The following Supporting Information is available for this article:

Fig. S1 Map of the twelve forest sites in Wexford and Manistee Counties, Manistee National Forest, Michigan, USA. All trees lay between 70-130m elevation. Pins are colored by rates of net N mineralization ($\mu g\ g^{-1}\ d^{-1}$). Insets: Michigan, USA with sites in red points; continental United States and location of Michigan, blue box. Satellite imagery, May 2018) is derived from Google Earth.
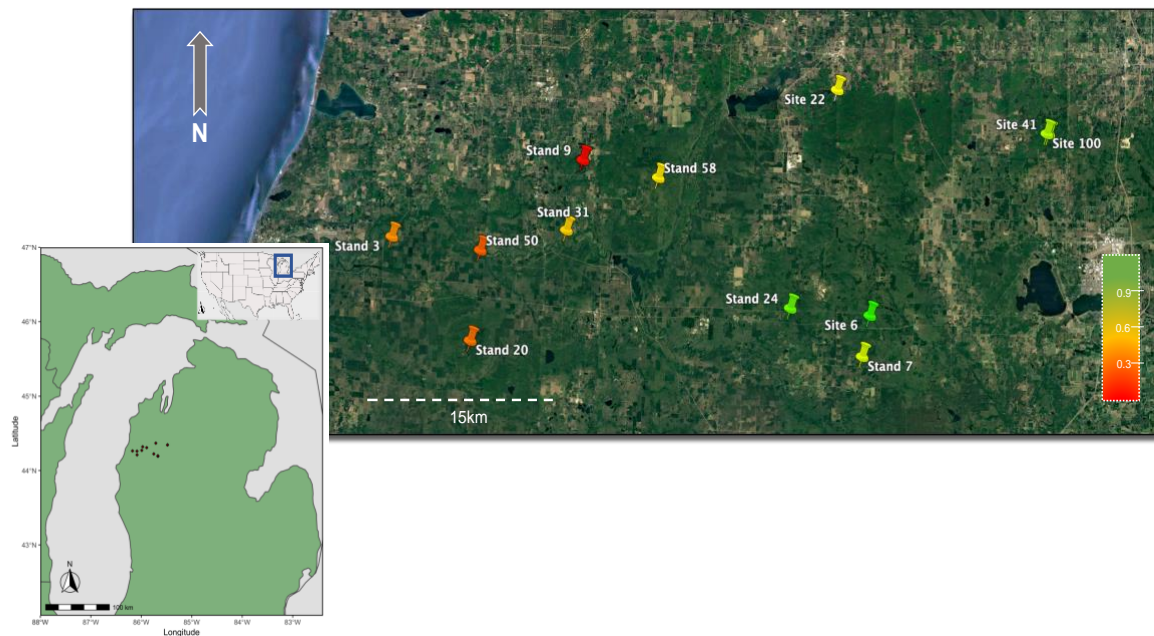
Fig S2 Relationship between May and August rates of net N mineralization. Repeated sampling occurred around the base of the same individual *Quercus rubra* trees ($R^2_{adj} = 0.58$. *P* <0.001). Black line is 1:1 plot. Shading around blue line depicts 95% confidence interval.
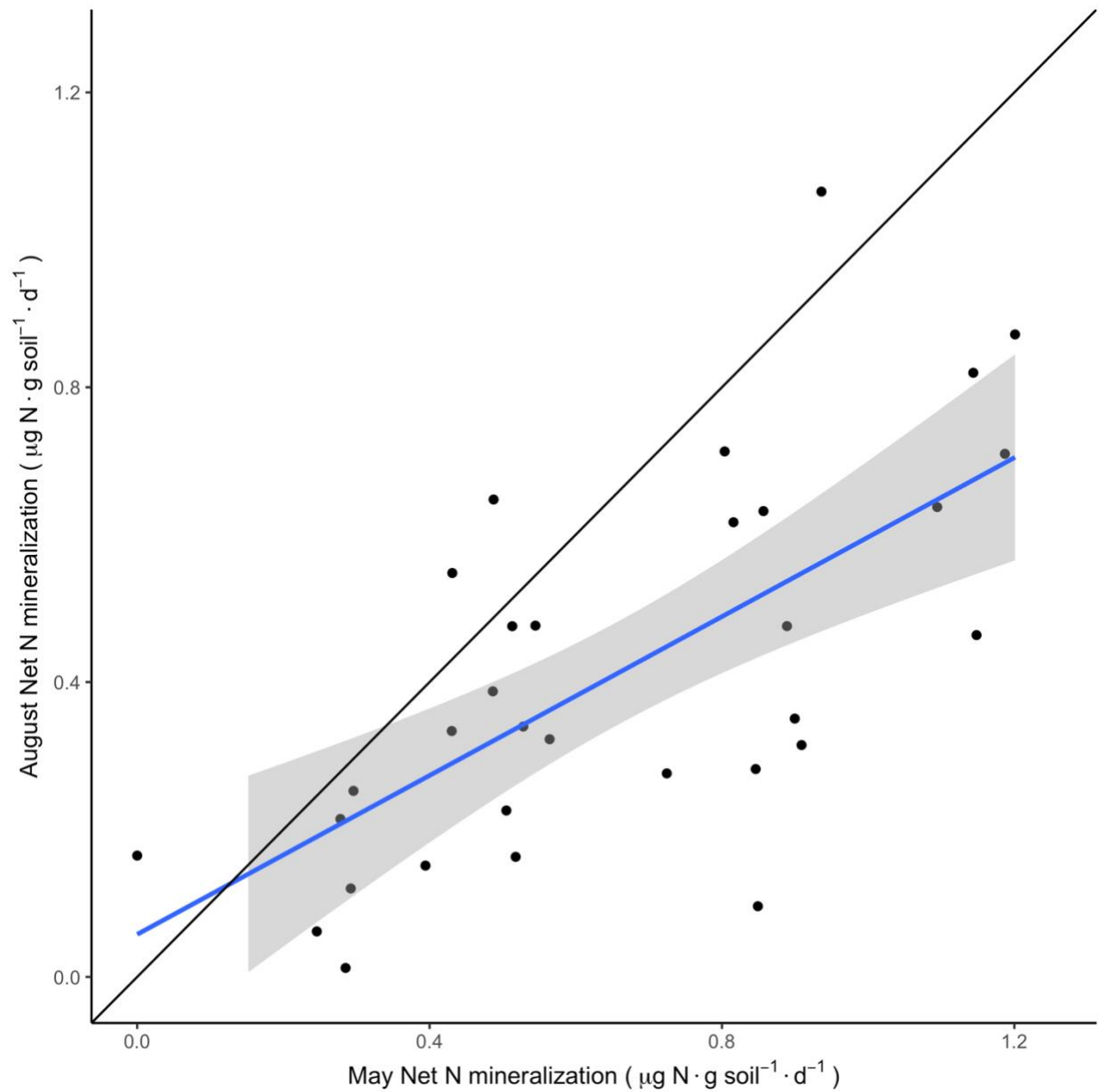
Fig. S3 Number of colonized ectomycorrhizal fungal root-tips encountered on focal *Quercus rubra* individuals across the soil gradient: $R^2_{adj} = 0.25$, $P < 0.001$. Shading around blue line depicts 95% confidence interval.
