Journal of Microbiological Methods 243 (2026) 107447

Contents lists available at ScienceDirect

Journal of Microbiological Methods

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

Manual curation of aquatic fungal ITS2 metabarcoding OTU matrices increases community identification on the genus and species level

Christiane Baschien [a,*], Steffen Carl [b]

[a] Leibniz-Institute DSMZ German Collection of Microorganisms and Cell Cultures, Inhoffenstraße 7B, 38124 Braunschweig, Germany
[b] German Environment Agency, Section IV 2.5 Trace Analysis, Artificial Streams and Ponds, Schichauweg 58, 12307 Berlin, Germany

ARTICLE INFO

Keywords:
Aquatic fungi
ITS rDNA
Metabarcoding
Sequence data
Taxonomic resolution

ABSTRACT

We conducted six metabarcoding studies from different aquatic habitats. It has been confirmed that fungal metabarcoding has the advantages of overcoming the cultivation bias, increasing the number of possible samples by high-throughput sequencing, and enabling the detection of yet uncultivated fungal taxa. However, the already well-known limitations, in particular the lack of taxonomic accuracy when using automated classification also occurred.

During the manual curation of pipeline-retrieved datasets, the quality of the automatic classification, specifically phylogenetic rank and current nomenclature, was verified. The manual curation step revealed that only 10–20% of the Operational Taxonomic Units (OTUs) were identified correctly by the automatic classification. Re-blasting of the remaining OTUs resulted in taxonomic and nomenclatural correction of a further 30–40% of OTUs. Approximately 50% of the OTUs resulted in insufficient identifications, which can be attributed to shortcomings in the database, including yet unidentified species, missing reference sequences, and the disadvantages of the genetic marker used.

Pipeline parameters that can be adjusted to improve the classification of taxa in metabarcoding studies include adjusting the cut-off threshold for sequence similarity, updating and refining databases, and aligning and collapsing of OTUs to species hypotheses. Here, we describe a method of manual data curation and its impact on improving the taxonomic resolution of a metabarcoding dataset. Shortcomings as well as advantages are discussed to illustrate the need for bioinformatic tools that support the process of data curation.

1. Introduction

Accurate species identification is a prerequisite for answering questions about function, resilience, interaction, and diversity effects in ecological studies of all fungal habitats. In addition, it is generally recognized that the isolation strategy and the limitations of the culture media significantly limit the information on fungal diversity.

For more than a decade, the diversity of fungi has been estimated using DNA metabarcoding (Jumpponen and Jones, 2009; Begerow et al., 2010; Větrovský et al., 2020). For DNA metabarcoding, one or more marker regions in the genome are PCR-amplified and sequenced using high-throughput sequencing platforms (HTS; Lindahl et al., 2013, Nilsson et al., 2019a). The partial ITS rDNA operon (White et al., 1990) is the most widely used marker for fungal metabarcoding (Schoch et al., 2012; Tedersoo et al., 2022).

The metabarcoding method has the advantage of analysing the fungal community in less-investigated habitats, such as freshwater ecosystems. The biodiversity of freshwater fungi, which are key drivers in ecosystem processes by initiating the decomposition of leaf litter in aquatic environments, has been investigated in hundreds of field and laboratory studies with meta-analyses conducted by Ferreira et al. (2015, 2023). These studies were accompanied by morphological and molecular biological methods to assess the biodiversity of aquatic hyphomycetes, especially in running waters (e.g., Nikolcheva et al., 2003; Casas et al., 2011; Duarte et al., 2015; Carl et al., 2022, 2025). Within the last decade, approximately 300 metabarcoding studies of aquatic habitats, including micro- and mesocosms, have been published (Google Scholar searches assessed Aug. 07, 2025). Many of these studies (e.g., Wurzbacher et al., 2016; Duarte et al., 2017; Matsuoka et al., 2019; French et al., 2024) revealed the limitation of fungal classification using a combination of different pipelines and classification databases, e.g., Mothur (https://mothur.org/), Qiime2 (https://qiime2.org/), Claident

* Corresponding author.
E-mail addresses: christiane.baschien@dsmz.de (C. Baschien), s.carl@gmx.de (S. Carl).

https://doi.org/10.1016/j.mimet.2026.107447
Received 11 January 2026; Received in revised form 26 February 2026; Accepted 26 February 2026
Available online 2 March 2026

(https://www.claident.org/), with datasets UNITE (https://unite.ut.ee/), SILVA (https://www.arb-silva.de/), and NCBI (https://www.ncbi.nlm.nih.gov/). However, the taxonomic resolution of classification on which conclusions in these studies were based usually did not exceed phylum, class, or order ranks. Classifications at higher ranks such as phylum, class, order, and family limit the ability to answer ecological questions, such as identifying changes in community composition, because inaccurate identification reduces the observed fungal diversity. Furthermore, the heterogeneity of the rDNA ITS marker sequences in many fungal genera (Bradshaw et al., 2023; Cedeño-Sanchez et al., 2024) affects the estimated number of taxa within a dataset. This can lead to an overestimation of species hypotheses, as for example in a study by Lazar et al. (2024). On the other hand, species numbers can be underestimated if the ITS rDNA region shows insufficient resolution for example in cases of horizontal gene transfer as in the genus *Fusarium* (Lieckfeldt and Seifert, 2000) or if there is extensive similarity in closely related species as, e.g., in the genus *Aspergillus* (Houbraken et al., 2021). The drawbacks of the ITS rDNA as a metabarcoding marker include intraspecific and intragenomic variability of this genetic region (Nilsson et al., 2008; Kauserud, 2023). In turn, the pitfalls of the metabarcoding method, combined with misinterpretation of results, lead to a decreased meaningfulness of many community-based studies.

HTS studies are prone to a range of biases that are caused by methodological and database-related shortcomings (reviewed in Lindahl et al., 2013 and Tedersoo et al., 2022). This starts with the careful choice of an appropriate DNA extraction method depending on the substrate probed, the careful choice of primer-sets, awareness of sequencing errors and biases of clustering and amplicon methods, as well as awareness of database incompleteness and the heterogeneous nucleotide structure of ITS operons (Paloi et al., 2022). Furthermore, PCR and sequencing errors can introduce or mask variations in ITS rDNA. In metabarcoding, this can lead to over- or underestimation of the number of taxa present in a sample. As already suggested by Tedersoo et al. (2022), the obtained OTU matrices using metabarcoding pipelines require manual curation to increase the quality of fungal classification. For the first time, this improvement has been demonstrated in Carl et al. (2022), where we successfully curated datasets obtained from Illumina MiSeq metabarcoding of aquatic fungal communities in leaf litter under both field and laboratory conditions.

The aim of the present paper is to show how the yield of information from metabarcoding studies can be improved through manual dataset curation. For this purpose, we describe in detail the manual curation of ITS2 metabarcoding OTU tables and demonstrate how this approach improves the taxonomic resolution of a metabarcoding dataset. We also want to emphasize, that the ambition of the present article is not to compare different bioinformatic tools such as the application of OTU clustering or the use of haplotype sequences like ASVs (amplicon sequence variants). Instead, it was our purpose to provide ecologists and mycologists with a protocol describing how to manually curate metabarcoding data and to open the discussion for new tools, that may facilitate to automatize this laborious step in the near future. To this end, we have drawn on our experience from six extensive HTS studies in lakes, streams, and microcosms, four of which have already been published (Heeger et al., 2018; Carl et al., 2022, 2025; Gonçalves et al., 2024).

