Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Forest mesophication alters soil fungal community structure and depletes genomic decay traits

Erica Houser
University of Minnesota
Peter Falb
University of Minnesota
Lang C. DeLancey
University of Wisconsin
Yue-Hua Hu
Chinese Academy of Sciences
Talia Michaud
University of Minnesota
Rebecca A. Montgomery
University of Minnesota
Andrew W. Ratz
University of Minnesota
Jonathan S. Schilling
University of Minnesota
Peter G. Kennedy
kennedyp@umn.edu

University of Minnesota

Article

Keywords:

Posted Date: August 14th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10490569/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

The widespread and sustained removal of fire from forests in eastern North America has led to ongoing changes in tree species composition, shifting dominance by pyrophytic species such as pine and oaks to more mesophytic species such as maples and cherries. While this shift, also known as mesophication, generally corresponds with a major shift in dominant tree mycorrhizal type, soil abiotic conditions, and nutrient availability, its effect of soil fungal community structure and associated functional traits is largely unknown. Using a space-for-time substitution approach, we sampled soils in a newly established ForestGEO plot in Minnesota, USA, to compare the soil fungal communities in three forest types (pine, oak, and maple-oak) representing different stages of mesophication. In each forest type, we quantified fungal community structure and soil organic matter (SOM) decay potential using high throughput sequencing, quantitative PCR, and metagenomics. We found that although total soil fungal community abundance did not vary across the three forest types, fungal richness and community composition shifted significantly across the mesophication gradient in a guild-specific manner. These structural shifts corresponded with changes in richness and composition of fungal metagenomic decay traits (carbohydrate-active enzyme genes, i.e. CAZymes). Specifically, less acidic soils and more mesophytic host tree abundances had lower SOM decay potential, both of which align with biotic and abiotic trends of mesophication. We conclude that mesophication is likely to significantly alter soil fungal community structure and functioning, with important consequences on soil carbon cycling.

Introduction

In the forests of eastern North America, the loss of cultural burning and the widespread 20th century management strategy of fire suppression have contributed to major shifts in tree species composition[1–3]. Sites historically dominated by fire-adapted, pyrophytic species, such as pines and oaks, have increasingly shifted towards a greater abundance of mesophytic species, such as maples, cherries, and basswoods that tolerate fire poorly[4–6]. This biotic transition is accompanied by multiple abiotic changes, such as increased understory shading, lower vapor pressure deficit, and higher soil moisture[4]. Further, the leaves of mesophytic species are typically less flammable than pyrophytic species[7], and their increasing presence reinforces a low fire state[4]. While changes in fire frequency have been the dominant driver of this compositional transition, broadly termed ‘mesophication’, ongoing changes in climate have also been influential[8]. Specifically, warmer conditions have allowed species that are less cold-tolerant and pyrophilic to establish and survive[9], thus hastening the rate at which mesophication is occurring. To date, a handful of studies have shown that mesophication-driven shifts in forest composition have significantly altered litter decomposition rates and soil nutrient dynamics[10–13], but the impacts of this change on soil microbial community structure and function is largely undocumented.

The shift in forest composition due to mesophication is generally associated with significant changes in dominant tree mycorrhizal type[14]. The mycorrhizal associated nutrient economy (MANE) framework posits that nutrient cycling in arbuscular mycorrhizal (AM) and ectomycorrhizal (EM) dominated forests

is fundamentally different due to variation in litter quality[15]. In AM systems, such as those dominated by maples and cherries, more labile litter decomposes quickly, and this rapid mineralization leads to macronutrients (e.g. nitrogen) to cycle in soil primarily in plant-available, inorganic forms (e.g. ammonium and nitrate). Conversely, in EM systems, such as those dominated by pines and oaks, less labile litter decomposes more slowly and thus nutrients cycle primarily in organic forms (e.g. amino acids and peptides). Belowground, AM forests tend to be dominated by fungi and bacteria that can effectively scavenge inorganic nutrients, while EM forests are co-dominated by EM and saprotrophic fungi, members of which both have the enzymatic capacities to break down organic matter directly[15]. Differences in litter chemistry between AM- and EM-dominated systems also affect other soil properties such as pH, with EM systems typically more acidic than AM systems[16]. The consequences of shifts in mycorrhizal types additionally impact the forms and stocks of carbon in soils[17]. For example, working in co-occurring AM- and EM-dominated eastern North American forests, Craig et al.[18] showed that EM stands accumulated more soil organic matter in surface soils, whereas AM stands stored more carbon in deeper mineral horizons. This pattern was attributed to the greater production and stabilization of microbial residues under the faster, more bacterially dominated decomposition characteristic of AM systems[18]. Since mesophication appears to generally favor AM trees[14,19], similar alterations to soil nutrient and carbon cycling and associated microbial community structure and function are likely, yet direct comparisons in forests where this process is occurring are currently lacking.

High throughput sequencing (HTS) has greatly expanded the capacity to characterize microbial community structure, but assessment of microbial community functioning with this method has multiple challenges. For fungi, one well recognized limitation of HTS is the uneven taxonomic coverage of the common ITS primers used to characterize soil fungal communities that typically underrepresent certain functional groups such as AM fungi[20]. To characterize these latter communities, it is often necessary to target a different barcoding region that better captures AM fungal richness and composition[21]. A second issue is the compositional nature of HTS data[22], which can provide relative but not absolute measures of both whole community and taxon-specific abundances[23]. Applying other methods such as quantitative PCR can better capture differences in absolute fungal abundances across habitats, although copies of fungal barcoding regions can vary widely across taxa[24], so directly linking these measures to ecological functioning remains challenging[25]. A third outstanding issue is the assignment of functional guilds, which is typically based on taxonomic matching to annotated databases at the genus or family level [26,27]. While these databases have expanded in coverage, there are still many fungi with poorly resolved identities and ecologies and thus their traits and functional potential remain unclassified during guild annotation.

A more direct assessment of soil microbial functional traits is the profiling of all genes via metagenomics[28]. As this approach does not rely on amplification of specific gene regions, it does not suffer from the same taxonomic biases that affect HTS-based datasets. Additionally, rather than inferring organismal functional trait potential based on taxonomy alone, metagenomics provides a direct

measure of the genes that are present, indicating its genomic functional potential (i.e., the genetic capacity of a community, independent of whether those genes are expressed). Despite these advantages, the application of metagenomics to characterize soil fungal community traits remains in a nascent stage. This is primarily because most of the genes present in metagenomic datasets belong to other groups of organisms, particularly prokaryotes[29,30], making fungal contributions harder to detect. Although the metagenomic bioinformatics pipelines for fungi and other eukaryotes have been consistently improving, there are many genes with limited annotation[31], which can complicate ecological inferences. For fungi, the carbohydrate-active enzymes (CAZymes), a suite of gene families involved in the degradation of organic matter[32], are arguably the most widely understood in terms of fungal metagenomic traits and their effects on ecosystem function. These gene families can be mapped to specific target substrates. For example, many of the Auxiliary Activities (AA) families are specifically adapted for the oxidative breakdown of lignin, a process with major consequences on ecosystem carbon sequestration[33]. While meta-transcriptomics and -proteomics are needed to measure which fungal genomic traits are utilized *in situ* (i.e. gene expression and activity)[34], documenting shifts in genomic functional potential that track changes in fungal community composition will help in identifying how fungal functional traits may be altered by shifting biotic and abiotic conditions [35,36].

To assess the potential for mesophication to alter the structure and function of soil fungal communities in the temperate forests of eastern North America, we sampled soils in a newly established ForestGEO plot in Minnesota, USA. Specifically, we used a space-for-time substitution approach to compare the soil fungal communities in three forest stand types that represent a transition from a more pyrophytic to a more mesophytic tree species composition. In each stand type, we quantified fungal community structure and functional potential using high throughput sequencing, quantitative PCR, and metagenomics to test three hypotheses: 1) fungal guild richness will change across stand types, 2) different fungal guilds are structured by different biotic or abiotic indicators of mesophication and 3) fungal community functional traits will shift as the chemistry of organic matter inputs change from pyrophytic to more mesophytic composition. Specifically, we expected plant pathogenic, saprotrophic, and AM fungi to be richer and EM fungi less rich with greater mesophication due to higher AM host tree abundance[37]. We also expected EM, AM, plant pathogenic fungal communities would more closely track biotic shifts (e.g. host composition and abundance), while saprotroph communities would be more strongly influenced by changes in soil abiotic conditions (e.g. pH, moisture, nutrient availability). Finally, we expected a depletion in genomic trait potential associated with lignin and complex organic matter degradation in locations with a higher mesophytic composition.

Methods

Study Site and Itasca ForestGEO Plot

The study was conducted at the boreal-temperate forest ecotone in Itasca State Park, located in north central Minnesota, USA at the headwaters of the Mississippi River (47.24°N, -95.21°W) on the ancestral

lands of the Ojibwe and Dakota peoples. The climate in the region is strongly continental, with cold winters and mild summers (mean annual temperature ~ 3.8°C) and a frost-free growing season typically beginning in May. Mean annual precipitation is ~ 640 mm, with roughly two-thirds falling as rain during the growing season. The park sits on the Itasca Moraine, which is characterized by a knob-and-kettle topography and includes a mosaic of soil types dominated by fine and coarse loams. Pollen records indicate a trajectory following glaciation from forests dominated by spruce to fire-maintained pine forests that persisted for roughly 2,500 years, followed then by more mesic deciduous forest expansion over the past ~ 4,000 years[38]. Recurrent fires that were an integral part of Ojibwe land management practices structured the park's forests, with fire-scar chronologies documenting 32 fires between 1650 and 1922[39]. The removal of native peoples and the fire suppression policies implemented in 1922 effectively removed this cultural and ecological process from the park except for limited amounts of prescribed burning that were reintroduced in the 1990s[38,40].

The study plot (https://forestgeo.si.edu/sites/north-america/itasca-state-park-isp) was established within Itasca State Park as part of the Forest Global Earth Observatory (ForestGEO) network[41]. The plot is centered at 47.22013°N, -95.16618°W at 1571 m elevation and encompasses a mix of fire-dependent pine and mesic hardwood forests characteristic of the park's upland habitats. The Natural Resource Conservation Service soil survey (https://websoilsurvey.nrcs.usda.gov/app/WebSoilSurvey.aspx) predicts the soil in the plot to be primarily the fine-loam Sol series. The 15 ha plot was established by a professional survey team in the summer of 2019 and measures 380 × 400 m. It is divided into 380 contiguous 20 × 20 m subplots (a 19 × 20 array) to allow for accurate geolocation of stems. Following the ForestGEO protocol[41], every free-standing woody stem ≥ 1 cm in diameter at breast height (DBH; 1.3 m) was tagged, measured, identified to species, and mapped to its spatial position within the plot. The initial census began in the summer of 2020 and was completed in the summer of 2024. Based on fire scar analyses in adjacent areas (D. Brumm, personal communication), the last major fire in the plot was likely in the 1880s, after which no burning has occurred.