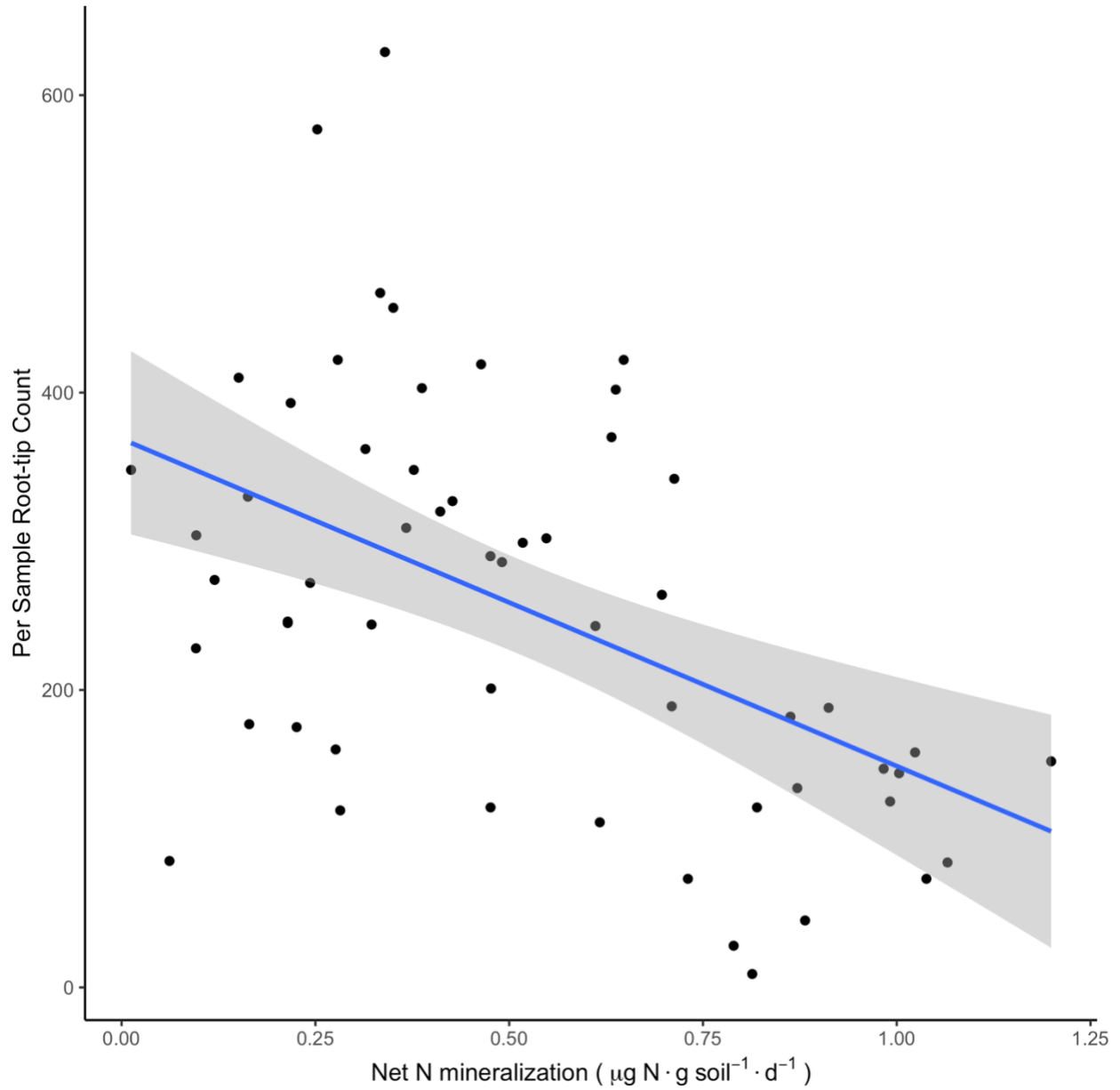
Fig S4 Freeze-dried weight of ectomycorrhizal fungal root-tips collected from focal *Quercus rubra* individuals across the soil gradient of net N mineralization rates $R^2_{adj} = 0.10$, $P < 0.01$. Shading around blue line depicts 95% confidence interval.
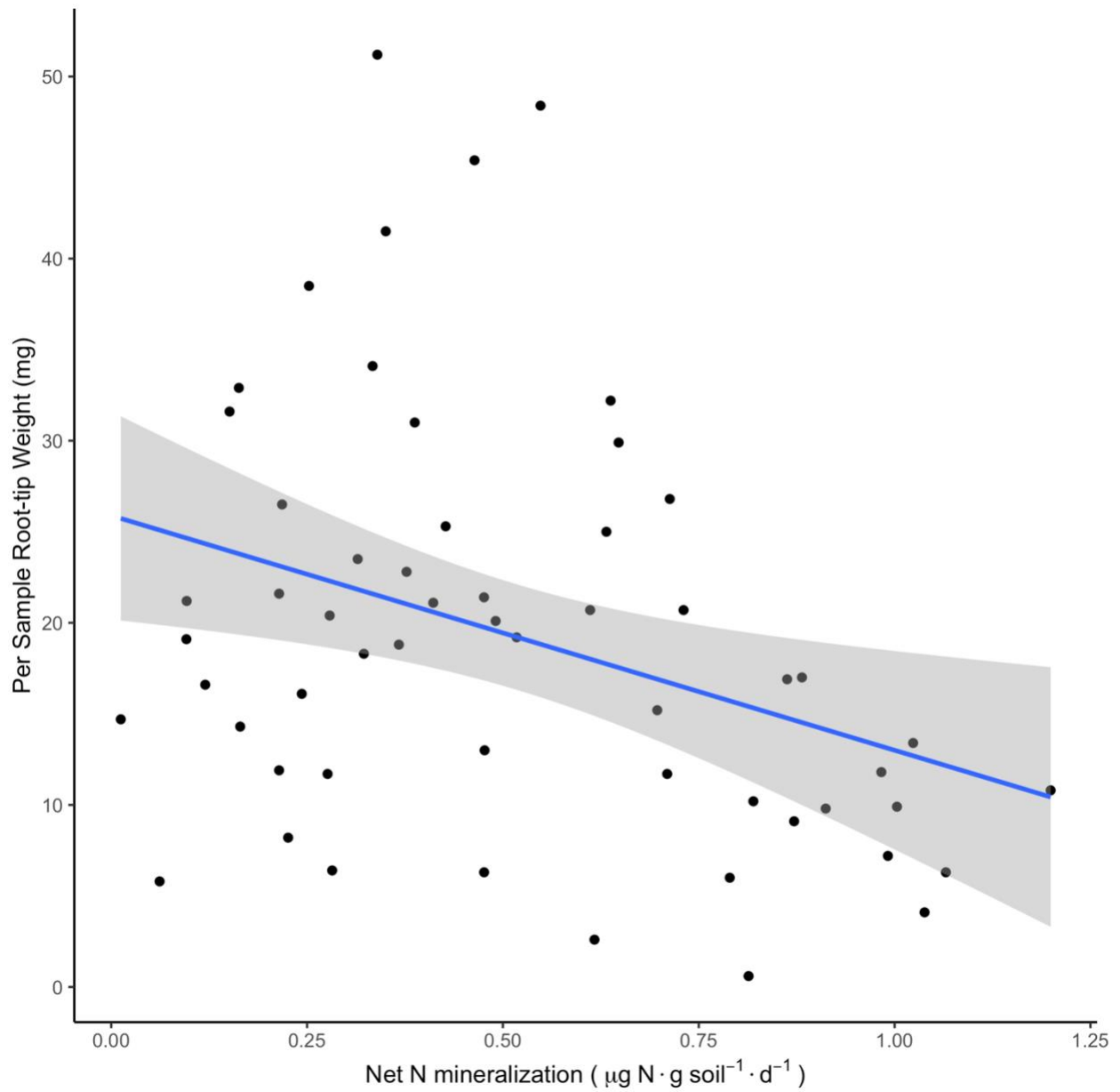
Fig S5 Indices of ectomycorrhizal (ECM) fungal alpha diversity across the soil inorganic N gradient. Y-axis depicts distinct scales for each panel, Observed ECM amplicon sequence variants derived from dada2, and Simpson Index.
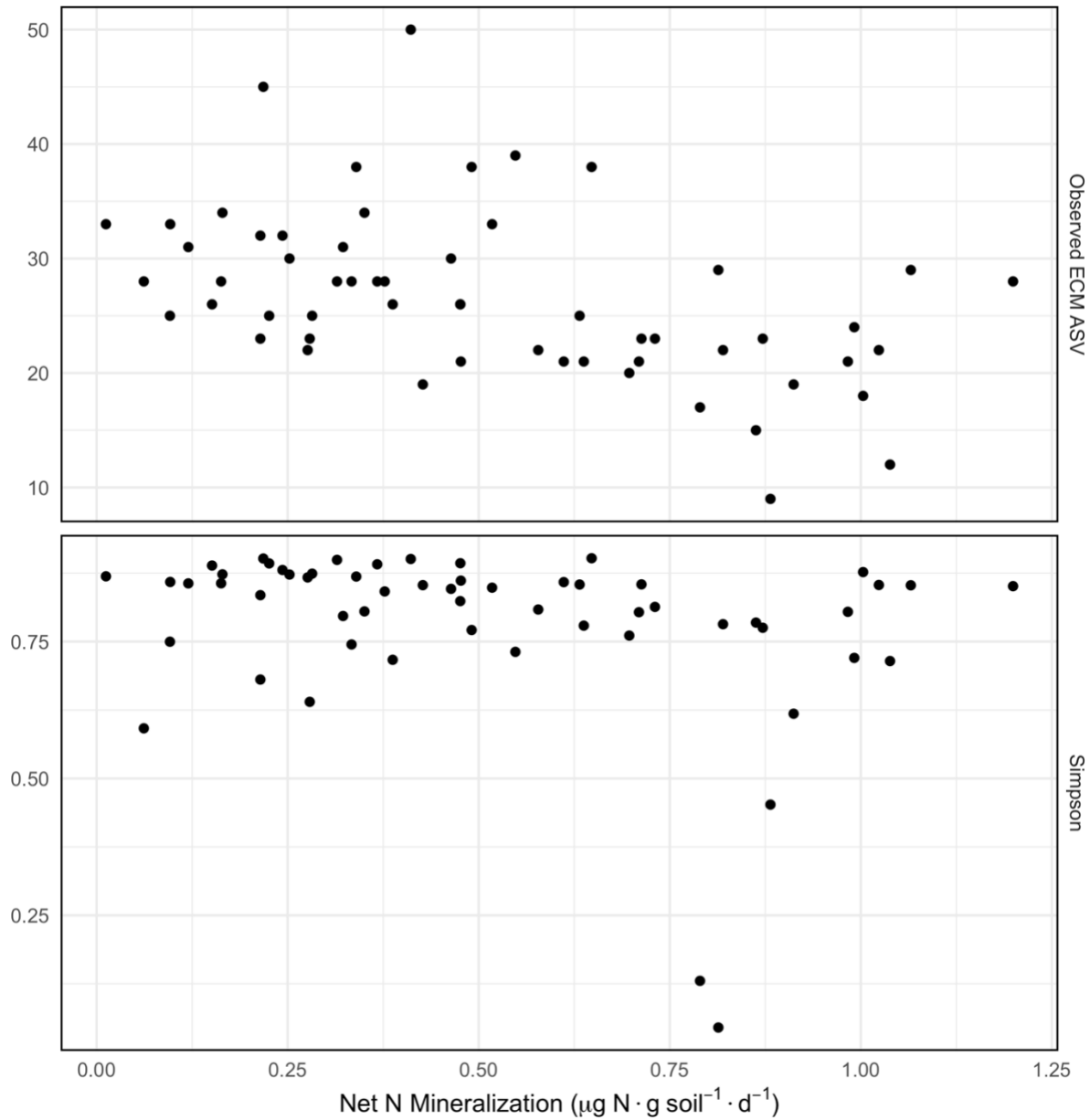
Fig S6 Metagenomic sequencing yield for each filtering step (color) across all samples. QC (red) represent quality filtered metagenomic reads, see main text. Kraken Unmapped (blue), represents reads that remain after Kraken filtering against plant and other databases containing contaminants, such that reads remaining represent putative fungal sequences. No significant relationships across the gradient of net N mineralization rates: linear regression: QC: $P = 0.36$. Kraken Unmapped: $P = 0.73$.