2. Methods

2.1. Origin of the exemplary dataset

To explain the curation method and show its efficacy to increase the taxonomic resolution, a small dataset of a reference stream was used from our most recent metabarcoding study (Carl et al., 2025). Briefly, black alder leaf litter was exposed in fine mesh bags (200 μm mesh width) in the Lieberoser Mühlenfließ, a mesotrophic woodland brook in the German Spreewald region (51°59′N, 14°21′E) for six weeks with weekly sampling in triplicate. Samples from weeks four, five, and six (i.e., $n = 9$) were taken as exemplary dataset for the present paper. Additionally, a mock community was included to analyse the recovering efficiency of fungal taxa particularly with regard to the delimitation of closely related species. For this, the DNA from 12 fungal reference strains (see Table S4) was pooled in equimolar DNA concentrations. Details about the sequencing method and the dataset can be found in Carl et al., 2025 and the following data repository: https://zenodo.org/records/13337703. Raw sequences were deposited in the European Nucleotide Archive under the Project number PRJEB77411 (https://www.ebi.ac.uk/ena/browser/view/PRJEB77411).

2.2. DNA extraction, Illumina sequencing, and bioinformatic processing

The procedure for DNA extraction, library preparation, sequencing, and sequence processing is described in Carl et al. (2022, 2025). Briefly, total DNA of leaf samples containing three freeze dried alder leaf discs ($d = 19$ mm) were extracted using the FastDNA SPIN Kit for Soil. Prior to PCR, extracts were diluted 50-fold with PCR-grade water. Amplicon libraries of the fungal ITS2 rDNA region were generated from the diluted samples in a 3-step-PCR workflow including preamplification, barcode- and index primer ligation by use of the ITS3tagmix and the ITS4ngs primers as published in Tedersoo et al. (2014, 2015) including the corresponding barcode and index primers with Illumina MiSeq adapters (see above-mentioned data repository for primer sequences, PCR conditions, and Table S1 for sample list). Preamplification was conducted in triplicate for each sample and the resulting PCR products were pooled in order to reduce the amplification bias (Lindahl et al., 2013; Tedersoo et al., 2022). The resulting pool of PCR products after clean-up and levelling of DNA concentrations (see Carl et al., 2022) was sequenced by use of a 600-cycle MiSeq reagent Kit v3 at a concentration of 4.4 pM (Illumina, San Diego, USA). Paired-end sequencing generated 2 × 300 bp single barcoded reads (see Carl et al., 2025 and the corresponding repositories mentioned in Section 2.1).

After sequencing, the demultiplexed samples were processed using the PIPITS pipeline (Version 2.5, Gweon et al., 2015, https://github.com/hsgweon/pipits/releases), including quality filtering with the fastx toolkit (version 0.0.14), extraction of the ITS2 sub region with ITSx (version 1.1b; Bengtsson-Palme et al., 2013), chimera filtering against the UNITE UCHIME database (Nilsson et al., 2015; dataset from June 28, 2017), clustering of OTUs with VSEARCH (version 2.14.2; https://github.com/torognes/vsearch/) including unique sequences (otherwise default options i.e., clustering of sequences with cluster_fast function at the 97% sequence similarity level), and their taxonomic assignment with the RDP classifier (version 2.0.2; Wang et al., 2007) according to the "species hypotheses" (SH, Kõljalg et al., 2013, 2020) of the trained UNITE database (version from February 04, 2020; Nilsson et al., 2019b).

2.3. Manual curation method and test for curation success

Taxonomic assignments of OTUs with the PIPITS implemented RDP classifier (Cole et al., 2009) SHs of the UNITE database were ranked according to the classification confidence value (CV, ranging from zero to one). Classification of OTUs was checked manually by re-blasting MiSeq derived sequences using a range of databases implemented in Mycobank (www.mycobank.org; Robert et al., 2013) and the curated reference sequences refseq_rna (O'Leary et al., 2016) as well as the default BLAST option (all nr databases) in GenBank and own reference data from cultures (DSMZ culture collection). First, the OTUs in the entire dataset were sorted in descending order from the highest to the lowest CV. Subsequently, the sequences of the OTUs in FASTA format were subjected to BLAST searches in GenBank to find the closest database hits with the sequences. Results deviating from the pipeline classification were curated. For each SH the following quality criteria were applied (Carl et al., 2022): (i) significant similarity to any BLAST-hit of a fungal taxon (preferably ≥97%), (ii) reasonable coverage of sequence

(preferably ≥95%), (iii) highest e-value (ratio between coverage and similarity of the sequence), (iv) reliably published sequence (reference database, isolate voucher, publication yes/no), (v) correct species and other rank names according to up to date taxonomy (Mycobank) and phylogenetic concepts (yes/no). Blast-hits leading to environmental sequences from uncultivated fungi were also retained. These could conceal undescribed species that were also found in other studies. Uninformative OTUs without meaningful hits or with ambiguous hits were flagged for later exclusion prior to community analysis.

In the next step, OTUs representing the same SH were checked for their actual percentage of sequence similarity. This check also included known ITS sequence variation of the particular taxon, if available, e.g., from literature or reference data based on available cultures. Thus, OTUs resembling the same SH were merged per sample to reduce overestimation of taxon richness due to amplification bias, sequencing errors, ITS heterogeneity by evolutionary variation or the presence of multiple rRNA operons (Heeger et al., 2018; Paloi et al., 2022). Merging of OTUs leading to the same taxon was performed after alignment of the relevant OTUs using the ARB software environment (version 7.0, 2021; Ludwig, 2004). For a better visualisation of the OTUs to be merged, sequences were aligned with sequences of reference strains and a maximum likelihood tree was calculated with the RAxML method (Stamatakis, 2014; settings: GTRGAMMA model, rapid bootstrapping, 1000 replications) as implemented in ARB. During curation of the dataset, phylogenetic rank and current nomenclature were corrected according to Mycobank.

To compare the accuracy of the automatic classification to manual curation, the Matthews Correlation Coefficient was applied (MCC, Matthews, 1975, Baldi et al., 2000, Chicco and Jurman, 2020), which correlates the observed (manual curation) and predicted (RDP classifier) classification results. A coefficient value of +1 represents the best prediction whereas −1 indicates total disagreement between automatic and manual classification.

3. Results

3.1. Taxonomic resolution before and after curation

The curated dataset finally comprised 1050 OTUs that were assigned to 157 different SHs (see Table S3). The species annotation of the RDP classifier against the UNITE database was checked manually for all SHs, irrespective of the confidence values that ranged between 0.0 and1.0. As a result, 66 (~ 42%) of the species annotations were changed, whereas only 33 (21%) led to the same taxon as compared to the PIPITS result (Fig. 1A). A total of 58 SHs (~ 37%) had no reasonable blast-hit (~ 21%), led to non-fungal organisms (~ 8%), or uncultured fungi (~ 8%, mainly *Cryptomycota* and *Chytridiomycota*), respectively. For the remaining 99 SHs (Fig. 1B), manual curation led to an increased taxonomic resolution with 51 SHs that could be identified down to the species level, 38 on genus level, and the rest on family (8) or order level (2).

