomycetes (98.3%), the remainder being basidiomycetes. Morphotyping in combination with Sanger sequencing of ITS allowed 96% of the ascomycete samples to be classified into 9 taxa, 6 at the species level (*Lophodermium seditiosum*, *L. pinastri*, *L. conigenum*, *Cenangium ferruginosum*, *Sydowia polyspora*, *Dothistroma septosporum*), 1 at the level of genus (*Preussia*) and 2 at the level of family (Xylariaceae, Pleosporaceae). Remaining unidentified isolates, classified as 'Other', represented singletons, basidiomycetes yeasts plus other low frequency taxa and comprised 5.2%, 8.5% and 6.7% of the total isolates at Glensaugh, Inverewe and Yair respectively.

4.1.2. Number of isolates obtained from culturing

Analysis of variance indicated a highly significant effect of site ($F_{(2,266)} = 53.35$, $P < 0.001$), but no effect of tree provenance or family within tree provenance on the total number of fungal isolates obtained from each tree. The mean number of fungal cultures isolated was significantly greater at Glensaugh (10.24, SE = 0.335) than at either Inverewe (6.36, SE = 0.356) or Yair (6.16, SE = 0.261).

4.1.3. Endophyte community composition based on data from culturing

The needle pathogen *L. seditiosum* was by far the most abundant taxon overall, and was the dominant taxon at both Glensaugh and Inverewe (Fig. 2A). The endophyte, *L.pinastri*, that fruits on dead needles, was also an important component of the community at all sites. The only other taxa at appreciable frequency were *Cenangium ferruginosum*, which reached a frequency of 24.9% at Yair, and a *Xylariaceae* taxon that was found at a frequency of 11.7% at Inverewe.

4.1.4. Endophyte community diversity based on data from culturing

The taxon richness as measured by culturing techniques was very similar among the three sites, with 10, 10 and 11 taxa isolated from

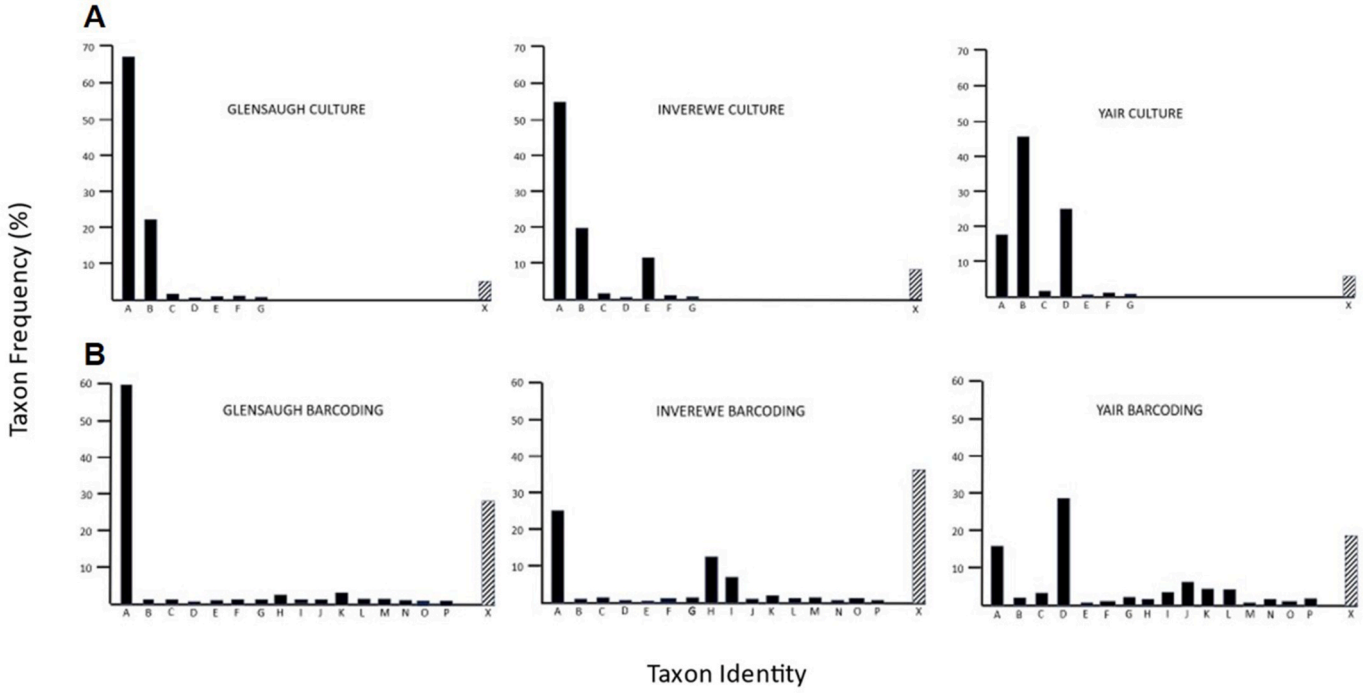

Fig. 2. Relative frequency (per cent) of fungal endophyte taxa in samples from a provenance/progeny trial of *P. sylvestris* replicated at three trial sites (Glensaugh, Inverewe and Yair). Shown are taxa whose frequency exceeded 1% and were identified using either a. culture (seven out of eleven identified taxa) or b. metabarcoding. Key to taxa: A – *Lophodermium seditiosum*, B – *L. pinastri*, C – *L. conigenum*, D – *Cenangium ferruginosum*, E – *Xylariaceae*, F - *Pleosporaceae*, G - *Sydowia polyspora*, H - *Helotiales* OTU2253, I – *Phaeomoniella*, J. - *Phaeotheca fissurella*, K – *Capnodiales*, L – *Chaetothyrales*, M - *Perusta inaequalis*, N - *Pyrenopeziza revincta*, O – *Cyclaneusma minus*, P – *Davidiellaceae*, X - Other.

Glensaugh, Inverewe and Yair respectively. However due to the dominance of *L. seditiosum* at Glensaugh, the Shannon diversity index for this site (1.04) was considerably lower than that for either Inverewe (1.40) or Yair (1.43).

4.1.5. Effect of site, provenance, and family within provenance on endophyte community composition

There were highly significant differences among sites in the relative proportions of the complete set of taxa identified in the culturing study (Table 1, $P < 0.001$). Within each site, the frequency of the most common endophyte taxon (*L. seditiosum* for Glensaugh and Inverewe, *L. pinastri* for Yair) was significantly different among provenances at Glensaugh ($P < 0.001$) and Yair ($P < 0.05$) but not at Inverewe. When tested within a site, the frequency of the most common endophyte taxon differed significantly among families within a provenance in 11 of the 15 $\chi 2$ tests ($P < 0.05$).

