Forest Ecosystems 12 (2025) 100254

Contents lists available at ScienceDirect

Forest Ecosystems

journal homepage: www.keaipublishing.com/cn/journals/forest-ecosystems

Ghost species form an important component of the epiphytic lichens in temperate forests

Jan Vondrák [a,b], Jiří Košnar [b], Stanislav Svoboda [a,b], Zdeněk Palice [a], Jaroslav Šoun [c,*], Jiří Kubásek [d], Pavel Říha [a], Jiří Malíček [a], Jan Rydlo [a], Jeňýk Hofmeister [e,f]

[a] Institute of Botany of the Czech Academy of Sciences, Zámek 1, CZ-25243 Průhonice, Czech Republic
[b] Department of Botany, Faculty of Science, University of South Bohemia, Branišovská 1160, CZ-37005 České Budějovice, Czech Republic
[c] Museum of Dr. Bohuslav Horák in Rokycany, Branch of Museum of West Bohemia, nám. J. Urbana 141, CZ-33701 Rokycany, Czech Republic
[d] Department of Experimental Plant Biology, Faculty of Science, University of South Bohemia, Branišovská 1160, CZ-37005 České Budějovice, Czech Republic
[e] Department of Forest Ecology, Faculty of Forestry and Wood Sciences, Czech University of Life Sciences in Prague, Kamýcká 129, CZ-16500 Praha – Suchdol, Czech Republic
[f] Global Change Research Institute of the Czech Academy of Sciences, Bělidla 986/4a, 603 00, Brno, Czech Republic

ARTICLE INFO

Keywords:
Biodiversity
DNA barcode
ITS
Environmental sampling
Global change
Lichen
Mitochondrial SSU
Taxonomic survey

ABSTRACT

Sequencing of environmental samples has great potential for biodiversity research, but its application is limited by the lack of reliable DNA barcode databases for species identifications. Such a database has been created for epiphytic lichens of Europe, allowing us to compare the results of environmental sequencing with standard taxonomic surveys. The species undetected by taxonomic surveys (what we term the ghost component) amount to about half of the species actually present in hectare plots of Central European forests. Some of these, which currently occur only as diaspores or weakly developed thalli, are likely to be favoured in the course of global change. The ghost component usually represents a larger fraction in managed forests than in old-growth unmanaged forests. The total species composition of different plots is much more similar than suggested by taxonomic surveys alone. On a regional scale, this supports the well-known statement that "everything is everywhere, but, the environment selects".

1. Introduction

In forests, as in other biomes, the dominant component of biodiversity is made up of microorganisms and organisms of small macroscopic size. Most studies deal with the soil biota (e.g. Hofmann et al., 2023; Leclerc et al., 2023; Raimbault et al., 2024) and the biota inside living and dead plant bodies (e.g. Runnel et al., 2024; Saine et al., 2024). However, epiphytic organisms are also a species-rich component (Dreyling et al., 2022; Hofmann et al., 2023), which receive less attention due to their lesser economic importance. Epiphytic communities in temperate forests consist mainly of relatively small organisms (i.e. algae, bryophytes, lichens and microfungi) which are known to be sensitive bioindicators useful for monitoring air pollution and current global changes (Dittrich et al., 2022), or the impact of management on forest ecosystems (Kaufmann et al., 2018).

The sensitive bioindicators are undoubtedly epiphytic lichens (Delves et al., 2023), which are traditionally studied by taxonomists using phenotypic characters. Many lichen species are, however, small and barely detectable in the field, or separable from other species, owing to their scarcity of phenotypic characters. As a result, biodiversity surveys of these organisms based on phenotypic characters fail to record many of them (Vondrák et al., 2024). This fact has been demonstrated by studies comparing traditional taxonomic surveys with more powerful environmental sequencing (e.g. Wright et al., 2019; Henrie et al., 2022; Robison et al., 2023; Dreyling et al., 2024). However, most ot these studies did not have access to a reference database for identifying DNA sequences (Kerr and Leavitt, 2023), so do not provide a direct comparison with taxonomic surveys.

We now have a reference database of DNA barcodes (ITS and mtSSU) of European epiphytic lichens called Martin7 (Vondrák et al., 2023), currently including 1,172 species, with which we can identify species from sequences obtained by environmental sequencing in Central European forests. The taxonomic survey together with environmental sequence data (120 samples) from twenty forest sites along the

* Corresponding author.
E-mail address: jsoun@zcm.cz (J. Šoun).

https://doi.org/10.1016/j.fecs.2024.100254
Received 6 June 2024; Received in revised form 24 September 2024; Accepted 24 September 2024

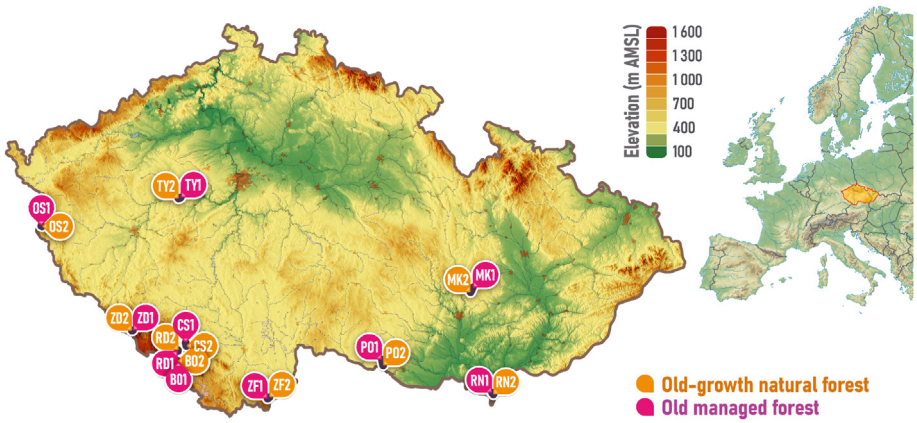

Fig. 1. Study sites in the Czech Republic in an altitudinal range 170–1,270 m.

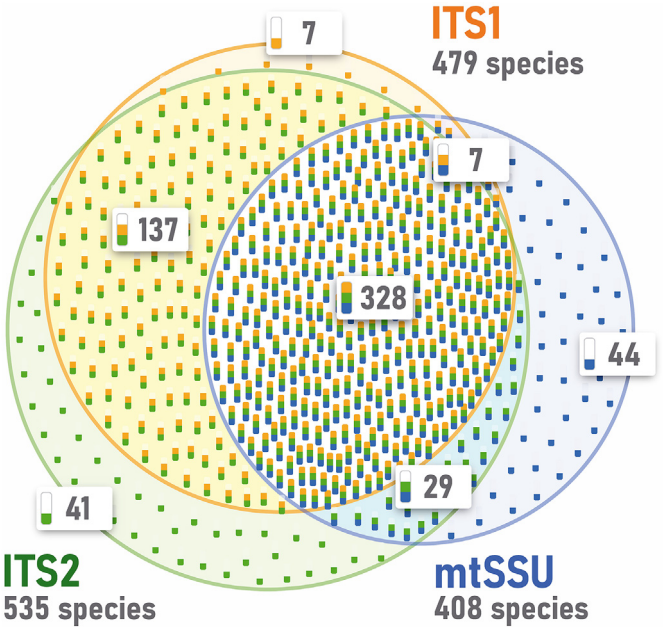

Fig. 2. Species detected by individual DNA barcodes and their combinations. Counted from the entire dataset.

altitudinal gradient in the Czech Republic (Fig. 1) allow us to address the specific question: What is the proportion of species detected only in environmental samples, not by taxonomic surveys? With this question, we delve deeper into two components within species lists, which we term "regular species/regular component" and "ghost species/ghost component". Regular species are those detectable by a traditional voucher-based taxonomic survey. Ghost species are not detectable by taxonomical survey for various reasons, including: (a) young colonizers present in a state of diaspores or unidentifiable initial thalli, (b) dying or dead species present in the form of unidentifiable thalli and (c) poorly known species that taxonomists do not recognize, and overlooked species. To explore these components of biodiversity at different spatial levels, we conducted an additional taxonomic survey and environmental sequencing on twenty individual tree trunks. Before we reached our conclusions about the "ghost component", we tested the reliability of data obtained from the detection of individual DNA barcodes. The results were satisfactory (see the first two sections of Results).

