What do we learn from cultures in the omics age? High-throughput sequencing and cultivation of leaf-inhabiting endophytes from beech (Fagus sylvatica L.) revealed complementary community composition but similar correlations with local habitat conditions

Abu Bakar Siddique¹, Anis Mahmud Khokon¹, Martin Unterseher²

¹ Ernst-Moritz-Arndt Universität Greifswald, Institut für Botanik und Landschaftsökologie, Soldmannstr. 15, 17487 Greifswald, Germany ² Evangelisches Schulzentrum Martinschule, Max-Planck-Str. 7, 17491 Greifswald, Germany

Corresponding author: Abu Bakar Siddique (biplobsau@gmail.com)

Abstract
Comparative simultaneous studies of environmental high-throughput sequencing (HTS) and cultivation of plant-associated fungi have rarely been conducted in the past years. For the present contribution, HTS and extinction culturing were applied for the same leaf samples of European beech (Fagus sylvatica) in order to trace both “real” environmental drivers as well as method-dependent signals of the observed mycobiomes. Both approaches resulted in non-overlapping community composition and pronounced differences in taxonomic classification and trophic stages. However, both methods revealed similar correlations of the fungal communities with local environmental conditions. Our results indicate undeniable advantages of HTS over cultivation in terms of revealing a good representation of the major functional guilds, rare taxa and biodiversity signals of leaf-inhabiting fungi. On the other hand our results demonstrate that the immense body of literature about cultivable endophytic fungi can and should be used for the interpretation of community signals and environmental correlations obtained from HTS studies and that cultivation studies should be continued at the highest standards, e.g. when sequencing facilities are not available or if such surveys are expanded into functional aspects with experiments on living isolates.

Key words
Cultivation, high-throughput sequencing, metabarcoding, fungal endophytes, biodiversity
Introduction

Fungal endophytes reside in the living tissues of plants without causing visible disease symptoms (e.g. Christian et al. 2015). Particular research interest is given to endophytes of the phyllosphere and other photosynthetic organs due to the enormous availability and environmental importance of leafy habitats (Lindow and Brandl 2003), the complex biochemical processes in these plant tissues and the generally close interconnectivity of leaf mycobiome with their hosts (reviewed in Rodriguez et al. 2009, Peršoh 2015). Within the last decade, significant progress has been made in unraveling plant-associated mycobiomes by using both cultivation (Collado et al. 2007, Untersee and Schnittler 2009, Gazis and Chaverri 2015) and direct high-throughput sequencing (HTS) techniques (Bálint et al. 2016). To date, most knowledge about endophyte richness, composition or host preferences is still based on traditional culturing approaches (Arnold 2007, Sieber 2007, Albrectsen et al. 2010, Untersee and al. 2013a). On the other hand, direct environmental assessment of mycobiomes provides unprecedented details of community diversity, composition, taxonomy and interactions (e.g. Peršoh 2013, Jumpponen and Brown 2014, Bálint et al. 2015, Eusemann et al. 2016, Purahong et al. 2016). These improvements go side by side with the ever increasing accuracy of reference sequence databases (Nilsson et al. 2014, 2015, Abarenkov et al. 2016).

It is well recognized that interpretation of diversity of endophytes and other fungi depends on the applied methods (Untersee 2011). On the one hand, cultivation data are often biased by under-sampling and the use of selective media. On the other hand, amplicon library preparation can differ strongly among studies (Salter et al. 2014) and HTS data are generally squeezed through complex and highly customisable bioinformatic pipelines, leading to variable data analysis (Meiser et al. 2014, Bálint et al. 2016), even if the same plant-fungus system is used in independent studies (Cordier et al. 2012, Siddique and Untersee 2016).

To date, published studies about the comparative assessment of fungal biodiversity are rare. Allmér et al. (2006) demonstrated the advantages and limitations of fruit body observation, mycelial cultivation and T-RFLP identification of wood-inhabiting fungi. Zhang et al. (2014) assessed microbial communities from deep-sea sediments with cultivation and environmental molecular cloning and identified two complementary assemblages. Similar conclusions were made by Langarica-Fuentes et al. (2014) who identified a biased composition of compost fungi by cultivation. Recently, HTS was rated superior over cultivation, given its ability to detect more obligate, slow growing and rare fungi (Al-Sadi et al. 2015, Oono et al. 2015). Whereas the general lack of congruence between mycobiomes generated by cultivation and HTS can be safely postulated nowadays (Pitkaranta et al. 2008), much less is known about environmental correlations of these differing data sets.