Table 1
Results of X^2 tests for differences in the relative proportions of all taxa (among sites) or the most abundant taxon (*Lophodermium seditiosum* for Glensaugh and Inverewe, *L. pinastri* for Yair) in endophyte communities assessed using culturing. Analyses were conducted among sites, among provenances within site, and among families within provenance at each site. n.s. non-significant, * - P<0.05, ** - P< 0.01, *** - P<0.001.

Among Sites	$X^2_{(14)} = 908.12$ ***		
Site	Glensaugh	Inverewe	Yair
Among Provenances	$X^2_{(4)} = 40.5$ ***	$X^2_{(4)} = 1.33$ n.s.	$X^2_{(4)} = 11.50$ *
Among Families within Provenance			
Glen Tanar	$X^2_{(7)} = 22.17$ **	$X^2_{(6)} = 12.19$ n.s.	$X^2_{(6)} = 13.30$ *
Beinn Eighe	$X^2_{(7)} = 16.84$ *	$X^2_{(7)} = 27.15$ ***	$X^2_{(6)} = 2.47$ n.s.
Glen Affric	$X^2_{(7)} = 19.14$ **	$X^2_{(7)} = 20.37$ **	$X^2_{(7)} = 20.15$ **
Rhidorroch	$X^2_{(7)} = 15.56$ *	$X^2_{(6)} = 4.91$ n.s.	$X^2_{(7)} = 17.18$ *
Glen Loy	$X^2_{(7)} = 50.24$ ***	$X^2_{(7)} = 9.16$ n.s.	$X^2_{(6)} = 17.58$ **

4.2. Data from DNA metabarcoding

4.2.1. Diversity of taxa detected by metabarcoding

Taxa detected by metabarcoding belonged to five different fungal phyla: Ascomycota, Basidiomycota, Zygomycota, Chitridiomycota and Glomeromycota (Fig. 3). The community was dominated by members of the Ascomycota, which accounted for 18 of the 20 most abundant genera and 98.4% of the amplicons from the top 20 genera.

4.2.2. Endophyte community composition measured with metabarcoding data

The total number of endophyte taxa detected by metabarcoding was extremely high (821), two orders of magnitude larger than the number of taxa identified through culturing. The vast majority of those 810 taxa not detected by culturing were present at a read frequency well below 1% (Supplementary Table 1). However, nine taxa detected only by metabarcoding had overall frequencies above 1%, and one, a taxon in the Helotiales, was found at a frequency as high as 13% at Inverewe. Despite the predominantly very low frequency of taxa detected only by metabarcoding, their very large number means that together they comprised a substantial fraction of the total endophyte community detected (32%, 66% and 55% of total reads at Glensaugh, Inverewe and Yair respectively).

Compared to data from culturing metabarcoding identified nine additional taxa, whose overall read frequency reached at least 1% (Fig. 2B). At Glensaugh and Inverewe the pathogen *L. seditiosum* shows the highest read frequency, and though not the dominant taxon at Yair, it is present at a relatively high frequency of 16%. This pattern shows congruence with the results from culture isolations (Fig. 2A). At Yair, but not elsewhere, a high proportion of reads come from *C. ferruginosum*, again reflecting the distribution of this taxon found in the culturing study. In contrast, the endophyte *L. pinastri*, which was isolated at a substantial frequency through culturing at all three sites, shows a uniformly very low number of metabarcoding reads; and the high number

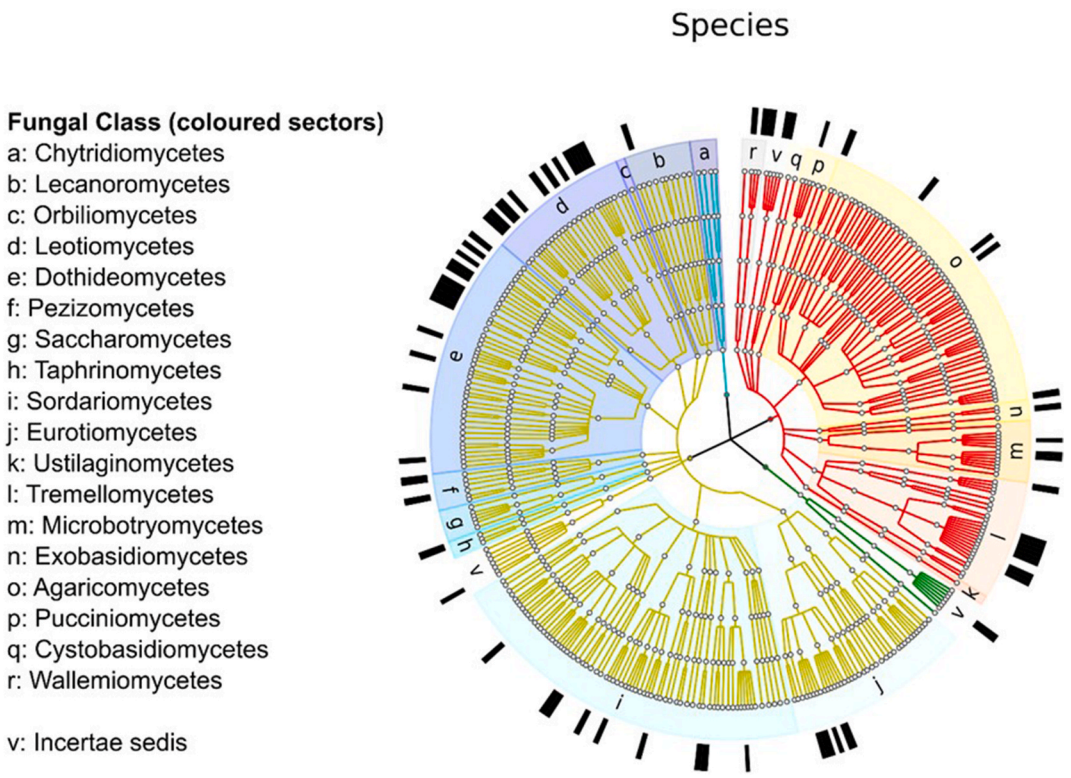

Fig. 3. Taxonomic tree of endophytes found in needles of *P. sylvestris* by metabarcoding. The colour of branches indicates phyla: Ascomycota – yellow; Basidiomycota – red; Zygomycota – green; Chytridiomycota – cyan. Sector colours indicate classes (see legend). Bars on the outer ring mark core taxa.

of cultures of a Xylariaceae taxon isolated at Inverewe is not reflected in a correspondingly high read count in metabarcoding data.