2. Material and methods

2.1. Field research & sampling

Ten pairs of 1-ha square plots were delineated in the Czech Republic to span the altitudinal gradient (170–1,270 m; Fig. 1) and the range of important forest communities. Abbreviations of plots refer to Fig. 1 and detailed information is available in Appendix S1: Table S1. The plots at altitudes >500 m cover forests dominated by *Picea abies* (BO), *Fagus sylvatica* (OS, ZD, ZF), *Abies alba* (CS), and a montane ravine forest with *Acer platanoides* and *A. pseudoplatanus* (RD). Forests at lower altitudes are mostly dominated by *Carpinus betulus*, *Fraxinus excelsior* and *Quercus petraea* (MK, PO, TY), the lowland flood-plain forest (RN) consisting of *Acer campestre*, *Carpinus betulus*, *Quercus robur* and *Fraxinus angustifolia*. Within the investigated forests, the plot pairs were represented by one plot located in an unmanaged old-growth forest stand in a nature reserve and the other in an old managed forest less than 5 km away. All plots were selected by the search for local hot-spots (Vondrák et al., 2018). When selecting the plots, we avoided the northern part of the country because of the high level of historic atmospheric pollution, which significantly impoverished the epiphytic biota. Therefore, most of the upper-altitude plots were selected in the Šumava Mts., an area with relatively low air pollution.

Epiphytic lichens (and semilichens in the sense of Vondrák et al., 2022) were sampled, where the word "epiphytic" is used in a broad sense and includes species occurring on all substrates composed of living or dead plants except epilithic/epigeic bryophytes and humus. Each plot was examined using both a classical taxonomic approach and environmental DNA sampling. The detailed taxonomic survey was performed by three lichenologists (Palice, Šoun, Vondrák) for approximately 8 h. Well-known species were identified directly in the field, while species requiring verification of identification by anatomical characters and secondary metabolite analysis were collected and subsequently archived in the herbarium PRA. Secondary metabolites were determined by TLC (thin layer chromatography) following Orange et al. (2010).

A minimum of five environmental samples, three by lichenologists and two by technicians (beforehand instructed about the ecological requirements of lichens), were collected from each plot; in total, 126 samples were obtained from the twenty plots. Environmental samples were collected from organic substrates, i.e., the bark of trunks 0–2 m in height, the surface of branches and twigs accessible from the ground, and all types of wood (e.g., logs, stumps, and snags up to 2 m in height). Individual samples were collected on the plots for ca. 2 h and were taken by scraping with a pre-cleaned knife into sterile 50 mL tubes. Each sample eventually contained about 30 mL of organic matter. Soil and

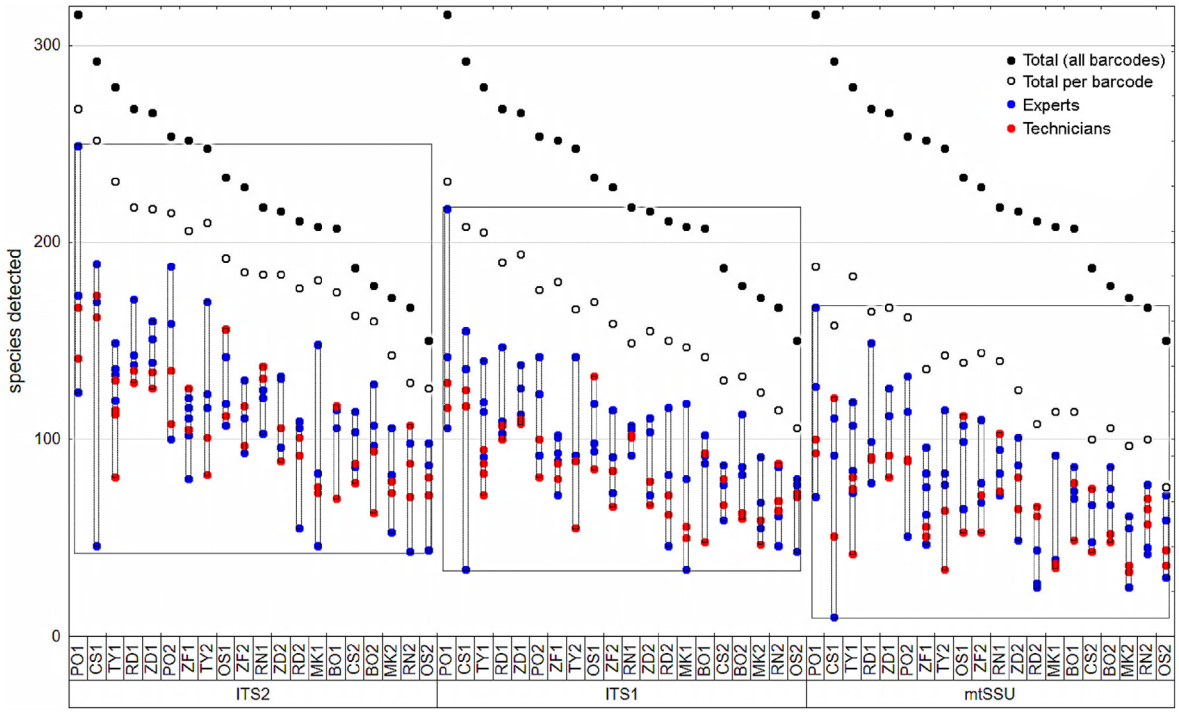

Fig. 3. Species detected in individual environmental samples from the twenty studied plots. Samples by lichenologists are blue dots, samples by technicians are red. Data shown separately for the DNA barcodes. Plots sorted by decreasing number of detected species in the all barcode dataset (black dots). (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

inorganic substrates were avoided. The rule of thumb was to prioritize communities of small crustose lichens over the collection of biomass-rich macrolichens and to cover the full diversity of microhabitats and substrates.

Additionally, taxonomic survey and environmental sampling were carried out on ten individual tree trunks in plot ZF1 and on ten trunks in ZF2. Tree trunks were surveyed over the entire area at a height of 0–2 m above the ground.