In October 2024, soil samples were taken from locations representing three of the dominant stand types present in the plot: pine, oak, and maple-oak (n = 10 per stand type, N = 30). Samples were located throughout the plot to reduce spatial autocorrelation across stand types (Fig. 1a). Within each stand type, soil samples were collected 1 m north of the subplot boundary stake by first removing all leaf litter and then pounding a sterile 50 mL centrifuge tube into the mineral soil to a depth of 10 cm. All samples were refrigerated at 4°C within 8 hours of sampling and stored at -20°C ahead of molecular and chemical analyses. For those analyses, samples were thawed and sieved through a 2 mm metal screen that was surface sterilized between samples. Aliquots for DNA extraction were weighed (~ 250 mg) and then extracted. Aliquots for carbon and nitrogen analyses were dried, ground, and analyzed on an Elemental Analyzer (ESC 4010 Elemental Analyzer, Valencia, CA, USA). Soil pH was measured using a 1:1 water protocol[42]. Throughout the 2024 growing season, soil moisture in the top 10 cm was measured with a Field Scout TDR 150 probe (Spectrum Technologies, Aurora, IL, USA) on 5 dates at 190 of the 380 plot boundary stakes. In June 2026, woody stems < 1cm dbh and herbaceous plants were surveyed in 245 1

m^2 plots including the 190 soil moisture stakes and additional stakes representing a range of stand types.

Molecular and Bioinformatic Analyses

Total genomic DNA from each soil sample was extracted using the DNeasy PowerSoil Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. To quantify total fungal abundances, we used quantitative PCR (qPCR). For the qPCR reactions, serial dilutions (undiluted, 1:8, 1:64) of DNA from each sample were amplified with the ITS1F-ITS2 primer pair[43,44] in a 6 µL reaction volume using KAPA HiFi polymerase on a Qiagen QS5 thermocycler. Amplification conditions consisted of 5 min at 95°C, followed by 25 cycles of 20 s at 98°C, 15 s at 55°C, and then 1:00 min at 72°C. Quantifications were based on C_t values compared to fungal standard curves and expressed as the log10 number of ITS copies per microliter.

High throughput sequencing (HTS) was used to characterize soil fungal community alpha and beta diversity according to the methods detailed in Gohl et al.[45]. The 5.8S and ITS4 primers, which broadly amplify fungi, were used to target the ITS2 region[46], while the WANDA and AML2 primers, which are specifically designed to amplify AM fungi, were used to target the 18S region[47,48]. In addition to all soil samples, a negative extraction control (no substrate added) and positive controls consisting of an 11-member synthetic ITS mock community[49], a 2-fungal member ZymoBIOMICS Microbial Community Standard mock community, and a 5-member AM fungal mock community (derived from Egan et al.[50]) were included. ITS and 18S amplicons were generated using 25 and 35 cycles, respectively. Clean, equimolar amplicons from each sample were sequenced using Illumina NextSeq v3 chemistry (2 x 300 paired-end bp) at the University of Minnesota Genomics Center. Raw sequences for each marker gene were deposited in the NCBI Short Read Archive (SRA), under the following BioProject numbers: PRJNAXXXX (ITS) and PRJNAXXXX (18S) (released upon article acceptance).

HTS sequences were processed using a custom in-house pipeline (https://github.com/peterhfalb/HTS_DADA2processing). First, adapter sequence artifacts and primers were trimmed from forward and reverse reads using cutadapt[51]. Next, reads were run through the DADA2 pipeline[52] in which they were filtered and trimmed (filterAndTrim; maxN = 0, maxEE = c(2,2)), dereplicated (derepFastq), and denoised (learnErrors, dada). Forward and reverse reads were then merged, with minimum overlap of ten base pairs (mergePairs; minOverlap = 10), and chimeras removed (removeBimeraDenovo; method = “consensus”). After denoising, amplicon sequence variants are then clustered at 97% similarity to generate operational taxonomic units (OTUs) using VSEARCH[53] (--cluster_size). A second chimera check using VSEARCH –uchime-denovo removed any chimeras missed with DADA2. To avoid oversplitting, OTUs were further clustered using the MUMU/LULU algorithm[54,55]. Finally, the RDP algorithm from DADA2[52] was used to assign taxonomy to the OTU sequences using the UNITE sh database[56] for ITS sequences, and the MaarjAM database for 18S AM sequences[57].

Both the raw ITS2 and 18S OTU datasets were quality filtered to account for common HTS artifacts[58]. First, 3x the sequence reads per OTU from negative controls were subtracted from all samples, such that known contaminant reads were completely removed from both negative and positive controls. To account for index bleed, all cells with 11 or fewer read counts were converted to zeros, which removed 1254 or 32.2% of OTUs. This cutoff was chosen because one member of the synmock mock community was observed in the Zymock mock community with 11 reads, which could only occur due to tag switching[59]. Similarly, the 18S mock community samples contained 5 AM fungal OTUs with > 700 reads, while the remaining had 4 or fewer, likely due to index bleed. Thus, following Lindahl et al.[60], all cells with 4 or fewer read counts were converted to zeros, resulting in the removal of 328 or 16.1% of OTUs. The ITS2 OTUs were filtered to retain only those assigned to Kingdom fungi, which removed 706 or 26.8% of OTUs. Because the MaarjAM database only contains AM fungi, to avoid forced matching of non-AM fungi OTUs, 18S OTUs were filtered to retain a genus bootstrap of > = 60%, removing 1556 or 91.3% of OTUs. Species accumulation curves were generated with vegan 'rarecurve' for both datasets and indicated that sufficient sampling depth was achieved for all samples, eliminating the need for rarefaction (Supplementary Fig. 1). Finally, ITS2 OTUs with > = 80% genus bootstrap confidence were assigned primary lifestyle with the FungalTraits database[27] (Supplementary Fig. 2).

Shotgun metagenomics were used to assess the functional potential of the soil fungal communities. NexteraXT libraries were created for each sample, which were then pooled and sequenced on an Aviti Freestyle High 2x150bp run. That run generated ~ 1 billion reads pass quality filters, which was lower than the expected output, so the pool was resequenced on an additional Aviti Freestyle Medium 2x150bp run. Data from the two runs were merged for a total of ~ 1.8 billion reads. Raw sequences for each sample were deposited in the NCBI SRA, under the following BioProject number: PRJNAXXXX (released upon article acceptance).

Metagenome *fastq* files were denoised and filtered according to the NMDC Metagenome Workflow[61], adapted for use on the University of Minnesota Agate supercomputer using a linux/SLURM-based submission system (https://github.com/peterhfalb/MetagenomeQC). In brief, BBDuk was used to trim Illumina adapters (k-mer, k = 23/11), remove PhiX174 spike-in reads, poly-G tails (≥ 5 nt), and paired-end read overlap artifacts[62]. Next, BBDuk was used to perform quality and length filtering: 3' ends were quality trimmed (Phred < 0), and reads were filtered with a mean quality less than three, greater than 3 ambiguous bases, or less than 51 base pairs, removing 0.216% (± 0.01% SD) of reads across all 30 samples (average of 207,405 reads removed per sample). After filtering, samples had on average 94.4 (± 17.6 SD) million reads per sample, with all reads having mean Phred quality score above 10. To normalize functional gene abundances to genome equivalents, we adapted the approach of Pellitier & Zak[35]. Near-single-copy ortholog groups were identified from OrthoDB v12[63] at the taxonomic level of Fungi (NCBI taxid 4751). Ortholog groups were retained if they were present in ≥ 90% of fungal species in the database and occurred as a single copy in ≥ 90% of those species, yielding 1,204 near-single-copy ortholog groups. Raw quality-filtered reads from each sample were mapped against this database using DIAMOND *blastx* (*--sensitive, e-value ≤ 1 × 10⁻⁵*) and a maximum of one target sequence per read[64].

Forward and reverse reads were mapped independently and results were combined. To account for within-group sequence redundancy, read counts were averaged across all member sequences within each ortholog group. The geometric mean of these per-group averages was then computed across all ortholog groups and used as the per-sample genome-equivalent estimate, resulting in an average of 1.67 (± 0.115 SD) genome-equivalents per sample.

To annotate our metagenomes for fungal SOM decay traits, we queried our metagenome reads against the CAZyme database[65]. First, we tried to assemble the metagenomes using MetaSPADES[66], and perform eukaryote gene prediction using MetaEuk[67]. However, because ~ 1% of the total DNA belonged to fungi, only a very small percentage of reads assembled into contigs identified as CAZymes, and the data was very sparse, resulting in poor assembly data quality. Thus, raw quality-filtered reads were instead mapped directly against the CAZy database (*CAZyDB.07242025.fa*) using DIAMOND *blastx* (v2.0.15, --sensitive, --max-target-seqs 1, e-value ≤ 1×10^{-5}), following the competitive all-kingdom approach of Bahram et al.[68]. Forward (R1) and reverse (R2) reads were mapped independently and merged. Hits were filtered to ≥ 50% amino acid identity and e-value ≤ 1×10^{-9}. Paired reads originating from the same insert were deduplicated by retaining the highest-bitscore hit per read-pair × CAZy family combination, resulting in an average 1.9 million reads (± 371,702 SD) per sample detected as CAZymes. Approximately 86.6% of CAZy sequences were matched to an NCBI taxid; the remainder were JGI MycoCosm sequences whose taxids are embedded directly in the sequence identifier. Taxids were expanded to full lineages using *taxonomizr::getTaxonomy()* against a locally built NCBI taxonomy database[69]. Hits were assigned to Fungi based on membership in fungal phyla, with 1.8% (± 0.137% SD) of reads mapping to fungi (average 34,968 ± 7659 SD fungal CAZyme reads per sample). Given the short-read lengths of 150bp, we did not try to assign taxonomy at the level of guild (ectomycorrhizal vs. saprotroph), as the polyphyletic nature of EM fungi makes it difficult to assign taxonomy below the phylum level using a *blast*-based analysis [70,71]. CAZyme gene family counts were normalized by dividing by per-sample genome-equivalent estimates from OrthoDB, yielding ratio values interpretable as the abundance of each gene family relative to a single fungal genome equivalent in the community (pseudo-genome count). We created a curated reference table mapping CAZyme gene families to the substrates (cellulose, hemicellulose, lignin, chitin, etc.) they act upon, and what from which type of biomass those substrates originate (plant, fungal, bacterial). CAZyme to gene family mappings were only created when CAZymes were known to have high substrate-specificity with those compounds. Gene family to substrate mappings were made based on previous fungal genomics studies[33,72–75], and/or the CAZypedia (https://www.cazypedia.org)[76]. Biomass types were selected based on the main components in and between cell walls in each of those organismal groups. Information on white-rot associated enzymes were taken from Floudas et al.[77]. See Supplementary Table 1 for more details on specific families to substrate mappings with paper source. Family-level pseudo-genome counts were summed within each origin and substrate category. Substrate categories and biomass origin types were modelled as response variables in the uni- and multi-variate analyses (see *Statistical Analyses*).