4.2.3. Endophyte community diversity measured with metabarcoding data

At the level of the total endophyte community present within each site, taxon richness was greatest at Glensaugh, intermediate at Yair, and lowest at Inverewe (586, 385 and 292 taxa respectively) (inset to

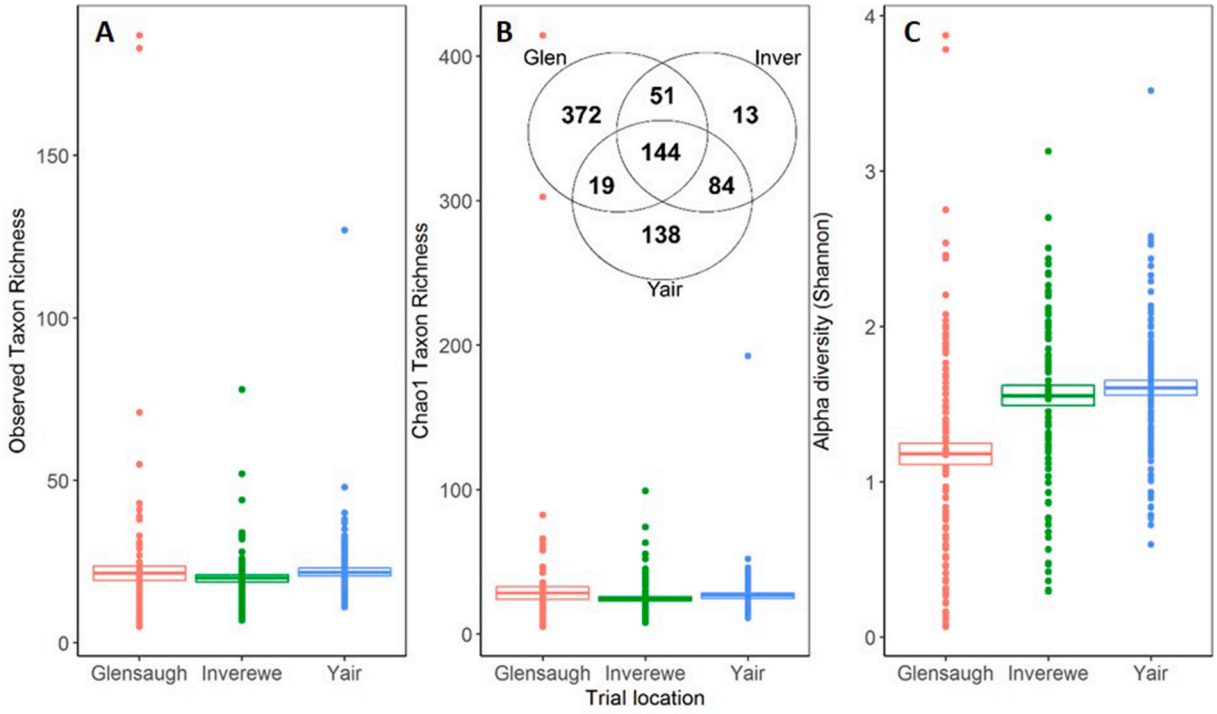

Fig. 4. Measures of taxon richness and α diversity of endophyte communities in sample trees at the three trial sites measured using data from metabarcoding. (A) Observed taxon richness, (B) Chao1 estimate of taxon richness and (C) Shannon diversity. The horizontal line and box indicate the mean ± 1 SE. The inset in (B) shows the number of taxa observed at each site and shared between trial locations. (A further breakdown of the results by provenance can be found in Suppl. Fig. 1).

Fig. 4B). Only 17% of taxa were present at all sites, with 19% present at two sites and 64% at one. In terms of Shannon diversity, the pattern was different. Glensaugh had a much lower diversity index (2.27) than either Inverewe (3.05) or Yair (2.95). This reflected the dominance of *L. seditiosum* at Glensaugh compared with the other sites.

At the level of the endophyte community present within individual trees at a site, there were no significant differences in taxon richness or Chao1 estimates of taxon richness among the sites with a mean of 21.22 (±1.0) and 26.73 (±4.7) taxa per tree respectively (Fig. 4A and B). However, community diversity within trees as measured by the Shannon index was lower at Glensaugh (mean 1.188 ± 0.068) compared with both Inverewe (1.551 ± 0.067) and Yair (1.605 ± 0.048) ($P < 0.001$) (Fig. 4C). Again, this reflects the dominance of *L. seditiosum* at Glensaugh.

4.2.4. Core and satellite species

When a piecewise regression of log(occurrence) vs log(abundance) was conducted for the 821 taxa resolved by metabarcoding a breakpoint was indicated where the mean square error (MSE) was minimised at log (occurrence) = 1.39, implying that core species were those occurring in more than 4 samples (Fig. 5A). The piecewise model provided a better fit than a simple linear model despite the additional parameters (likelihood-ratio test $p < 0.001$, AICpiecewise = 1945.8 < 1962.6 AIClinear). The species abundance distribution for the core fungal endophyte community from Scots pine needles was well described by a lognormal distribution (Kolmogorov – Smirnov D = 0.08, $p > 0.05$) (Fig. 5B). That for the satellite species was significantly different from an exponential distribution (K–S D = 0.41, $p < 0.05$), though it provided a much better fit than to a lognormal distribution (AICexp = 1018 < AIClnorm = 1586).

The piecewise model differentiated 120 core species from the remaining satellite or opportunistic species (Supplementary Table 2). Fig. 3 indicates the distribution of the core taxa at genus level across the fungal taxonomy. While there were clusters within some classes such as the Leotiomycetes or Dothideomycetes, core species were only missing from fungal classes that were generally not well represented. Of the 120 core species, 11 occurred at only one of the three trial locations, 15 species were shared between Glensaugh and Inverewe, 18 species between Glensaugh and Yair, while eight species were shared between Yair and Inverewe. Sixty-eight core species were found at all three trial locations.

4.2.5. Effect of site, provenance, and family within provenance on endophyte community composition

Probability values derived from Monte Carlo resampling show that there were highly significant differences in endophyte community composition, as measured by rarefied read abundance, among sites (Table 2, $P < 0.001$).

Within sites, communities differed significantly among provenances at Inverewe ($P < 0.05$) and Yair ($P < 0.05$) but not at Glensaugh. There was no evidence that endophyte communities differed among families within provenances at Inverewe, but at Glensaugh and Yair significant

Table 2
Results of general linear models run in mvabund (Wang et al., 2012) using rarefied amplicon abundance data to test for differences in endophyte community composition among sites, among provenances within each site, and among families within provenance at each site. Probability levels derived from Monte Carlo resampling simulations are shown. n.s. - non-significant, * - P<0.05, ** - P< 0.01, *** - P<0.001.