2.2. Sequencing of environmental samples (Extended version in Appendix S2: SI materials and methods)

Genomic DNA was extracted using DNeasy Plant Mini Kit (QIAGEN, Venlo, Netherlands) and purified using PowerClean Pro DNA Clean-Up Kit (QIAGEN, Venlo, Netherlands). The three barcodes (ITS1, ITS2 and mtSSU) were amplified using Combi Taq polymerase (Top-Bio, Praha, Czech Republic) and sample-specific uniquely tagged primers (tag sequences No. 1–142, attached to both forward and reverse primers). Cycling conditions are summarized in Appendix S1: Table S2. For each amplified sample, a negative control (an aliquot of PCR mixture without template) was subjected to PCR cycling and the absence of amplification was confirmed using agarose electrophoresis. Additionally, negative controls evaluated by sequencing were simultaneously prepared using uniquely tagged primers (tag sequence No. 143, attached to both forward and reverse primers) and common universal reagents used for processing of given batch of amplified samples. Amplifications were purified using NucleoSpin Gel and PCR Clean-up kit (Macherey Nagel, Düren, Germany) and pooled at equimolar ratio. The barcode pools were purified using Agencourt AMPure XP beads (Beckman Coulter Brea, California, USA) and sent for Illumina library preparation and paired-end 2 × 250 bp sequencing performed at SEQme Company (Dobříš, Czech Republic).

2.3. Bioinformatic analysis (Extended version in Appendix S2, SI materials and methods)

As some of the sequences in the reference database Martin7 (Vondrák et al., 2023) are incomplete, the data were processed in order to include barcode regions with highest coverage in the database. Therefore, regions starting from reverse primers were selected for ITS1 (primer ITS2) and mtSSU (primer mrSSU3R) barcodes, whereas region starting from forward primer (5.8S-Fun) was selected for ITS2 barcode.

The dataset included 110,565,274 paired-end reads. For ITS1 and ITS2 barcodes, reads were assembled using FLASH v1.2.11 (Magoč and Salzberg, 2011). The sequences were demultiplexed using SEED v.2.0 (Větrovský et al., 2018). Quality trimming was performed using VSEARCH v1.11.1 (Rognes et al., 2016). The sequences were oriented to start from reverse primer (ITS2) in ITS1 barcode or from forward primer (5.8S-Fun) in ITS2 barcode, respectively. Primer sequences were trimmed, yielding 13,723,340 (ITS1) and 11,435,674 (ITS2) sequences. As the length of both ITS barcodes varies greatly among fungi (approx. 145–695 bp in ITS1, and 267–511 bp in ITS2, respectively; Taylor et al., 2016), assembly of short Illumina reads fails to assemble complete sequence in taxa with barcode length exceeding ca. 400 bp. Therefore, non-assembled reads were also taken into consideration. Such reads were demultiplexed and filtered for target reads starting with desired primer (ITS2 primer for ITS1, 5.8S-Fun primer for ITS2, respectively). The sequences were quality trimmed and primer sequences were removed, yielding 818,190 (ITS1) and 109,067 (ITS2) non-assembled sequences. For mtSSU barcode dataset, the considerable amplicon length (approx. 800–1,100 bp) did not allow read assembly. The reads were demultiplexed and filtered for target reads starting with desired mrSSU3R primer. The sequences were quality trimmed and primer sequences were removed, yielding 2,092,392 mtSSU sequences.

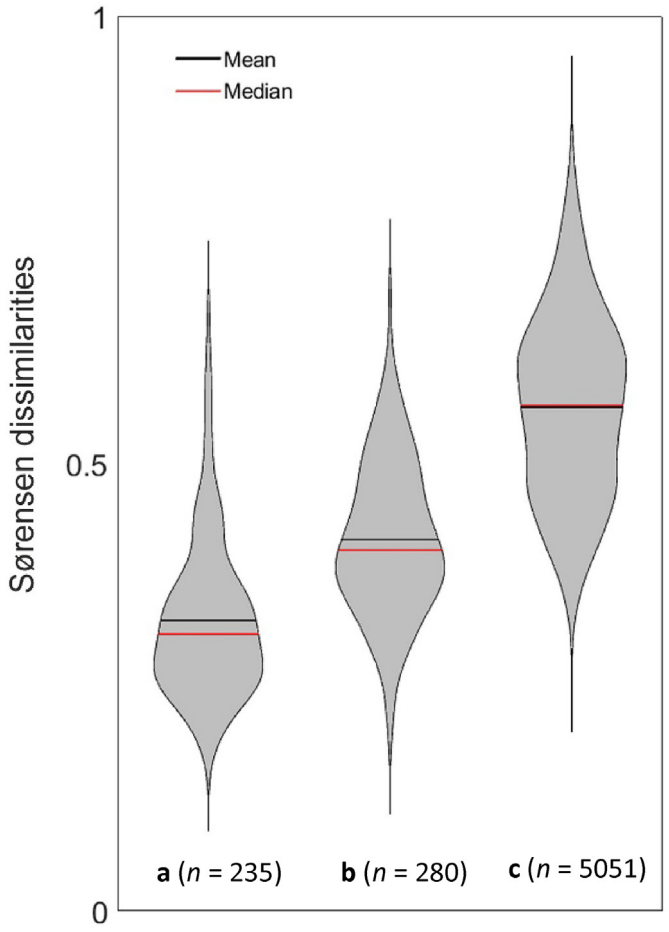

Fig. 4. Data distribution of Sørensen dissimilarity indices between environmental samples: **a**, between samples within plots; **b**, between samples in paired plots (natural vs. old managed); **c**, between samples from unpaired plots.

The sequences were BLAST identified with the sequences of lichen taxa from Martin7 database extracted into three barcode databases (ITS1, ITS2, and mtSSU) covering homologous regions as processed Illumina sequences. In taxa represented by multiple identical sequences, only one sequence was retained for the database. In taxa represented by multiple otherwise identical sequences but differing in length, the most complete (i.e. longest) sequence was selected. In case when two or more taxa showed identical sequence, the taxa were labelled as a group of indistinguishable taxa (see Appendix S1: Table S3). No sequences <100 bp were used unless they showed a unique genotype not shared with different taxa. In ITS2 database, another five sequences >100 bp but <200 bp were not included as they showed identity with taxa from other genera. The sequences in ITS1 and ITS2 databases were shortened to 180 and 220 bp, respectively. Uniform length in both ITS barcodes was necessary to avoid false positive detections of taxa represented by longer Martin7 sequences, because our trial BLAST using non-shortened sequences revealed significant artificial detection favouring congeners with database sequence longer than true target taxa. BLAST identification was performed using SSU pipeline (Vasar et al., 2017). The following criteria were required for a BLAST match: sequence similarity ≥97%; alignment length not differing from the length of the shorter of the query and subject sequences by >5%; and a BLAST e-value <1e–50. The negative controls evaluated by sequencing yielded negligible number of sequences (70 for ITS1; 10 for ITS2; and 803 for mtSSU, respectively), which all proved to be of non-lichen origin based on BLAST identification against Nucleotide collection (nr/nt) database at NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastn&BLAST_SPEC=GeoBlast&PAGE_TYPE=BlastSearch).

