Broad geographical and ecological diversity from similar genomic toolkits in the ascomycete genus *Tetracladium*

Jennifer L. Anderson[1*] and Ludmila Marvanová[2]

1. Systematic Biology, Department of Organismal Biology, Uppsala University, Norbyvägen 18D, 752 36 Uppsala, Sweden

2. Czech Collection of Microorganisms, Institute of Experimental Biology, Faculty of Science, Masaryk University, Tvrdého 14, 625 00 Brno, Czech Republic

* Corresponding author. J. Anderson, jennifer.anderson@ebc.uu.se

Anderson https://orcid.org/0000-0002-0713-6897

ABSTRACT

The ascomycete genus *Tetracladium* is best known for containing aquatic hyphomycetes, which are important decomposers in stream food webs. However, some species of *Tetracladium* are thought to be multifunctional and are also endobionts in plants. Suprisingly, *Tetracladium* sequences are increasingly being reported from metagenomics and metabarcoding studies of both plants and soils world-wide. It is not clear how these sequences are related to the described species and little is known about the non-aquatic biology of these fungi. Here, the genomes of 24 *Tetracladium* strains, including all described species, were sequenced and used to resolve relationships among taxa and to improve our understanding of ecological and genomic diversity in this group. All genome-sequenced *Tetracladium* fungi form a monophyletic group. Conspecific strains of *T. furcatum* from both aquatic saprotrophic and endobiont lifestyles and a putative cold-adapted clade are identified. Analysis of ITS sequences from water, soil, and plants from around the world reveals that multifunctionality may be widespread through the genus. Further, frequent reports of these fungi from extreme environments suggest they may have important but unknown roles in those ecosystems. Patterns of predicted carbohydrate active enzymes (CAZyme) and secondary metabolites in the *Tetracladium* genomes are more similar to each other than to other ascomycetes, regardless of ecology, suggesting a strong role for phylogeny shaping genome content in the genus. *Tetracladium* genomes are enriched for pectate lyase domains (including PL3-2), GH71 α-1,3-glucanase domains and CBM24 α-1,3-glucan/mutan binding modules, and both GH32 and CBM38, inulinase and inulin binding modules. These results indicate that these fungi are well-suited to digesting pectate and pectin in leaves when living as aquatic hyphomycetes, and inulin when living as root endobionts. Enrichment for α-1,3-glucanase domains may be associated with interactions with biofilm forming microorganisms in root and submerged leaf environments.

INTRODUCTION

The freshwater fungi traditionally called "aquatic hyphomycetes" are best known as ecologically important decomposers in stream ecosystems. The group is polyphyletic, its members are classified in various orders and families of Ascomycetes and Basidiomycetes. These fungi thrive on leaves and other plant debris that enter streams and other water bodies from terrestrial plants. In the process, carbon and nutrients from recalcitrant plant compounds, including cellulose and lignin, become accessible to diverse consumers in the stream food web (reviewed in (1, 2)). Up to 99% of the carbon in some streams enters as plant debris that is largely inaccessible to aquatic consumers without the degradative activity of these fungi (3, 4). Aquatic hyphomycete adaptations to stream environments include the production of asexual propagules (conidia) while under water. These spores are often branched, typically tetraradiate, or sigmoid, readily detachable from the fungus (*e.g.* from conidiophores) and passively dispersed in water. The conidia strongly adhere to surfaces upon contact. These features facilitate dispersal and colonization of new substrates (5-8). Less well known is that some aquatic hyphomycetes are dual niche (9) or multifunctional (10) species that differ in mode of nutrition or lifestyle in different environments, and are both decomposers of dead plant material in water (saprotrophs) and endobionts in terrestrial or aquatic plants (6, 9, 11-15). Although it has been known since at least 1939 (16) that some aquatic hyphomycetes could be isolated from terrestrial sources, these fungi are still intensively studied in their aquatic context due to their massive presence ecological importance, especially in streams.