In this study, we investigated endophytic phyllosphere fungi with dilution-to-extinction cultivation (and ITS barcoding) and Illumina sequencing of the same DNA region and from the same material. In accordance with existing knowledge, we expected lower OTU (operational taxonomic unit) richness in the cultivation data and a
preferential isolation of ubiquitous, primary saprobic taxa. Consequently, we hypoth-
esized that the cultivable mycobiome exhibits different ecological signals compared
with the mycobiome obtained by HTS.

**Materials and methods**

**Study sites and sampling**

The samples were obtained from an experimental site established in 2013 (Unterseher
et al. 2016), consisting of two plots of 4 years-old *Fagus sylvatica* trees from a differ-
ent origin. The tree seeds originated from the Lustian lowland area (central eastern
Germany, HKG 81005) and were grown in a nursery in northern Germany (near
Hamburg) for two years. (Unterseher et al. 2016). The plots are located at the same
slope of the mountain massif “Untersberg” in Bavaria, Germany at 517 m asl. (above
sea level, respectively) (Valley site; Lat: 47.712946 Long: 13.040101) and at 975 m asl.
(Mountain site; Lat: 47.683158 Long: 13.002102) (Siddique and Unterseher 2016).
It is humus-rich with a well developed topsoil layer (A horizon) at the valley site,
while the A horizon is only weakly developed at the mountain site. The ground and
understory vegetation of the mountain site was mainly composed of *Acer pseudoplat-
nus*, *Picea abies*, *Daphne mezereum*, *Cardamine (= Dentaria) enneaphyllos*, *Helleborus
niger* and *Hepatica nobilis* (Siddique and Unterseher 2016). At the valley site, ground
vegetation was different with dominance of *Acer pseudoplatanus*, *Mentha spp.*, *Petasites
hybridus*, *Equisetum sylvaticum* and *Rubus* sp (Unterseher et al. 2016). Five trees each
from each site were selected randomly in October 2014, exactly one year after plant-
ing. Ten symptomless green leaves per tree were removed and processed as described in
Unterseher et al. (2016).

**Cultivation of endophytes using the dilution-to-extinction method**

Isolation of endophytes followed the dilution-to-extinction cultivation (Solis et al.
2016). In brief, samples were blended into tiny particles and filtered. Smallest particles
(Ø < 0.2 mm) were resuspended and strongly diluted before plating onto malt extract
agar (MEA, 1.5%) containing 48-well plates (Carl Roth, Karlsruhe, Germany). The
plates were inspected regularly for four weeks. Emerging colonies were transferred into
Petri dishes containing the same growth medium.

**DNA extraction and ITS sequencing from axenic cultures**

Instead of classifying the fungal cultures according to macroscopic and microscopic
characters (Khoyratty et al. 2015) and selecting only a few representative strains per
morphotype for downstream processing, genomic DNA was extracted from all isolates using a traditional, chloroform-based protocol (e.g. Solis et al. 2016). The ITS region was amplified with the primer pair V9G - ITS4 (de Hoog and Gerrits van den Ende 1998) using approved amplification kits (MangoTaq; Bioline, Germany) and cycling conditions (Solis et al. 2016). Unpurified PCR products were shipped to Beckman Coulter Genomics (Takeley, UK) for sequencing. Sequences were discarded if their corresponding chromatograms showed pronounced signs of ambiguous base calling after end trimming.

**Library preparation for high-throughput Illumina sequencing**

Total genomic DNA was extracted with the Charge Switch® gDNA Plant Kit (Invitrogen, Germany) from the same fresh leaf particle mass that was used for cultivation. Library preparation consisted of two consecutive amplification steps in order to add sample-specific tag combination for multiplexing. Please refer to Siddique and Unterseher (2016) and Unterseher et al. (2016) for the detailed description of this procedure. Amplicons were sequenced in pair-end mode on an Illumina MiSeq platform (Illumina Inc.) at the Genetics Section, Biocentre of the LMU Munich, Germany.