Among Sites	0.001***		
Site	Glensaugh	Inverewe	Yair
Among Provenances	0.093 n.s.	0.013*	0.021*
Among Families within Provenance			
Glen Tanar	0.038*	0.127 n.s.	0.019*
Beinn Eighe	0.018*	0.077 n.s.	0.042*
Glen Affric	0.363 n.s.	0.153 n.s.	0.030*
Rhidorroch	0.031*	0.089 n.s.	0.004**
Glen Loy	0.289 n.s.	0.120 n.s.	0.111 n.s.

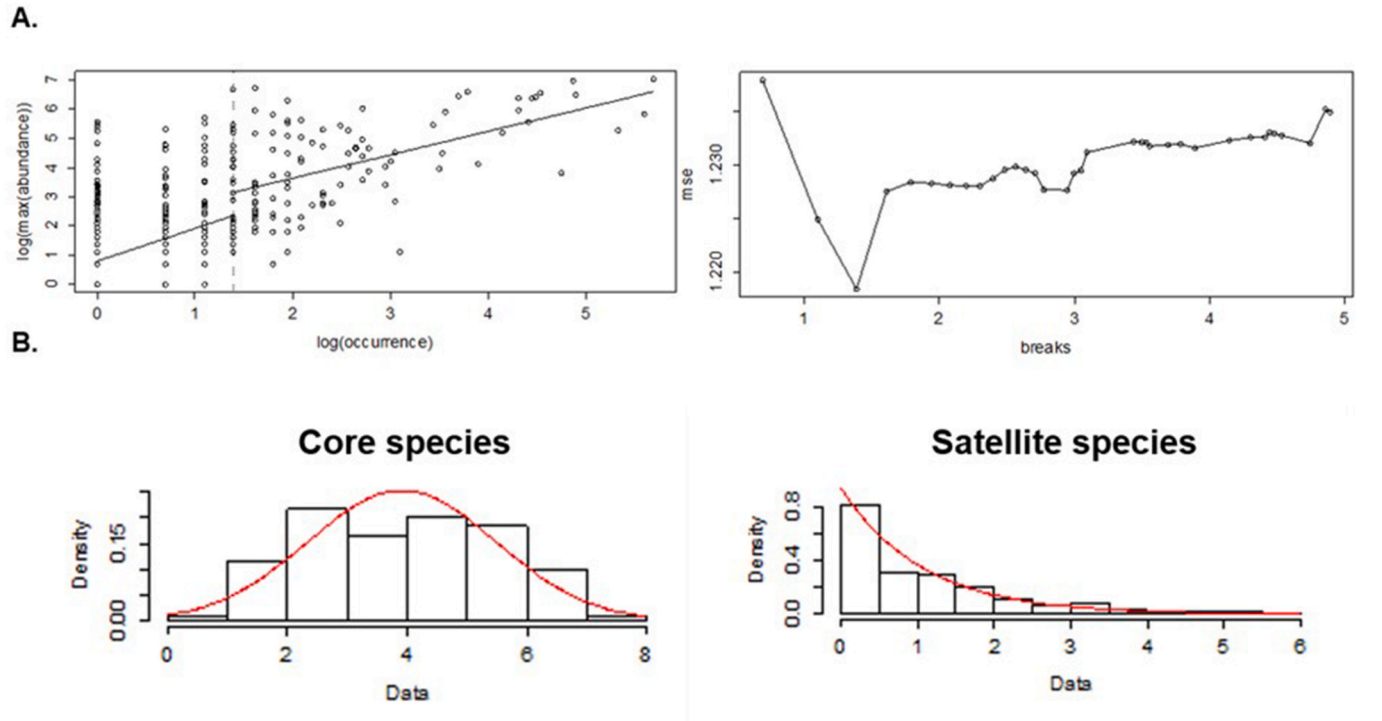

Fig. 5. (A) Split regression of log(abundance) on log(occurrence) for 821 needle fungal endophyte taxa in sample trees, showing break point at log(occurrence) = 1.39 where the mean square error (mse) is minimised. (B) Probability density plots for fit of core species to lognormal distribution and for fit of the satellite species to an exponential function.

differences in endophyte communities among families were found for more than half the provenances ($P < 0.05$).

5. Discussion

The design of our experiment, in which a set of *P. sylvestris* provenances and families were replicated at three sites, allowed us to separate the influence of both environment and host genotype on the composition of endophyte communities establishing within needles of *P. sylvestris*. Our results demonstrate that both environment and host genotype have significant effects on the endophyte communities. These conclusions are supported by data derived from both culturing and metabarcoding.

The influence of the local environment on endophyte community composition is demonstrated by comparison among the three sites which differ in both their biotic and climatic characteristics. Our analyses show significant differences in the abundance, community composition and diversity of the endophyte populations among these sites using both culture and metabarcoding data. This is in accordance with previous studies which involve comparisons of site effects on endophyte communities (Müller and Hallaksela, 1998; Helander et al., 2006; Jumponen and Jones, 2009; Matsumura and Fukuda, 2013; Millberg et al., 2015; Nguyen et al., 2016). However, none of these studies used replicated common garden experiments, and therefore could not rule out the possibility that endophyte community differences between sites were at least partially accounted for by genetic differences between host populations at these different sites.

A major cause of these differences may be that the sites vary with respect to their endophyte spore rain (Seabloom et al., 2019). The rain of endophyte spores compatible with *P. sylvestris* is expected to be greatest at Glensaugh, adjacent to a *P. sylvestris* plantation, while at Inverewe and Yair the local endophyte spores will primarily come from the non-native species, *Picea sitchensis* and *P. contorta*. This could account for the greater rate of endophyte isolations, larger number of taxa identified, and distinctness of the endophyte community at Glensaugh compared with the other two sites. The important influence on endophyte communities of spore rain from the same or closely related species has previously been noted by Hata and Futai (1996) and has been shown for the endophyte recruitment of non-native host-tree species (Gundale et al., 2016). Other factors that may be important in determining differences in endophyte communities among trees from different sites are climatic differences affecting the success of endophyte infection, and variation in host phenotype among the three sites caused by differences in growing conditions.

While environmental variation among sites has a very large influence on the endophyte communities establishing within *P. sylvestris* needles, our results also demonstrate that genetic variation among host trees influences the endophyte communities that establish. Given the limited data available from cultures it was only possible to test for variation in community composition by analysing differences in frequency of the most common taxon. However, for metabarcoding data we could analyse for changes in frequency across all endophyte taxa.