The genus *Tetracladium* de Wild. (Ascomycetes, Leotiomycetes) was described in 1893 as "*parmi les algues, les diatomées et les débris de végétaux supérieurs dans les étangs, les fossés*" [among algae, diatoms and higher plant debris in ponds, ditches] (17). Since that time, conidia of *Tetracladium* have been recorded in aquatc habitats world-wide (18). The genus has grown to include eleven species, most of which produce distinctive, characteristically shaped conidia (19) (Fig. 1). The three most recent additions to the genus, *T. ellipsoideum*, *T. globosum*, and *T. psychrophilum* are the exceptions (20). These species were described from glacial soil from the Tibet Plateau and have elliptical and globose spores in the first two cases, and no reported spores in the third. These are the first species in the genus described from soil. However, it has been known since 1970 that *Tetracladium* could be isolated from terrestrial sources including sterilized roots (*T. marchalianum* (21)) and one species, *T. nainitalense*, was described from endophytic isolates (11). *Tetracladium*-like fungi, fungi identified as *Tetracladium* based on ITS similarity without morphological information, are increasingly reported to be among the most frequently found and abundant taxa in culture and metagenomics based studies of soil fungi, rhizosphere fungi, and endophytes world-wide. These

studies represent diverse ecosystems, soils, and associated plants, ranging from domesticated crops (carrot (22) oilseed rape (23), ginseng (24), lettuce (25) and wheat (26, 27) to wild orchids *(28)* and mosses (29), from agricultural fields (30) to glacial and subglacial soils (31, 32), from sea level (33) to high altitudes (above 2600 m.a.s.l (34)), and spanning the globe from the Arctic (35) to the Antarctic (29, 36). Three psychrophilic species have been described from glacial soils from the Tibet Plateau (20)—one of the most extreme environments on earth; fungal extremophiles are of ecological and industrial interest due to their roles in ecosystem functioning and the secondary metabolites they produce (33, 37). Although hundreds of ITS sequences for *Tetracladium*-like fungi are available in public databases, it is not yet known how the fungi from these diverse environments are related to each other whether they fit within the genus *Tetracladium* as understood morphologically and phylogenetically at present. The generic type species, *T. marchalianum*, was neotypified in 1989 (38), but lack of the neotype ITS sequence also hinders taxonomic efforts within the genus.

Compared to the aquatic lifestyle of *Tetracladium* species (2), little is known about the biology or ecological importance of *Tetracladium* living outside of water. *In vitro*, some species are known to degrade lignin (39), pectin (40), startch (41), solubilize phosphate (42), have antimicrobial activity (43, 44), and to increase metal tolerance in host plants (45). Correlation studies have associated *Tetracladium* with increased plant growth (46), however, clear patterns are not yet emerging from which the roles or impact of *Tetracladium* as endophytes or in soil can be inferred (47). From studies in other fungal functional groups, such as pathogens and non-aquatic hyphomycete endophytes, there is increasing data available about secondary metabolite (SM, *e.g.* (48)) and carbohydrate active enzyme (CAZymes, *e.g.* (49)) profiles, as predicted from genomes, that may be associated with different lifestyles. By comparing the predicted molecular repertoires for *Tetracladium* both among strains from different ecological niches within the genus and to those of better studied functional groups it may be possible to gain insight into the ecology of *Tetracladium*. This reverse ecology approach to understanding *Tetracladium* biology has not before been possible; to date there are no publicly available *Tetracladium* genomes. Relatively few molecular tools exist for these species, they include microsatellites (50), mass spectrometry profiles (51), and taxon-specific fluorescence in situ hybridization probes (52).

The i) ecological importance of *Tetracladium* species in streams; ii) global reports of *Tetracladium*-like species in soils and as endophytes; iii) paucity of information about the non-aquatic roles of these fungi; iv) emerging reports of *Tetracladium* species from extreme environments; all signal a need to better understand fungi callasified in the genus *Tetracladium*. Moreover, these features

raise an exciting suite of questions and opportunities for study. How have *Tetracladium* species evolved to tolerate environmental extremes? What are the features of genomes and their regulation that enable multifunctionality and broad environmental tolerance or adaptation in these species? Can the same genotypes or species thrive as endophytes and aquatics? How is multifunctionality distributed across the *Tetracladium* phylogeny? It is not possible to address all these questions in a single study, but here the foundations are laid to address them and to support advances across all aspects of *Tetracladium* biology. In this study, genomes of all known species of *Tetracladium* and several *Tetracladium*-like species were sequenced, with more than one representative strain per species when available. Phylogenomics was used to resolve evolutionary relationships within the genus. To gain insight into the biology of these fungi from their genomes, secondary metabolites and carbohydrate-active enzymes were predicted from the genomes to compare among species and with profiles for plant pathogens and other functional groups. Further, analysis of the ITS sequences of *Tetracladium*-like fungi from studies world-wide were analyzed together with the strains sequenced herein to test and illustrate the geographic and ecological diversity of the genus.