2.4. Analyses of biodiversity data

All data for the analyses are available in Dryad (https://doi.org/10.5061/dryad.fqz612k0g). We calculated dissimilarities in species composition among all plots from the results of taxonomic survey and environmental samples using Sørensen dissimilarity index as a measure of total beta-diversity (Baselga, 2010). Additionally, we computed two additive components of total beta-diversity: i) Simpson dissimilarity which represents species composition turnover and ii) remnant part of dissimilarity accounted for nestedness (details in Baselga, 2010). We tested the congruence between dissimilarities in species composition obtained from taxonomic survey and environmental samples, taking into account the geographic distances and altitudinal differences of respective plots. For this purpose, we used a partial Mantel test in vegan package (Oksanen et al., 2022), for which significance was assessed with 9999 permutations. Using a partial Mantel test, we also tested the potential relationship between dissimilarities in species composition and the two environmental variables: geographic distances and altitudinal differences among plots. Analyses were performed in R (R Core Team, 2013), using the packages 'betapart' (Baselga et al., 2015) and 'vegan' (Oksanen et al., 2022).

3. Results

3.1. Species detection by DNA barcodes

From the whole dataset, 593 species of epiphytic lichens were detected using all three barcodes, of which ITS2 detected 535, ITS1 479 and mtSSU 408 species. 328 species were represented by all barcodes, while 44 species were detected only by mtSSU, 41 only by ITS2 and 7 only by ITS1 (Fig. 2). The most common reason for mtSSU-only detection is the lack of reference sequences for ITS, while the most common reason for ITS2-only detection is a combination of the absence of mtSSU reference and likely low amplification for ITS1 (Appendix S1, Table S4). At the plot level, detection of species richness using individual barcodes correlates very closely with detection from the entire set of barcodes (Fig. 3).

3.2. Similarity of environmental samples is clearly higher within sites than between sites

Environmental samples had demonstrably the lowest dissimilarities within plots (Sørensen indices with quartiles 0.26–0.36), slightly higher among paired plots (0.36–0.47), and by far the highest among unpaired plots (0.48–0.63). The data distribution is summarized on the violin chart (Fig. 4), while the heatmap (Fig. 5) shows the dissimilarities among samples across the entire data set. Individual samples are most similar within plots and between paired plots (distinct blue areas along the diagonal in Fig. 5). However, some samples break this rule by having exceptionally low numbers of species (black arrows on the right Fig. 5). For example, two samples from plots MK1 and CS1 probably became mouldy and only common species were sequenced from them, causing a high level of nestedness and also a high dissimilarity with most other samples. Clearly, the variation between samples is overwhelmingly due to the species turnover and not to nestedness. This means that the samples have a balanced number of species (mean 130 species; quartiles 105–153), except for the exceptionally species-poor samples (the mouldy ones from MK1 and CS1), or conversely the exceptionally species-rich sample from the PO1 (red arrows on the right in Fig. 5).

The higher similarities in species composition of samples within plot pairs and especially within individual plots indicate the reliability of the environmental data obtained. Similarly, the stratification of similarities as a function of elevation (Fig. 5) is consistent with the assumption that upland species would be predominantly in samples from higher

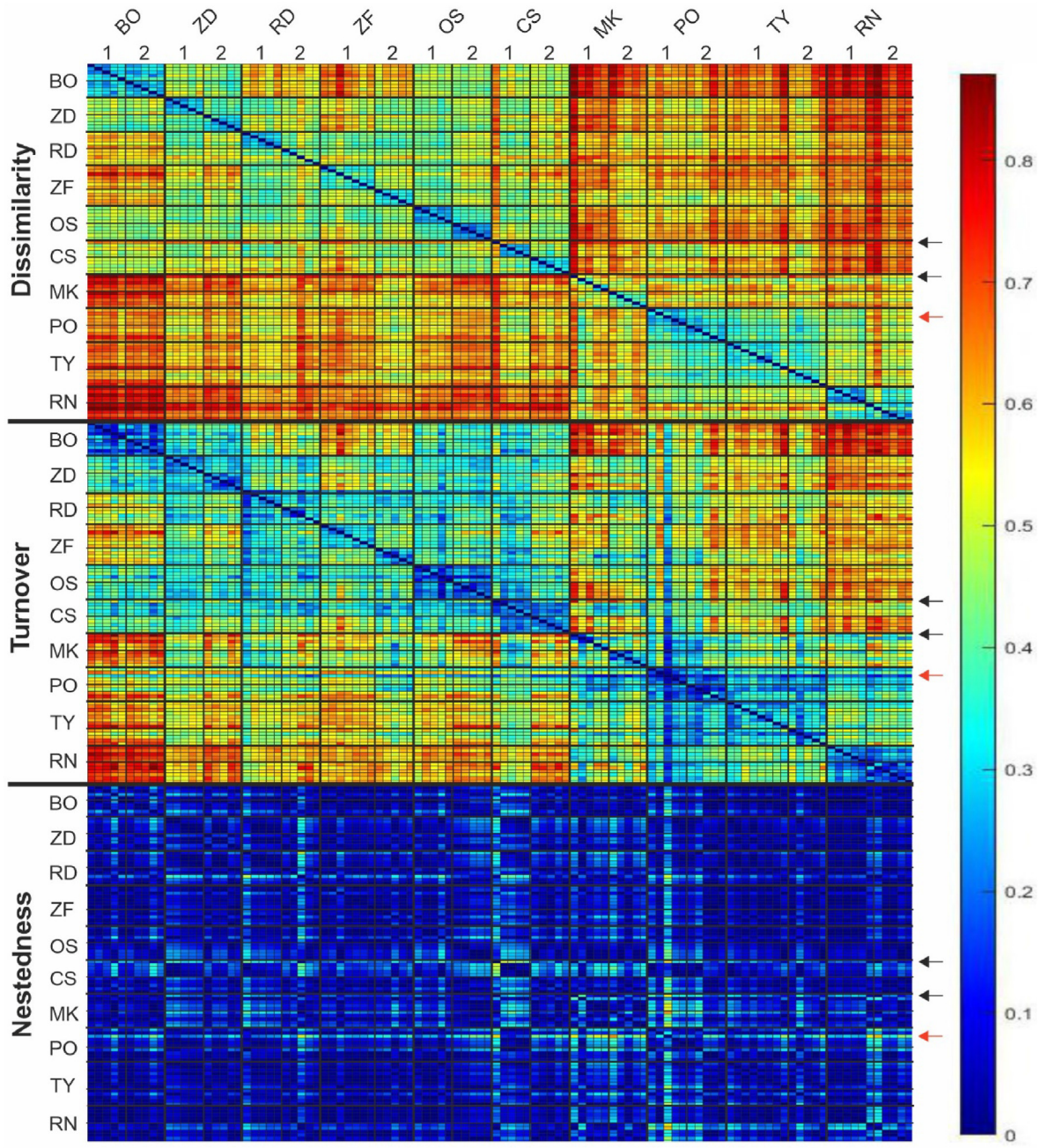

Fig. 5. Sørensen dissimilarities, turnover and nestedness between 106 environmental samples from the twenty studied plots ordered by decreasing altitude. Low altitudinal MK-RN are significantly distinct from upland BO-CS plots. 1 – unmanaged forest stand, 2 – old managed forest. Black arrows on the right – exceptionally species poor samples, red arrows on the right – exceptionally species rich sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

elevations, whereas lowland species would be the opposite.