For both datasets, differences in the composition of endophyte communities among provenances were supported by moderate evidence at two sites, but evidence was weaker at the third site. In addition, variation in endophyte community among families within provenances was found for 11 of the 15 tests conducted using culture data, and 7 of the 15 tests using metabarcoding data. Following the principles of graduated evidence-based reporting (Muff et al., 2021), we can conclude that, though an effect is not apparent for every site or provenance, the overall support for host genotype influencing the composition of needle fungal endophyte communities is moderate to strong. Our results concur with previous research that has provided evidence for a significant effect of a host's genotype on various aspects of its fungal endophyte community (Cordier et al., 2012; Rajala et al., 2013; Balint et al., 2013, 2015).

Our finding that genetic effects on endophyte community composition are only apparent at certain sites could be due to lack of statistical power to observe such an effect. It could also signal that there may be limited opportunity for genetic variation to influence endophyte community composition due to local environmental conditions. Variation in abiotic factors such as humidity and air movements can cause both spatial and temporal localisation of particular endophyte taxa. If they are not present in the spore rain, there can be no host genetic influence on their frequency in the community.

Some idea of the number and identity of taxa able to infect *P. sylvestris* needles consistently and across a diversity of sites can be obtained from our analysis of core and satellite species. This classified 120 taxa as core species, 68 of which were recorded in needles from all three sites, 41 from two sites and 11 from single sites where they were present in almost all samples at high abundance. This presence of numerous core taxa distributed across all three sites suggests that absence of an appropriate spore rain or of conditions for infection are unlikely to be major impediments to detecting host genetic effects on endophyte community composition in our experiment.

In this study we used the complementary approaches of culturing and metabarcoding to quantify the endophyte communities of *P. sylvestris* needles. The two approaches have different strengths, limitations and biases and provide contrasting windows on the endophyte community. Here we explore the extent to which results from the two methods agree, how and why they may differ, and ascertain the benefits that can be gained by conducting both forms of assessment on the same endophyte community.

The first point to note is that the 11 morphotypes and largely unidentified 'other' taxa in the culture dataset form a similar community, in terms of diversity, to previous culture studies of Scots pine needles. Just a few species dominate, while a long tail of 'other' species occur at much lower frequency (Helander et al., 1994; Helander, 1995; Kowalski 1982, 1993; Taylor et al., 2019). The 11 taxa at high frequency in the culture samples were all detected under metabarcoding, providing a form of quality control for the metabarcoding study. Furthermore, in both datasets the dominant taxon was *Lophodermium seditiosum*, the frequency of which declined from Glensaugh to Inverewe to Yair. This species is known as a primary pathogen of *P. sylvestris* causing damage on seedlings in forest nurseries, and on stressed mature trees (Minter, 1981; Minter and Millar, 1980; Diwani and Millar, 1987). The dominance of this pathogen in healthy needles (up to 60% of the endophyte community) is striking proof that the presence of a pathogen within a needle, even at high frequency, is insufficient on its own to incite disease; the appropriate environmental conditions must be present for transition from latent to active pathogen. A high frequency of *L. seditiosum* in healthy needles has previously been detected in natural populations of *P. sylvestris* in Scotland and is characteristic of young trees similar in age to those sampled in this study (Taylor et al., 2019).

Another feature common to both datasets was the scarcity of *Cenangium ferruginosum* at Glensaugh and Inverewe, but its presence at relatively high frequency at Yair in both culture and metabarcoding samples (24.9% and 29.8% respectively). Together with the data on *L. seditiosum*, these results provide reassurance that culturing and metabarcoding are providing a degree of congruence in descriptions of the composition of the endophyte community of *P. sylvestris* (Taylor et al., 2016).

Despite these agreements, there are also some striking differences that require explanation. The most obvious of these is that the total number of taxa detected by culturing (11) were an order of magnitude lower than that detected by metabarcoding (821). The vast majority of the additional taxa identified using metabarcoding were present at frequencies of less than 1%, though a Helotiales taxon, absent from the cultured samples, was found at a frequency of 12.9% in the metabarcoding of a sample from Inverewe. Likely explanations for this difference are that additional taxa detected by metabarcoding were either at too low a frequency to be detected in the limited number of culture samples, that the taxa cannot be cultured under the conditions used, or

that though they are culturable, they are overgrown and their presence obscured by more vigorous taxa. Other metabarcoding studies concur that hyperdiversity, with many taxa present at very low frequency, is a common feature of fungal endophyte communities that is not revealed by culturing methods (Arnold et al., 2007; Millberg et al., 2015; Christian et al., 2017; Dissanayake et al., 2018; Jayawardena et al., 2018).

The other clear discrepancy between the cultured and metabarcoding datasets was that two taxa, *L. pinastri* and a member of the Xylariaceae, were detected at far higher frequency under culturing than under metabarcoding (approximately 20x for *L. pinastri*). This could be due to poor PCR amplification of the ITS sequences of these taxa. However, this can be ruled out for *L. pinastri*, because all three cryptic species of this taxon are identical to *L. seditiosum* at both the forward and reverse ITS primer sequences used in this study (Reignoux et al., 2014) and the latter showed no PCR inhibition. A more likely explanation is that the taxa occur at low frequency in the needles *in vivo*, but rapidly colonise dead needles on culture plates and overgrow other taxa, thereby leading to overrepresentation in the culture samples.

While the analysis above highlights limitations of culturing for quantifying endophyte communities, it should be remembered that metabarcoding techniques are themselves subject to bias. Differential amplification of ITS sequences among taxa, differential ITS copy number among taxa, delineation of OTUs across taxa on the basis of a single clustering threshold, and the inability to link ITS sequences to species hypotheses above a certain level of confidence have all been described as potential problems (Schirmer et al., 2015; Thomas et al., 2016; Zinger et al., 2019; Nilsson et al., 2019b).

Our results indicate a need for caution when comparing results from culture-based studies of endophyte communities with those from metabarcoding studies. Differences between the endophyte communities may reflect idiosyncrasies associated with the different assessment techniques rather than true biological differences. However, datasets collected using a common methodology should be legitimate for studying factors that influence endophyte community diversity and composition. We can conclude that there is added value in undertaking cultural alongside metabarcoding assessments of endophyte communities, not least for quality control purposes (for cross checking retrieved species). Furthermore, when results from both methods are in agreement concerning the factors that influence endophyte community composition (as here) this lends strength to the conclusions.

The endophyte community living in a host needle represents an aspect of the extended phenotype of the tree (Partida-Martinez and Heil, 2011; Whitham et al., 2003). Here we have presented evidence that variation in this extended phenotype can be under host genetic control. A possible consequence for the endophyte community may be an increase in functional redundancy and phylogenetic relatedness because favoured fungal taxa will tend to share common traits that adapt them to the *P. sylvestris* environment, and/or benefit the host (Shade and Handelsman, 2012). A possible consequence for the host tree is that natural selection could act on ecologically important variation in the endophyte community (Mejía et al., 2014; Christian et al., 2017). Further research is now required to explore these potential ecological and evolutionary consequences of host control of endophyte community composition.