In the next step, OTUs representing the same taxon with 97–100% sequence similarity were merged per sample to reduce overestimation of taxa due to sequencing errors or heterogeneity of multiple rRNA operons. In total, 17 of such "inflated species hypotheses" were merged (see Table S2 for full dataset). Exemplarily, the merging of 54 OTUs from one sample resembling the species *Flagellospora curvula* with a CV range from 0.44 to 0.85 is shown in Fig. 2, with the highest CV class being represented by most of the sequence reads. The maximum likelihood tree based on *F. curvula* OTUs shows the result of the heterogeneity of ITS2 sequences simulating several species (Fig. 3). Consequently, the low bootstrap values along with the substitutions per site show that there is no statistical support for several "new" *Flagellospora* species. The sequence similarity within the *F. curvula* OTUs averaged 4.23%. Furthermore, taxonomic identification is supported by the inclusion of reference strains of Leotialean *Flagellospora* species. Here, the average similarity of all OTUs to three *F. curvula* reference strains was 2.98%, while it was higher for all other *Flagellospora* species included (see distance matrix in Table S6 and the corresponding alignment in Table S7).

3.2. Curation success and accuracy as calculated with MCC

Taking all results of the RDP classifier into account (i.e., 157 SHs; see Table S3), the annotation was correct for ~26% of the SHs including some correct classifications of uncultured fungi, whereas it led to higher ranks compared to manual blast-hits in ~21% of the cases. 53% of the SHs were annotated to lower ranks (~ 12%) or to completely different taxa (~ 41%).

The accuracy tested by MCC differed clearly between the RDP classifier using the UNITE database and the manual curation (Fig. 4). The RDP classifier performed best in bootstrap confidence values that ranged between 0.2 and 0.29 and 0.5–0.59. However, the pipeline never reached the accuracy level of the manual curation, which was lowest in the CV range between 0.6 and 0.69. In the range between 0.8 and 1 the manual curation reached complete accuracy.

3.3. Sequencing and classification success of mock community strains

Using the present metabarcoding approach, ten out of twelve species from the mock community could be retrieved (Table 1). The classification was accurate in delimiting species belonging to the same genera in case of *Cladosporium sphaerospermum* and *C. halotolerans*, *Aspergillus*

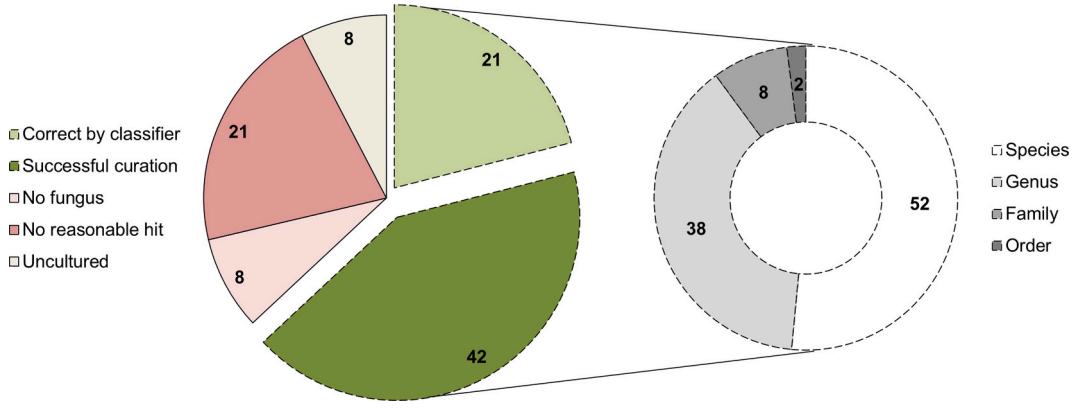

Fig. 1. Curation success. A Distribution of categories differentiated during curation of the metabarcoding dataset. Numbers represent percentages of species hypotheses (SH; $n = 157$) B Taxonomic resolution after curation of the dataset. Numbers represent remaining taxa ($n = 99$) after exclusion of uninformative SHs as shown by the lines.

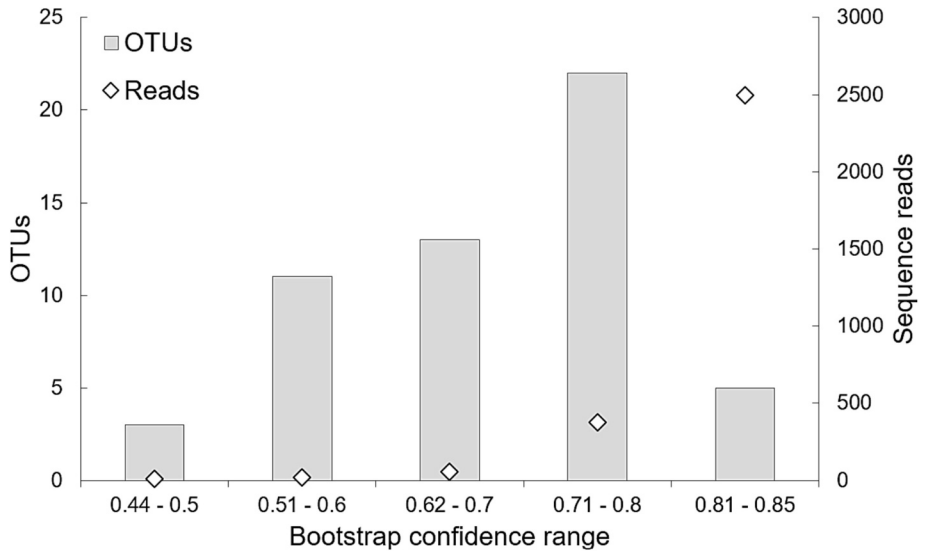

Fig. 2. Confidence value distribution of 54 OTUs with a total of 2958 ITS2 sequence reads representing the same species hypothesis *Ascomycota*_sp_SH1235051.08FU (curated as *Flagellospora curvula*).

penicillioides and *A. restrictus*, as well as *Chaetomium elatum* and *Ch. globosum*. Nine of 10 species were classified correctly by the pipeline. However, there were also some erroneous classifications: *Cladosporium_domesticum*_SH1966695.08FU was classified instead of *C. sphaerospermum*, which both belong to the same species complex. Fungi_sp_SH1162020.08FU was identified by the pipeline instead of *A. penicillioides*. Wrongly classified *Chaetomium afropilosum*_SH2475999.08FU belongs to the *Chaetomium globosum* species complex. It was identified by the classifier instead of *Ch. globosum*, which occurred as a singleton, while *C. afropilosum* was inflated by multiple OTUs. *Epicoccum nigrum* was classified as Fungi_sp_SH1547083.08FU. *Sarocladium strictum* and *Rhizopus stolonifer* were not retrieved at all. In addition, the different OTUs of the inflated SHs had varying sequence read numbers (see Table S5 for full dataset). Most of these OTUs have only one to few reads but often there is one "main" OTU represented with most of the reads and, if identified correctly, higher CV values compared to the other OTUs (cf. Fig. 2).