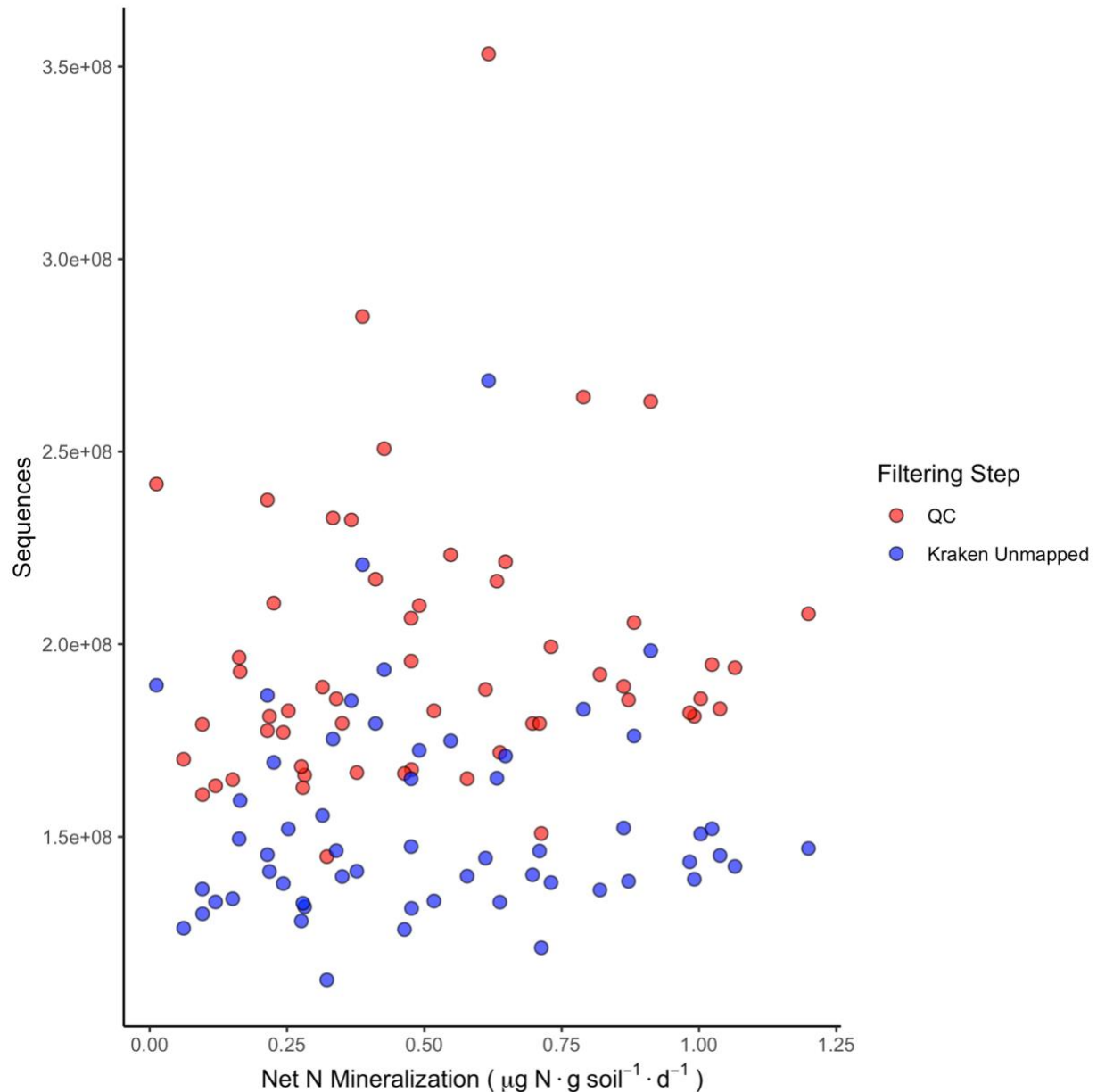
Fig S7 Average number of single copy fungal gene counts per million kraken2 filtered metagenomic sequences. The geometric mean number of single copy genes was calculated from 1312 gene families derived from the Ortho DB v.9 database, divided by the sum of all reads in that sample and then multiplied by 1e6. Error bars represent SE, and may be interpreted as a measure of variance in genome completeness at the community level. No significant variation in the yield of single copy fungal gene counts across the gradient of net N mineralization (P = 0.17).
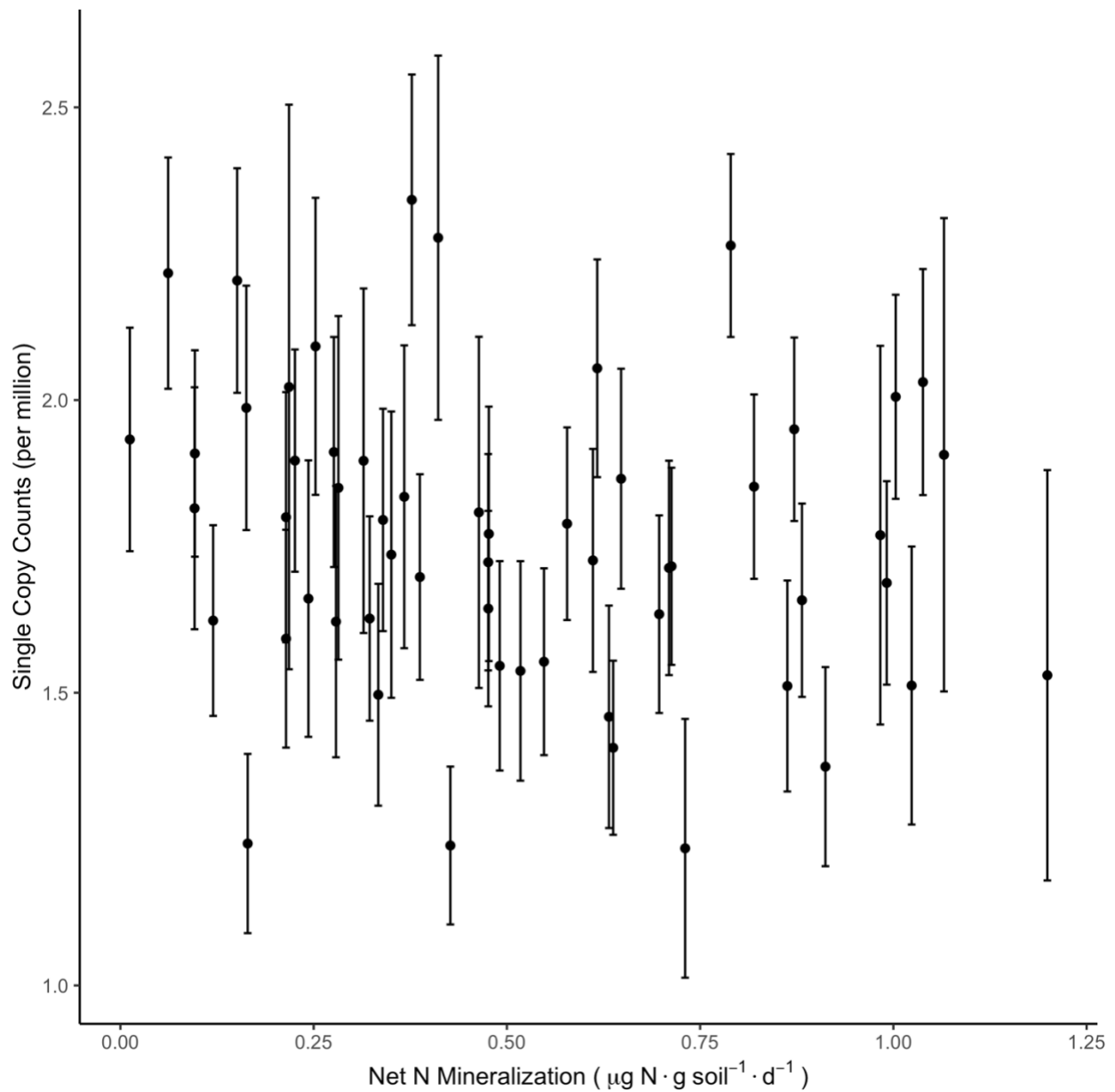
Fig S8. The number of fungal genomes estimated using metagenomic estimates (x-axis), and their relationship with amplicon based (ITS2) estimates (y-axis) of the observed number of amplicon sequence variants (ASV) for ectomycorrhizal (ECM) fungal communities (top) and all fungal ASV (bottom). No significant relationships were detected for either panel. Note distinct y-axis.