Statistical analysis

Plant community structure

All statistical analyses were run in R version 4.5.1[78]. To assess the degree of difference in woody plant community composition by stand type, analyses of similarity (ANOSIM) with function 'pairwise.anosim' from package OmicFlow[79] and beta dispersion tests using the 'betadisper' function in vegan[80] were performed on a Bray-Curtis dissimilarity matrix of plant communities across the 30 sample locations. For those analyses, woody plant stems were filtered to include only living stems identified to at least the genus level. The pairwise Bray-Curtis dissimilarity values were calculated on the relative basal area of each woody species within a 10 m radius of each stake to ensure non-overlapping communities among adjacent stakes. Compositional differences were visualized with non-metric multidimensional scaling (NMDS) on Bray-Curtis dissimilarity matrix with function 'metaMDS' from vegan (k = 2, trymax = 100). To determine whether abiotic soil variables and summary metrics for the plant community differed by stand type, one-way analyses of variance (ANOVAs) by stand type were performed with function 'aov' in base R. Means and 95% confidence intervals were extracted from these models, and pairwise significant differences were determined with function 'glht' from multcomp using Tukey HSD comparisons.

Fungal community structure

Guild-level fungal richness was calculated as total OTUs based on ITS2 sequences assigned as 'ectomycorrhizal', 'plant pathogen', 'saprotroph' (any subtype of saprotroph was combined), and 18S read counts for AM fungi, respectively. Single linear models of guild-level fungal richness were run with each soil and plant predictor, including pH, C:N (carbon to nitrogen) ratio, nitrogen (%), volumetric water content (%), ectomycorrhizal basal area (cm^2/m^2), arbuscular mycorrhizal basal area (cm^2/m^2), ectomycorrhizal stem density ($stems/m^2$), arbuscular mycorrhizal stem density ($stems/m^2$), herb richness, and total herb cover (%) (Supplementary Table 2). Volumetric water content was calculated as the average across 2024 samplings for a subset of the stakes that were sampled for molecular analyses (N = 17, pine = 5, oak = 6, maple-oak = 6). Basal area metrics were calculated from census data, and stem density metrics were calculated as the sum of census and 2026 woody seedling surveys stem density values. To assess overall differences in community composition, PERMANOVA was performed with the function 'pairwise.adonis' from package 'pairwiseAdonis' with Bonferroni correction on Bray-Curtis dissimilarity matrices of fungal guilds. Bray-Curtis dissimilarity matrices were calculated with 'vegdist' that were previously Hellinger-transformed with 'decostand' from vegan. Compositional differences were visualized with non-metric multidimensional scaling (NMDS) on the Bray-Curtis dissimilarity matrices with function 'metaMDS' from vegan (trymax = 100). NMDS were fit to k = 2 dimensions, but if stress > 0.2, they were refit to 3 dimensions. To assess differences in the dispersion of guild-specific fungal communities across stand types, beta dispersion tests by stand type were run using the 'betadisper' function in vegan[80]. To assess whether compositional and dispersional differences were correlated with environmental variables, environmental variables were fitted to the NMDS ordination using 'envfit' from vegan. Environmental variables included the same soil and plant predictors as above (Supplementary

Table 2). Finally, to determine which fungal OTUs are characteristic of each stand type, indicator species analysis was performed on Hellinger-transformed community data with function 'multipatt' (func = 'IndVal', nperm = 9999) from indicspecies with Benjamini-Hochberg correction for multiple hypothesis testing.

Fungal metagenomic traits

For each CAZyme category (substrate types, biomass origins, and total counts), three nested models were fit. The model family was selected by diagnostic tests applied to the residuals of a preliminary linear regression: if the proportion of zeros exceeded 15% or a Shapiro-Wilk test on model residuals rejected normality, a Tweedie compound-Poisson-Gamma GLM (log-link; glmmTMB) was fit[81]. Otherwise, we used a Gaussian linear model. Fungal ITS copy number, ITS OTU richness, and saprotrophic relative abundance were each divided by their standard deviation prior to model fitting, so coefficients represent responses per one SD change in the predictor. Models compared were: M1: CAZyme response variable ~ pH + C:N ratio, M2: M1 + ITS copies + ITS richness + saprotroph relative abundance, and M3: M2 + AM basal area ($cm^2\ m^{-2}$). We selected the 'best-supported' model using the 2-ΔAIC rule[82]: the simplest model (M1 preferred over M2, over M3) whose AIC fell within 2 units of the minimum AIC across all three. Model fit was summarized by marginal R^2 (the proportion of variance explained by the fixed effects). We calculated standardized regression coefficients (β ± 95% CI) from M3 for all predictors (β × SD(x) / SD(y)) and used Spearman rank correlations to estimate individual relationships between CAZyme families and ecological variables. We calculated correlations between every annotated CAZyme family with ≥ 20 total pseudo-genome count, and nine ecological predictors: soil pH, C:N ratio, volumetric water content, fungal ITS copy number, ITS OTU richness, and two mesophication indicators (AM basal area [$cm^2\ m^{-2}$] and stem count). We also calculated Spearman correlations between total CAZyme pseudo-genome count as well as richness.

Differences in category-level (substrate and origin) CAZyme pseudo-genome counts between stand types were tested using generalized linear models. The model family was selected by diagnostic tests described above. Pairwise contrasts between stand types were estimated via the *emmeans* package[83] with Tukey-HSD adjustment[84]. CAZyme 'community composition' was characterized using non-metric multidimensional scaling of Bray-curtis dissimilarities calculated on the pseudo-genome count values for each CAZyme family (metaMDS[80], vegan). NMDS analyses were performed on annotated CAZyme families with known substrates. To assess the potential drivers of compositional variance, we ran PERMANOVAs (adonis2, vegan, 999 permutations), with the following predictors: stand type, soil pH, C:N ratio, volumetric water content (June 2026), AM stem density, total basal area, stand type, fungal ITS copy number, ITS OTU richness, and saprotrophic relative abundance. To test for pairwise differences between stand type CAZyme composition, we ran pairwise PERMANOVA comparisons among all stand type pairs, correcting for multiple comparisons using the Bonferroni method. AM basal area was fitted with the function ordisurf and displayed as contour lines. Group standard error bars are shown for each stand type using the function ordihull. For the 30 most abundant CAZyme families (via mean overall

pseudo-genome counts), we calculated the average and standard error for each CAZyme family across samples within each stand type (Supplemental Fig. 3).

Results

Plant community structure

Across the entire Itasca ForestGEO plot, a total of 27,452 living stems > = 1cm DBH were present (all multi-stemmed individuals were counted as 1 stem), including 28 woody plant species, 16 genera, and 11 families, as well as 2188 standing dead stems (Fig. 1a). Mean stand density was 1716 stems/ha and mean basal area was 28.5 m^2/ha. Dominant canopy trees included *Acer saccharum* (48% stems, 15% basal area), *Acer rubrum* (11% stems, 17% basal area), *Ostrya virginiana* (14% stems, 4% basal area), *Fraxinus pennsylvanica* (4% stems, 3% basal area), *Fraxinus nigra* (3% stems, 3% basal area), *Quercus rubra* (4% stems, 27% basal area), *Quercus macrocarpa* (1% stems, 6% basal area), *Pinus strobus* (0.2% stems, 8% basal area), and *Pinus resinosa* (0.3% stems, 7% basal area). The size and stem distributions vary considerably among genera, with pines dominating larger size classes, oaks dominating intermediate size classes, and maples and ashes dominating smaller size classes (Fig. 1c). The survey of woody stems < 1cm dbh and herbaceous plants found 15 tree species, 15 shrub/ground layer species, 3 vining species, and 57 herbaceous species.

Across the 30 locations sampled for soil fungal community structure and metagenomic decay traits, there were 21 woody plant species that varied in abundance across the three stand types. In the pine, oak, and maple-oak stand types, pine represented 72%, 6%, and 0%, oak represented 10%, 59%, and 41%, and maple represented 15%, 23%, and 47% of the basal area, respectively (Fig. 1d). The stand types differed significantly in woody plant composition (ANOSIM, R = 0.35, p = 0.001, Fig. 1b), indicating they were sufficiently differentiated across the intended mesophication gradient. Beta dispersion tests showed significantly higher dispersion of the woody plant communities in the pine compared to the oak stand type (F = 6.98, p = 0.001). The three stand types also differed in the abiotic and biotic variables used as predictor variables in the community and genomic trait analyses. The maple-oak stand type had significantly higher volumetric water content ($F(2, 27) = 4.82$, $p = 0.017$), lower EM basal area ($F(2, 27) = 5.87$, $p = 0.008$), and higher AM stem density ($F(2, 27) = 3.48$, $p = 0.045$) than pine stand type, while the pine stand type had significantly lower AM basal area ($F(2, 27) = 5.52$, $p = 0.010$) than both the oak and maple-oak stand type (Fig. 1e), indicating that the intended mesophication gradient was captured.

Fungal community structure

Fungal community abundances did not vary significantly across stand types (Fig. 2a). In total, there were 548 saprotrophic, 63 plant pathogenic, 207 EM, and 145 AM fungal OTUs. The richness of fungal guilds did vary significantly across stand types; oak had the highest EM fungal OTU richness ($F_{(2,27)} = 14.1$, $p = 0.0006$), maple-oak had the highest AM fungal OTU richness ($F_{(2,17)} = 3.66$, $p = 0.039$), and plant pathogenic fungal OTU richness increased from pine to maple-oak to oak stand types ($F_{(2,27)} = 17.4$, p =

0.0001, Fig. 2a). Comparing guilds across different abiotic and biotic predictor variables, plant pathogenic fungal OTU richness increased with herbaceous plant richness (estimate = 1.13, R^2 = 0.2, p = 0.015) and decreased with EM stem density (estimate = -0.88, R^2 = 0.19, p = 0.015), and AM fungal OTU richness increased with total herb cover (estimate = 0.17, R^2 = 0.21, p = 0.012) and AM basal area (estimate = 0.44, R^2 = 0.13, p = 0.049).