MATERIALS AND METHODS

Strains and sequencing

Twenty-four fungal strains, including all described species in the genus *Tetracladium*, were obtained for genome sequencing (see Table S1 for strain information and Nagoya Protocol status). *Tetracladium marchalianum* JA005 (full designation JA005.1) and *T. setigerum* JA001 are new collections from the Fyris River, Sweden. The remaining strains were obtained from the Czech Collection of Microorganisms (CCM), CABI, China General Microbiological Culture Collection (CGMCC), the Yeast Culture Collection at the Universidad de Chile, from Dai Hirose at Nihon University, Japan, and from Magdalena Grudzinska-Sterno from the Swedish University of Agricultural Sciences.

Cultures were maintained on potato dextrose agar (Fluka, cat #70139) at room temperature with ambient light or at 4°C in the dark (CGMCC strains) with backup cultures frozen at -80°C in 18% glycerol. Strains were grown in flasks with 50-200 mL of malt extract peptone medium (17 g/L malt extract, 2.5 g/L bacto-peptone) at room temperature with shaking for up to one month to obtain sufficient mycelium for DNA extraction. Mechanical disruption via grinding in liquid nitrogen or bead beating (Mini-Beadbeater Biospec 25/s, for 3 s or less) was performed on fresh or freeze dried mycelium before DNA extraction using the Zymo Quick DNA Fungal/Bacterial Miniprep Kit (D6005) as directed.

Sequencing libraries were prepared at the SNP&SEQ Technology Platform in Uppsala, Sweden, from 1 μg DNA using the TruSeq PCRfree DNA sample preparation kit (FC-121-3001/3002, Illumina Inc.) targeting an insert size of 350 bp as directed. Libraries were then sequenced paired-end with a 150 bp read length on an Illumina HiSEQX Ten, using v2.5 sequencing chemistry.

Genome assembly and annotation

Sequence reads were quality and adapter trimmed using Trimmomatic v0.36 (53). Bases with quality scores below 3 were removed from the start and end of reads. Sequence falling below a quality score of 20 in a sliding window of size 4 and adapter sequences were removed. Resulting sequences greater than 36 bp were retained for genome assembly. Assembly was performed in in Spades v3.11.1 (54) using the "careful" option to reduce mismatches and short indels. The resulting scaffolds were used for further analysis. Assembled genomes were evaluated in QUAST V4.5.4 (55) and BUSCO v3.0.2b (56) using the Pezizomycotina odb9 database to test for completeness. Average depth of coverage was estimated in QualiMap v2.2.1 (57) after read mapping in BWA v0.7.17 (58).

Genomes were annotated using an iterative MAKER v3.01.2-beta (59) annotation pipeline. In the initial run, protein and EST evidence from *Botrytis cinerea* B05.10 (ASM83294v1), *Glarea lozoyensis* (ATCC 20868), *Phialocephala subalpine* (PAC v1), and *Sclerotinia sclerotiorum* (ASM14694v1), accessed from Genbank (28-29 May 2018), were used as evidence to train the Hidden Markov Model. Species specific repeat libraries generated using RepeatModeler v1.0.8 (60), using default settings and the NCBI search engine, were used for repeat masking. Outputs from this run were used to support *ab initio* gene predictions in GenMark-ES v4.33 (61) using the fungus algorithm option. The GenMark outputs were then used in MAKER for gene annotation on scaffolds larger than 1000 bp.

Putative gene functions and protein domains for the annotated genes were identified using Blastp (62) against the UniProtKB/Swiss-Prot annotated non-redundant database (63) (downloaded 29 October 2018) to return no more than one hit per gene with an evalue threshold of $1x10^{-6}$ and InterProScan v5.30-69.0 (64) with mapping to Gene Ontology and member database signatures (without using the pre-calculated match lookup service). These data were used to update the annotations in MAKER.