**Processing of Illumina reads and Sanger sequences**

Demultiplexing and quality filtering of Illumina reads relied on QIIME (Navas-Molina et al. 2013; see Suppl. material 1) and is also detailed in Unterseher et al. (2016) and Eusemann et al. (2016). Extraction of ITS1 (forward R1 Illumina reads) was done with ITSx (Bengtsson-Palme et al. 2013) followed by reference-based chimera checking (Nilsson et al. 2015), open-reference OTU picking (complete-linkage clustering at 97% similarity; Rideout et al. 2014), selection of representative sequences and taxon assignment (Kõljalg et al. 2013). These last steps were also performed with all high-quality Sanger sequences to guarantee best comparability. Final quality filtering of HTS OTUs consisted of the removal of unique (occurring in only one sample) and rare OTUs (having less than five reads, cf. Brown et al. 2015). In contrast, all OTUs from cultivation data were retained, knowing that they belonged to true fungi. The reasons for using only the ITS1 region for subsequent analyses is comprehensively explained in recent papers (e.g. Unterseher et al. 2016, Eusemann et al. 2016).

**Biodiversity analysis and assessment of functional guilds**

The analysis of fungal biodiversity comprised the assessment of OTU richness and further indicators of diversity (Fisher’s Alpha, Shannon index and three numbers of Hill’s series of diversity, the latter considering different levels of rarity) (Hill 1973). Given the strong positive correlation of OTU richness and read numbers (data not shown, but compare
What do we learn from cultures in the omics age?

Siddique and Unterseher (2016)) as well as the strongly different sequencing depth between the two approaches, read counts were standardised (divided) by sample totals.

Community composition was assessed with PCoA (principal coordinate analysis) and NMDS (non-metric multidimensional scaling) and tested with PERMANOVA (permutational multivariate analysis of variance). Functional guild analysis was performed according to Nguyen et al. (2016). The entire biodiversity analysis was performed with R (freely available at www.r-project.org, last accessed 11/2016). The corresponding command script and necessary input files are available as “Suppl. material 2”. Curated Sanger sequences were taxonomically annotated as far as possible and made available in ENA/GenBank through accession numbers LT604837–LT604881. High-throughput data are available under SRA accession number SRX1211311.

Results

Basic data exploration, diversity and community composition

Data volumes differed strongly between Illumina sequencing and cultivation. Illumina sequencing resulted in 597 OTUs from 170480 curated ITS1 reads and cultivation revealed 70 OTUs from 438 culture-based Sanger sequences with the same settings for OTU clustering. The combined data set comprised 630 OTUs (+ 33 OTUs compared with Illumina data). Thirty-seven OTUs were detected with both methods (see Table S1 on Suppl. material 3 and Suppl. material 4).

An insignificant trend of lower fungal diversity at the mountain site across all indexes was observed for HTS data (Fig. 1A). Richness analysis of cultivation data was partly contradictory with significant differences of Fisher’s Alpha between valley and mountain samples (p <0.01, see Table S2 and S3 “Suppl. material 3”) but nearly identical accumulation curves (Fig. 1B). In addition and contrary to HTS data, Hill numbers N2 and N3 were significantly different for cultivation data (Fig. 1B, for statistical details see Table S2 and S3; “Suppl. material 3”). The accumulation curves for HTS data (Fig. 1A) revealed a clearly lower fungal richness for mountain than for valley samples, whereas the cultivation data failed to show such differences (Fig. 1B).

The analysis of community composition with non-metric multidimensional scaling (NMDS), principal coordinate analysis (PCoA) and PERMANOVA discovered a significant influence of the isolation method (df = 1, F = 7.58, R² = 0.30, p= 0.001) as well as of locality (df = 1, F = 2.87, R² = 0.14, p= 0.014) (Fig. 2). The differences between valley and mountain site were more pronounced for HTS (F = 3.94, R² = 0.33, p= 0.007) than for cultivation data (F = 1.96, R² = 0.20, p= 0.027) (Fig. 2).