There were significant compositional differences among stand types for all four fungal guilds, although the contrasts by stand type varied: pine saprotrophic fungal communities (F = 2.22, R^2 = 0.14, p = 0.001, Fig. 2b) and EM fungal communities (F = 1.72, R^2 = 0.11, p = 0.001, Fig. 2d) were distinct from oak and maple-oak, all three stand types for plant pathogenic fungal communities were distinct (F = 3.67, R^2 = 0.21, p = 0.001, Fig. 2c), and only pine and oak stand types for AM fungal communities were distinct (F = 1.85, R^2 = 0.12, p = 0.002, Fig. 2e). Soil volumetric water content was the most consistent environmental correlate (saprotrophic: R^2 = 0.37, p = 0.010; plant pathogenic: R^2 = 0.51, p = 0.010, AM: R^2 = 0.56, p = 0.010), followed by EM basal area (saprotrophic: R^2 = 0.32, p = 0.020; plant pathogenic: R^2 = 0.25, p = 0.00; EM: R^2 = 0.46, p = 0.010), AM basal area (saprotrophic: R^2 = 0.26, p = 0.040; EM: R^2 = 0.21, p = 0.040), and AM stem density (pathogens: R^2 = 0.25, p = 0.030; AM: R^2 = 0.24, p = 0.040). By contrast, soil pH was only correlated with plant pathogenic fungal composition (R^2 = 0.22, p = 0.010). Beta dispersion did not differ significantly by stand type for any guild. Indicator species analyses associated nearly all significant OTUs with the oak stand type: *Mucor ramosissimus* (Indval = 1, p = 0.0001) and *Mucor moelleri* (IndVal = 0.87, p = 0.0001) for saprotrophic fungi, *Ganoderma applanatum* (Indval = 1, p = 0.0001), *Cytospora sp.* (Indval = 0.92, p = 0.0001), *Ramularia sp.* (Indval = 0.82, p = 0.0003), and *Fusarium oxysporum* (Indval = 0.81, p = 0.0004) for pathogenic fungi, *Cortinarius sp.* (Indval = 0.88, p = 0.0002) and *Scleroderma areolatum* (Indval = 0.85, p = 0.0002) for EM fungi, and *Glomus* VTX00448 (Indval = 0.84, p = 0.0002) for AM fungi. The lone fungal indicator for the pine stand type was *Cryptosporopsis sp*. (IndVal = 0.92, p = 0.0001), with the maple-oak site having no significant indicator OTUs from any guild.

Fungal metagenomic traits

The most abundant fungal CAZymes in all stand types were enzymes putatively involved in cell wall synthesis[76,85] (GT22, GT2, GT34, and GT4; Supplementary Fig. 3). Samples from the maple-oak stand type generally had the lowest per-enzyme family counts, especially for GT22, GT34 and GT38, all which are likely involved in cell-wall synthesis and likely represent a general measure of fungal productivity[76]. After aggregating CAZyme families into substrate and biomass targets, there were no differences in total pseudo-genome counts of CAZyme among stand types (Gaussian LMs; Supplemental Fig. 4a-g; Supplementary Table 3). The compositional analysis of CAZyme families did, however, reveal significant effects of AM stem count on the CAZyme pool (R^2 = 0.122, F = 4.173, p = 0.029), with marginally significant effects of C:N ratio (R^2 = 0.0726, F = 2.478, p = 0.096). Variance explained and p-values are reported in Supplementary Table 3.

Multivariate analyses found that total CAZyme counts were best predicted by a model that included soil abiotic, fungal, and plant variables (R^2 = 0.408, AIC = 571.1; Fig. 3a). The abundance of CAZymes targeting plant and fungal biomass were best predicted by models that included AM basal area as a predictor (R^2 = 0.420, AIC = 506.1; R^2 = 0.405, AIC = 382.2, respectively; Fig. 3a). AM basal area was also statistically informative for predicting enzyme abundances involved cellulose, hemicellulose, chitin, and lignin decomposition. In contrast, the abundance of enzymes targeting peptidoglycan (a primary component of bacterial cell walls) was not well predicted by AM basal area, or the richness and composition of the fungal community (R^2 = 0.097, AIC = 328.0; R^2 = 0.201, AIC = 128.7, respectively). Across all models, AM basal area had strong negative relationships with CAZyme pseudo-genome counts, with the strongest effects on plant cell wall degradation enzymes (-0.46 ± 0.38), particularly for lignocellulolytic enzymes (-0.48 ± 0.39). Soil pH was always found to be negatively related to CAZyme counts, with the strongest negative impacts on plant cell wall degradation enzymes (cellulose: -0.50 ± 0.41, and hemicellulose: -0.64 ± 0.40), although fungal biomass degradation enzymes were not associated with pH. Further, ITS copies (fungal abundance) were always negatively related to CAZyme counts independent of changes in saprotroph relative abundance, especially for chitin (-0.41 ± 0.37), cellulose (-0.42 ± 0.39), and hemicellulose decomposition (-0.40 ± 0.38). In contrast, C:N ratios had consistently positive relationships with CAZyme pseudo-genome counts, with the strongest effect on plant biomass enzymes (0.42 ± 0.38), but all estimates had high uncertainty.

Univariate Spearman rank correlations found that soil pH was generally strongly negatively correlated with the total richness of CAZyme genes, but not CAZyme counts of total and individual families (Fig. 3b). Enzyme families whose abundance was most closely linked to pH were positively correlated, and included enzymes for cellulose, hemicellulose, lignin, and chitin degradation (GH6, GH51, AA5, and GH81, respectively). Average volumetric water content had strong negative correlations with many CAZyme counts, and total counts were also negatively correlated. ITS copies and saprotroph relative abundance had minimal but generally positive correlations with CAZyme pseudo-genome counts, but very few relationships were significant. ITS richness was largely negatively correlated with the abundance of genes involved in plant biomass degradation, with minimal and no relationship with fungal and bacterial biomass degradation enzymes, respectively. AM basal area and AM stem count both significantly negatively correlated with total CAZyme abundance, and the abundance of most individual gene families. AM stem count had consistently stronger relationships with CAZyme pseudo-genome count, particularly for those enzymes involved in lignocellulolytic decomposition (Fig. 3b).

Discussion

Both the mesophication and MANE frameworks provide conceptual links between the tree species composition and functioning of forests[4,15]. While each is well recognized in its field (forest and soil ecology, respectively), to our knowledge this is the first study to characterize the structure and functional potential of the soil fungal communities that represent key mechanistic parts of both frameworks. Using a combination of molecular-based analyses, we show that changes in forest dominance from pyrophytic

species such as pines and oaks to mesophytic species such as maples influence both the composition and metagenomic decay traits of the fungal communities. The Itasca ForestGEO plot represented an ideal study site as a forest in active transition; from one historically dominated by pines when fire was common on the landscape[38] to one where mesophytic species, particularly maples, now account for more than half of all stems in the plot. That shift is largely consistent with other forests in eastern North America, where similar trends have been well documented[3,4]. Consistent with this plot-level trajectory, woody plant composition differed significantly across the three stand types sampled, with the maple-oak stand type having the highest AM basal area and stem density. Importantly, these compositional shifts were also accompanied by abiotic conditions predicted under both frameworks (e.g. high soil moisture and soil N concentrations in the maple stand type), indicating that these stand types captured not only a shift in tree mycorrhizal dominance but also associated changes in soil conditions that are directly relevant to the structure and functioning of fungal communities. More broadly, because mesophication is closely tied to land management practices (particularly fire suppression)[2,86], documenting how both the structure and functioning of soil fungal communities shifts along mesophication gradients could help inform restoration efforts, such as the use of prescribed fire. Similarly, because the dominant mycorrhizal type can strongly impact forest carbon cycling, characterizing fungal community structure and function in response to mesophication also helps predict carbon storage potential in soils.

Our first hypothesis predicted that changes in forest composition would change the richness of both types of mycorrhizal fungi as well as that of plant pathogenic fungi. We found mixed support for these predictions; AM fungi was highest in the maple-oak stand type, where AM basal area was significantly higher, which was broadly consistent with expectations of the MANE framework as well as the obligate nature of the AM symbiosis[15]. The lack of differences in AM fungal OTU richness between the maple-oak and pine stand types, however, indicates that AM fungi are not necessarily impoverished in pine forests, but rather more likely associated with abundance of understory herbaceous plants[87]. For EM fungi, OTU richness was highest in the oak stand type. Previous studies have shown that relative to pines, oak EM fungal communities are often more species rich[88,89]. The decline in EM fungal richness in the maple-oak stand type likely reflects the decrease in EM basal area, which is typically a strong correlate of EM richness[90]. Plant pathogenic fungi also had the highest richness in the oak stand type, which aligns with other studies showing that oaks often have diverse pathogenic communities[91]. Notably, saprotrophic fungi were the only guild for which richness did not differ among stand types. Since many saprotrophic fungi are known to have specific substrate preferences[92], it appears that shifts in saprotrophic fungal composition across the forest types resulted largely in substitutions in the stand type-level richness of this guild.

Along with changes in guild richness by stand type, we also hypothesized that different guilds would be structured by different biotic and abiotic variables. Specifically, we predicted that AM, EM, and pathogenic fungal communities would change in composition to track host compositional shifts across the mesophication gradient. Conversely, we expected that saprotrophic fungal community composition would be more strongly structured by changing soil abiotic conditions. As with our first hypothesis, we

found mixed support for this second hypothesis. Saprotrophic fungal composition was most strongly associated with abiotic conditions, principally volumetric water content. The pine stand type, which was the driest, differed significantly in saprotrophic fungal community composition from those of the oak and maple stand types. However, both the AM and plant pathogenic fungal communities were also most strongly structured by soil moisture, weakening support for a distinction between host-related biotic versus soil-related abiotic factors as differential drivers of fungal guild composition. We believe, however, that the recurrence of soil moisture as the top predictor across three of the four guilds is concordant with rather than contradictory to the mesophication framework. Specifically, higher soil moisture is itself a recognized feature of mesophytic forests[4] and we suspect the soil moisture signal on fungal guild composition likely represents a downstream consequence of the vegetation shift rather than an independent abiotic filter (although our space-for-time design cannot formally distinguish between the two - see below). Nonetheless, the presence of several indicator species by stand type suggests that host abundance does have some influence on guild-specific composition. For example, in the oak stand type, *Scleroderma areolatum* is a known EM fungal associate of oaks[93]. Conversely, in the pine stand type, one indicator OTU, *Cryptosporiopsis* sp, may in fact be a pathogen of the understory shrub *Corylus cornuta*[94], which was highly abundant in the pine stand type.

In support of our third hypothesis, fungal genomic traits linked to degradation of both plant and microbial substrates, including lignin, were negatively associated with mesophication and the abundance of AM trees in particular (Fig. 3ab). We also consistently found negative relationships between pH and the abundance and richness of fungal gene families related to plant biomass degradation, similar to previous investigations of fungal decay traits that have also found negative relationships between pH and fungal CAZyme richness[95,96] (Fig. 3a,b). However, at global scales the negative effect of pH on fungal CAZymes is small relative to those of AM tree abundance and $\delta^{15}N$[97,98] (a proxy for nitrification and N availability), both of which are strongly negatively related to CAZyme richness[95]. This suggests that a primary factor regulating fungal decay traits is the dominance of AM trees and the availability of inorganic N[4,15,95], which typically decreases fungal:bacterial ratios, promotes bacterial-dominated decomposition, and reduces fungal mining of organic matter for nutrient acquisition[95,99,100]. In addition to changes in inter-guild abundance, the decreases in fungal CAZyme decay traits in mesophytic stands could also be due to trait changes within fungal guilds, specifically within EM fungi, which have been found to be depleted in CAZymes for soil organic matter degradation (ie. lignin, chitin, cellulose) in sites rich in inorganic N[35]. While we did not directly measure the relative enrichment in CAZyme gene families within EM fungi (see below), it is possible that the EM fungal taxa in the maple-oak locations may be more adapted for inorganic N acquisition, and less so for complex soil organic matter degradation (i.e. absorbers rather than miners[101]). Alternatively, these shifts could indicate a switch towards EM and AM fungi priming bacterial decomposition of organic matter for organic N acquisition that wouldn't be detected in fungal CAZymes[102–106]. Finally, we also observed lower abundances of CAZymes under high total fungal abundance (ITS copies), especially for plant biomass degrading enzymes (Fig. 3a). This trend was not associated with increases in the relative abundance of either of the saprotrophic or EM

fungal guilds, suggesting that in general, when fungi are abundant, the species in those communities are relatively depleted in soil organic matter degradation potential (Fig. 3a). This is consistent with the well-documented negative relationship between nutrient availability and decomposition rates, and we speculate that in resource-rich environments (e.g. high N availability) with high fungal productivity, fungi may invest minimally in the costly genetic machinery for organic N-mining of plant tissues[107–111].