Single copy orthologs and phylogenomics

Single copy orthologs (SCO) present in all Tetracladium strains and 27 additional ascomycetes (Table S2) were selected from ortholog groups identified by OrthoMCL v2.0.9 (65) (with inflation index 2). Annotations for *Cadophora malorum* and *Articulospora tetracladia* are not available at NCBI and were generated for use herein as described above. Before alignment, SCO sequences were processed in PREQUAL v1.02 (66) using default settings to remove sequence stretches with no evidence of homology. The SCO were then aligned (MAFFT v7.407 G-INS-I) with a variable scoring matrix and α_{max}=0.0.8 (67) and high-entropy regions were trimmed (BMGE v1.12 (68) with the BLOSUM95 similarity matrix). The resulting dataset (all-SCO) and two subsets of the data were used for analyses. The two subsets contained SCOs from the all-SCO dataset that are also in the BUSCO fungi (fungi-SCO) or ascomycete (asco-SCO) odb9 databases. Phylogenetic analyses were performed with maximum likelihood (ML) in IQ-TREE v1.6.8 (69). For the all-SCO dataset, a guide tree inferred under the X+G+C60 model (X = best-fit model inferred by BIC in ModelFinder (70); for +C60 see (71)). The guide tree and best-fit model from that analysis were used to infer the final tree with the posterior mean site frequency (PMSF) model (71) which better accounts for heterogeneity of amino acid frequencies across sites. Branch support was evaluated using ultrafast bootstrap (UFBoot, (72)) and SH-like approximate likelihood ratio test (SH-aLRT), each with 1000 replicates. The fungi-SCO and asco-SCO datasets were both analyzed using the best-fit models identified by ModelFinder and support estimated using 1000 UFBoot and SH-aLRT replicates. The fungi-SCO dataset was additionally analyzed using the FreeRate model (X+R5). Lifestyle, ecology and other designations for the fungi in this study are based on the actual sample sequenced, this can differ from the generalized biology of the species (Tables S1, S2). Trees were visualized and annotated using iTOL (73) with *Tuber borchii* as the root.

ITS phylogeny

ITS sequences including the term *Tetracladium* were retrieved from GenBank and aligned in Geneious 10.2.4 (http://www.geneious.com). Sequences containing data spanning the ITS region (ITS1, 5.8s, and ITS2) were trimmed using BMGE (DNAPAM1 matrix). Using a distance matrix produced in Geneious, sequences with identical bases and alignments were identified and removed from the dataset. Sequences with excessive uncalled bases (>30) or poor alignments were also removed. The strains sequenced herein may be doubly represented in the dataset if not identified as duplicates in this process. Information about each sequence was manually searched to determine the origin and lifestyle when possible. The ITS-region ML phylogeny was determined in IQTree with

the best-fit model and branch support estimated using 1000 non-parametric bootstrap replicates and visualized with *Penicillium antarcticum* as the root.

Carbohydrate active enzymes (CAZymes) and secondary metabolites

To identify genes involved in recognition, metabolism, and synthesis of complex carbohydrates, the proteins annotated in MAKER for each strain were compared against existing CAZyme-related databases using the dbCAN2 meta server (74) (accessed online at http://bcb.unl.edu/dbCAN2, September-October 2019) with HMMER, DIAMOND, and Hotpep and the default settings. Only CAZymes predicted by at least two tools were used in further analyses. To count the domains identified, each prediction identifier for multi-family/multi-domain genes were counted separately not together as a new combined category. Secondary metabolite clusters were identified in genomes using the antiSMASH 5.0 fungal version public web server (75) (accessed online at https://fungismash.secondarymetabolites.org, August-September 2019) including use of the *KnownClusterBlast* function. Cluster analysis and visualization were done in R (76) on scaled (*scale*) count data using *pheatmap*. Differences in numbers and types of predicted CAZymes and SM among the genus *Tetracladium*, other Leotiomycetes, and the remaining taxa were evaluated using ANOVA in R, with taxon-group as the main effect, and post hoc analyses with Tukey's HSD ($p \leq 0.05$) as implemented in *agricolae*.

RESULTS

Strains

Four strains from this study have been deposited in the CCM culture collection and will be publicly available after peer reviewed publication; JA005 (CCM 9013), JA001 (CCM 9012), AFCN889 (CCM 9011), and AFCN900 (CCM 9010). *Tetracladium sp.* 82A210 failed to revive from preserved and long term cultures at the time the cultures were deposited. Efforts to revive this strain are ongoing and it will be deposited in the CCM if successful.