3.4. Step by step protocol

Parameters that can be adjusted to improve the classification of taxa in metabarcoding studies include adjusting the cut-off threshold for sequence similarity, updating and refining of databases, improving clustering algorithms (https://github.com/hsgweon/pipits, https://unite.ut.ee/) as well as aligning and collapsing of OTUs to species and taxon hypotheses (Kõljalg et al., 2013, 2020).Here, we present a step-by-step protocol for the curation of fungal sequence data derived from metabarcoding:

1) Use the classification table and the OTU dataset retrieved from your metabarcoding pipeline
2) Check for unclassified and misidentified species, e.g., Fungi sp., other high ranks, or cryptic species
3) Sort the table and check at least all sequences of SHs with very low bootstrap confidence range (CV-value)
4) Re-blast the sequences of questionable SHs using RefSeq/default GenBank, Mycobank, EUKARYOME (Tedersoo et al., 2024), specific databases, and/or internal databases; also, re-blast SHs with high CVs to confirm or curate their classification. Document at least the closest blast-hit (cf. Table S3).
5) Remove superfluous OTUs due to small sequence errors and intraspecific variability by merging of OTUs leading to the same SH; check validity of merging performing alignments of the merged OTU sequences
6) Check for the validity of taxonomic assignment using Mycobank and/or Index Fungorum
7) Correct the names in the classification table prior to further analyses
8) Document all curation steps in suitable tables (cf. Supplements to the present publication).

4. Discussion

Metabarcoding methods paved the way to perform large scale mycobiome studies for all kinds of habitats possible. The advantages of successful curation include providing more realistic data about taxonomic information, species traits, and guilds thus providing valuable information about ecological functions. The 97% cut-off used in this study is widely used and might be necessary to avoid overestimation of biodiversity, given the heterogeneity of ITS. On the other hand, too low cut-off for sequences of undescribed taxa can lead to an underestimation of biodiversity (Lücking et al., 2020; Stadler et al., 2020; Wilson et al., 2023). Since it is not possible to distinguish between unknown species and biases of the method (primer, database, etc.), the more conservative approach is preferable to avoid misinterpretation of community data. Sometimes it is difficult to tell if a questionable OTU is unknown or, e.g., a sequence artefact. The phenomenon of so-called "inflated species" due to small sequence differences in the range of PCR and sequencing errors (Anslan et al., 2018; Tedersoo et al., 2022) can, in combination with imperfect clustering algorithms, lead to an overestimation of taxon richness. Moreover, this also compromises the relative read proportions of taxa, which are often used as proxies for their abundance in a sample. Therefore, the determination of abundances based on sequence reads should be either avoided or at least regarded with extra caution (cf. Carl et al., 2025). One method to avoid the contamination of the dataset by PCR duplicates is the use of Unique Molecular Identifiers (UMI, Smith et al., 2017). On the other hand, keeping low read occurrences with, e.g., 1—10 sequence reads, such as singletons (single sequence read) and unique taxa (OTUs that occur only once within a dataset, but with more than one read) can be useful for the estimation of biodiversity, depending on the habitat probed. In aquatic leaf litter for example, a single specimen can colonise a large fraction of a leaf disc (Shearer and Lane, 1983).

Classification pipelines, especially clustering and other algorithms, are currently undergoing continuous development (e.g., Hakimzadeh

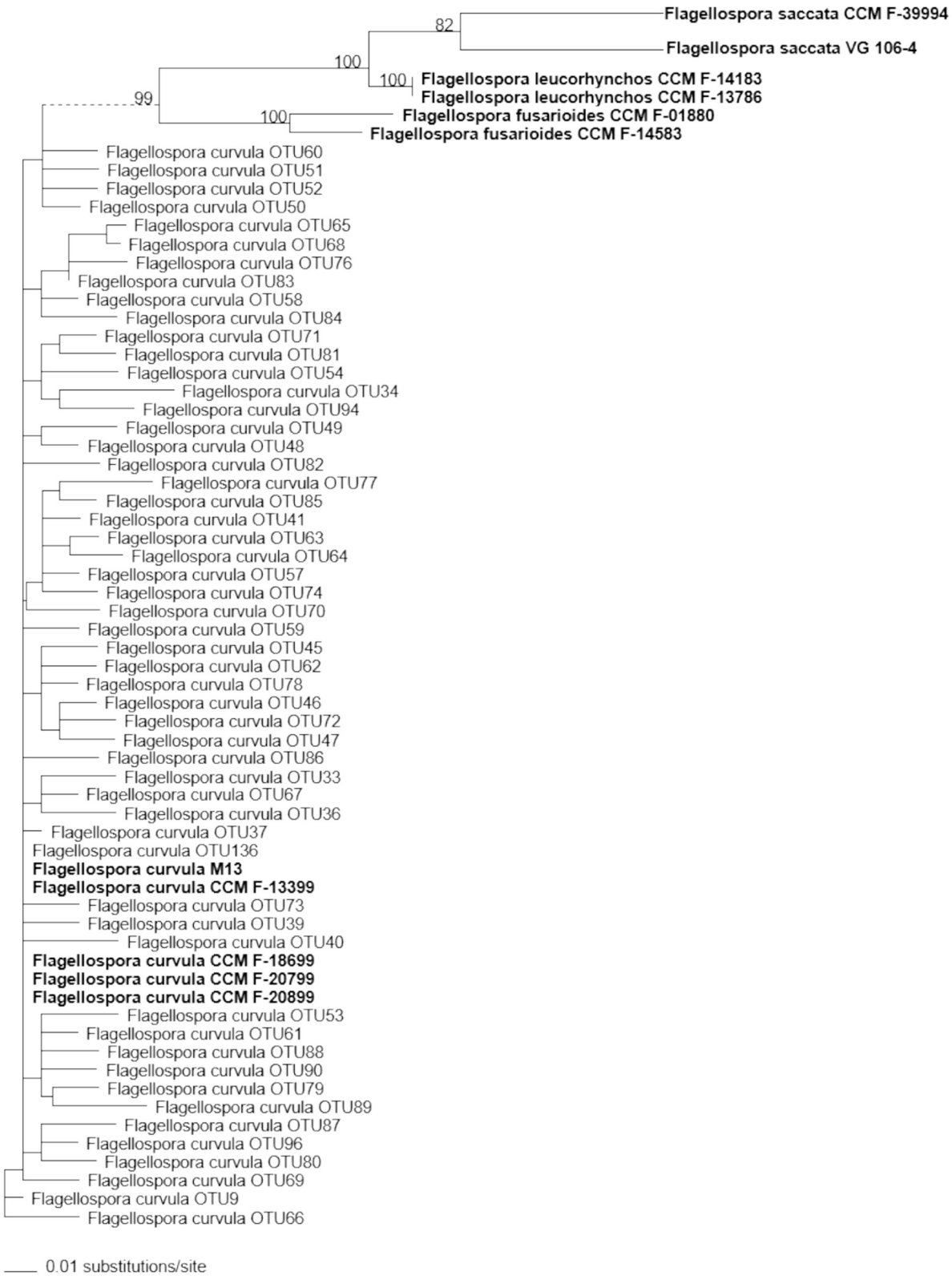

Fig. 3. Maximum likelihood tree of 54 ITS2 region based OTUs representing one species hypothesis and sequences of reference strains (in bold). Numbers at branches represent bootstrap percentages from 1000 replicates. Log Likelihood (ln*L*) of the tree = − 1602.856356.