While the results provide important insights regarding the impact of mesophication of soil fungal community structure and functioning, we acknowledge a range of important limitations. First, our space-for-time substitution assumes that the three stand types differ principally in their stage of mesophication, and that the pine stand type we sampled approximates the past condition of the oak and maple-oak stand types rather than a persistently distinct habitat. We took steps to limit this concern by restricting the sampling for this study to upland positions located at comparable elevations and on the same soil series. Nonetheless, initial abiotic differences among spatial substitutes are a recognized source of error in chronosequence inference[112]. As such, repeated sampling of these same locations over time, paired with experimental manipulations such as reintroduced fire or reciprocal litter transfers, would help separate the effects of mesophication itself from the site conditions that predisposed each stand type to its current location. A second issue is our designation of stand types, which while broadly representative of different mesophication stages, clearly contain varying mixtures of the major tree genera across the plot. To address this, we analyzed the effects of changes in woody plant composition using both categorical (e.g. stand type) and continuous (e.g. percentage EM tree basal area) approaches, which generated similar but not identical results. We also recognize that a key stand type, i.e. one fully dominated by maple representing a late stage of mesophication, was not included, so projecting the observed trends onto future mesophication stages should be done with caution. Finally, we acknowledge that compared to other fields using ‘omics technologies, fungal metagenomics lags significantly behind, with no unified consensus on methodologies, and a limited set of shared tools for analysis[28–30]. Although metagenomics allows unbiased, untargeted sampling across all genomic DNA in a sample, the resulting DNA is highly fragmented and only represents the most abundant taxa (genomically), which is about 90% prokaryotic[29,30]. Additionally, given the short length of our metagenomics sequences (150bp), which provide limited phylogenetic information, we were unable to make any direct conclusions about the relative contributions of EM versus saprotrophic fungi towards CAZyme abundances[70].

In conclusion, our results indicate that the biotic and abiotic changes accompanying mesophication restructure both the composition and the functional potential of soil fungal communities. While a reduced capacity for complex organic matter degradation as AM trees replace EM pines and oaks might intuitively suggest greater carbon retention in soils, we believe the implications for soil carbon storage are more nuanced. Specifically, the ongoing mesophication of eastern North American forests may not simply increase or decrease soil carbon storage[18], but rather reorganize its vertical distribution and chemical form, with important consequences for the persistence of that carbon in soils under a changing climate[113]. Resolving the net outcome will require pairing the community and genomic

approaches deployed here with repeated depth-resolved measurements of soil carbon stocks as forest composition changes from pyrophytic to mesophytic dominance.

Declarations

Acknowledgements

We gratefully acknowledge the many members of the Itasca ForestGEO initial census team (in alphabetical order by last name): Collin Adams, Justice Alvarado, Lex Bancos, Briana Beatty, Catherine Glenn-Stone, Hakeem Hamshari, Josie Johnston, Soren Matheny, Achala Narayanan, Eduardo Pérez-Pazos, Benjamin Tanner, and Kathleen Thompson; Louis Peterson at the Minnesota Department of Natural Resources for assistance in locating the plot and facilitating ongoing research; the Itasca Biological Station and Laboratories (IBSL) staff for logistical and technical assistance; an IBSL Seed-To-Root Grant for financial support in establishing the plot; John Fredell, Becca Dallinger, John Short, and the University of Minnesota College of Biological Sciences for financial support and assistance with the IBSL American Indian Intern Fund Internship.

References

1. Knott, J. A., Desprez, J. M., Oswalt, C. M. & Fei, S. Shifts in forest composition in the eastern United States. *For. Ecol. Manag.* 433, 176–183 (2019).
2. Abrams, M. D. & Nowacki, G. J. Examining the heritage and legacy of Indigenous land management in oak and pine forests of northeastern United States. *Int. J. Ecol. Environ. Sci.* 47, 27 (2021).
3. Brose, P. H., Dey, D. C. & Waldrop, T. A. *The Fire-Oak Literature of Eastern North America: Synthesis and Guidelines*. 1–98 https://research.fs.usda.gov/treesearch/46252 (2014) doi:10.2737/NRS-GTR-135.
4. Nowacki, G. J. & Abrams, M. D. The Demise of Fire and “Mesophication” of Forests in the Eastern United States. *BioScience* 58, 123–138 (2008).
5. Ziegler, S. S. The Past and Future of White Pine Forests in the Great Lakes Region. *Geogr. Compass* 4, 1179–1202 (2010).
6. Hanberry, B. B., Palik, B. J. & He, H. S. Comparison of historical and current forest surveys for detection of homogenization and mesophication of Minnesota forests. *Landsc. Ecol.* 27, 1495–1512 (2012).
7. Kane, J. M., Kreye, J. K., Barajas-Ramirez, R. & Varner, J. M. Litter trait driven dampening of flammability following deciduous forest community shifts in eastern North America. *For. Ecol. Manag.* 489, 119100 (2021).
8. Nowacki, G. J. & Abrams, M. D. Is climate an important driver of post-European vegetation change in the Eastern United States? *Glob. Change Biol.* 21, 314–334 (2015).

9. Boisvert-Marsh, L., Périé, C. & De Blois, S. Shifting with climate? Evidence for recent changes in tree species distribution at high latitudes. *Ecosphere* 5, 1–33 (2014).
10. Babl-Plauche, E. K., Alexander, H. D., Siegert, C. M., Willis, J. L. & Berry, A. I. Mesophication of upland oak forests: Implications of species-specific differences in leaf litter decomposition rates and fuelbed composition. *For. Ecol. Manag.* 512, 120141 (2022).
11. Alexander, H. D. & Arthur, M. A. Increasing Red Maple Leaf Litter Alters Decomposition Rates and Nitrogen Cycling in Historically Oak-Dominated Forests of the Eastern U.S. *Ecosystems* 17, 1371–1383 (2014).
12. Kreye, J. K., Varner, J. M., Hamby, G. W. & Kane, J. M. Mesophytic litter dampens flammability in fire-excluded pyrophytic oak–hickory woodlands. *Ecosphere* 9, e02078 (2018).
13. Fernández-Alonso, M. J., Curiel Yuste, J., Kitzler, B., Ortiz, C. & Rubio, A. Changes in litter chemistry associated with global change-driven forest succession resulted in time-decoupled responses of soil carbon and nitrogen cycles. *Soil Biol. Biochem.* 120, 200–211 (2018).
14. Jo, I., Fei, S., Oswalt, C. M., Domke, G. M. & Phillips, R. P. Shifts in dominant tree mycorrhizal associations in response to anthropogenic impacts. *Sci. Adv.* 5, eaav6358 (2019).
15. Phillips, R. P., Brzostek, E. & Midgley, M. G. The mycorrhizal-associated nutrient economy: a new framework for predicting carbon–nutrient couplings in temperate forests. *New Phytol.* 199, 41–51 (2013).
16. Hobbie, S. E. Plant species effects on nutrient cycling: revisiting litter feedbacks. *Trends Ecol. Evol.* 30, 357–363 (2015).
17. Yan, G. *et al.* Forest carbon stocks increase with higher dominance of ectomycorrhizal trees in high latitude forests. *Nat. Commun.* 15, 5959 (2024).
18. Craig, M. E. *et al.* Tree mycorrhizal type predicts within-site variability in the storage and distribution of soil organic matter. *Glob. Change Biol.* 24, 3317–3330 (2018).
19. Palus, J. D., Goebel, P. C., Hix, D. M. & Matthews, S. N. Structural and compositional shifts in forests undergoing mesophication in the Wayne National Forest, southeastern Ohio. *For. Ecol. Manag.* 430, 413–420 (2018).
20. Berruti, A., Desirò, A., Visentin, S., Zecca, O. & Bonfante, P. ITS fungal barcoding primers versus 18S AMF-specific primers reveal similar AMF-based diversity patterns in roots and soils of three mountain vineyards. *Environ. Microbiol. Rep.* 9, 658–667 (2017).
21. Öpik, M., Davison, J., Moora, M. & Zobel, M. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. *Botany* 92, 135–147 (2014).
22. Gloor, G. B., Macklaim, J. M., Pawlowsky-Glahn, V. & Egozcue, J. J. Microbiome Datasets Are Compositional: And This Is Not Optional. *Front. Microbiol.* 8, (2017).
23. Amend, A. S., Seifert, K. A. & Bruns, T. D. Quantifying microbial communities with 454 pyrosequencing: does read abundance count? *Mol. Ecol.* 19, 5555–5565 (2010).

24. Lofgren, L. A. *et al.* Genome-based estimates of fungal rDNA copy number variation across phylogenetic scales and ecological lifestyles. *Mol. Ecol.* 28, 721–730 (2019).
25. Baldrian, P. *et al.* Estimation of fungal biomass in forest litter and soil. *Fungal Ecol.* 6, 1–11 (2013).
26. Nguyen, N. H. *et al.* FUNGuild: An open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol.* 20, 241–248 (2016).
27. Põlme, S. *et al.* FungalTraits: a user-friendly traits database of fungi and fungus-like stramenopiles. *Fungal Divers.* 105, 1–16 (2020).
28. Shelton, B. R. *et al.* Fungal ecology in the age of 'omics. *New Phytol.* 250, 822–833 (2026).
29. Ren, Y. *et al.* Benchmarking metabarcoding and metagenomics pipelines in taxonomic profiling for soil microbial eukaryotes. *Soil Biol. Biochem.* 218, 110144 (2026).
30. Belliardo, C. *et al.* Improvement of eukaryotic protein predictions from soil metagenomes. *Sci. Data* 9, 311 (2022).
31. Mondo, S. J. & Grigoriev, I. V. A genomic perspective on fungal diversity and evolution. *Nat. Rev. Microbiol.* 23, 718–733 (2025).
32. Hage, H. & Rosso, M.-N. Evolution of Fungal Carbohydrate-Active Enzyme Portfolios and Adaptation to Plant Cell-Wall Polymers. *J. Fungi* 7, 185 (2021).
33. Floudas, D. *et al.* The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* 336, 1715–1719 (2012).
34. Zak, D. R. *et al.* Exploring the role of ectomycorrhizal fungi in soil carbon dynamics. *New Phytol.* 223, 33–39 (2019).
35. Pellitier, P. T. & Zak, D. R. Ectomycorrhizal fungal decay traits along a soil nitrogen gradient. *New Phytol.* 232, 2152–2164 (2021).
36. Argiroff, W. A., Zak, D. R., Upchurch, R. A., Pellitier, P. T. & Belke, J. P. Fungal community composition and genetic potential regulate fine root decay in northern temperate forests. *Mol. Ecol.* 32, 2005–2021 (2023).
37. Eagar, A. C. *et al.* Arbuscular Mycorrhizal Tree Communities Have Greater Soil Fungal Diversity and Relative Abundances of Saprotrophs and Pathogens than Ectomycorrhizal Tree Communities. *Appl. Environ. Microbiol.* 88, e01782-21 (2022).
38. Hansen, H. L., Ness, D. D. & Kurmis, V. *The Ecology of Upland Forest Communities and Implications for Management in Itasca State Park, Minnesota*. (Agricultural Experiment Station, University of Minnesota, St. Paul, 1974).
39. Frissell Jr, S. S. The Importance of Fire as a Natural Ecological Factor in Itasca State Park, Minnesota. *Quat. Res.* 3, 397–407 (1973).
40. Webb, S. L. & Scanga, S. E. Windstorm Disturbance Without Patch Dynamics: Twelve Years of Change in a Minnesota Forest. *Ecology* 82, 893–897 (2001).
41. Davies, S. J. *et al.* ForestGEO: Understanding forest diversity and dynamics through a global observatory network. *Biol. Conserv.* 253, 108907 (2021).