Genomes

Twenty-four *Tetracladium* and *Tetracladium*-like strains, including all described species, were sequenced and the resulting raw data deposited as ENA Study PRJEB36440 (available after peer review). Data for each genome, including contigs, trimming, assembly, and coverage statistics as well as sample and read number details are provided in Table S3. The genomes assembled into 266 – 6,765 scaffolds (median = 378). Scaffold N50 values ranged from 185,943 – 2,082,751 bp (median = 894,004 bp). Total assembled genome sizes ranged from 34.3 – 43.6 Gb. Genome-wide average GC

content in the assemblies was 46.13%. These values exclude scaffolds smaller than 500 bp. The *Tetracladium sp.* T11Df assembly had the lowest depth of coverage (69X) and had more than 5-times more scaffolds than the next highest, T. *maxilliforme* CCM F-529 with 1032. Depth of coverage ranged from 69 – 178X (average = 128X). These high depths of coverage were not targeted, rather they are the by-product of small genome size and high sequencer output; the 24 barcoded genomes were pooled and sequenced in one lane on an Illumina HiSeqX Ten. All genomes are highly complete as assessed in BUSCO. Of the 3,156 genes included in the Pezizomycotina odb9 database, 20 - 34 were missing from the genome of any strain. The genomes had 98.3 – 98.7% complete BUSCOs, and 97.8 – 98.5% were complete and single copy.

Annotations and Orthologs

The number of annotated gene models ranged from 7769 in *T. sp.* T11Df to 9799 in *T. sp.* 82A210 (mean 8768, median 8694; Table S3). To predict gene function, the annotated genes for each strain were compared to the UniProtKB/Swiss-Prot annotated non-redundant protein database. These analyses returned similarity-based information for an average of 6810 (median 6759) predicted genes per strain. An average of 8681 protein signatures per strain (median 8613) were identified by comparing the annotated gene models to the InterProScan database. The gff files containing all annotation data will be available after peer review. A total of 1820 SCO present in all 51 genomes (all-SCO) included in this study were identified from the output of OrthoMCL.

Phylogeny

The all-SCO dataset contained 794,205 aligned amino acid sites, 434,869 were parsimony-informative (Table S4). Analysis using the PMSF model starting with the best fit model LG+F+I+G4 and a guide tree (Fig. S1) as input, resulted in a tree with all branches fully supported (100% UFBoot and SH-aLRT)(Fig 2). All known and putative *Tetracladium* strains in this analysis form a monophyletic clade with three distinct groups. Group A contains all *Tetracladium* species with typical Tetracladium conidia, including *T. sp.* F-10008 which also produces conidia typical for Tetracladium when submerged (19). *Tetracladium sp.* 82A210, conidia yet unknown, was isolated as an endophyte of wheat and is resolved within *T. furcatum*. Group B includes the moss endobionts from the Antarctic, *T. sp*. AFCN889 and *T. sp*. AFCN900, which appear to be conspecific and the three soil species from the Tibet Plateau, *T. ellipsoideum*, *T. psychrophilum*, and *T. globosum*. Group C is monotypic, containing only the Antarctic yeast *T. sp.* T11Df. Only two branches in the initial guide tree were not supported (UFBoot values <95% and SH-aLRT <80%), one within Group A, separating *T.*

furcatum from *T. maxilliforme*, and one within the Leotiomycetes separating *Phialocephala scopiformis* and *Articulospora tetracladia* (Fig. S1).

To confirm the robustness of the phylogenetic relationships inferred from analyses of the large all-SCO dataset additional analyses were performed using subsets of the data with more complex substitution models than were computationally feasible for full dataset. The asco-SCO dataset contained 182,311 parsimony informative amino acid sites (339,460 total) from 709 SCOs (Table S4). ModelFinder identified LG+F+R6 as the best fit model. The resulting tree had a highly similar topology to the all-SCO result, with uncertain resolution for *T. furcatum/T. maxilliforme* and *P. scopiformis/A. tetracladia* (Fig. S2).