et al., 2025) in order to avoid the bioinformatic biases summarized in the present article. Furthermore, databases used for classification should be curated continuously such as the RefSeq dataset of GenBank sequences (reference sequences from ex-type cultures and other reference cultures) and the UNITE SH database (most recent publication Abarenkov et al., 2024). The phylogenetic tree in Fig. 3 shows a small analysis with reference sequences as an example how a phylogenetic framework identifies differently assigned OTUs to one species hypothesis. However, it is not scalable to do this for all OTUs unless machine learning methods will be able to do this. Therefore, the most up-to-date

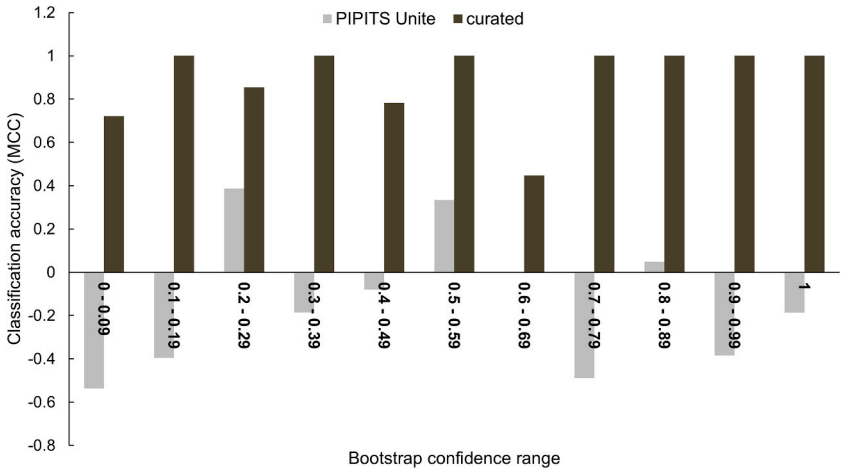

Fig. 4. Classification accuracy as calculated by the Matthews Correlation Coefficient (MCC) for bootstrap confidence values from classification of ITS2 sequences before (grey) and after (black) manual curation.

Table 1
Sequencing and classification success of mock community strains. Strain designations of the mock species are listed in Table S4.

Mock species	RDP Classifier result	OTUs	Reads
Cladosporium sphaerospermum	*Cladosporium_sphaerospermum*_SH1572915.08FU	24	1519
	*Cladosporium_sphaerospermum*_SH1744613.08FU	12	1033
	*Cladosporium_domesticum*_SH1966695.08FU	1	1
Cladosporium halotolerans	*Cladosporium_halotolerans*_SH1966692.08FU	2	6
Aspergillus penicillioides	*Aspergillus_penicillioides*_SH1160644.08FU	2	192
		1	3
	*Aspergillus_penicillioides*_SH1531574.08FU	1	12
	Fungi_sp_SH1162020.08FU		
Aspergillus restrictus	*Aspergillus_restrictus*_SH2105248.08FU	2	57
Acremonium fusidioides	*Acremonium_fusidioides*_SH1560627.08FU	3	502
Chaetomium elatum	*Chaetomium_elatum*_SH1615843.08FU	11	909
Chaetomium globosum	*Chaetomium_globosum*_SH2596458.08FU	1	1
	*Chaetomium_afropilosum*_SH2475999.08FU	12	426
Ceriporiopsis gilvescens	*Ceriporiopsis_gilvescens*_SH1599722.08FU	3	44
Wallemia sebi	*Wallemia_sebi*_SH1605562.08FU	6	147
Epicoccum nigrum	Fungi_sp_SH1547083.08FU	8	1546
Sarocladium strictum	–	–	–
Rhizopus stolonifer	–	–	–

databases available should be used in metabarcoding analyses and we highly recommend to perform alignments for particularly questionable species that are poorly covered in databases. Furthermore, we recommend to increase information of metabarcoding classification success by implementing the sequences of isolates derived from the probed habitat in the dataset of the classifier, as for example done in Salis et al. (2023). A reference sequence set of properly identified isolates (e.g., ex-type and authentic cultures) can be used to train the classifier data set which is very helpful for identification of fungi from metabarcoding data. The curation effort will vary depending on improvement and availability of pipelines and databases applied, e.g., latest pipeline version, latest update of databases such as UNITE and the relatively new rRNA reference database EUKARYOME (Tedersoo et al., 2024).

There is an urgent need to use additional and/or different fungal gene markers (e.g., Bai et al., 2020) to improve the identification of e.g., *Aspergillus* species (using calmodulin) or beta-tubulin for the identification of *Penicillium* species. Recently, a diagnostic study of human pathogenic fungi showed that the partial translation elongation factor alpha gene (tef1) outperforms the ITS1 region as a metabarcoding marker (Weaver et al., 2024). The drawbacks of the ITS rDNA region as a taxonomic marker have been discussed in numerous publications (see Introduction) and cause much of the current data that has to be curated. This leads us to the downside of the curation approach: The curation of hundreds of SHs can take several weeks, depending on the taxon richness of the habitat investigated. Furthermore, the curator needs to be an expert in fungal taxonomy, at least for the taxa of the investigated habitat. It is this taxonomic expertise that is often missing in ecological studies and is also difficult to automate (e.g., in machine learning methods). However, we hope that skilled bioinformaticians will be able to cooperate with taxonomists to achieve a solution for this gap found in all sequence processing pipelines we know.

CRediT authorship contribution statement

Christiane Baschien: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Steffen Carl:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization.

Ethical statement

No ethical statement was reported.

Funding

This research did not receive any specific grant from funding

agencies in the public, commercial, or not-for-profit sectors.

Declaration of competing interest

The authors have declared that no competing interests exist.

Acknowledgements

We would like to thank all colleagues from the Leibniz Institute German Collection of Microorganisms and Cell Cultures (DSMZ) involved in this study. Special thanks go to Carola Plagge and Carola Berg for assistance with the lab work as well as Cathrin Sproer, Boyke Bunk, and their colleagues from the sequencing department for conducting the Illumina MiSeq runs and providing bioinformatic resources.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mimet.2026.107447.

Data availability

We have shared the data providing links to data repositories and EMBL

References

Abarenkov, K., Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., May, T.W., Frøslev, T.G., Pawlowska, J., Lindahl, B., Põldmaa, K., Truong, C., Vu, D., Hosoya, T., Niskanen, T., Piirmann, T., Ivanov, F., Zirk, A., Peterson, M., Cheeke, T.E., Ishigami, Y., Jansson, A.T., Jeppesen, T.S., Kristiansson, E., Mikryukov, V., Miller, J.T., Oono, R., Ossandon, F.J., Paupério, J., Saar, I., Schigel, D., Suija, A., Tedersoo, L., Kõljalg, U., 2024. The UNITE database for molecular identification and taxonomic communication of fungi and other eukaryotes: sequences, taxa and classifications reconsidered. Nucleic Acids Res. 52, D791–D797. https://doi.org/10.1093/nar/gkad1039.

Anslan, S., Nilsson, R.H., Wurzbacher, C., Baldrian, P., Tedersoo, L., Bahram, M., 2018. Great differences in performance and outcome of high-throughput sequencing data analysis platforms for fungal metabarcoding. MycoKeys 39, 29–40. https://doi.org/10.3897/mycokeys.39.28109.