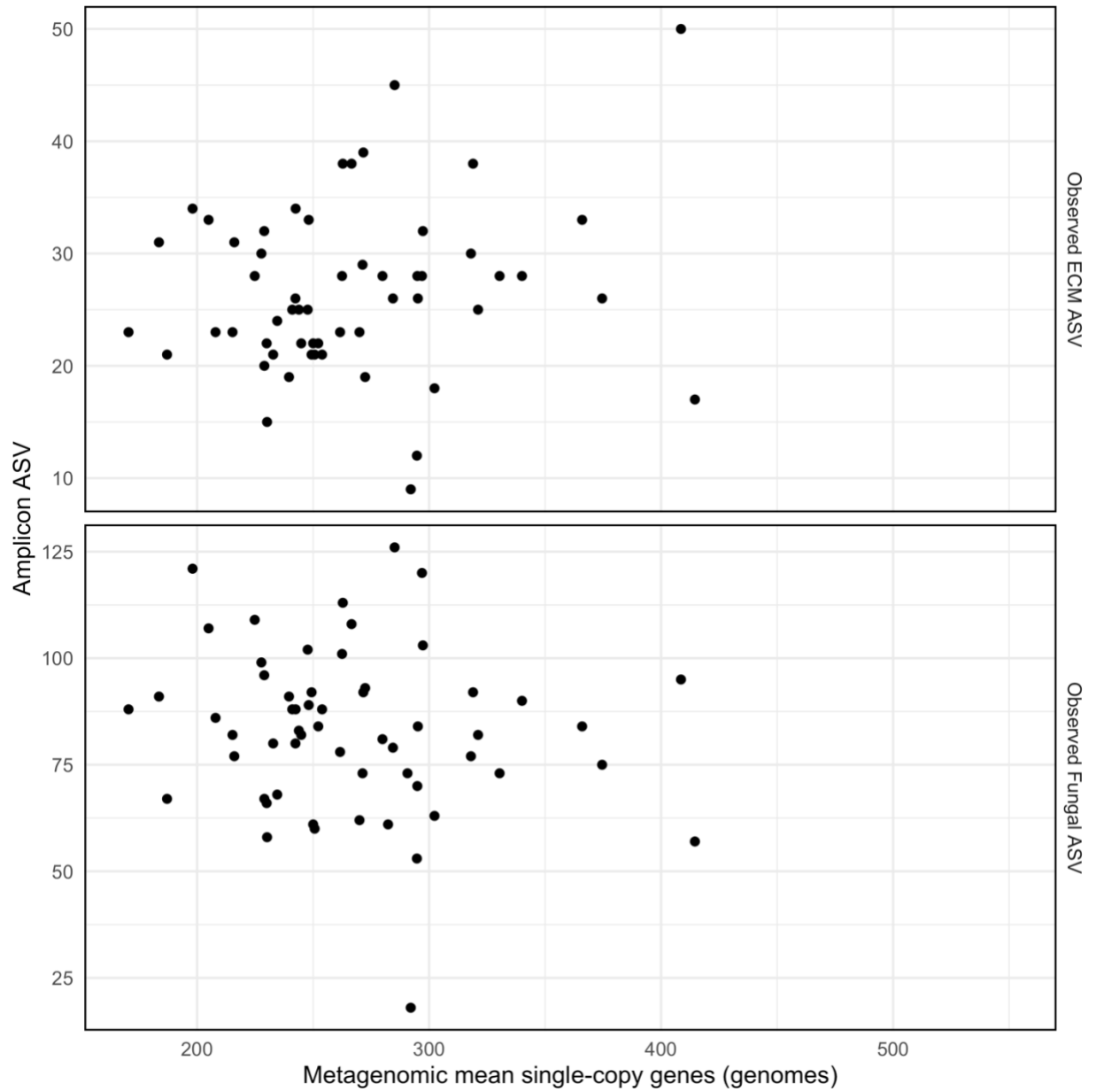
Fig S9. TITAN2 output depicting change points for fungal gene families that respond negatively (green) and positively (blue) to increasing bulk soil C (%C). Peaks along the soil gradient (x-axis) depict location of greatest shift in relative abundance of respective gene family. Purity and reliability = 0.95 for plotting purposes.
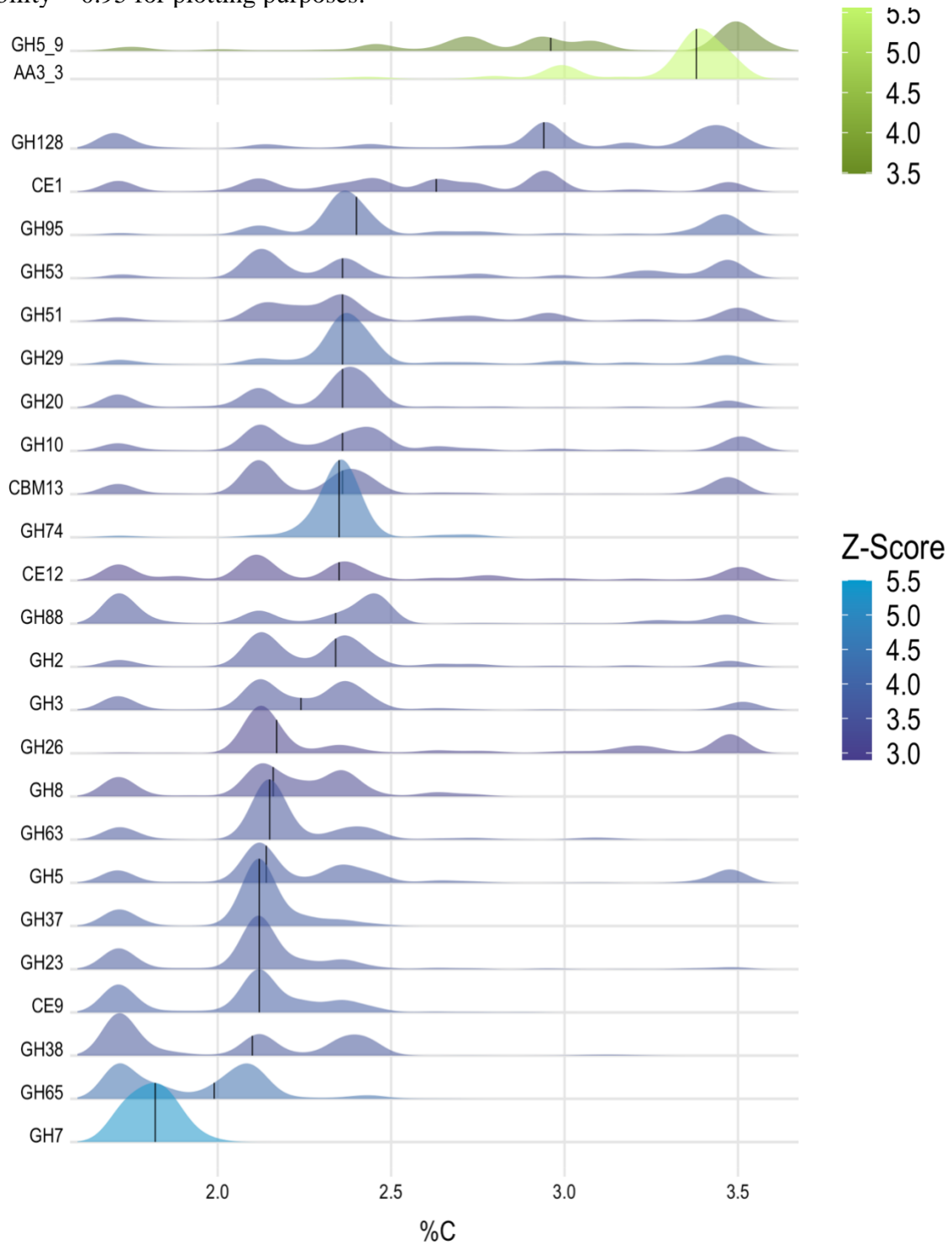
Fig S10 Change points for the fungal gene families that responded positively to increasing soil water availability. Purity and reliability = 0.95, for plotting purposes. Peak along the soil gradient (x-axis) depicts location of greatest shift in relative abundance of individual gene families.
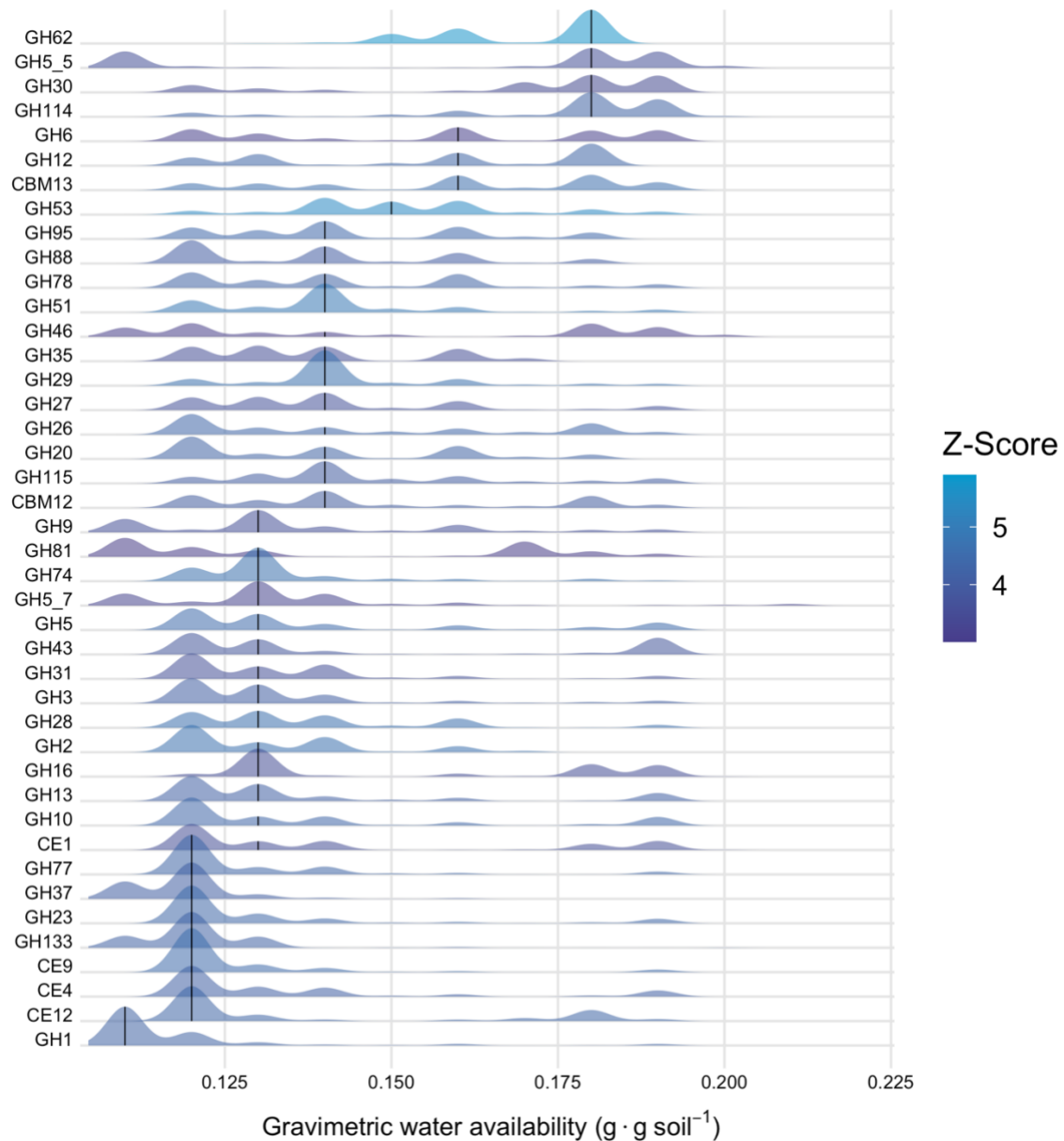
Fig S11. Change points for the fungal gene families that responded negatively to increasing soil water availability. Purity and reliability = 0.95, for plotting purposes. Peak along the soil gradient (x-axis) depicts location of greatest shift in relative abundance.
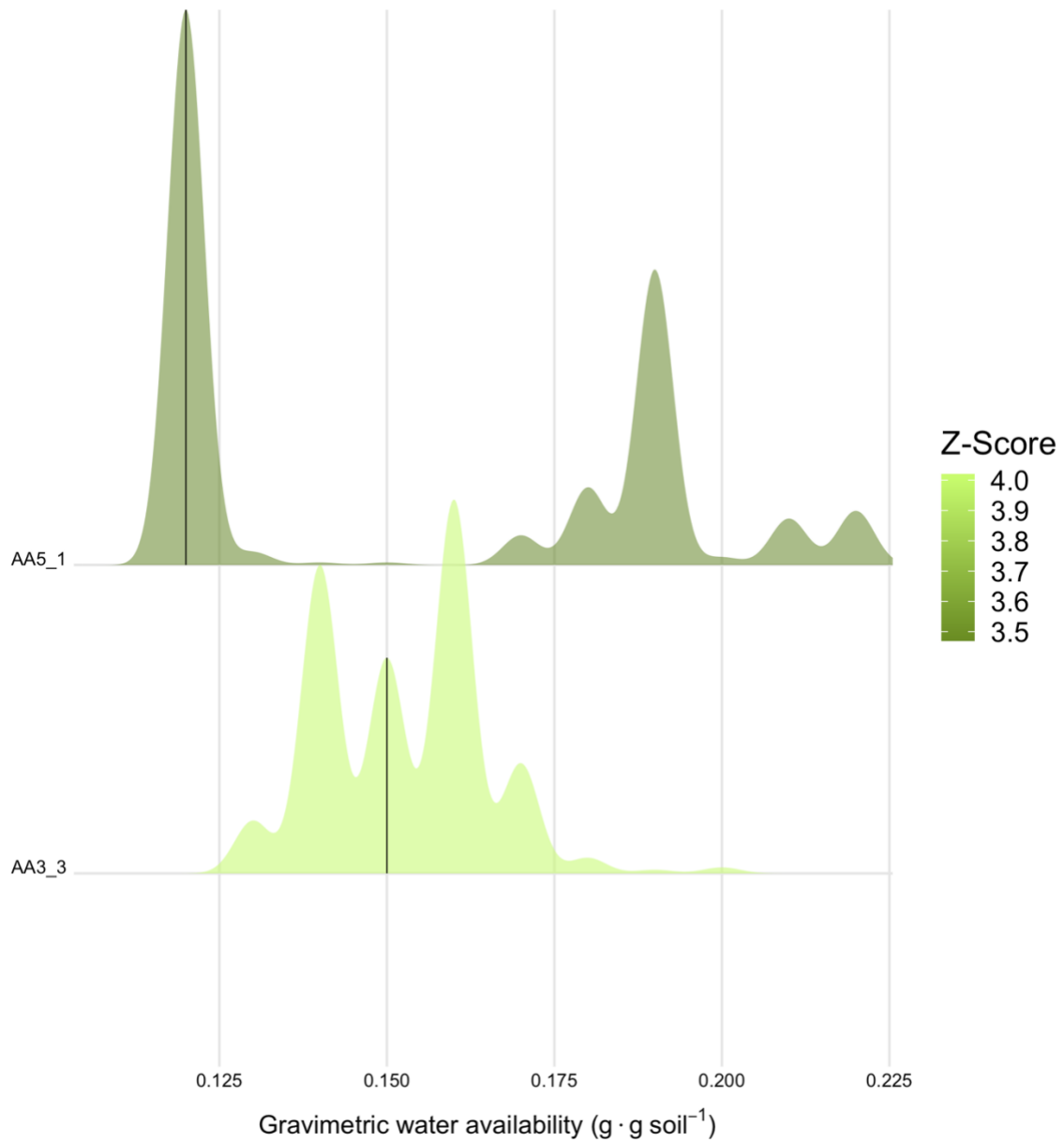