Taxonomic composition of HTS and cultivation data

Three of the five most abundant orders from Illumina data were also most abundant in cultivation data (Capnodiales – both methods, Helotiales – both methods, Saccharomy-
cetales – Illumina only, Pleosporales – both methods; all Ascomycota) (Fig. 3). In addition, HTS revealed further and abundant orders from both Asco- and Basidiomycota, which were not recovered during cultivation (Fig. 3), such as the yeast fungi Saccharomyccetales (Ascomycota) and Malasseziales (Basidiomycota). The Xylariales (Ascomycota, present with one isolate) was the only order from the cultivation data that was not detected with HTS.

Composition of trophic guilds revealed by HTS and cultivation

The five main guilds (pathotrophs, patho-saprotrophs, patho-symbiotrophs, saprotrophs and symbiotrophs) were all detected by HTS. Cultivation largely failed to detect pathotrophs (including patho-saprotrophs and -symbiotrophs). The relative abundance of saprotrophs was clearly higher in cultivation than in HTS data (Fig. 4B, C).

When analysing the influence of locality for the occurrence of different ecological guilds, it turned out that the abundance of pathotrophs was significantly higher in leaves of mountain trees than of valley trees (Fig. 5A). Saprotrophs were also more abundant in leaves of mountain trees, whereas symbiotrophs were more abundant in valley than in mountain trees (Fig. 5A, B).

Results from HTS and cultivation data were congruent in as much as saprotrophs and symbiotrophs revealed similar abundance patterns for both methods.
What do we learn from cultures in the omics age?

Figure 2. Principal coordinate analysis (PCoA) of fungal leaf-inhabiting endophytes of beech display strongly differing assemblages obtained with Illumina sequencing and cultivation. Both methods revealed differing mycobiomes from valley and from mountain leaves, although these differences were less pronounced for cultivation data. Abbreviations: IM = Illumina data from mountain samples, IV = Illumina data from valley samples, CM = cultivation data from mountain samples, CV = cultivation data from valley samples

Discussion

The methodology of biodiversity assessment influences the interpretation of community composition and community ecology

The most abundant orders were the same for both cultivation and HTS, namely Capnodiales, Helotiales and Pleosporales (Fig. 3). Ascomycota clearly dominated the cultivation data thus confirming results of many earlier cultivation studies (e.g. U’ Ren et al. 2012, Scholtysik et al. 2013, Unterseher et al. 2013b). Many of own isolates, such as those with highest sequence similarities to the genus Mycosphaerella and its anamorphs, are often described in literature as frequent asymptomatic inhabitants of living leaves. In this study, taxa generally known as saprotrophs were the dominant trophic
Figure 3. Abundance distribution of the 20 most abundant orders of fungal leaf-inhabiting endophytes of beech on a logarithmic scale. Three of the five most abundant orders from high-throughput sequencing were also most abundant in cultivation data.

The compositional difference in the two mycobiomes also corresponded to the presence of parasitic taxa (Taphrinales, Erysiphales) and yeast-like fungi (Saccharomycetales and Tremellales) in the HTS data, whereas the cultivation data were devoid of fungi with obligate parasitic, biotrophic or pathogenic lifestyle. The latter guilds usually cannot be cultivated, and yeasts are often detected only during cultivation studies when growth of
What do we learn from cultures in the omics age?

Figure 4. Relative abundance of fungal leaf-inhabiting endophytes of beech among the five main trophic guilds as revealed by analysis with FUNGuild (Nguyen et al. 2016). A compares the two methods for each trophic guild and unassigned data. B displays the trophic guilds and unassigned taxa for Illumina data, C for cultivation data. Abbreviations in B and C: U = Unassigned, P = Pathotrophs, PSa = Patho-Saprotrophs, PSy = Patho-Symbiotrophs, Sa = Saprotrophs, Sy = Symbiotrophs.

Figure 5. Relative abundance distribution of fungal leaf-inhabiting endophytes of beech among the five main trophic guilds as revealed by analysis with FUNGuild (Nguyen et al. 2016). A compares the two localities for each trophic guild on the basis of Illumina data B compares the two localities for each trophic guild on the basis of cultivation data.