Bai, L., He, L., Yu, P., Luo, J., Yang, M., Wei, X., 2020. Molecular characterization of mycobiota and *aspergillus* species from *Eupolyphagasinensis* walker based on high-throughput sequencing of ITS1 and CaM. J. Food Qual. 2020, 172. https://doi.org/10.1155/2020/1752415.

Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F., Nielsen, H., 2000. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 16, 412–424. https://doi.org/10.1093/bioinformatics/16.5.412.

Begerow, D., Nilsson, H., Unterseher, M., Maier, W., 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. Appl. Microbiol. Biotechnol. 87, 99–108. https://doi.org/10.1007/s00253-010-2585-4.

Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., De Sousa, F., Amend, A., Jumpponen, A., Unterseher, M., Kristiansson, E., Abarenkov, K., Bertrand, Y.J.K., Sanli, K., Eriksson, K.M., Vik, U., Veldre, V., Nilsson, R.H., 2013. Improved software detection and extraction of ITS1 and ITS 2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. Methods Ecol. Evol. 4, 914–919. https://doi.org/10.1111/2041-210X.12073.

Bradshaw, M.J., Aime, M.C., Rokas, A., Maust, A., Moparthi, S., Jellings, K., Pane, A.M., Hendricks, D., Pandey, B., Li, Y., Pfister, D.H., 2023. Extensive intragenomic variation in the internal transcribed spacer region of fungi. iScience 26, 107317. https://doi.org/10.1016/j.isci.2023.107317.

Carl, S., Mohr, S., Sahm, R., Baschien, C., 2022. Laboratory conditions can change the complexity and composition of the natural aquatic mycobiome on *Alnus glutinosa* leaf litter. Fungal Ecol. 57–58, 101142. https://doi.org/10.1016/j.funeco.2022.101142.

Carl, S., Baschien, C., Mohr, S., Kusebauch, B., Bader, S., Contardo-Jara, V., Sahm, R., 2025. Effects of the fungicide penconazole on the leaf litter associated aquatic mycobiome in artificial stream channel and flask experiments. Ecotoxicol. Environ. Saf. 295, 118160. https://doi.org/10.1016/j.ecoenv.2025.118160.

Casas, J.J., Gessner, M.O., López, D., Descals, E., 2011. Leaf-litter colonisation and breakdown in relation to stream typology: insights from Mediterranean low-order streams: leaf-litter breakdown in relation to stream typology. Freshw. Biol. 56, 2594–2608. https://doi.org/10.1111/j.1365-2427.2011.02686.x.

Cedeño-Sanchez, M., Cheng, T., Lambert, C., Kolarík, M., Kuhnert, E., Cox, R.J., Kalinowski, J., Verwaaijen, B., Stadler, M., 2024. Unraveling intragenomic polymorphisms in the high-quality genome of *Hypoxylaceae*: a comprehensive study of the rDNA cistron. Mycol. Prog. 23, 5. https://doi.org/10.1007/s11557-023-01940-2.

Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics 21, 6. https://doi.org/10.1186/s12864-019-6413-7.

Cole, J.R., Wang, Q., Cardenas, E., Fish, J., Chai, B., et al., 2009. The ribosomal database project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res. 37 (Suppl. 1), D141–D145.

Duarte, S., Bärlocher, F., Trabulo, J., Cássio, F., Pascoal, C., 2015. Stream-dwelling fungal decomposer communities along a gradient of eutrophication unraveled by 454 pyrosequencing. Fungal Divers. 70, 127–148. https://doi.org/10.1007/s13225-014-0300-y.

Duarte, S., Cássio, F., Pascoal, C., Bärlocher, F., 2017. Taxa-area relationship of aquatic fungi on deciduous leaves. PLoS One 12, e0181545. https://doi.org/10.1371/journal.pone.0181545.

Ferreira, V., Castagneyrol, B., Koricheva, J., Gulis, V., Chauvet, E., Graça, M.A.S., 2015. A meta-analysis of the effects of nutrient enrichment on litter decomposition in streams. Biol. Rev. 90, 669–688. https://doi.org/10.1111/brv.12125.

Ferreira, V., Graça, M.A., Elosegi, A., 2023. A meta-analysis of drought effects on litter decomposition in streams. Hydrobiologia 850 (8), 1715–1735.

French, L.C., Jusino, M.A., Chambers, R.M., Skelton, J., 2024. Community structure and functional diversity of aquatic fungi are correlated with water quality: insights from multi-marker analysis of environmental DNA in a coastal watershed. Environ. DNA 6, e576. https://doi.org/10.1002/edn3.576.

Gonçalves, S., Feckler, A., Pollitt, A., Baschien, C., Michael, J., Schreiner, V.C., Zubrod, J. P., Bundschuh, M., 2024. Elevated fungicide and nutrient concentrations change structure but not function of aquatic microbial communities. Environ. Toxicol. Chem. 43, 1300–1311. https://doi.org/10.1002/etc.5863.

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. https://doi.org/10.1111/2041-210X.12399.

Hakimzadeh, A., Mikryukov, V., Metsoja, M., Tedersoo, L., Anslan, S., 2025. Are we throwing away good data? Evaluation of chimera detection algorithms on long-read amplicons reveals high false-positive rates across algorithms. PeerJ 13, e20456.

Heeger, F., Bourne, E.C., Baschien, C., Yurkov, A., Bunk, B., Spröer, C., Overmann, J., Mazzoni, C.J., Monaghan, M.T., 2018. Long-read DNA metabarcoding of ribosomal RNA in the analysis of fungi from aquatic environments. Mol. Ecol. Resour. 18, 1500–1514. https://doi.org/10.1111/1755-0998.12937.

Houbraken, J., Visagie, C.M., Frisvad, J.C., 2021. Recommendations to prevent taxonomic misidentification of genome-sequenced fungal strains. Microbiol. Resour. Announc. 10. https://doi.org/10.1128/MRA.01074-20 e01074–20.

Jumpponen, A., Jones, K.L., 2009. Massively parallel 454 sequencing indicates hyperdiverse fungal communities in temperate *Quercus macrocarpa* phyllosphere. New Phytol. 184, 438–448. https://doi.org/10.1111/j.1469-8137.2009.02990.x.

Kauserud, H., 2023. ITS alchemy: on the use of ITS as a DNA marker in fungal ecology. Fungal Ecol. 65, 101274. https://doi.org/10.1016/j.funeco.2023.101274.

Kõljalg, U., Nilsson, R.H., Abarenkov, K., Tedersoo, L., Taylor, A.F.S., Bahram, M., Bates, S.T., Bruns, T.D., Bengtsson-Palme, J., Callaghan, T.M., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G.W., Hartmann, M., Kirk, P.M., Kohout, P., Larsson, E., Lindahl, B.D., Lücking, R., Martín, M.P., Matheny, P.B., Nguyen, N.H., Niskanen, T., Oja, J., Peay, K.G., Peintner, U., Peterson, M., Põldmaa, K., Saag, L., Saar, I., Schüßler, A., Scott, J.A., Senés, C., Smith, M.E., Suija, A., Taylor, D.L., Telleria, M.T., Weiss, M., Larsson, K., 2013. Towards a unified paradigm for sequence-based identification of fungi. Mol. Ecol. 22, 5271–5277. https://doi.org/10.1111/mec.12481.