Fig S12. Relative sequence abundance of non-ectomycorrhizal fungal genera present at greater than 0.1% relative sequence abundance. Red lines indicate GAM fits and 95% confidence interval. Note distinct y-axis scales
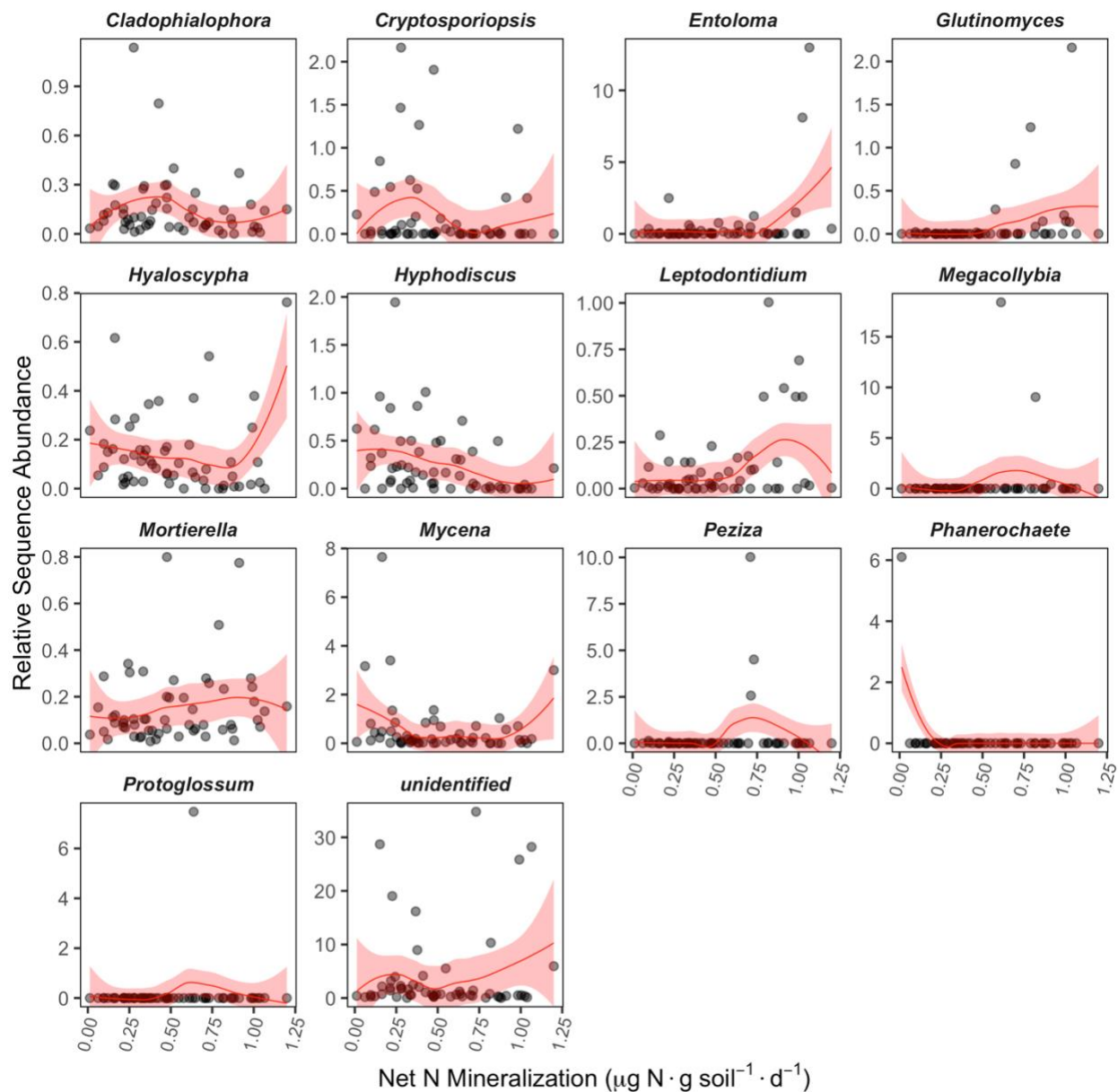
Fig S13. Generalized dissimilarity model (GDM) results depicting shifts in the rate of change (slope) in the relative abundance of fungal decay gene families along the respective gradients on shown on the x-axis (panels). Two fungal community dissimilarity matrices were here incorporated as predictors (A & B), along with two abiotic soil measures (C & D). See also Figure 2 in the main text. The maximum height of the regression line indicates the relative proportion of variance explained by each variable. Note that 'Putative Endophytic' communities studied here comprise less than ~5% of sequence abundance across samples.
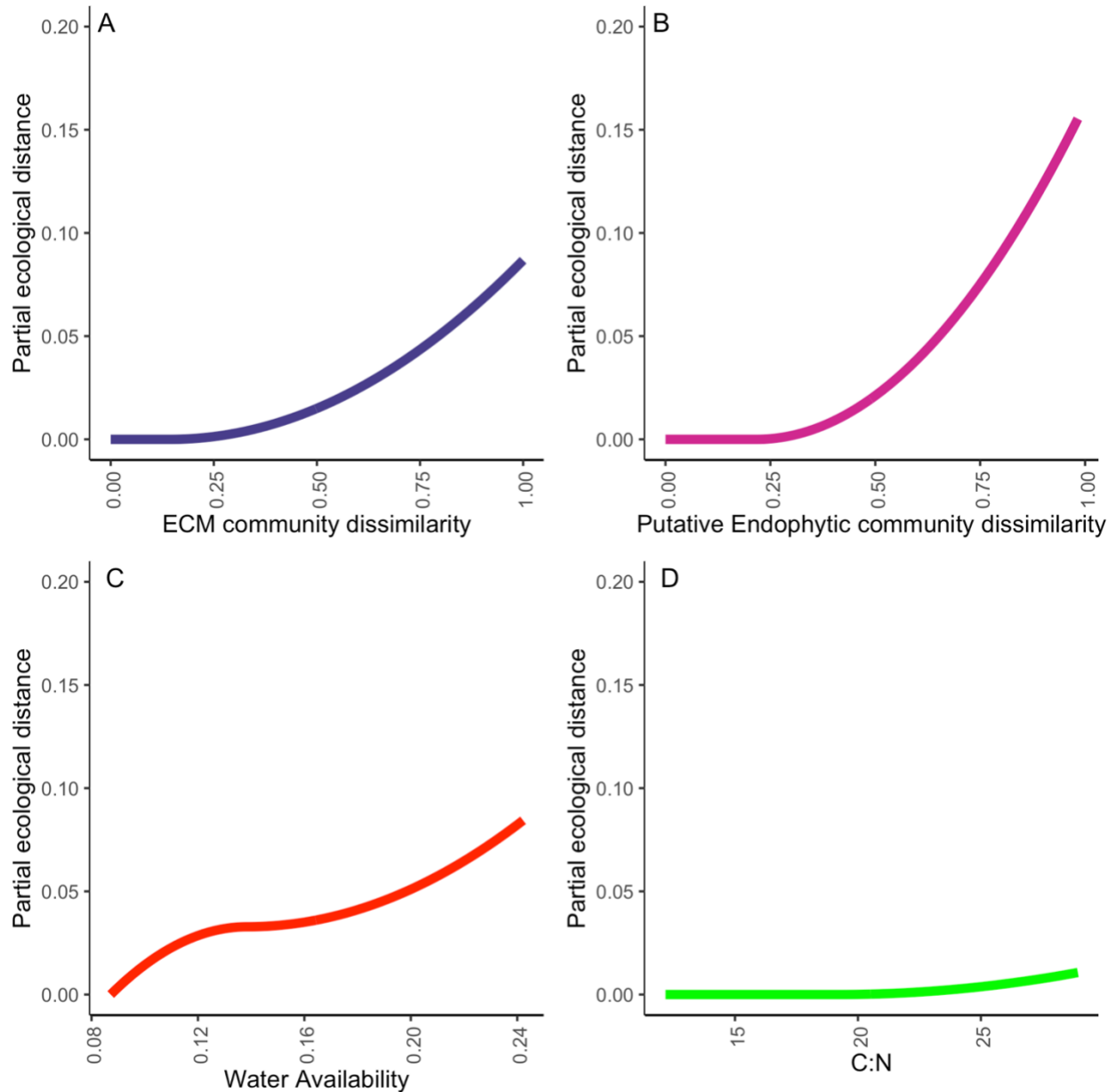
Table S1. Gene families, enzymes they encode and putative substrates. Sourced from (CAZy: http://www.cazy.org; http://peroxibase.toulouse.inra.fr/) and (CAZypedia.org)