Acknowledgements

This work was part of PROTREE, a project supported by grant BB/L012243/1, funded jointly by the Biotechnology and Biological Sciences Research Council, the Department for Environment, Food and Rural affairs, the Economic and Social Research Council, the Forestry Commission, the Natural Environment Research Council and the Scottish Government, under the Tree Health and Plant Biosecurity Initiative.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.funeco.2022.101162.

References

Albrectsen, B.R., Witzell, J., 2012. Disentangling functions of fungal endophytes in forest trees. In: Silva, A.P., Sol, M. (Eds.), Fungi: Types, Environmental Impact and Role in Disease. Nova Science Publishers, Inc., New York, pp. 235–246.
Albrectsen, B.R., Siddique, A.B., Decker, V.H.G., Unterseher, M., Robinson, K.M., 2018. Both plant genotype and herbivory shape aspen endophyte communities. Oecologia 187, 535–545.
Amend, A.S., Seifert, K.A., Bruns, T.D., 2010. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? Mol. Ecol. 19, 5555–5565.
Arnold, A.E., Lutzoni, F., 2007. Diversity and host range of foliar fungal endophytes: are tropical leaves biodiversity hotspots? Ecology 88, 541–549.
Arnold, A.E., Henk, D.A., Eells, R.L., Lutzoni, F., Vilgalys, R., 2007. Diversity and phylogenetic affinities of foliar fungal endophytes in loblolly pine inferred by culturing and environmental PCR. Mycologia 99, 185–206.
Arnold, A.E., Mejia, L.C., Kyllo, D., Rojas, E.I., Maynard, Z., Robbins, N., Herre, E.A., 2003. Fungal endophytes limit pathogen damage in a tropical tree. In: Proceedings of the National Academy of Sciences USA, 100, pp. 15649–15654.
Balint, M., Bartha, L., O'Hara, R.B., Olson, M.S., Otte, J., Pfenninger, M., Robertson, A.L., Tiffin, P., Schmitt, I., 2015. Relocation, high-latitude warming and host genetic identity shape the foliar fungal microbiome of poplars. Mol. Ecol. 24, 235–248.
Balint, M., Tiffin, P., Hallstrom, B., O'Hara, R.B., Olson, M.S., Fankhauser, J.D., Piepenbring, M., Schmitt, I., 2013. Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). PLoS One 8, e53987.
Bowman, E.A., Arnold, A.E., 2018. Distribution of ectomycorrhizal and foliar endophytic fungal communities associated with Pinus ponderosa along a spatially constrained elevational gradient. Am. J. Bot. 105, 687–699.
Bullington, L.S., Larkin, B.G., 2015. Using direct amplification and next-generation sequencing technology to explore foliar endophyte communities in experimentally inoculated western white pines. Fungal Ecology 17, 170–178.
Cadotte, M.W., Tucker, C.M., 2017. Should environmental filtering be abandoned? Trends Ecol. Evol. 32, 429–437.
Christian, N., Herre, E.A., Mejia, L.C., Clay, K., 2017. Exposure to the leaf litter microbiome of healthy adults protects seedlings from pathogen damage. In: Proceedings of the Royal Society B, 284, 20170641.
Christian, N., Sedio, B.E., Florez-Buitrago, X., Ramírez-Camejo, L.A., Rojas, E.I., Mejía, L. C., Palmedo, S., Rose, A., Schroeder, J.W., Herre, E.A., 2020. Host affinity of endophytic fungi and the potential for reciprocal interactions involving host secondary chemistry. Am. J. Bot. 107, 219–228.
Cordier, T., Robin, C., Capdevielle, X., Desprez-Loustau, M.-L., Vacher, C., 2012. Spatial variability of phyllosphere fungal assemblages: genetic distance predominates over geographic distance in a European beech stand (*Fagus sylvatica*). Fungal Ecology 5, 509–520.
Crawley, M.J., 2007. The R Book. John Wiley & Sons, New York.
Delignette-Muller, M.L., Dutang, C., 2015. Fitdistrplus: an R package for fitting distributions. J. Stat. Software 64, 1–34.
Dissanayake, A.J., Purahong, W., Wubet, T., Hyde, K.D., Zhang, W., Xu, H., Zhang, G., Fu, C., Liu, M., Xing, Q., Li, X., Yan, J., 2018. Direct comparison of culture-dependent and culture-independent molecular approaches reveal the diversity of fungal endophytic communities in stems of grapevine (*Vitis vinifera*). Fungal Divers. 90, 85–107.
Diwani, S.A., Millar, C.S., 1987. Pathogenicity of three Lophodermium species on *Pinus sylvestris* L. For. Pathol. 17, 53–58.
Elamo, P., Helander, M.L., Saloniemi, I., Neuvonen, S., 1999. Birch family and environmental conditions affect endophytic fungi in leaves. Oecologia 118, 151–156.
Fort, T., Pauvert, C., Zanne, A.E., Ovaskainen, O., Caignard, T., Barret, M., Compant, S., Hampe, A., Delzon, S., Vacher, C., 2021. Maternal effects shape the seed mycobiome in *Quercus petraea*. New Phytol. 230, 1594–1608.
Gardes, M., Bruns, T.D., 1993. ITS primers with enhanced specificity for basidiomycetes-application to the identification of mycorrhizae and rusts. Mol. Ecol. 2, 113–118.
Gundale, M.J., Almeida, J.P., Wallander, H., Wardle, D.A., Kardol, P., Nilsson, M.-C., Fajardo, A., Pauchard, A., Peltzer, D.A., Routsalainen, S., Mason, B., Rosenstock, N., 2016. Differences in endophyte communities of introduced trees depend on the phylogenetic relatedness of the receiving forest. J. Ecol. 104, 1219–1232.
Gweon, H.S., Oliver, A., Taylor, J., Booth, T., Gibbs, M., Read, D.S., Griffiths, R.I., Schonrogge, K., 2015. PIPITS: an automated pipeline for analyses of fungal internal transcribed spacer sequences from the Illumina sequencing platform. Methods Ecol. Evol. 6, 973–980.
Hata, K., Futai, K., 1996. Variation in fungal endophyte populations in needles of the genus *Pinus*. Can. J. Bot. 74, 103–114.
Hawkes, C.V., Kjøller, R., Raaijmakers, J.M., Riber, L., Christensen, S., Rasmussen, S., Christensen, J.H., Dahl, A.B., Westergaard, J.C., Nielsen, M., Brown-Guedira, G., 2021. Extension of plant phenotypes by the foliar microbiome. Annu. Rev. Plant Biol. 72, 823–846.
Helander, M.L., Sieber, T., Petrini, O., Neuvonen, S., 1994. Endophytic fungi in Scots pine needles: spatial variation and consequences of simulated acid rain. Can. J. Bot. 72, 1108–1113.
Helander, M.L., 1995. Responses of pine needle endophytes to air pollution. New Phytol. 131, 223–229.
Helander, M., Wäli, P., Kuuluvainen, T., Saikkonen, K., 2006. Birch leaf endophytes in managed and natural boreal forests. Can. J. For. Res. 36, 3239–3245.