Kõljalg, U., Nilsson, H.R., Schigel, D., Tedersoo, L., Larsson, K.-H., May, T.W., Taylor, A. F.S., Jeppesen, T.S., Frøslev, T.G., Lindahl, B.D., Põldmaa, K., Saar, I., Suija, A., Savchenko, A., Yatsiuk, I., Adojaan, K., Ivanov, F., Piirmann, T., Pöhönen, R., Zirk, A., Abarenkov, K., 2020. The taxon hypothesis paradigm—on the unambiguous detection and communication of taxa. Microorganisms 8, 1910. https://doi.org/10.3390/microorganisms8121910.

Lazar, A., Griffiths, R.I., Goodall, T., Norton, L.R., Mushinski, R.M., Bending, G.D., 2024. Regional scale diversity and distribution of soil inhabiting *Tetracladium*. Environ. Microbiome 19, 111. https://doi.org/10.1186/s40793-024-00646-6.

Lieckfeldt, E., Seifert, K.A., 2000. An evaluation of the use of ITS sequences in the taxonomy of the *Hypocreales*. Stud. Mycol. 45, 35–44.

Lindahl, B.D., Nilsson, R.H., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjøller, R., Kõljalg, U., Pennanen, T., Rosendahl, S., Stenlid, J., Kauserud, H., 2013. Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. New Phytol. 199, 288–299. https://doi.org/10.1111/nph.12243.

Lücking, R., Aime, M.C., Robbertse, B., Miller, A.N., Ariyawansa, H.A., Aoki, T., Cardinali, G., Crous, P.W., Druzhinina, I.S., Geiser, D.M., Hawksworth, D.L., Hyde, K. D., Irinyi, L., Jeewon, R., Johnston, P.R., Kirk, P.M., Malosso, E., May, T.W., Meyer, W., Öpik, M., Robert, V., Stadler, M., Thines, M., Vu, D., Yurkov, A.M., Zhang, N., Schoch, C.L., 2020. Unambiguous identification of fungi: where do we stand and how accurate and precise is fungal DNA barcoding? IMA Fungus 11, 14. https://doi.org/10.1186/s43008-020-00033-z.

Ludwig, W., 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32, 1363–1371. https://doi.org/10.1093/nar/gkh293.

Matsuoka, S., Sugiyama, Y., Sato, H., Katano, I., Harada, K., Doi, H., 2019. Spatial structure of fungal DNA assemblages revealed with eDNA metabarcoding in a forest river network in western Japan. Metabarcoding Metagenomics 3, e36335. https://doi.org/10.3897/mbmg.3.36335.

Matthews, B.W., 1975. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochimica et Biophysica Acta (BBA) Protein Struct. 405, 442–451. https://doi.org/10.1016/0005-2795(75)90109-9.

Nikolcheva, L.G., Cockshutt, A.M., Bärlocher, F., 2003. Determining diversity of freshwater Fungi on decaying leaves: comparison of traditional and molecular approaches. Appl. Environ. Microbiol. 69, 2548–2554. https://doi.org/10.1128/AEM.69.5.2548-2554.2003.

Nilsson, R.H., Kristiansson, E., Ryberg, M., Hallenberg, N., Larsson, K.-H., 2008. Intraspecific *ITS* variability in the kingdom *Fungi* as expressed in the international sequence databases and ITS implications for molecular species identification. Evol. Bioinforma. 4, EBO.S653. https://doi.org/10.4137/EBO.S653.

Nilsson, R.H., Tedersoo, L., Ryberg, M., Kristiansson, E., Hartmann, M., Unterseher, M., Porter, T.M., Bengtsson-Palme, J., Walker, D.M., De Sousa, F., Gamper, H.A., Larsson, E., Larsson, K.-H., Kõljalg, U., Edgar, R.C., Abarenkov, K., 2015. A comprehensive, automatically updated fungal ITS sequence dataset for reference-based chimera control in environmental sequencing efforts. Microbes Environ. 30, 145–150. https://doi.org/10.1264/jsme2.ME14121.

Nilsson, R.H., Anslan, S., Bahram, M., Wurzbacher, C., Baldrian, P., Tedersoo, L., 2019a. Mycobiome diversity: high-throughput sequencing and identification of fungi. Nat. Rev. Microbiol. 17, 95–109. https://doi.org/10.1038/s41579-018-0116-y.

Nilsson, R.H., Larsson, K.-H., Taylor, A.F.S., Bengtsson-Palme, J., Jeppesen, T.S., Schigel, D., Kennedy, P., Picard, K., Glöckner, F.O., Tedersoo, L., Saar, I., Kõljalg, U., Abarenkov, K., 2019b. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. Nucleic Acids Res. 47, D259–D264. https://doi.org/10.1093/nar/gky1022.

O'Leary, N.A., Wright, M.W., Brister, J.R., Ciufo, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse, B., Smith-White, B., Ako-Adjei, D., Astashyn, A., Badretdin, A., Bao, Y., Blinkova, O., Brover, V., Chetvernin, V., Choi, J., Cox, E., Ermolaeva, O., Farrell, C. M., Goldfarb, T., Gupta, T., Haft, D., Hatcher, E., Hlavina, W., Joardar, V.S., Kodali, V.K., Li, W., Maglott, D., Masterson, P., McGarvey, K.M., Murphy, M.R., O'Neill, K., Pujar, S., Rangwala, S.H., Rausch, D., Riddick, L.D., Schoch, C., Shkeda, A., Storz, S.S., Sun, H., Thibaud-Nissen, F., Tolstoy, I., Tully, R.E., Vatsan, A. R., Wallin, C., Webb, D., Wu, W., Landrum, M.J., Kimchi, A., Tatusova, T., DiCuccio, M., Kitts, P., Murphy, T.D., Pruitt, K.D., 2016. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Res. 44, D733–D745. https://doi.org/10.1093/nar/gkv1189.

Paloi, S., Luangsa-ard, J.J., Mhuantong, W., Stadler, M., Kobmoo, N., 2022. Intragenomic variation in nuclear ribosomal markers and its implication in species delimitation, identification and barcoding in fungi. Fungal Biol. Rev. 42, 1–33. https://doi.org/10.1016/j.fbr.2022.04.002.

Robert, V., Vu, D., Amor, A.B.H., Van De Wiele, N., Brouwer, C., Jabas, B., Szoke, S., Dridi, A., Triki, M., Daoud, S.B., Chouchen, O., Vaas, L., De Cock, A., Stalpers, J.A., Stalpers, D., Verkley, G.J.M., Groenewald, M., Dos Santos, F.B., Stegehuis, G., Li, W., Wu, L., Zhang, R., Ma, J., Zhou, M., Gorjón, S.P., Eurwilaichitr, L., Ingsriswang, S., Hansen, K., Schoch, C., Robbertse, B., Irinyi, L., Meyer, W., Cardinali, G., Hawksworth, D.L., Taylor, J.W., Crous, P.W., 2013. MycoBank gearing up for new horizons. IMA Fungus 4, 371–379. https://doi.org/10.5598/imafungus.2013.04.02.16.