Table S2. GDM model output. Tab 1: GDM output for ECM communities, Tab: 2 GDM output for variables contributing to ECM metagenomic decay potential, Tab 3:GDM output for variables contributing to ECM metagenomic decay potential with both ECM and non-ECM fungi as predictors.

Table S3. Number of Raw, QC'd and Kraken filtered reads (Fungal hits). Methods for compiling single copy counts and Total CAZy and Peroxibase Gene Counts can be found in *Supplementary Methods*. ID represents individual tree stems, corresponding to plots depicted in Figure S1.

Table S4. Linear mixed model output including spatially explicit correlation matrix to account for potential non-independence among samples. Regressions for rates of Net N mineralization, Soil C (%), and Gravimetric water availability.

Methods S1:

Soil chemical analyses

Following field collection, soil and ectomycorrhizal root-cores were stored on ice and immediately transported to the laboratory. Soil was collected around the dripline of focal trees for assessment of mineralization rates. Samples were collected identically in both May and August 2018 around the same trees. Soil inorganic N was extracted from fresh sieved (2mm) soil using 2M KCl, followed by a 14-day aerobic incubation assay in order to measure rates of soil inorganic N mineralization (Vitousek, 1982). NO_3^- and NH_4^+ in soil extracts were analyzed colorometrically (AQ2; Seal Analytical, Mequon, WI, USA). Eight of the soil incubations for the May sampling were disrupted, and August mineralization rates are reported throughout. Total free primary amines (TFPA) (primarily amino acids and amino sugars) was measured using unincubated soil extracted with 2M KCl following (Darrouzet-Nardi *et al.*, 2013) using a Synergy HT microplate reader (Bio-Tek INC., Winooski, VT, USA). TFPA is expressed as µmol leucine equivalents, because leucine was used as an analytical standard; estimates of TFPA

availability may be considered relative indices of labile organic N availability in soil solution. Soil was dried 105°C and total C and N contents (% of dry mass) were determined using combustion analysis on a LECO TruMac CN analyzer (LECO Corporation, St. Joseph, MI, USA). Soil pH was determined for 2:1 deionized water-soil slurries with an Accumet 15 pH meter (Fisher Scientific, Waltham, MA, USA).