We did not observe any clear differences between the results from samples taken by technicians versus those from lichenologists. The variability plot (Fig. 3) shows that the performances of technicians and lichenologists are very similar in terms of number of species. In terms of beta-diversity, the samples from lichenologists differ in species composition from those taken by technicians, but these differences are not significantly higher than the differences between samples from lichenologists only or from technicians only.

3.3. Taxonomic survey versus environmental sampling

The species richness revealed in the environmental samples greatly exceeds the numbers of species detected by the taxonomic survey. The entire dataset (gamma diversity from twenty study plots) includes 643 species, of which 157 (i.e. about 25%) species were detected by environmental sequencing alone. This 25% represents the ghost component (as defined above) of the total identified biodiversity. The remaining 75%, i.e. 486 species, represents the regular component (as defined above). Of this component, 48 species were identified by taxonomic survey only. Their absence in the environmental sample data is due mainly to the absence of reference barcode sequences (Appendix S1: Table S4).

At the 1-ha scale, the total number of species was about 220 species (mean 227, range 150–316), and of this the ghost component made up about half (Fig. 6). This proportion is lower and more balanced in the

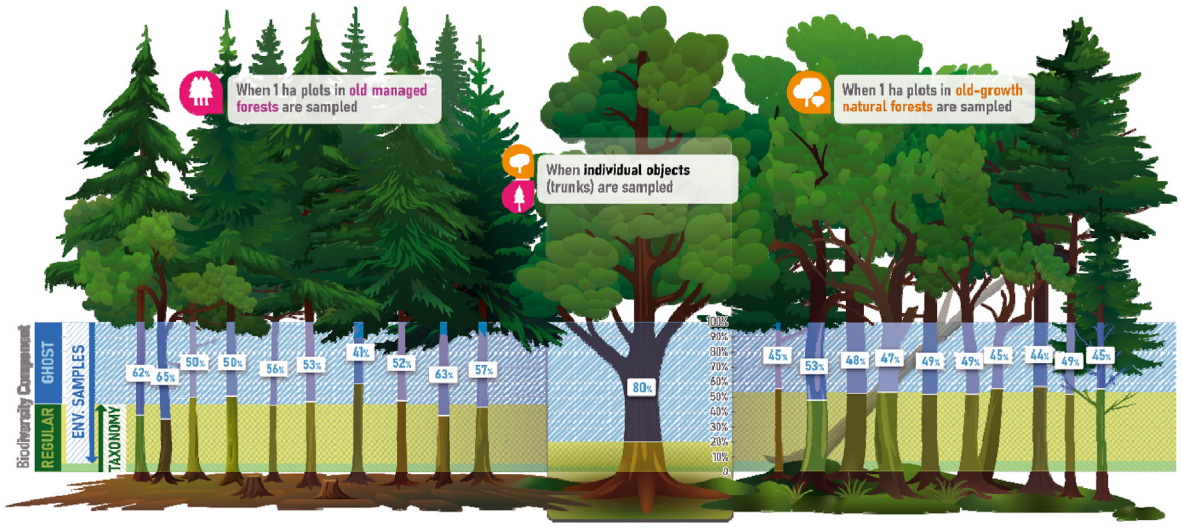

Fig. 6. Proportion of the ghost component and the regular component of biodiversity in 1-ha plots in old managed forests (left) and in old growth natural forests (right). Each tree trunk represents an individual plot, ordered from the northernmost. The average proportion of both components on individual objects (tree trunks) is in the middle.

case of unmanaged old-growth forest stands (mean 47%, range 44%–53%; Fig. 6, right) than in the case of old managed forests (mean 55%, range 41%–65%; Fig. 6, left). In nine out of ten plots of unmanaged forests, the regular component dominates over the ghost component, whereas in old managed forests the proportion is reversed (Appendix S3: Fig. S1).

At the scale of individual tree trunks, the ratio of ghost species is much higher than in plots (Fig. 6, middle) and highly variable (Appendix S3: Fig. S1). It appears to be slightly more balanced in unmanaged forests (mean 82%, range 71%–91%) than in old managed forests (mean 79%, range 60%–92%).

3.4. The ghost component mitigates differences in species composition

The ghost component tends to offset differences based on regular diversity. This is clearly visible at the scale of 1-ha plots (Fig. 7a) and even more so at the scale of individual tree trunks (Fig. 7b), where the ghost component complements the regular component so that the resulting total species composition varies much less between plots or tree trunks. Thus, the dissimilarity of plots based on individual components is generally higher than the dissimilarity of plots based on the full observed species richness (Fig. 7a and b, top). A similar pattern is seen in the species turnover (Fig. 7a and b, middle), but nestedness is generally low and lowest for the ghost component (Fig. 7a and b, bottom).

The ghost component also blurs the distinction between unmanaged forests and old managed forests, to some extent. Within the pairs of plots (old-growth unmanaged vs. old managed), the regular component and the total species richness are always distinctly higher in unmanaged old-growth forest stands, whereas the differences in the ghost component are less noticeable (Appendix S3: Fig. S2). In two cases, the ghost component is even higher in old managed forests.

3.5. Beta-diversity increases with geographical and altitudinal distances

In the dataset from taxonomic surveys, we found, unsurprisingly, a strong correlation between dissimilarity in species composition and geographic and altitudinal distances (Appendix S3: Figs. S3a and c). However, a similarly strong correlation was also evident in the environmental sample data (Appendix S3: Figs. S3b and d). In both cases, the turnover component of beta-diversity (i.e. differences in species composition) is responsible for these relationships (Appendix S3: Fig. S4). The nestedness component, reflecting mainly differences in species richness, is negligible in most cases.

4. Discussion

4.1. Everything is everywhere, but, the environment selects

The hypothesis that everything is everywhere, but, the environment selects (Baas Becking, 1934) has been tested and to some extent supported at various geographic and taxonomic levels in recent years (e.g. Fuhrman, 2009; Fondi et al., 2016). In terms of epiphytic lichens, the second part 'environment selects' is apparently valid in general, while the first part 'everything is everywhere' is supported at local and regional levels. Our most important result is that adding the ghost component of biodiversity to the regular component yields much higher similarities (i.e., lower dissimilarities in Fig. 7) between plots than does the regular component alone. This suggests the existence of a kind of "Central European pool of species" that are able to disperse regionally and persist in the form of diaspores or weakly developed thalli in sites/microhabitats where they do not currently find suitable conditions for the development of typical phenotypes. However, 'everything is everywhere' does not seem to work on a supra-regional scale, and significant limits are visible even on a regional scale.