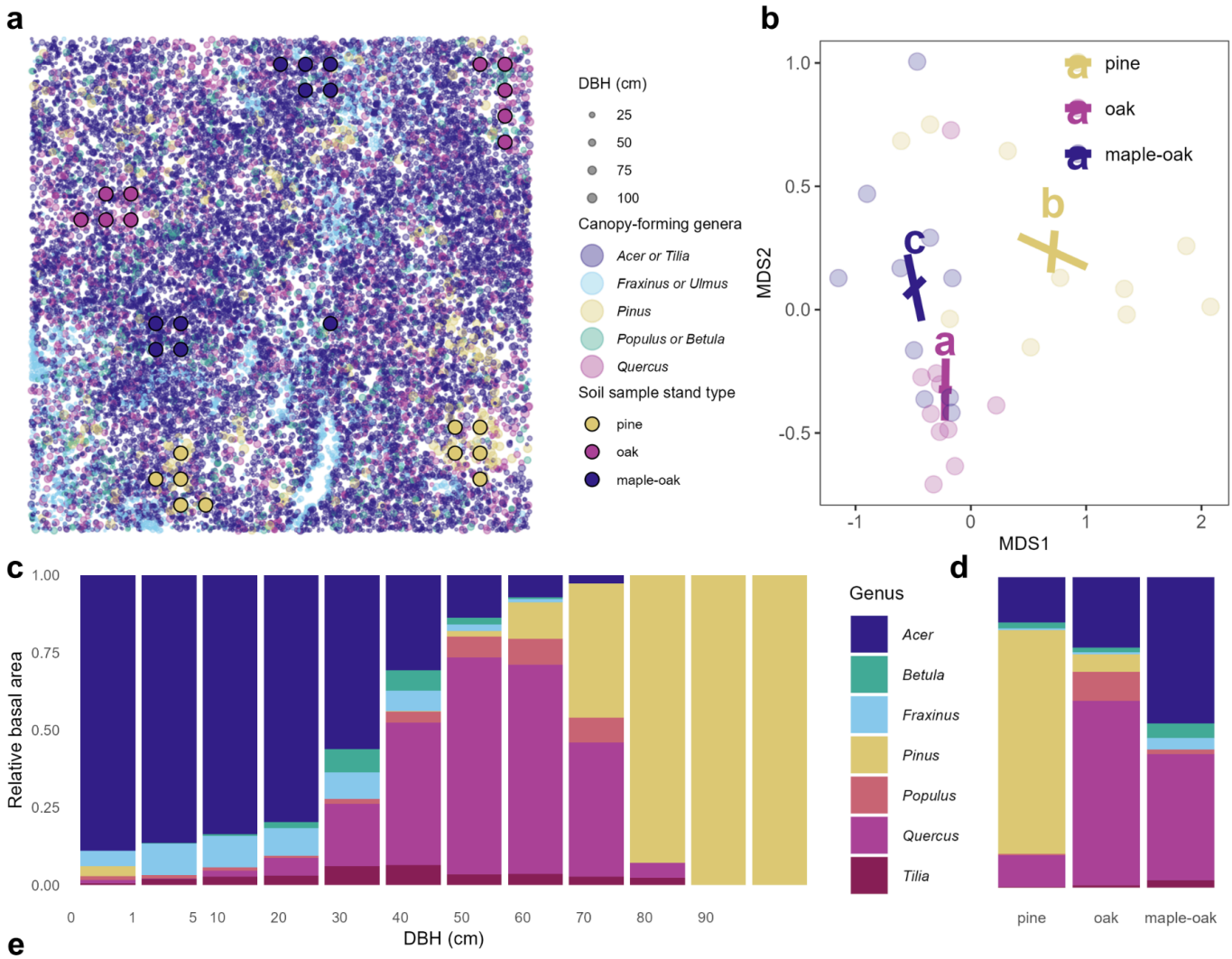
42. SOIL SURVEY LABORATORY *METHODS MANUAL Version 4.0*. (2004).
43. White, T. J., Bruns, T., Lee, S. & Taylor, J. 38 - AMPLIFICATION AND DIRECT SEQUENCING OF FUNGAL RIBOSOMAL RNA GENES FOR PHYLOGENETICS. in *PCR Protocols* (eds Innis, M. A., Gelfand, D. H., Sninsky, J. J. & White, T. J.) 315–322 (Academic Press, San Diego, 1990). doi:10.1016/B978-0-12-372180-8.50042-1.
44. Gardes, M. & Bruns, T. D. ITS primers with enhanced specificity for basidiomycetes - application to the identification of mycorrhizae and rusts. *Mol. Ecol.* 2, 113–118 (1993).
45. Gohl, D. M. *et al.* Systematic improvement of amplicon marker gene methods for increased accuracy in microbiome studies. *Nat. Biotechnol.* 34, 942–949 (2016).
46. Taylor, D. L. *et al.* Accurate Estimation of Fungal Diversity and Abundance through Improved Lineage-Specific Primers Optimized for Illumina Amplicon Sequencing. *Appl. Environ. Microbiol.* 82, 7217–7226 (2016).
47. Lee, J., Lee, S. & Young, J. P. W. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiol. Ecol.* 65, 339–349 (2008).
48. Dumbrell, A. J. *et al.* Distinct seasonal assemblages of arbuscular mycorrhizal fungi revealed by massively parallel pyrosequencing. *New Phytol.* 190, 794–804 (2011).
49. Palmer, J. M., Jusino, M. A., Banik, M. T. & Lindner, D. L. Non-biological synthetic spike-in controls and the AMPtk software pipeline improve mycobiome data. *PeerJ* 6, e4925 (2018).
50. Egan, C. P. *et al.* Using mock communities of arbuscular mycorrhizal fungi to evaluate fidelity associated with Illumina sequencing. *Fungal Ecol.* 33, 52–64 (2018).
51. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* 17, 10 (2011).
52. Callahan, B. J. *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583 (2016).
53. Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* 4, e2584 (2016).
54. Mahé, F. mumu: post-clustering curation tool for metabarcoding data. (2026).
55. Frøslev, T. G. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat. Commun.* 8, 1188 (2017).
56. Abarenkov, K. *et al.* 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 (2024).
57. Öpik, M. *et al.* The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytol.* 188, 223–241 (2010).
58. Tedersoo, L. *et al.* Best practices in metabarcoding of fungi: From experimental design to results. *Mol. Ecol.* 31, 2769–2795 (2022).

59. Carlsen, T. *et al.* Don't make a mista(g)ke: is tag switching an overlooked source of error in amplicon pyrosequencing studies? *Fungal Ecol.* 5, 747–749 (2012).
60. Lindahl, B. D. *et al.* Fungal community analysis by high-throughput sequencing of amplified markers – a user's guide. *New Phytol.* 199, 288–299 (2013).
61. Clum, A. *et al.* DOE JGI Metagenome Workflow. *mSystems* 6, 10.1128/msystems.00804 – 20 (2021).
62. Bushnell, B. BBTools/BBMap: A Fast, Accurate, Splice-Aware Aligner. (2014).
63. Tegenfeldt, F. *et al.* OrthoDB and BUSCO update: annotation of orthologs with wider sampling of genomes. *Nucleic Acids Res.* 53, D516–D522 (2025).
64. Buchfink, B., Reuter, K. & Drost, H.-G. Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* 18, 366–368 (2021).
65. Drula, E. *et al.* The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577 (2022).
66. Nurk, S., Meleshko, D., Korobeynikov, A. & Pevzner, P. A. metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* 27, 824–834 (2017).
67. Levy Karin, E., Mirdita, M. & Söding, J. MetaEuk—sensitive, high-throughput gene discovery, and annotation for large-scale eukaryotic metagenomics. *Microbiome* 8, 48 (2020).
68. Bahram, M. *et al.* Structure and function of the global topsoil microbiome. *Nature* 560, 233–237 (2018).
69. Sherrill-Mix, S. taxonomizr: Functions to Work with NCBI Accessions and Taxonomy. https://doi.org/10.32614/CRAN.package.taxonomizr (2025).
70. Carr, R. & Borenstein, E. Comparative Analysis of Functional Metagenomic Annotation and the Mappability of Short Reads. *PLoS ONE* 9, e105776 (2014).
71. Strullu-Derrien, C., Selosse, M.-A., Kenrick, P. & Martin, F. M. The origin and evolution of mycorrhizal symbioses: from palaeomycology to phylogenomics. *New Phytol.* 220, 1012–1030 (2018).
72. Cantarel, B. L., Lombard, V. & Henrissat, B. Complex Carbohydrate Utilization by the Healthy Human Microbiome. *PLOS ONE* 7, e28742 (2012).
73. López-Mondéjar, R. *et al.* Metagenomics and stable isotope probing reveal the complementary contribution of fungal and bacterial communities in the recycling of dead biomass in forest soil. *Soil Biol. Biochem.* 148, 107875 (2020).
74. Pellitier, P. T. & Zak, D. R. Ectomycorrhizal fungal decay traits along a soil nitrogen gradient. *New Phytol.* 232, 2152–2164 (2021).
75. Zhao, Z., Liu, H., Wang, C. & Xu, J.-R. Correction: Comparative analysis of fungal genomes reveals different plant cell wall degrading capacity in fungi. *BMC Genomics* 15, 6 (2014).
76. Drula, E. *et al.* The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577 (2022).
77. Floudas, D. *et al.* The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* 336, 1715–1719 (2012).

78. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing (2025).
79. Gusinac, A. agusinac/OmicFlow. (2026).
80. Oksanen, J. *et al.* vegan: Community Ecology Package. (2025).
81. McGillycuddy, M., Popovic, G., Bolker, B. M. & Warton, D. I. Parsimoniously Fitting Large Multivariate Random Effects in glmmTMB. *J. Stat. Softw.* 112, 1–19 (2025).
82. Sutherland, C. *et al.* Practical advice on variable selection and reporting using Akaike information criterion. *Proc. R. Soc. B Biol. Sci.* 290, 20231261.
83. Lenth, R. V. *et al.* emmeans: Estimated Marginal Means, aka Least-Squares Means. (2026).
84. Searle, S. R., Speed, F. M. & Milliken, G. A. Population Marginal Means in the Linear Model: An Alternative to Least Squares Means. *Am. Stat.* 34, 216–221 (1980).
85. Alshareef, S. A. Metabolic analysis of the CAZy class glycosyltransferases in rhizospheric soil fungiome of the plant species *Moringa oleifera*. *Saudi J. Biol. Sci.* 31, 103956 (2024).
86. Abrams, M. D., Nowacki, G. J. & Hanberry, B. B. Oak forests and woodlands as Indigenous landscapes in the Eastern United States. *J. Torrey Bot. Soc.* 149, 101–121 (2021).
87. Korb, J. E., Johnson, N. C. & Covington, W. W. Arbuscular mycorrhizal propagule densities respond rapidly to ponderosa pine restoration treatments. *J. Appl. Ecol.* 40, 101–110 (2003).
88. Adamo, I., Dashevskaya, S. & Alday, J. G. Fungal Perspective of Pine and Oak Colonization in Mediterranean Degraded Ecosystems. *Forests* 13, 88 (2022).
89. Pérez-Rosas, B., Gómez-Hernández, M. & Gándara, E. Variación en diversidad y composición de especies macrofúngicas a través de diferentes tipos de vegetación en Oaxaca, México. *Bot. Sci.* 100, 827–836 (2022).
90. Cortese, A. M. *et al.* Ectomycorrhizal fungal diversity may be influenced by arbuscular mycorrhizal trees in mixed temperate forests. *Mycorrhiza* 36, 24 (2026).
91. Moricca, S. *et al.* Endemic and Emerging Pathogens Threatening Cork Oak Trees: Management Options for Conserving a Unique Forest Ecosystem. *Plant Dis.* 100, 2184–2193 (2016).
92. Ye, L. *et al.* Substrate Preference Determines Macrofungal Biogeography in the Greater Mekong Sub-Region. *Forests* 10, 824 (2019).
93. Ortega, A., Lorite, J. & Valle, F. Mycorrhizal macrofungi diversity (Agaricomycetes) from Mediterranean Quercus forests; a compilation for the Iberian Peninsula (Spain and Portugal). *Nova Hedwig*. 1–31 (2010) doi:10.1127/0029-5035/2010/0091-0001.
94. Patejuk, K. *et al.* First report of Cryptosporiopsis tarraconensis causing leaf lesion of Corylus avellana in Central Europe (Poland). *Ann. For. Sci.* 81, 1 (2024).
95. Bahram, M. *et al.* Plant nutrient-acquisition strategies drive topsoil microbiome structure and function. *New Phytol.* 227, 1189–1199 (2020).
96. Bahram, M. *et al.* Structure and function of the global topsoil microbiome. *Nature* 560, 233–237 (2018).

97. Garten, C. T. Variation in Foliar 15N Abundance and the Availability of Soil Nitrogen on Walker Branch Watershed. *Ecology* 74, 2098–2113 (1993).
98. Garten Jr., C. T. & Miegroet, H. V. Relationships between soil nitrogen dynamics and natural 15N abundance in plant foliage from Great Smoky Mountains National Park. *Can. J. For. Res.* 24, 1636–1645 (1994).
99. Cheeke, T. E. *et al.* Dominant mycorrhizal association of trees alters carbon and nutrient cycling by selecting for microbial groups with distinct enzyme function. *New Phytol.* 214, 432–442 (2017).
100. Van Der Heijden, M. G. A., Bardgett, R. D. & Van Straalen, N. M. The unseen majority: soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems. *Ecol. Lett.* 11, 296–310 (2008).
101. Jörgensen, K. *et al.* A trait spectrum linking nitrogen acquisition and carbon use of ectomycorrhizal fungi. *New Phytol.* 246, 2425–2434 (2025).
102. Allison, S. D., Weintraub, M. N., Gartner, T. B. & Waldrop, M. P. Evolutionary-Economic Principles as Regulators of Soil Enzyme Production and Ecosystem Function. in *Soil Enzymology* (eds Shukla, G. & Varma, A.) 229–243 (Springer, Berlin, Heidelberg, 2011). doi:10.1007/978-3-642-14225-3_12.
103. Dijkstra, F. A., Carrillo, Y., Pendall, E. & Morgan, J. A. Rhizosphere priming: a nutrient perspective. *Front. Microbiol.* 4, (2013).
104. Brzostek, E. R., Dragoni, D., Brown, Z. A. & Phillips, R. P. Mycorrhizal type determines the magnitude and direction of root-induced changes in decomposition in a temperate forest. *New Phytol.* 206, 1274–1282 (2015).
105. Gorka, S. *et al.* Rapid Transfer of Plant Photosynthates to Soil Bacteria via Ectomycorrhizal Hyphae and Its Interaction With Nitrogen Availability. *Front. Microbiol.* 10, (2019).
106. Fernandez, C. W. & See, C. R. The pH influence on ectomycorrhizal nitrogen acquisition and decomposition. *New Phytol.* 246, 867–875 (2025).
107. Berg, B. & Matzner, E. Effect of N deposition on decomposition of plant litter and soil organic matter in forest systems. 5, (1997).
108. Craine, J. M., Morrow, C. & Fierer, N. Microbial Nitrogen Limitation Increases Decomposition. *Ecology* 88, 2105–2113 (2007).
109. Hagedorn, F., Spinnler, D. & Siegwolf, R. Increased N deposition retards mineralization of old soil organic matter. *Soil Biol. Biochem.* 35, 1683–1692 (2003).
110. Magill, A. H. & Aber, J. D. Long-term effects of experimental nitrogen additions on foliar litter decay and humus formation in forest ecosystems. *Plant Soil* 203, 301–311 (1998).
111. Neff, J. C. *et al.* Variable effects of nitrogen additions on the stability and turnover of soil carbon. *Nature* 419, 915–917 (2002).
112. Kreyling, J. Space-for-time substitution misleads projections of plant community and stand-structure development after disturbance in a slow-growing environment. *J. Ecol.* 113, 68–80 (2025).
113. Hicks Pries, C. E. *et al.* The Deep Soil Organic Carbon Response to Global Change. *Annu. Rev. Ecol. Evol. Syst.* 54, 375–401 (2023).

Figures



e

Compartment	Variable	Pine	Oak	Maple-oak
Soil	pH	5.6±0.3^{a}	5.7±0.3^{a}	5.7±0.3^{a}
	C:N ratio	19.2±1.6^{a}	20.2±1.5^{a}	19.8±1.5^{a}
	Nitrogen (%)	0.3±0.1^{a}	0.2±0^{a}	0.3±0^{a}
	Volumetric water content (%)	**23±2.6^{a}**	**26.3±2.8^{ab}**	**28.6±2.6^{b}**
Plant	Ectomycorrhizal basal area (cm^2/m^2)	**38.9±11.1^{a}**	**26.4±11^{ab}**	**13.1±11.1^{b}**
	Ectomycorrhizal stem density ($stems/m^2$)	2.6±1.9^{a}	0.4±1.8^{a}	1.5±1.8^{a}
	Arbuscular mycorrhizal basal area (cm^2/m^2)	**6.9±2.8^{a}**	**7.4±2.8^{a}**	**12.7±2.8^{b}**
	Arbuscular mycorrhizal stem density ($stems/m^2$)	**6.3±5.9^{a}**	**12.2±5.8^{ab}**	**16.8±5.8^{b}**
	Herb richness	7±1.5^{a}	8.8±1.5^{a}	6.9±1.5^{a}

Figure 1

(a) Stem map with soil sampling locations of the Itasca ForestGEO plot. Some genera were grouped by ecological niche. **(b)** Non-metric dimensional scaling of Bray-Curtis dissimilarity on woody species by basal area. Letters indicate results from ANOSIM by stand type, with centroids showing standard error.

Relative basal area of canopy-forming woody genera by **(c)** diameter class across the whole plot and **(d)** stand type for locations sampled for fungi (n = 30). Spruce (*Picea*), elm (*Ulmus*), and cherry (*Prunus*) were excluded as they represented less than 1% of stems. **(e)** Table of summary values for abiotic soil variables and plant community for each stand type. Superscripts indicate significant differences between stand types by one-way ANOVA; ± indicate 95% confidence intervals.

a

Guild	Variable	Pine	Oak	Maple-oak
	Fungal abundance (ITS2 copies)	13251.8±5955.2^{a}	8999.7±5955.2^{a}	18168.7±5955.2^{a}
Saprotroph	OTU richness	100.6±12.3^{a}	113.5±12.3^{a}	105.4±12.3^{a}
	Relative abundance (% ITS reads)	35.2±5.5^{a}	30±5.6^{a}	38.9±5.6^{a}
Plant pathogen	OTU richness	**12.6±2.6^{a}**	**23±2.6^{b}**	**17±2.6^{c}**
	Relative abundance (% ITS reads)	2.9±1.5^{a}	4.3±1.6^{a}	2.7±1.5^{a}
Ectomycorrhizal	OTU richness	**19.1±2.6^{a}**	**27.3±2.6^{b}**	**19±2.6^{a}**
	Relative abundance (% ITS reads)	19.8±7.5^{a}	24±7.5^{a}	16.4±7.5^{a}
Arbuscular mycorrhizal	OTU richness	**31.3±3.6^{ab}**	**28.5±3.6^{a}**	**35.1±3.6^{b}**