The fungi-SCO dataset was the smallest, with 35,673 parsimony informative sites (72,086 total) from 164 SCOs (Table S4). The best-fit model for the data was LG+F+R4. When analyzed using this model there was again poor resolution for *T. furcatum/T. maxilliforme* and *P. scopiformis/A. tetracladia* (Fig. S3). There was also insufficient support for the branch to *T. palmatum* and *T. setigerum*. Using the empirical mixture model (C60), all branches that were unsupported with the simpler LG+F+R4 model met the threshold for support, but the branch separating *Coleophoma cylindrospora* from *Scytalidium lignicola* was not supported (Fig. S3). Overall the phylogenies obtained in these analyses were highly consistent. Within the *Tetracladium* and *Tetracladium*-like clade, only the relationships of *T. furcatum* and *T. maxilliforme* to each other and the rest of Group A varied among analyses.

ITS Phylogeny

To evaluate how ecological diversity is distributed within the genus, a phylogeny including 198 ITS DNA was produced and visualized in with information about the sequence (Fig. 3, Table S5). To further highlight diversity in the genus, orchid and challenging/extreme-environment associated sequences are also indicated. The later, identified as “cold” in Fig. 3, are from high latitude, high altitude, alpine, and glacier associated locales. The analysis included 229 sequences, with seven outgroups and the 24 strains sequenced herein. Of 402 total sites, 144 were parsimony informative and the best fit model for the data was TIM3e+R3. Branch support was determined using standard non-parametric bootstrapping (not UFBoot); bootstrap values >75% are considered well-supported using this approach. Conservatively, only branches with bootstrap values >85% are highlighted in the resulting figure (Fig. 3). As expected, this ITS tree does not reflect relationships between the species of *Tetracladium,* but similar ITS sequences from different sources and lifestyles are clustered. Only one *Tetracladium*-like sequence, JX630692, is not similar to species in Groups A, B, and C. When

compared to other ascomycetes in GenBank, this sequence does not return hits to any culture identified *Tetracladium* species. Thus, the "*Tetracladium*" identity of this sequence is suspect.

CAZymes

Cluster analysis based on the predicted number of carbohydrate-binding modules (CBM) and catalytic domains of each CAZyme class by species did not reveal any discernable ecological grouping (Fig. 4, Table S6). Clustering of the majority of the *Tetracladium* species reflects phylogeny, not ecology, because the only non-*Tetracladium* aquatic hyphomycete in the study does not cluster with *Tetracladium* species and the *Tetracladium* species cluster together regardless of ecology. The exception is the Antarctic yeast, *T. sp.* T11Df, which clusters with taxonomically and ecologically diverse species.

Like other Leotiomycetes (490.8 $\pm$ 161), *Tetracladium* genomes (487.6 $\pm$ 47.9) contain more CAZyme and CBM domains than other ascomycetes (329.1 $\pm$ 102.6; Fig. 5A; $F_2 = 6.97$, $p < 0.002$; Tukey's HSD, $p \leq 0.05$). Likewise, Leotiomycetes and *Tetracladium* genomes have more predicted carbohydrate esterase (CE), glycoside hydrolase (GH), and glycosyltransferase (GT) domains than other ascomycetes (Fig. 5C; CE: $F_2 = 8.5$, $p < 0.0001$; GH: $F_2 = 5.9$, $p = 0.005$; GT: $F_2 = 6.3$, $p = 0.004$; Tukey's HSD, $p \leq 0.05$). However, *Tetracladium* genomes have fewer predicted Auxiliary Activities (AA) domains than other Leotiomycetes, and so are in line with other ascomycetes ($F_2 = 6.7$, $p = 0.0003$; Tukey's HSD, $p \leq 0.05$). Note that the taxon-groups differ in sample size(*Tetracladium* = 24, Leotiomycetes = 18, Other ascomycetes = 8) and represent diversity at different taxonomic levels. Also, only domains predicted by HMMER are presented here, summaries of all HMMER results (Table S6) and results from Hotpep and DIAMOND (Table S7) are provided. These numbers are influenced by the contents of the databases and the genome assemblies and are thus "predicted" values.

Tetracladium genomes are specifically enriched for polysaccharide lyase (PL) domains relative to other Leotiomycetes and ascomycetes (Fig. 5C; PL: $F_2 = 81.78$, $p < 0.0001$; Tukey's HSD, $p \leq 0.05$). *Tetracladium* genomes have 34 $\pm$ 4.98 (mean $\pm$ 1 standard deviation) PL domains per genome, which is about 3 times the number in other Leotiomycetes (9.95 $\pm$ 9.97), and 6 times as many as the other ascomycetes (5.5 $\pm$ 3.8). This difference in part reflects the higher copy number of pectate lyase (77) PL3-2 in the *Tetracladium* genomes (9.3 $\pm$ 1.4) than the Leotiomycetes (1.8 $\pm$ 2.2) and other ascomycetes (0.9 $\pm$ 0.8; $F_2 = 134.8$, $p < 2e\text{-}16$; Tukey's HSD, $p \leq 0.05$; Table S6).