Helander, M., Ahlholm, J., Sieber, T.N., Hinneri, S., Saikkonen, K., 2007. Fragmented environment affects birch leaf endophytes. New Phytol. 175, 547–553.
Hoyer, A.K., Hodkinson, T.R., 2021. Hidden fungi: combining culture-dependent and-independent DNA barcoding reveals inter-plant variation in species richness of endophytic root fungi in *Elymus repens*. Journal of Fungi 7, 466.
Ihrmark, K., Bodeker, I.T., Cruz-Martinez, K., Friberg, H., Kubartova, A., Schenck, J., Strid, Y., Stenlid, J., Brandstrom-Durling, M., Clemmensen, K.E., et al., 2012. New primers to amplify the fungal ITS2 region–evaluation by 454-sequencing of artificial and natural communities. FEMS (Fed. Eur. Microbiol. Soc.) Microbiol. Ecol. 82, 666–677.
Jayawardena, R.S., Witoon Purahong, W., Zhang, W., Wubet, T., Hong, X., Liu, M., Zhao, W., Hyde, K.D., Liu, J.H., Yan, J., 2018. Fungal Divers. 90, 1–84.
Johnston, P.R., Park, D., Smissen, R.D., 2017. Comparing diversity of fungi from living leaves using culturing and high-throughput environmental sequencing. Mycologia 109, 643–654.
Jumpponen, A., Jones, K.L., 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. New Phytol. 184, 438–448.
Ko Ko, T.W., Stephenson, S.L., Bahkali, A.H., Hyde, K.D., 2011. From morphology to molecular biology: can we use sequence data to identify fungal endophytes? Fungal Divers. 50, 113–120.
Koljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F., Bahram, M., Bates, S. T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., et al., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22, 5271–5277.
Kowalski, T., 1982. Fungi infecting *Pinus sylvestris* needles of various ages. Eur. J. For. Pathol. 12, 182–190.
Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., Schloss, P.D., 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl. Environ. Microbiol. 79, 5112–5120.
Kraft, N.J.B., Adler, P.B., Godoy, O., James, E.C., Fuller, S., Levine, J.M., 2015. Community assembly, coexistence and the environmental filtering metaphor. Funct. Ecol. 29, 592–599.
Lamit, L.J., Lau, M.K., Sthultz, C.M., Wooley, S.C., Whitham, T.G., Gehring, C.A., 2014. Tree genotype and genetically based growth traits structure twig endophyte communities. Am. J. Bot. 101, 467–478.
Larena, I., Salazar, O., González, V., Julián, M.C., Rubio, V., 1999. Design of a primer for ribosomal DNA internal transcribed spacer with enhanced specificity for ascomycetes. J. Biotechnol. 75, 187–194.
Lazarević, J., Menkis, A., 2020. Fungal diversity in the phyllosphere of *Pinus heldreichii* H. Christ—an endemic and high-altitude pine of the Mediterranean Region. Diversity 12, 172.
Magurran, A.E., Henderson, P.A., 2003. Explaining the excess of rare species in natural species abundance distributions. Nature 422, 714–716.
Matsumura, E., Fukuda, K., 2013. A comparison of fungal endophytic community diversity in tree leaves of rural and urban temperate forests of Kanto district, eastern Japan. Fungal Biol. 117, 191–201.
Mejía, L.C., Rojas, E.I., Maynard, Z., Bael, S.V., Arnold, A.E., Hebbar, P., Samuels, G.J., Robbins, N., Herre, E.A., 2008. Endophytic fungi as biocontrol agents of *Theobroma cacao* pathogens. Biol. Control 46, 4–14.
Mejía, L.C., Herre, E.A., Sparks, J.P., Winter, K., García, M.N., Van Bael, S.A., Stitt, J., Shi, Z., Zhang, Y., Guiltinan, M.J., Maximova, S.N., 2014. Pervasive effects of a dominant foliar endophytic fungus on host genetic and phenotypic expression in a tropical tree. Front. Microbiol. 5, 479.
Millberg, H., Boberg, J., Stenlid, J., 2015. Changes in fungal community of Scots pine (*Pinus sylvestris*) needles along a latitudinal gradient in Sweden. Fungal Ecology 17, 126–139.
Minter, D.W., Millar, C.S., 1980. Ecology and biology of three *Lophodermium* species on secondary needles of *Pinus sylvestris*. Eur. J. For. Pathol. 10, 169–180.
Minter, D.W., 1981. *Lophodermium* on pines. Mycological Papers 147, 1–54.
Muff, S., Nilsen, E.B., O'Hara, R.B., Nater, C.R., 2021. Rewriting results sections in the language of evidence. Trends Ecol. Evol. S0169534721002846.
Müller, M.M., Hallaksela, A.M., 1998. Diversity of Norway spruce needle endophytes in various mixed and pure Norway spruce stands. Mycol. Res. 102, 1183–1189.
Nguyen, D., Boberg, J., Ihrmark, K., Stenstrom, E., Stenlid, J., 2016. Do foliar fungal communities of Norway spruce shift along a tree species diversity gradient in mature European forests? Fungal Ecology 23, 97–108.
Nilsson, R.H., et al., 2019a. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res. 47, 259–264.
Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., Tedersoo, L., 2019b. Mycobiome diversity: high-throughput sequencing and identification of fungi. Nat. Rev. Microbiol. 17, 95–109.
Oita, S., Carey, J., Kline, I., Ibáñez, A., Yang, N., Hom, E.F.Y., Carbone, I., U'Ren, J.M., Arnold, E., 2021. Methodological approaches frame insights into endophyte richness and community composition. Microb. Ecol. 82, 21–34.
Oksanen, J., Blanchet, G.B., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P., Henry, M., Stevens, H., et al., 2016. vegan: community Ecology Package. R package version 2.3-5. http://CRAN.R-project.org/package=vegan.
Oono, R., Lefèvre, E., Simha, A., Lutzoni, F., 2015. A comparison of the community diversity of foliar fungal endophytes between seedling and adult loblolly pines (*Pinus taeda*). Fungal Biol. 119, 917–928.
Partida-Martinez, L.P., Heil, M., 2011. The microbe-free plant: fact or artifact? Front. Plant Sci. 2, 100.