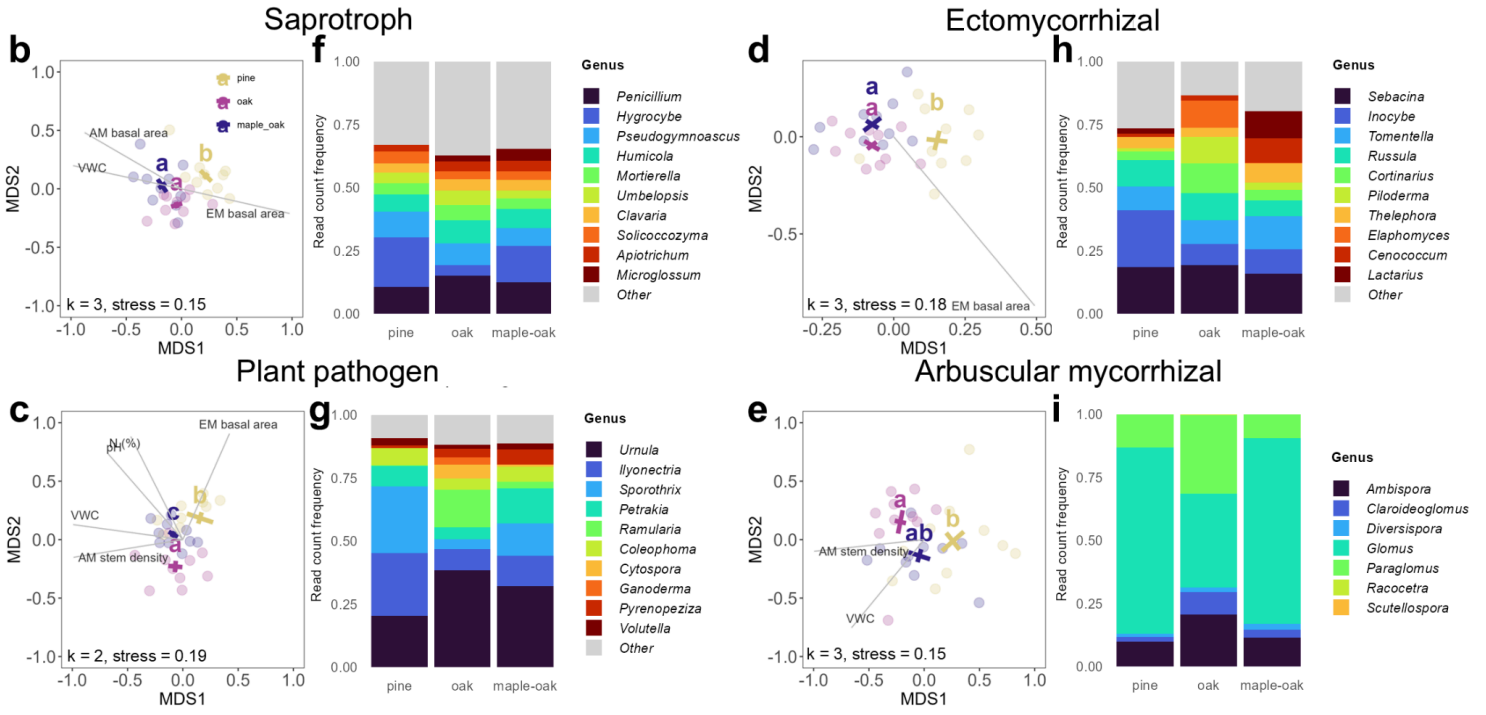


Figure 2

(a) Fungal abundance and fungal guild OTU richness and relative abundance. Superscripts indicate significant differences between stand types by one-way ANOVA; ± indicate 95% confidence intervals. Non-metric dimensional scaling of Bray-Curtis dissimilarity on **(b)** saprotroph, **(c)** plant pathogen, and **(d)** ectomycorrhizal fungal communities by ITS2 reads, and **(e)**arbuscular mycorrhizal fungal community by 18S reads. Letters indicate results from PERMANOVA by stand type, with centroids showing standard error. Vectors represent significant biotic and abiotic predictors of fungal community composition from

envfit ($p < 0.05$). Stacked bar plots of the relative read count abundances of the 10 most abundant **(f)** saprotroph, **(g)**plant pathogen, **(h)** ectomycorrhizal, and **(i)** arbuscular mycorrhizal genera by stand type. The arbuscular mycorrhizal community contained only 8 genera.

a

			Diagnostics		Model	R^2			AIC			Standardized β ± 95% CI (M3)					
Response	n	Family	Zeros	S-W p	Best	M1: *abiotic*	M2: *abiotic + fungal*	M3: *M2 + AM BA*	M1: *abiotic*	M2: *abiotic + fungal*	M3: *M2 + AM BA*	pH	C:N	ITS copies	ITS richness	Saprotroph *rel. abund.*	AM basal area *(cm² m⁻²)*
Total																	
Total CAZy	30	Gaussian	0%	0.3983	M3	0.096	0.244	0.408	575.8	576.4	571.1	-0.42±0.41	+0.39±0.38	**-0.43±0.39**	-0.14±0.37	+0.16±0.36	**-0.50±0.39**
Origin																	
Plant Biomass	30	Gaussian	0%	0.7220	M3	0.137	0.278	0.420	510.0	510.6	506.1	**-0.56±0.41**	**+0.42±0.38**	**-0.41±0.38**	-0.15±0.37	+0.14±0.36	**-0.46±0.38**
Fungal Biomass	30	Gaussian	0%	0.2995	M3	0.120	0.289	0.405	385.9	385.5	382.2	-0.15±0.41	+0.38±0.39	-0.38±0.39	-0.25±0.37	+0.19±0.37	**-0.42±0.39**
Bacterial Biomass	30	Gaussian	3%	0.2949	M1	0.201	0.269	0.326	128.7	132.0	131.6	-0.45±0.44	+0.26±0.41	*+0.06±0.41*	*-0.12±0.40*	*+0.17±0.39*	*-0.29±0.41*
Substrate																	
Cellulose	30	Gaussian	0%	0.7217	M3	0.113	0.255	0.400	413.2	413.9	409.5	**-0.50±0.41**	+0.40±0.39	**-0.42±0.39**	-0.15±0.37	+0.17±0.37	**-0.47±0.39**
Hemicellulose	30	Gaussian	0%	0.5867	M3	0.165	0.303	0.438	440.6	441.2	436.8	**-0.64±0.40**	+0.38±0.37	**-0.40±0.38**	-0.17±0.36	+0.18±0.36	**-0.45±0.38**
Lignin	30	Gaussian	0%	0.2417	M3	0.128	0.231	0.383	453.4	455.6	451.1	-0.39±0.42	+0.41±0.39	-0.35±0.39	-0.11±0.38	+0.15±0.37	**-0.48±0.39**
Chitin	30	Gaussian	0%	0.2822	M3	0.130	0.355	0.455	359.8	356.9	353.8	-0.09±0.39	+0.38±0.37	**-0.41±0.37**	-0.33±0.36	+0.19±0.35	-0.39±0.37
Glucans	30	Gaussian	0%	0.2130	M1	0.097	0.174	0.303	328.0	331.3	328.3	-0.26±0.44	+0.33±0.42	*-0.30±0.42*	*-0.10±0.40*	*+0.18±0.40*	*-0.44±0.42*
Peptidoglycan	30	Gaussian	3%	0.2949	M1	0.201	0.269	0.326	128.7	132.0	131.6	-0.45±0.44	+0.26±0.41	*+0.06±0.41*	*-0.12±0.40*	*+0.17±0.39*	*-0.29±0.41*

Note: All betas from M3 (full model; Wald tests). Grey italic β = predictor not in best AIC model; bold β = p < 0.05 (M3 Wald). Green AIC = best model (ΔAIC < 2 rule). Gaussian β: sd(x)/sd(y) R^2: r^2 or cor(fitted,obs)².

b

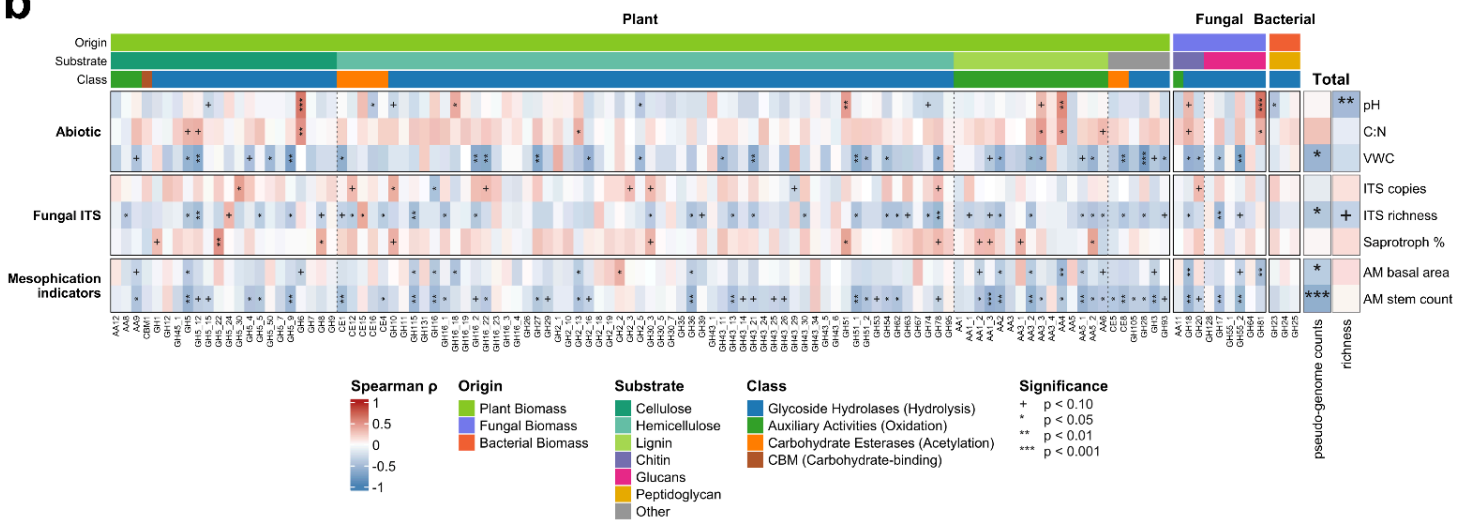


Figure 3

(a) Model outputs predicting CAZyme abundance of total CAZymes, and total abundance of gene families targeting known substrate and origin groups. All models were run with a Gaussian linear model, after diagnostics indicated <15% zeros, and gaussian residuals (Shapiro-Wilks normality test; p>0.05). Three models were compared: an *abiotic*, *abiotic + fungal*, and *abiotic + fungal + stand type indicator* model. Abiotic only models (M1) included pH and carbon to nitrogen ratios as predictors. M2 (*abiotic + fungal*) added ITS copies, richness and saprotroph relative abundance as predictors. M3 had all predictors, including arbuscular mycorrhizal tree basal area, an indicator of mesophication. Models were compared based on the AIC value, and the best-supported model was selected as the simplest model

with AIC within 2 units of the minimum across M1–M3. **(b)** Heatmap showing the strength of Spearman rank correlations between pseudo-genome count for each CAZyme family and nine environmental and stand-structure predictors: soil pH, carbon to nitrogen ratio (C:N), volumetric water content (VWC) from 2024, ITS copies, ITS OTU richness, saprotroph relative abundance, and arbuscular mycorrhizal tree basal area and stem count. Only families with ≥ 20 total hits and known substrate annotations were included; family–predictor pairs with fewer than four complete observations were excluded. Correlations between total fungal CAZyme counts and richness are also shown on the right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigures.docx
- SupplementalTable1CAZymeSubstrateMap.xlsx
- SupplementalTable2FungalDiversityEnvironmentalfits.xlsx
- SupplementalTable3CAZymesExtraStats.xlsx