Salis, R.K., Schreiner, V.C., Rozenberg, A., Ohler, K., Baudy-Groh, P., Schäfer, R.B., Leese, F., 2023. Effects of fungicides on aquatic fungi and bacteria: a comparison of morphological and molecular approaches from a microcosm experiment. Environ. Sci. Eur. 35, 62. https://doi.org/10.1186/s12302-023-00768-7.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding Consortium, Fungal Barcoding Consortium Author List, Bolchacova, E., Voigt, K., Crous, P.W., Miller, A.N., Wingfield, M.J., Aime, M.C., An, K.-D., Bai, F.-Y., Barreto, R.W., Begerow, D., Bergeron, M.-J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, A.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, A., Cubas, P., Cummings, C., Damm, U., De Beer, Z.W., De Hoog, G.S., Del-Prado, R., Dentinger, B., Diéguez-Uribeondo, J., Divakar, P.K., Douglas, B., Dueñas, M., Duong, T.A., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliegerova, K., Furtado, M., García, M.A., Ge, Z.-W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.-D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.-M., Hoffmann, K., Hofstetter, V., Högnabba, F., Hollingsworth, P.M., Hong, S.-B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Kõljalg, U., Kovács, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, A.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, A. R., Methven, A.S., Meyer, W., Moncalvo, J.-M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W., Raja, H.A., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramírez, J.M., Schmitt, I., Schüßler, A., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.-O., Sung, G.-H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.A., Urbina, H., Vágvölgyi, C., Vialle, A., Vu, T.D., Walther, G., Wang, Q.-M., Wang, Y., Weir, B.S., Weiß, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, A., Zamora, J.-C., Zhang, N., Zhuang, W.-Y., Schindel, D., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for *Fungi*. Proc. Natl. Acad. Sci. 109, 6241–6246. https://doi.org/10.1073/pnas.1117018109.

Shearer, C.A., Lane, L.C., 1983. Comparison of three techniques for the study of aquatic hyphomycete communities. Mycologia 75, 498–508. https://doi.org/10.1080/00275514.1983.12023713.

Smith, T., Heger, A., Sudbery, I., 2017. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. Genome Res. 27, 491–499. https://doi.org/10.1101/gr.209601.116.

Stadler, M., Lambert, C., Wibberg, D., Kalinowski, J., Cox, R.J., Kolařík, M., Kuhnert, E., 2020. Intragenomic polymorphisms in the ITS region of high-quality genomes of the Hypoxylaceae (Xylariales, Ascomycota). Mycol. Prog. 19, 235–245. https://doi.org/10.1007/s11557-019-01552-9.

Stamatakis, A., 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. https://doi.org/10.1093/bioinformatics/btu033.

Tedersoo, L., Bahram, M., Põlme, S., Kõljalg, U., Yorou, N.S., Wijesundera, R., Ruiz, L.V., Vasco-Palacios, A.M., Thu, P.Q., Suija, A., Smith, M.E., Sharp, C., Saluveer, E., Saitta, A., Rosas, M., Riit, T., Ratkowsky, D., Pritsch, K., Põldmaa, K., Piepenbring, M., Phosri, C., Peterson, M., Parts, K., Pärtel, K., Otsing, E., Nouhra, E., Njouonkou, A.L., Nilsson, R.H., Morgado, L.N., Mayor, J., May, T.W., Majuakim, L., Lodge, D.J., Lee, S.S., Larsson, K.-H., Kohout, P., Hosaka, K., Hiiesalu, I., Henkel, T. W., Harend, H., Guo, L., Greslebin, A., Grelet, G., Geml, J., Gates, G., Dunstan, W., Dunk, C., Drenkhan, R., Dearnaley, J., De Kesel, A., Dang, T., Chen, X., Buegger, F., Brearley, F.Q., Bonito, G., Anslan, S., Abell, S., Abarenkov, K., 2014. Global diversity and geography of soil fungi. Science 346, 1256688. https://doi.org/10.1126/science.1256688.

Tedersoo, L., Anslan, S., Bahram, M., Põlme, S., Riit, T., Liiv, I., Kõljalg, U., Kisand, V., Nilsson, H., Hildebrand, F., Bork, P., Abarenkov, K., 2015. Shotgun metagenomes and multiple primer pair-barcode combinations of amplicons reveal biases in metabarcoding analyses of fungi. MycoKeys 10, 1–43. https://doi.org/10.3897/mycokeys.10.4852.

Tedersoo, L., Bahram, M., Zinger, L., Nilsson, R.H., Kennedy, P.G., Yang, T., Anslan, S., Mikryukov, V., 2022. Best practices in metabarcoding of fungi: from experimental design to results. Mol. Ecol. 31, 2769–2795. https://doi.org/10.1111/mec.16460.

Tedersoo, L., Hosseyni Moghaddam, M.S., Mikryukov, V., Hakimzadeh, A., Bahram, M., Nilsson, R.H., Yatsiuk, I., Geisen, S., Schwelm, A., Piwosz, K., Prous, M., Sildever, S., Chmolowska, D., Rueckert, S., Skaloud, P., Laas, P., Tines, M., Jung, J.-H., Choi, J.H., Alkahtani, S., Anslan, S., 2024. EUKARYOME: the rRNA gene reference database for identification of all eukaryotes. Database 2024, baae043. https://doi.org/10.1093/database/baae043.

Větrovský, T., Morais, D., Kohout, P., Lepinay, C., Algora, C., AwokunleHollá, S., Bahnmann, B.D., Bílohnědá, K., Brabcová, V., D'Alò, F., Human, Z.R., Jomura, M., Kolařík, M., Kvasničková, J., Lladó, S., López-Mondéjar, R., Martinović, T., Mašínová, T., Meszárošová, L., Michalčíková, L., Michalová, T., Mundra, S., Navrátilová, D., Odriozola, I., Piché-Choquette, S., Štursová, M., Švec, K., Tláskal, V., Urbanová, M., Vlk, L., Voříšková, J., Žifčáková, L., Baldrian, P., 2020. GlobalFungi, a global database of fungal occurrences from high-throughput-sequencing metabarcoding studies. Sci. Data 7, 228. https://doi.org/10.1038/s41597-020-0567-7.

Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl. Environ. Microbiol. 73, 5261–5267. https://doi.org/10.1128/AEM.00062-07.

Weaver, D., Novak-Frazer, L., Palmer, M., Richardson, M., Bromley, M., Bowyer, P., 2024. Development of a novel mycobiome diagnostic for fungal infection. BMC Microbiol. 24, 63. https://doi.org/10.1186/s12866-024-03197-5.

White T J, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal genes for phylogenetics, in: Innis A I, Gelfand D H, Sninsky J J, White T J (Eds) PCR Protocols. Academic Press, pp. 315-322doi:https://doi.org/10.1016/B978-0-12-372180-8.50042-1.

Wilson, A.W., Eberhardt, U., Nguyen, N., Noffsinger, C.R., Swenie, R.A., Loucks, J.L., Perry, B.A., Herrera, M., Osmundson, T.W., DeLong-Duhon, S., Beker, H.J., Mueller, G.M., 2023. Does one size fit all? Variations in the DNA barcode gaps of macrofungal genera. J. Fungi 9, 788. https://doi.org/10.3390/jof9080788.

Wurzbacher, C., Warthmann, N., Bourne, E., Attermeyer, K., Allgaier, M., Powell, J.R., Detering, H., Mbedi, S., Grossart, H.-P., Monaghan, M., 2016. High habitat-specificity in fungal communities in oligo-mesotrophic, temperate Lake Stechlin (North-East Germany). MycoKeys 16, 17–44. https://doi.org/10.3897/mycokeys.16.9646.