Peršoh, D., 2013. Factors shaping community structure of endophytic fungi-evidence from the *Pinus-Viscum*-system. Fungal Divers. 60, 55–69.
Porras-Alfaro, A., Bayman, P., 2011. Hidden fungi, emergent properties: endophytes and microbiomes. Annu. Rev. Phytopathol. 49, 291–315.
Qian, X., Duan, T., Sun, X., Zheng, Y., Wang, Y., Hu, M., Yao, H., Ji, N., Lv, P., Chen, L., Shi, M., 2018. Host genotype strongly influences phyllosphere fungal communities associated with *Mussaenda pubescens* var. *alba* (Rubiaceae). Fungal Ecology 36, 141–151.
R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
Rajala, T., Velmala, S.M., Tuomivirta, T., Haapanen, M., Müller, M., Pennanen, T., 2013. Endophyte communities vary in the needles of Norway spruce clones. Fungal Biol. 117, 182–190.
Reignoux, S.N., Green, S., Ennos, R.A., 2014. Molecular identification and relative abundance of cryptic *Lophodermium* species in natural populations of Scots pine, *Pinus sylvestris* L. Fungal Biol. 118, 835–845.
Ridout, M., Newcombe, G., 2015. The frequency of modification of *Dothistroma* pine needle blight severity by fungi within the native range. For. Ecol. Manag. 337, 153–160.
Rim, S.O., Roy, M., Jeon, J., Montecillo, J.A.V., Park, S.C., Bae, H., 2021. Diversity and communities of fungal endophytes from four Pinus species in Korea. Forests 12, 302.
Rodriguez, R.J., White Jr., J.F., Arnold, A.E., Redman, R.S., 2009. Fungal endophytes: diversity and functional roles. New Phytol. 182, 314–330.
Romeralo, C., Martín-García, J., Martínez-Álvarez, P., Muñoz-Adalia, E.J., Gonçalves, D. R., Torres, E., Witzell, J., Diez, J.J., 2022. Pine species determine fungal microbiome composition in a common garden experiment. Fungal Ecology 56, 101137.
Salmela, M.J., Cavers, S., Wachowiak, W., Cottrell, J.E., Iason, G.R., Ennos, R.A., 2010. Understanding the evolution of native pinewoods in Scotland will benefit their future management and conservation. Forestry 83, 535–545.
Schirmer, M., Ijaz, U.Z., D'Amore, R., Hall, N., Sloan, W.T., Quince, C., 2015. Insight into biases and sequencing errors for amplicon sequencing with the Illumina MiSeq platform. Nucleic Acids Res. 43, e37.
Seabloom, E.W., Condon, B., Kinkel, L., Komatsu, K.J., Lumibao, C.Y., May, G., McCulley, R.L., Borer, E.T., 2019. Effects of nutrient supply, herbivory, and host community on fungal endophyte diversity. Ecology 100, e02758.
Shade, A., Handelsman, J., 2012. Beyond the Venn diagram: the hunt for a core microbiome. Environ. Microbiol. 14, 4–12.
Sun, X., Guo, L.D., 2012. Endophytic fungal diversity: review of traditional and molecular techniques. Mycology 3, 65–76.
Taudiere, A., Bellanger, J.M., Carcaillet, C., Hugot, L., Kjellberg, F., Lecanda, A., Lesne, A., Moreau, P.A., Scharmann, K., Leide, S., Richard, F., 2018. Diversity of foliar endophytic ascomycetes in the endemic Corsican pine forests. Fungal Ecology 36, 128–140.
Taylor, D.L., Walters, W.A., Lennon, N.J., Bochicchio, J., Krohn, A., Caporaso, J.G., Pennanen, T., 2016. Accurate estimation of fungal diversity and abundance through improved lineage-specific primers optimized for Illumina amplicon sequencing. Appl. Environ. Microbiol. 82, 7217–7226.
Taylor, J., Ellis, C., Ennos, R., Hollingsworth, P., 2019. Tissue age and tree age effects on endophyte species abundance of *Pinus sylvestris* (Scots pine) in an ancient semi-natural woodland in Scotland. Sydowia 71, 35–46.
Telford, A., Cavers, S., Ennos, R.A., Cottrell, J.E., 2014. Can we protect forests by harnessing variation in resistance to pests and pathogens? Forestry 88, 3–12.
Thomas, A.C., Deagle, B.E., Eveson, J.P., Harsch, C.H., Trites, A.W., 2016. Quantitative DNA metabarcoding: improved estimates of species proportional biomass using correction factors derived from control material. Molecul. Ecol. Resour. 16, 714–726.
U'Ren, J.M., Lutzon, F., Miadlikowska, J., Zimmerman, N.B., Carbone, I., May, G., Arnold, A.E., 2019. Host availability drives distributions of fungal endophytes in the imperilled boreal realm. Nature Ecol. Evolut. 3, 1430–1437.
van der Gast, C.J., Walker, A.W., Stressmann, F.A., Rogers, G.B., Scott, P., Daniels, T.W., Carroll, M.P., Parkhill, J., Bruce, K.D., 2011. Partitioning core and satellite taxa from within cystic fibrosis lung bacterial communities. ISME J. 5, 780–791.
Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73, 5261–5267.
Wang, X., Eijkemans, M.J., Wallinga, J., Biesbroek, G., Trzcinski, K., Sanders, E.A., Bogaert, D., 2012. Multivariate approach for studying interactions between environmental variables and microbial communities. PLoS One 7, e50267.
Wang, Y., Naumann, U., Wright, S.T., Warton, D.I., 2012. mvabund- an R package for model-based analysis of multivariate abundance data. Methods Ecol. Evol. 3, 471–474.
Warton, D.I., Wright, S.T., Wang, Y., 2012. Distance-based multivariate analyses confound location and dispersion effects. Methods Ecol. Evol. 3, 89–101.
Whitham, T.G., Young, W.P., Martinsen, G.D., Gehring, C.A., Schweitzer, J.A., Shuster, S. M., Wimp, G.M., Fischer, D.G., Bailey, J.K., Lindroth, R.L., Woolbright, S., Kuske, C. R., 2003. Community and ecosystem genetics: a consequence of the extended phenotype. Ecology 84, 559–573.
Zinger, L., Bonin, A., Alsos, I.G., Bálint, M., Bik, H., Boyer, F., Chariton, A.A., Creer, S., Coissac, E., Deagle, B.E., De Barba, M., 2019. DNA metabarcoding—need for robust experimental designs to draw sound ecological conclusions. Mol. Ecol. 28, 1857–1862.