Ectomycorrhizal root-tip processing and DNA extraction:

Ectomycorrhizal root-tips were isolated from root-cores within 12 days of field sampling. Definitive ectomycorrhizal tips were sampled after visual confirmation of ectomycorrhizal mantle and high turgor (Agerer 2001). Individually colonized *Q. rubra* root-tips were counted and excised, and rinsed in 2% CTAB supplemented with 0.8% ß-mercaptoethanol, and then stored and frozen in fresh 2% CTAB at -80C. CTAB was removed, and then freeze-dried at -50C. Lyophilized root-tips were then weighed using a microbalance. DNA was extracted from the totality of each sample ectomycorrhizal root-tip sample, using two or three individual extraction columns; each extraction utilized ~10 mg of lyophilized root-tip per extraction, so as not to bias extraction efficiencies. Each lysis tube contained 800μl of Buffer AP1 and 4μl of RNase A from a Qiagen DNeasy Plant Mini Kit. Tubes were vortexed and placed in a 65°C waterbath for 20 minutes. DNA was then extracted using the Qiagen DNeasy Plant Mini Kit following manufacturers recommended protocol. Extraction replicates were combined for each sample and DNA recovery was assessed using gel electrophoresis. Assessment of DNA quality was conducted using a Nanodrop Spectrophotometer (Thermo Fisher). The Quant-iT PicoGreen dsDNA Assay Kit (Thermo Fisher) and a BioTek SynergyHT Multi-Detection Microplate Reader were used to quantify DNA concentrations prior to PCR. Samples were split into DNA pools for amplicon sequencing and shotgun-metagenomic sequencing. Freeze-thaw cycles were carefully limited throughout the extraction, storage and sequencing steps in order to prevent shearing.

PCR and Amplicon Bioinformatic Processing:

The ITS2 region was amplified using Illumina dual-indexed primers 5.8S Fun and ITS4 Fun (Taylor *et al.*, 2016). The forward and reverse primer each contained the appropriate Illumina Nextera adaptor, linker sequence and error correcting Golay barcode for use with the Illumina MiSeq platform. All PCRs were performed in triplicate following Taylor et al. (2016), using Phusion High Fidelity DNA Polymerase and master mix (New England BioLabs, Ipswich MA,

USA). Each PCR contained 6 µl High Fidelity Phusion 5 × buffer, 0.75 µl each primer (10 µm initial concentration), 0.42 µl dNTPs (20 mmol^{-1} initial concentration of each dNTP), 1.5 µl of template DNA (mean concentration 3.76 ng/µl, SD=2.82) and 0.23 µl of Taq (2 U/µl) brought to a final volume of 20 µl with molecular-grade water. PCR conditions consisted of an initial denaturation step at 94°C for 3 min, followed by 27 cycles of the following: 30 s at 94°C, 45 s at 57°C and 90 s at 72°C followed by a final extension step of 72°C for 10 min.

Illumina sequencing generated a total of 27 274 716 raw reads that were demultiplexed. The first 10bp each forward read was trimmed due to low sequence quality, but sequences were not truncated (Pauvert *et al.*, 2019). The DADA2 pipeline was implemented in QIIME 2 in order to denoise sequences, detect and remove chimeras and remove PhiX contaminants and infer exact representative sequences named amplicon sequence variants (ASV) (Callahan *et al.*, 2016). A maximum of 2 expected errors (MaxEE = 2) was allowed. ASV were inferred using a total 6,869,462 of filtered forward sequences (mean= 5.64 x 10^5, *SD*= 1.39 x 10^5 sequences per sample). ASV were assigned taxonomy using the dynamic (97-99% sequence similarity) UNITE database (v.8)(Nilsson *et al.*, 2019) and the scikit-learn naive Bayes machine-learning algorithm (Bokulich *et al.*, 2018). This dynamic classification system captures known variation among fungal clades in delimited species sequence similarity (Garnica *et al.*, 2016). Samples were rarefied to 24,021 sequences, the second lowest number of quality sequences recovered per sample, and collapsed at the finest possible taxonomic level using the *taxa collapse* command in QIIME 2. Taxa that could not be assigned to Fungi, and appeared less than twice across all samples were removed.

We used the DEEMY (characterization and DEtermination of EctoMYcorrhizae) database (http://www.deemy.de/) to gather morphological information on the exploration type (hyphal foraging distance) and rhizomorph occurrence of ectomycorrhizal taxa present in our dataset at more than 0.5% relative abundance. When fungal taxa present in our study were not represented in the DEEMY database, congeners were surveyed and, if 90% of the entries agreed, consensus trait values were assigned to that taxon (Moeller *et al.*, 2014). This classification system is supported by the fact that foraging-related functional traits for fungal hyphae are typically conserved at the genus level (Agerer, 2006). This also allowed incorporation of taxa that could only be identified to Genus level. Only taxa producing 'abundant rhizomorphs' were

recorded as 'rhizomorphic'. Long-distance foraging types were rare in our study system, composing less than 7% of ectomycorrhizal-derived sequences in each sample (SE = 1.38).

Metagenomic Sequencing and Processing:

Prior to metagenomic sequencing library preparation, DNA extracts were quantified (Agilent 4200 TapeStation; Santa Clara, CA). 40ng of input DNA was used for library construction, however six of the 60 samples had lower total DNA yield. For these samples, the totality of all DNA was used. Libraries were then custom sheared using a Covaris S2 Focused-Ultrasonicator Woburn, MA, USA), to a target of 200 bp (duty =10%, intensity = 5, cycles/burst = 200, time =120 seconds); previous trials confirmed the efficacy of these settings. Libraries were prepared using the NEB Next Ultra 2 DNA Library Prep kit with seven cycles of PCR. 59 out of the 60 samples successfully yielded libraries suitable for sequencing. Sequencing was conducted using a full S4 flow cell of the Illumina NovaSeq 6000 instrument.

In total, 23 203 326 006 sequences were generated. Reads were then dereplicated, adapters trimmed, sequence Q >20 retained, and reads shorter than 40 bp were removed using BBDuk (jgi.doe.gov). 23 177 098 622 paired-end reads passed initial quality filtering. We then used an additional filtering step to remove non-fungal sequences using Kraken2 paired-end mode with default parameters (Wood *et al.*, 2019). Sequences were mapped against the standard Kraken2 database containing bacterial, archaeal and UniVec reads (containing sequencing adapters, linkers, and primer sequences), and further supplemented with sequences obtained from published *Quercus rubra* (Konar *et al.*, 2017) and *Qurcus lobata* genomes (Sork *et al.*, 2016) in order to remove plant sequences (contaminants). All mapped reads were removed. On average, 21.7% of sequences per sample were removed during this Kraken2 filtering step, and the mean number of sequences remaining in each sample after Kraken mapping was 307 041 274

Filtered reads were mapped to functional reference gene databases CAZy (accessed March 2019) (Lombard *et al.*, 2014) and Peroxibase (accessed February 2019) (Fawal *et al.*, 2013). Translated reads were mapped to CAZy using 'sensitive' mode in DIAMOND v. 0.9.31, with an -e value: $1e^{-4}$, following best practices for unmerged reads (Treiber *et al.*, 2020). BWA-MEM was used to map sequences to representatives downloaded from Peroxibase, using default settings (Li & Durbin, 2009). The number of mapped reads was averaged for unmerged forward and reverse reads for each reference gene to avoid double counting, the geometric mean of all mapped reads were then averaged across all reference sequences for a given gene family. Note

that gene counts were not normalized to gene size (for example, average length of each gene in CAZy) because this was unnecessary: comparisons of CAZy relative abundances were primarily to environmental parameters and not to each other, and the multivariate models used here are insensitive to absolute magnitudes.

We tabulated the number of near-single copy genes, as a proxy for the number of Dikaryotic fungal genomes present in each sample, using the OrthoDB v.9 orthologous ancestral gene database, which comprised 1312 near-single copy Dikaryotic gene variants (Kriventseva *et al.*, 2019). Filtered forward and reverse reads for each sample were mapped to the Dikaryotic OrthoDB database of 1312 orthologs using 'sensitive' DIAMOND as above. Mapped reads to each ortholog were averaged to prevent double-counting. Dikaryotic near single-copy genes were chosen because the majority (> 90% sequences) are Dikaryotic. Instead of relying on a single arbitrarily chosen house-keeping gene that may not be at true single-copy in complex environmental samples, we calculated the geometric mean number of 'single-copy' genes present across all orthologs and the standard error of orthologous gene counts for each sample.