(1) Absence of exotic species. In one hundred and twenty environmental samples from Central Europe, "non-Central European" species were absent. In other words, lichens known to be restricted to non-Central European regions (e.g. Mediterranean or EU-oceanic) were absent, only with the exception of the Macaronesian/western European *Lecania falcata* (Sérusiaux et al., 2012), which is however a microlichen, and may simply have been overlooked in Central Europe.
(2) Almost complete absence of species extinct from the Czech Republic. The most recent Czech Red list (Malíček, 2023) reports on 59 epiphytic lichens extinct from the Czech Republic, of which only three occurred in our samples in very low abundances: *Cladonia cyanipes*, *Cliostomum corrugatum*, and *Physconia detersa*. However, these species have recently been recorded in some Central European regions.
(3) Species lists from environmental samples correspond to local conditions. For example, upland species are almost exclusively

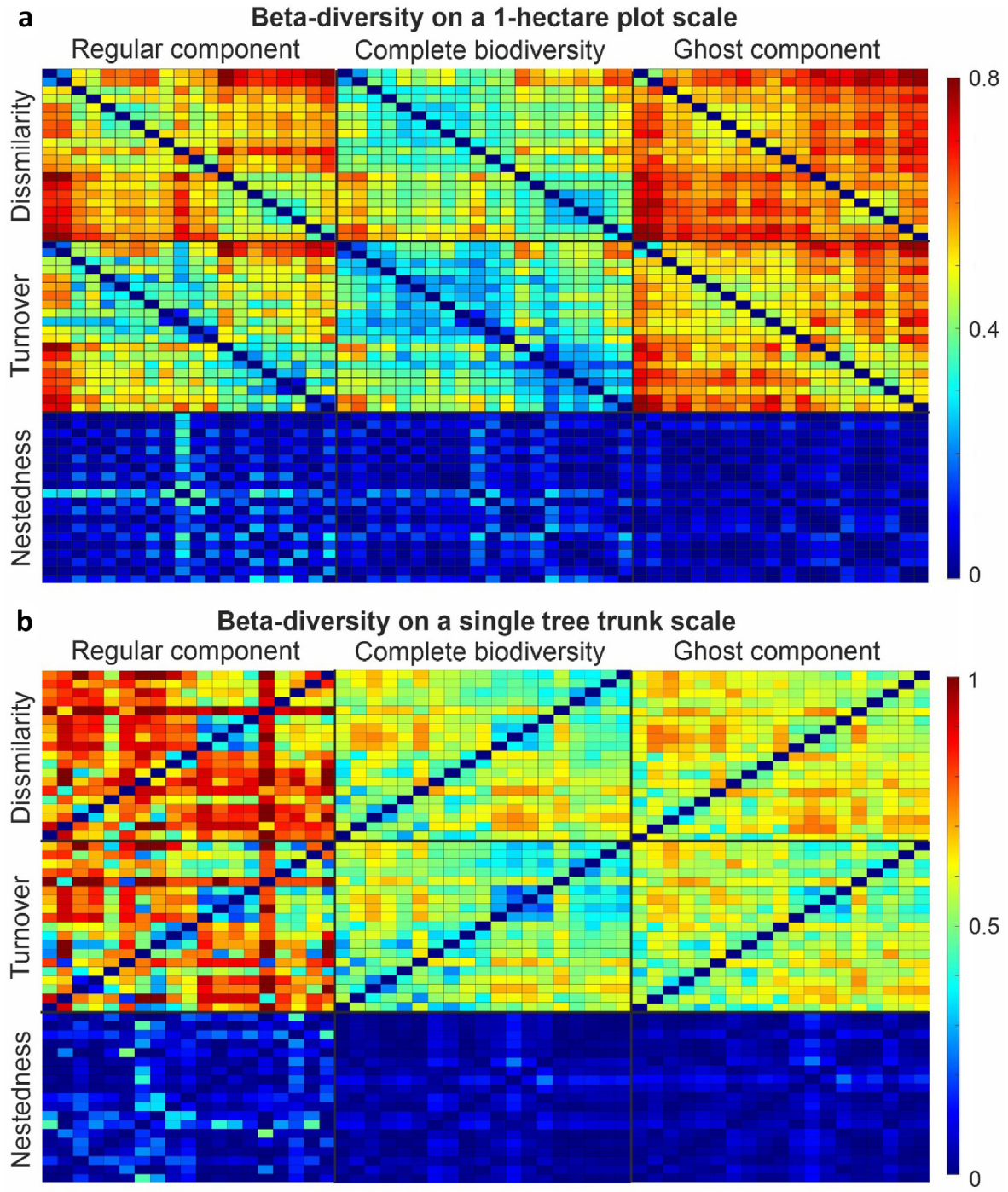

Fig. 7. Beta-diversity (dissimilarity) between the twenty studied plots ordered by decreasing altitude **(a)** and the twenty individual tree trunks **(b)**. Overall Sørensen dissimilarity is plotted here along with its two components: turnover (reflecting species composition) and nestedness (reflecting species richness). Entire biodiversity, the ghost component and the regular component are shown as separate data sources.

found in upland plots and lowland species in lowland plots. Plots of mountain spruce forest are very different from the lowland plots of thermophilous broadleaf forest (Appendix S3: Fig. S3d).

(4) Rarely spotted species are also rare as ghosts. Species considered to be regionally rare (Malíček et al., 2023) and species rarely recorded by taxonomic surveys are usually rarely represented in environmental samples. Conversely, species abundantly recorded by taxonomic surveys are also abundant in environmental samples (Appendix S3: Fig. S5).

4.2. Single tree trunks have significantly more ghosts

It is generally accepted for epiphytes that by reducing the size of the study area, fewer species will be detected, but more accurate data on species composition will be obtained (McCune and Lesica, 1992). This is probably true for the regular component of biodiversity, but the opposite is true for the ghost component, which increases with reducing sampling area (Note that we are talking here about the ghost component as a fraction of the total number of species, not about the number of species in

the ghost component). The enormous increase in the ghost component when moving from the 1-ha level to the individual tree level can be explained by the existence of a "local species pool" that includes a number of species commonly found in a given forest site in the form of diaspores or poorly developed thalli, but which only rarely, under specific circumstances, develop into the adult phenotype (e.g. only in hardly accessible, well-lit canopies). When exploring a 1-ha plot, taxonomists are able to detect numerous species that rarely develop into the adult phenotype (often on a single tree in the plot) and being detected by taxonomical survey, these species become a part of the regular component. By contrast, when exploring a single tree, the large part of the local species pool is present only in the state of ghost component in relation to a particular tree trunk – such species are either not developed yet and wait for their chance, or they are invisible again in the process of dying out. Only a limited number of species find a suitable habitat for the development of an adult phenotype on specific trees (of a specific age, character of the bark, moisture and light conditions, etc.).

4.3. In managed forests, epiphytes wait for suitable conditions in the form of ghost species

It is well known that managed forests are impoverished in epiphytes compared to old-growth unmanaged forests (e.g. Nascimbene et al., 2010; Strengbom et al., 2011; Malíček et al., 2019). The main reasons are for example lower structural heterogeneity and availability of veteran trees and coarse woody debris in managed forests (e.g. Hofmeister et al., 2016; Janssen et al., 2019; Kozák et al., 2023). Our data also show that plots in managed forests are poorer in species compared to plots in unmanaged forests (Appendix S3: Fig. S2), but have a relatively higher proportion of ghost species (Fig. 6). This suggests that older managed forests have already accumulated a large number of species, probably coming from nearby sources (in our case unmanaged old-growth forests). However, a significant proportion of these species cannot yet find suitable conditions (e.g. moisture, microhabitats) to develop into the adult phenotype. This is a strong argument for protecting old managed forests to fulfil their potential for rare epiphytic biota.