345 *Tetracladium* genomes also contain more CBMs (24.1 ± 6.2), than other Leotiomycetes (15.9 ± 9.1)
346 and ascomycetes (10.6 ± 5.2; CBM: F_2 = 12.89, p < 0.0001; Tukey's HSD, p ≤ 0.05). Most CBMs (75-
347 91%) were associated with glycoside-hydrolases (GH). Up to three CBMs were found to co-occur
348 flanked by GH32 (GH32+CBM38+CBM38+CBM38+GH32). *Tetracladium* genomes contain more
349 copies of GH32 (4.3 ± 1.4) and CBM38 (2.5 ± 1.6) than the other taxon-groups (Leotiomycetes: GH32
350 2.6 ± 1.7, CBM38 0.8 ± 1.1; ascomycetes: GH32 2.3 ± 1.8, CBM38 0.3 ± 0.5; F_2 = 8.4, p = 0.0007;
351 Tukey's HSD, p ≤ 0.05). GH32 family enzymes can function as invertases that convert sucrose into
352 fructose and glucose and act on inulin and fructose (77) and CBM38 has inulin-binding function (77).
353 The GH32+CBM38+CBM38+CBM38+GH32 conformation is unique within the genus *Tetracladium* in
354 this study. Versions of this CAZyme with one and two CBM38 between the flanking GH32s were also
355 predicted for some Leotiomycetes (Table S8). All *Tetracladium* genomes except *T. ellipsoideum* and
356 *T. sp.* T11Df had at least one GH32+CMB38 predicted CAZyme.
357
358 Copy number of CBM24 also contributes to the difference in CBM among taxon groups; there are 1-
359 15 copies (7.6 ± 4.7) in each *Tetracladium* genome, while the Leotiomycetes (0-11, 3.4 ± 4.0) and
360 ascomycete genomes (0-5, 2 ± 1.8) contain fewer (Table S6). CBM24 has α-1,3-glucan/mutan binding
361 function (77). In *Tetracladium*, CBM24 was almost always predicted in 1-3 copies in association with
362 GH71 an α-1,3-glucanase (77). GH71 is enriched in *Tetracladium* genomes (7.7 ± 3.1) relative to the
363 other Letiomycetes (4 ± 2.4) and ascomycetes (3.9 ± 2.5; F_2 = 10.39, p = 0.0002, Tukey's HSD, p
364 ≤ 0.05) . The high copy number of GH71, and its association with 1-3 copies of CBM24 per predicted
365 CAZyme, explains the abundance of CBM24 in the *Tetracladium* genomes.
366
367 Strains of the same species have similar predicted CAZyme repertoires overall (Table S9). The same
368 copy number was predicted for 63-93% of the predicted CAZyme domains in the genomes of the the
369 six species represented by more than one strain. This includes *T. sp*. 82A210 within *T. furcatum*, and
370 *T.sp*. AFCN889 and *T. sp*. AFCN900 as one species. *Tetracladium furcatum* had the lowest similarity
371 among strains for copy number, which is consistent with the phylogenetics results; F11883 and
372 82A210 are more similar to each other than to F06983. F11883 and 82A210 are 91% identical for
373 copy number. Chromosome level assemblies are required to determine absolute numbers present in
374 the genomes. However, consistent counts between related genomes can aid interpretation overall
375 and give a first estimate of variation within species. In most cases where counts differed between
376 strains within species the count differed by ±1 (75%, Table S9). Larger differences in copy numbers
377 can be seen for specific CAZymes within species in some cases. For example, GH71 (above) is
378 predicted in 4 copies in *T. marchalianum* F26399 and in 13 copies in *T. marchalianum* JA005 (Table

S6). As expected, CBM24 copy number also differed between these two strains (3 and 10 respectively).