Filamentous fungi is slowed down with low-temperature incubation. In this study, HTS retained a wide range of taxa (compare Al-Sadi et al. 2015) and all guilds available in the FUNGuild reference data base (Nguyen et al. 2016) as expected (Figs 4, 5).

A poor comparability of cultivation and HTS data, as it is presented here, was recently reported for a microbiome study (Eevers et al. 2016). It was caused by the fundamentally different sample coverage with the most abundant OTU counting 28621 reads for HTS (Mycosphaerellaceae, *Sphaerulina*) and 102 isolates for cultivation (Hyaloscyphaceae, *Lachnum*) (Suppl. material 4).
The two methods revealed consistent signals of both community data to environmental conditions

On the one hand side our results clearly demonstrate the limitations and biases of cultivation approaches for comprehensive biodiversity assessments. On the other hand, the results did not meet our expectations (see above), because significant correlations to environmental parameters (here, it was the difference between valley and mountain samples) were still recognized. The present cultivation data are in concordance with similar studies (Unterseher et al. 2013b) and confirm general knowledge about the community ecology of leaf endophytes (Cordier et al. 2012, Zimmermann and Vitousek 2012, Meng et al. 2013, Glynou et al. 2015, Rojas-Jimenez et al. 2016). Moreover, similar results were observed in the present comparative study on the basis of HTS data (see also Siddique and Unterseher 2016, Unterseher et al. 2016).

Conclusions

Our results clearly justify the co-existence of cultivation and high-throughput approaches. Despite the fast improvement and diversification of HTS technologies with many undeniable advantages in microbial biodiversity assessments (Bálint et al. 2016), cultivation should be retained at highest standards (e.g. Gazis and Chaverri 2015), given, among others, the availability of living cultures for genome, metabolome or bioprospecting analyses (Gazis et al. 2016, Kolarik et al. 2016).

Our results suggest that the immense body of literature about cultivable endophytic fungi can and should be consulted for the interpretation of community signals obtained from HTS studies.

Acknowledgements

This study was funded by the Deutsche Forschungsgemeinschaft (DFG) as a part of the project grant to MU (UN262/9-1). The authors thank Andreas Brachmann (Munich) and Derek Peršoh (Bochum) for profound help with Illumina sequencing.

References

What do we learn from cultures in the omics age?


What do we learn from cultures in the omics age?


Unterseher M, Siddique AB, Brachmann A, Peršoh D (2016) Diversity and composition of the leaf mycobiome of beech (Fagus sylvatica) are affected by local habitat conditions and leaf biochemistry. PLoS ONE 11: e0152878. https://doi.org/10.1371/journal.pone.0152878


Supplementary material 1

Bioinformatics pipeline
Authors: Abu Bakar Siddique, Anis Mahmud Khokon, Martin Unterseher
Data type: scripts, spreadsheet
Explanation note: This file provides all steps and commands necessary for quality filtering and demultiplexing of raw paired fastq sequences.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.
Supplementary material 2

Biodiversity workflow in R
Authors: Abu Bakar Siddique, Anis Mahmud Khokon, Martin Unterseher
Data type: Scripts, measurement
Explanation note: Bundle of files for biodiversity analysis in R. All necessary input files and a commented script of R-commands are provided.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 3

Common OTU lists and Statistical analysis
Authors: Abu Bakar Siddique, Anis Mahmud Khokon, Martin Unterseher
Data type: measurements
Explanation note: This file contains detected OTUs in both methods and biodiversity analysis (GLM and t-test)
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.

Supplementary material 4

Master data sheet
Authors: Abu Bakar Siddique, Anis Mahmud Khokon, Martin Unterseher
Data type: spreadsheet
Explanation note: Spreadsheet file containing information about read abundances of operational taxonomic units (OTUs) and sample metadata. Here, data were prepared for subsequent biodiversity analysis in R.
Copyright notice: This dataset is made available under the Open Database License (http://opendatacommons.org/licenses/odbl/1.0/). The Open Database License (ODbL) is a license agreement intended to allow users to freely share, modify, and use this Dataset while maintaining this same freedom for others, provided that the original source and author(s) are credited.