4.4. Our results are robust, with some qualifications

We consider that the main results presented here are robust, and a good representation of reality in Central European forests. We think it likely that they will apply over a much broader geographical region than Central Europe. It would be premature to speculate whether they apply to forests globally, or to non-forest vegetation, but the possibility does not seem unreasonable. However, the actual numbers and ratios cited here are specific to this study, as they depend on several methodological steps and decisions. The following three are probably the most important: (1) intensity and quality of the field research, (2) depth of the environmental sequencing and (3) completeness and reliability of the reference DNA barcode database. The first involves e.g. size of team, unequal taxonomic skills of workers, time management, conditions during fieldwork (For example, numerous tiny lichens become unrecognizable when wetted by rain). The second and third aspects strongly influence detection from environmental samples. We obtained tens to hundreds of thousands sequence reads per sample per barcode. Rather different results are expected when employing millions or more reads per sample that ought to detect more rare species, but can also be more susceptible to contaminations. The state of the DNA barcode database (Martin7; Vondrák et al., 2023), which at present is good but not complete, also inevitably influences the results.

Data availability

Data are provided as private-for-peer review via the following link https://datadryad.org/stash/share/pvfeA6Yfq3NF3Pm97DBQbtQfG1HOC9AgaVvXN1Zh5cM.

Funding

This work was supported by the Technology Agency of the Czech Republic (Grant Nos. SS01010270, SS06010420), and by a long-term research development grant RVO (Grant No. 67985939).

CRediT authorship contribution statement

Jan Vondrák: Writing – original draft, Methodology, Data curation, Conceptualization. **Jiří Košnar:** Formal analysis. **Stanislav Svoboda:** Data curation. **Zdeněk Palice:** Data curation. **Jaroslav Šoun:** Writing – review & editing, Data curation. **Jiří Kubásek:** Visualization. **Pavel Říha:** Visualization. **Jiří Malíček:** Data curation. **Jan Rydlo:** Data curation. **Jeňýk Hofmeister:** Formal analysis.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Linda in Arcadia suggested the term "ghost species" and kindly revised the manuscript. Anna Mácová, Václav Pouska, Zuzana Sejfová and Monika Vondráková contributed in collecting of environmental samples.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fecs.2024.100254.

References

Baas Becking, L.G.M., 1934. Geobiologie of inleiding tot de milieukunde. W.P. Van Stockum & Zoon, The Hague.

Baselga, A., 2010. Partitioning the turnover and nestedness components of beta diversity. Global Ecol. Biogeogr. 19, 134–143. https://doi.org/10.1111/j.1466-8238.2009.00490.x.

Baselga, A., Orme, D., Villeger, S., De Bortoli, J., Leprieur, F., 2015. Package "betapart" – Partitioning Beta Diversity into Turnover and Nestedness Components, 1.3. https://cran.r-project.org/web/packages/betapart/.

Delves, J., Lewis, J.E.J., Ali, N., Asad, S.A., Chatterjee, S., Crittenden, P.D., Jones, M., Kiran, A., Pandey, B.P., Reay, D., Sharma, S., Tshering, D., Weerakoon, G., van Dijk, N., Sutton, M.A., Wolseley, P.A., Ellis, C.J., 2023. Lichens as spatially transferable bioindicators for monitoring nitrogen pollution. Environ. Pollut. 328, 121575. https://doi.org/10.1016/j.envpol.2023.121575.

Dittrich, S., Lang, R., Albrecht, B.M., Stetzka, K.M., von Oheimb, G., 2022. Vertical distribution of cryptogamic epiphytes on trees in central German alluvial hardwood forests: relevance for bioindication and nature conservation. Herzogia 35, 443–461. https://doi.org/10.13158/heia.35.2.2022.443.

Dreyling, L., Boch, S., Lumbsch, H.T., Schmitt, I., 2024. Surveying lichen diversity in forests: a comparison of expert mapping and eDNA metabarcoding of bark surfaces. MycoKeys 106, 153–172. https://doi.org/10.3897/mycokeys.106.117540.

Dreyling, L., Schmitt, I., Dal Grande, F., 2022. Tree size drives diversity and community structure of microbial communities on the bark of Beech (*Fagus sylvatica*). Front. For. Glob. Change 5, 858382. https://doi.org/10.3389/ffgc.2022.858382.

Fondi, M., Karkman, A., Tamminen, M.V., Bosi, E., Virta, M., Fani, R., Alm, E., McInerney, J.O., 2016. "Every gene is everywhere but the environment selects": global geolocalization of gene sharing in environmental samples through network analysis. Genome Biol. Evol. 8, 1388–1400. https://doi.org/10.1093/gbe/evw077.

Fuhrman, J.A., 2009. Microbial community structure and its functional implications. Nature 459, 193–199. https://doi.org/10.1038/nature08058.

Henrie, J.R., Thomson, B.M., Yungfleisch, A.A., Kerr, M., Leavitt, S.D., 2022. Characterizing crustose lichen communities—DNA metabarcoding reveals more than meets the eye. Diversity 14, 766. https://doi.org/10.3390/d14090766.

Hofmann, B., Dreyling, L., Dal Grande, F., Otte, J., Schmitt, I., 2023. Habitat and tree species identity shape aboveground and belowground fungal communities in central European forests. Front. Microbiol. 14, 1067906. https://doi.org/10.3389/fmicb.2023.1067906.

Hofmeister, J., Hošek, J., Malíček, J., Palice, Z., Syrovátková, L., Steinová, J., Černajová, I., 2016. Large beech (*Fagus sylvatica*) trees as 'lifeboats' for lichen diversity in central European forests. Biodivers. Conserv. 25, 1073–1090. https://doi.org/10.1007/s10531-016-1106-x.

Janssen, P., Bergès, L., Fuhr, M., Paillet, Y., 2019. Do not drop OLD for NEW: conservation needs both forest continuity and stand maturity. Front. Ecol. Environ. 17, 370–371. https://doi.org/10.1002/fee.2086.

Kaufmann, S., Hauck, M., Leuschner, C., 2018. Effects of natural forest dynamics on vascular plant, bryophyte, and lichen diversity in primeval *Fagus sylvatica* forests and comparison with production forests. J. Ecol. 106, 2421–2434. https://doi.org/10.1111/1365-2745.12981.

Kerr, M., Leavitt, S.D., 2023. A custom regional DNA barcode reference library for lichen-forming fungi of the Intermountain West, USA, increases successful specimen identification. J. Fungi 9, 741. https://doi.org/10.3390/jof9070741.

Kozák, D., Svitok, M., Zemlerová, V., Mikoláš, M., Lachat, T., Larrieu, L., Paillet, Y., Buechling, A., Bače, R., Keeton, W.S., Vítková, L., Begovič, K., Čada, V., Dušátko, M., Ferenčík, M., Frankovič, M., Gloor, R., Hofmeister, J., Janda, P., Kameniar, O., Knír, T., Majdanová, L., Mejstřík, M., Pavlin, J., Ralhan, D., Rodrigo, R., Roibu, C.-C., Synek, M., Vostarek, O., Svoboda, M., 2023. Importance of conserving large and old trees to continuity of tree-related microhabitats. Conserv. Biol. 37, e14066. https://doi.org/10.1111/cobi.14066.