Statistical Analyses:

We visualized shifts in ectomycorrhizal community composition using NMDS with Hellinger transformed Bray-Curtis distance matrices; the relative abundance of hyphal morphotypes were correlated with NMDS axes using envfit in *vegan v 2.5-6*. The responses of individual ectomycorrhizal genera to changes in soil inorganic N availability (net N mineralization rates) was conducted using Threshold Indicator Taxa ANalysis (TITAN) using the 'TITAN2' package in R (Baker & King, 2010). TITAN combines bootstrapped indicator species analysis (IndVal scores) and change-point analysis (Baker & King 2010). Fungal genera with known ectomycorrhizal trophic status were included, while those that occurred in less than five samples, with a frequency of less than five were removed following (Baker & King, 2010). We used relative cutoff and threshold scores of 0.85 after Hellinger transformation of the ASV table. Indvals were calculated using the relative abundance obtained by the ratio of summed abundance in each partition to the total, to address skew.

We tested which environmental variables were most strongly correlated with ectomycorrhizal community composition using generalized dissimilarity modeling (GDM) (Ferrier *et al.*, 2007). GDM can accommodate non-linear responses to environmental gradients, and identify where along gradients community change is slow/rapid. Modeling assemblages of

fungal taxa along environmental gradients was accomplished by transforming predictor variables by fitting I-spline functions to the environmental variables and polynomial pieces are connected using knots (Ferrier *et al.*, 2007). We used Bray-Curtis dissimilarity matrices and three I-splines for each predictor (Ferrier *et al.*, 2007; Qin *et al.*, 2020). Environmental predictors initially included in the model, included net N mineralization rates, pH, soil C and N, C:N, total free primary amines (TFPA), gravimetric soil moisture, and Bray-Curtis transformed plant overstory dissimilarity matrices. We separately assessed plant overstory using the frequency of stems of a particular species, as well as summing the DBH of each overstory species. We used backwards model selection using *gdm.varImp* to iteratively remove variables that resulted in less than 0.5% change in model deviance (*nperm* = 250) (Bouma-Gregson *et al.*, 2019; Qin *et al.*, 2020). We performed additional testing on the selected GDM to confirm significance of remaining predictors (*nperm* =500). After finalizing model fit, we estimated the proportion of model variance uniquely attributable to soil inorganic N availability, by calculating the difference in the deviance explained by a GDM containing final variables and a model with all variables except soil inorganic N availability. We then converted this difference to a percentage by dividing by the deviance explained by the finalized GDM *sensu* (Gossner *et al.*, 2016). We identified a core set of decay gene families (unscaled) that were responsive to elements of the studied soil gradient deemed significant using GDM (net N mineralization rates, C, and soil water availability). We used purity and reliability scores > 0.8, metrics that are based on the robustness of the sign and magnitude of gene responses when resampled using 1000 bootstraps (Baker and King 2010). This threshold is more stringent than other 'omic studies using IndVal's to identify microbial gene families and transcripts that respond to environmental treatments (Malik *et al.*, 2020).

Gene families underlying Figure 1: (Chitin: GH18,GH19, GH20, GH46, AA10, AA11) Cellulose (AA8,AA9, AA11, AA12 GH1, GH12, GH43,GH6,GH7, GH9, GH5_5, GH5_22, AA3_1, AA3_2). Hemicellulose: (CE1, CE12, CE16, CE5, CE8, GH10,GH11, GH115, GH2, GH27, GH28, GH29, GH35, GH5_7, GH51, GH53, GH63, GH74, GH78, GH95, GH3). Lignin: (AA3_1, AA3_1, AA3_2, AA3_3, AA3_4, AA1, AA2, AA9, MnP, DyPrx, LiP).

Supplementary References:

Agerer R. **2006**. Fungal relationships and structural identity of their ectomycorrhizae. *Mycological Progress* **5**: 67–107.

Baker ME, King RS. **2010**. A new method for detecting and interpreting biodiversity and ecological community thresholds. *Methods in Ecology and Evolution* **1**: 25–37.

Bokulich NA, Dillon MR, Bolyen E, Kaehler BD, Huttley GA, Caporaso JG. **2018**. q2-sample-classifier: machine-learning tools for microbiome classification and regression. *Journal of open research software* **3**.

Bouma-Gregson K, Olm MR, Probst AJ, Anantharaman K, Power ME, Banfield JF. **2019**. Impacts of microbial assemblage and environmental conditions on the distribution of anatoxin-a producing cyanobacteria within a river network. *The ISME Journal* **13**: 1618–1634.

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. **2016**. DADA2: High-resolution sample inference from Illumina amplicon data. *Nature Methods* **13**: 581–583.

Darrouzet-Nardi A, Ladd MP, Weintraub MN. **2013**. Fluorescent microplate analysis of amino acids and other primary amines in soils. *Soil Biology and Biochemistry* **57**: 78–82.

Fawal N, Li Q, Savelli B, Brette M, Passaia G, Fabre M, Mathé C, Dunand C. **2013**. PeroxiBase: a database for large-scale evolutionary analysis of peroxidases. *Nucleic Acids Research* **41**: D441–D444.

Ferrier S, Manion G, Elith J, Richardson K. **2007**. Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**: 252–264.

Garnica S, Schön ME, Abarenkov K, Riess K, Liimatainen K, Niskanen T, Dima B, Soop K, Frøslev TG, Jeppesen TS, *et al.* **2016**. Determining threshold values for barcoding fungi: lessons from Cortinarius (Basidiomycota), a highly diverse and widespread ectomycorrhizal genus. *FEMS Microbiology Ecology* **92:** fiw045.

Gossner MM, Lewinsohn TM, Kahl T, Grassein F, Boch S, Prati D, Birkhofer K, Renner SC, Sikorski J, Wubet T, *et al.* **2016**. Land-use intensification causes multitrophic homogenization of grassland communities. *Nature* **540**: 266–269.

Konar A, Choudhury O, Bullis R, Fiedler L, Kruser JM, Stephens MT, Gailing O, Schlarbaum S, Coggeshall MV, Staton ME, *et al.* **2017**. High-quality genetic mapping with ddRADseq in the non-model tree Quercus rubra. *BMC Genomics* **18**: 417.

Kriventseva EV, Kuznetsov D, Tegenfeldt F, Manni M, Dias R, Simão FA, Zdobnov EM. **2019**. OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs. *Nucleic Acids Research* **47**: D807–D811.

Li H, Durbin R. **2009**. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760.

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. **2014**. The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Research* **42**: D490–D495.

Malik AA, Swenson T, Weihe C, Morrison EW, Martiny JBH, Brodie EL, Northen TR, Allison SD. **2020**. Drought and plant litter chemistry alter microbial gene expression and metabolite production. *The ISME Journal* **14:** 2236–2247.

Moeller HV, Peay KG, Fukami T. **2014**. Ectomycorrhizal fungal traits reflect environmental conditions along a coastal California edaphic gradient. *FEMS Microbiology Ecology* **87**: 797–806.

Nilsson RH, Larsson K-H, Taylor AFS, Bengtsson-Palme J, Jeppesen TS, Schigel D, Kennedy P, Picard K, Glöckner FO, Tedersoo L, *et al.* 2019. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Research* **47**: D259–D264.

Pauvert C, Buée M, Laval V, Edel-Hermann V, Fauchery L, Gautier A, Lesur I, Vallance J, Vacher C. **2019**. Bioinformatics matters: The accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology* **41**: 23–33.

Qin C, Zhu K, Chiariello NR, Field CB, Peay KG. **2020**. Fire history and plant community composition outweigh decadal multi-factor global change as drivers of microbial composition in an annual grassland. *Journal of Ecology* **108**: 611–625.

Sork VL, Fitz-Gibbon ST, Puiu D, Crepeau M, Gugger PF, Sherman R, Stevens K, Langley CH, Pellegrini M, Salzberg SL. **2016**. First Draft Assembly and Annotation of the Genome of a California Endemic Oak. *Genes/Genomes/Genetics* **6**: 3485–3495.

Taylor DL, Walters WA, Lennon NJ, Bochicchio J, Krohn A, Caporaso JG, Pennanen T. **2016**. Accurate Estimation of Fungal Diversity and Abundance through Improved Lineage-

Specific Primers Optimized for Illumina Amplicon Sequencing (D Cullen, Ed.). *Applied and Environmental Microbiology* **82**: 7217–7226.

Treiber ML, Taft DH, Korf I, Mills DA, Lemay DG. **2020**. Pre- and post-sequencing recommendations for functional annotation of human fecal metagenomes. *BMC Bioinformatics* **21**: 74.

Vitousek P. **1982**. Nutrient Cycling and Nutrient Use Efficiency. *The American Naturalist* **119**: 553–572.

Wood DE, Lu J, Langmead B. **2019**. Improved metagenomic analysis with Kraken 2. *Genome Biology* **20**: 257.