Leclerc, L., Calderón-Sanou, I., Martinez-Almoyna, C., Paillet, Y., Thuiller, W., Vincenot, L., Kunstler, G., 2023. Beyond the role of climate and soil conditions: living and dead trees matter for soil biodiversity in mountain forests. Soil Biol. Biochem. 187, 109194. https://doi.org/10.1016/j.soilbio.2023.109194.

Magoč, T., Salzberg, S.L., 2011. FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics 27, 2957–2963. https://doi.org/10.1093/bioinformatics/btr507.

Malíček, J., 2023. Červený seznam lišejníků ČR dle DaLiBora pro rok 2023 [Redlist of lichens in the Czech Republic, version 2023]. https://dalib.cz/data/redlist. (Accessed 8 August 2023).

Malíček, J., Palice, Z., Bouda, F., Knudsen, K., Šoun, J., Vondrák, J., Novotný, P., 2023. Atlas českých lišejníků [Atlas of Czech lichens]. https://dalib.cz/. (Accessed 15 September 2023).

Malíček, J., Palice, Z., Vondrák, J., Kostovčík, M., Lenzová, V., Hofmeister, J., 2019. Lichens in old growth and managed mountain spruce forests in the Czech Republic: assessment of biodiversity, functional traits and bioindicators. Biodivers. Conserv. 28, 3497–3528. https://doi.org/10.1007/s10531-019-01834-4.

McCune, B., Lesica, P., 1992. The trade-off between species capture and quantitative accuracy in ecological inventory of lichens and bryophytes in forests in Montana. Bryologist 95, 296–304. https://doi.org/10.2307/3243488.

Nascimbene, J., Marini, L., Nimis, P.L., 2010. Epiphytic lichen diversity in old-growth and managed *Picea abies* stands in Alpine spruce forests. For. Ecol. Manag. 260, 603–609. https://doi.org/10.1016/j.foreco.2010.05.016.

Oksanen, J., Simpson, G.L., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Solymos, P., Stevens, M.H.H., Szoecs, E., Wagner, H., Barbour, M., Bedward, M., Bolker, B., Borcard, D., Carvalho, G., Chirico, M., de Caceres, M., Durand, S., Evangelista, H.B.A., FitzJohn, R., Friendly, M., Furneaux, B., Hannigan, G., Hill, M.O., Lahti, L., McGlinn, D., Ouellette, M., Cunha, E.R., Smith, T., Stier, A., Braak, C.J.F.T., Weedon, J., 2022. Package 'vegan' – Community Ecology Package 2.6–4. https://cran.r-project.org/web/packages/vegan/index.html. (Accessed 15 September 2023).

Orange, A., James, P.W., White, F.J., 2010. Microchemical Methods for the Identification of Lichens. British Lichen Society, London.

R Core Team, 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/. (Accessed 15 September 2023).

Raimbault, A., Brin, A., Manzi, S., Savoie, J.M., Gandois, L., Oliva, P., Fogliani, O., Roy-Camille, C., Gratacap, L., Roy, M., 2024. Influence of habitat fragmentation and habitat amount on soil fungi communities in ancient forests. Landsc. Ecol. 39, 19. https://doi.org/10.1007/s10980-024-01821-3.

Robison, A., Adams, P., Leavitt, S.D., 2023. Discovering neglected lichen diversity with DNA-based inventories: metabarcoding lichen-forming fungi in Bryce Canyon National Park, Utah, USA. Plant Fungal System 68, 395–410. https://doi.org/10.35535/pfsyst-2023-0007.

Rognes, T., Flouri, T., Nichols, B., Quince, Ch, Mahé, F., 2016. VSEARCH: a versatile open source tool for metagenomics. PeerJ 4, e2584. https://doi.org/10.7717/peerj.2584.

Runnel, K., Lõhmus, P., Küngas, K., Marmor-Ohtla, L., Palo, A., Pütsepp, G., Tamm, H., Tammekänd, I., Lõhmus, A., 2024. Aerial eDNA contributes vital information for fungal biodiversity assessment. J. Appl. Ecol. 1–12. https://doi.org/10.1111/1365-2664.14691.

Saine, S., Penttilä, R., Furneaux, B., Monkhouse, N., Zakharov, E.V., Ovaskainen, O., Abrego, N., 2024. Natural deadwood hosts more diverse pioneering wood-inhabiting fungal communities than restored deadwood. Restor. Ecol. 32, e14056. https://doi.org/10.1111/rec.14056.

Sérusiaux, E., Boom, P.P.G. van den, Brand, M.A., Coppins, B.J., Magain, N., 2012. *Lecania falcata*, a new species from Spain, the Canary Islands and the Azores, close to *Lecania chlorotiza*. Lichenologist 44, 577–590. https://doi.org/10.1017/S0024282912000308.

Strengbom, J., Dahlberg, A., Larsson, A., Lindelöw, Å., Sandström, J., Widenfalk, O., Gustafsson, L., 2011. Introducing intensively managed spruce plantations in Swedish forest landscapes will impair biodiversity decline. Forests 2, 610–630. https://doi.org/10.3390/f2030610.

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. https://doi.org/10.1128/AEM.02576-16.

Vasar, M., Andreson, R., Davison, J., Jairus, T., Moora, M., Remm, M., Young, J.P.W., Zobel, M., Öpik, M., 2017. Increased sequencing depth does not increase captured diversity of arbuscular mycorrhizal fungi. Mycorrhiza 27, 761–773. https://doi.org/10.1007/s00572-017-0791-y.

Větrovský, T., Baldrian, P., Morais, D., 2018. Seed 2: a user-friendly platform for amplicon high-throughput sequencing data analyses. Bioinformatics 34, 2292–2294. https://doi.org/10.1093/bioinformatics/bty071.

Vondrák, J., Malíček, J., Palice, Z., Bouda, F., Berger, F., Sanderson, N., Acton, A., Pouska, V., Kish, R., 2018. Exploiting hot-spots; effective determination of lichen diversity in a Carpathian virgin forest. PLoS One 13, e0203540. https://doi.org/10.1371/journal.pone.0203540.

Vondrák, J., Svoboda, S., Košnar, J., Malíček, J., Šoun, J., Frolov, I., Svensson, M., Novotný, P., Palice, Z., 2023. Martin7: a reference database of DNA barcodes for European epiphytic lichens and its taxonomic implications. Preslia 95, 311–345. https://doi.org/10.23855/preslia.2023.311.

Vondrák, J., Svoboda, S., Malíček, J., Palice, Z., Kocourková, J., Knudsen, K., Mayrhofer, H., Thüs, H., Schultz, M., Košnar, J., Hofmeister, J., 2022. From Cinderella to Princess: an exceptional hotspot of lichen diversity in a long-inhabited central-European landscape. Preslia 94, 143–181. https://doi.org/10.23855/preslia.2022.143.

Vondrák, J., Svoboda, S., Malíček, J., Šoun, J., Košnar, J., Svensson, M., Timdal, E., Machač, J., Palice, Z., 2024. Combining environmental DNA data and taxonomic surveys provides an unprecedented understanding of lichen diversity and accelerates the discovery of new species. Preslia 96 (in press).

Wright, B., St Clair, L.L., Leavitt, S.D., 2019. Is targeted community DNA metabarcoding suitable for biodiversity inventories of lichen-forming fungi? Ecol. Indicat. 98, 812–820. https://doi.org/10.1016/j.ecolind.2018.11.061.