Secondary Metabolites

Tetracladium genomes contain 24.7 ± 3.6 SM clusters (Fig. 5B) detected by antiSMASH which is fewer than other Leotiomycetes (43 ± 17.5), but in line with other ascomycetes (31.8 ± 15; F_2 = 11.53, $p < 0.0001$; Tukey's HSD $p < 0.05$). All SM data are available in Table S9. *Tetracladium* genomes each contain one betalactone cluster (1.0 ± 0.2) , except *T. sp*. T11Df with none, setting them apart from the other taxon-groups (Fig. 5D; Leotiomycetes 0.4 ± 0.5; ascomycetes 0.5 ± 0.8; F_2 = 9.3, $p < 0.001$; Tukey's HSD $p < 0.05$). Indole clusters are predicted in only 50% of *Tetracladium* genomes and never more than 1 per genome (0.5 ± 0.5), which is fewer than in Leotiomycete genomes (1.3 ± 1.1) but not different from other ascomycetes (1.3 ± 1.2; F_2 = 5.1, $p < 0.01$; Tukey's HSD $p < 0.05$). More nonribosomal peptides (NRPS) are found in Leotiomycete (8.2 ± 4.6) and other ascomycete (6.5 ± 4.1) genomes than in *Tetracladium* genomes (3.2 ± 0.8; F_2 = 12.7, $p < 0.0001$; Tukey's HSD $p < 0.05$). All Leotiomycete genomes (9.2 ± 3.7), including *Tetracladium* (8.4 ± 1.2), contain more NRPS-like SM clusters than the other ascomycetes (5.3 ± 2.8; F_2 = 6.4, $p = 0.004$; Tukey's HSD $p < 0.05$). Fewer Type 1 polyketide synthase (T1PKS) are predicted in *Tetracladium* genomes (6.8 ± 2.0) than in leotiomycetes (17.2 ± 9.1), but neither group differs from the ascomycetes (11.6 ± 5.1; F_2 = 15.66, $p < 0.0001$; Tukey's HSD $p < 0.05$). Both Leotiomycete (5.1± 2.0) and other ascomycete genomes (5.5 ± 2.7) contain more terpene SM clusters than *Tetracladium* genomes (3.4 ± 0.5; F_2 = 7.7, $p = 0.001$; Tukey's HSD $p < 0.05$). The frequencies of ribosomally synthesized and posttranslationally modified peptides (fungal-RIPPs), phosphonates, siderophores and Type III polyketide synthases (T3PKS) clusters were low and did not differ among taxon groups (Table S10).

While *Tetracladium* genomes each contain around 24 SM clusters detectable by antiSMASH only 13% returned a BLAST match for most similar known cluster in the MiBIG database(78). In total 18 SM clusters were identified to type (Fig. S4, Table S10). A nonribosomal peptide synthetase (NPS), Dimethylcoprogen, was the only SM predicted in all *Tetracladium* genomes (100% similarity (75)). This cluster is found in only 2 of 18 (11%) of the other Leotiomycete genomes and 3 of 8 (38%) other ascomycetes. All other identified SM clusters were Group or taxon specific within *Tetracladium* (Table S10). Unique to some members of Group A: Solanapyrone, Hexadehydro-astechrome, Cytochalasin, Phyllostictine A/phyllostictine B, Citreoviridin, Clavaric acid, Aureobasidin A1 and Chaetoglobosins. Unique to Group some members of Group B: Clapurines, Naphthopyrone,

Shearinine D, PR toxin, Pyranonigrin E, Azanigerone. Unique to Group C: Duclauxin. Depudecin (Groups A and B), Brefeldin (B and C) were also predicted.

DISCUSSION

Relationships among *Tetracladium* and *Tetracladium*-like fungi

This study presents the first phylogeny with all species of *Tetracladium* described to date. The eleven described species and all newly sequenced *Tetracladium*-like (putative) strains/species form a monophyletic group with three partitions (Fig 2, Groups A, B, and C). Whereas interspecific relationships within *Tetracladium* were unresolved in analyses using 18S (79), 28S (20), and ITS+28S (80) data and few taxa, relationships within the genus are largely stable across analyses and datasets herein. Within the *Tetracladium* and *Tetracladium*-like group only the branching of *T. furcatum* and *T. maxilliforme* varied between analyses. *Tetracladium furcatum* and *T. maxilliforme* are well supported separate species, but it is unclear whether they are sister species or whether *T. furcatum* alone is sister to the rest of Group A.

The majority of species and strains in Group A originate from submerged plant debris or foam that forms on rivers and produce stereotypical *Tetracladium*-shaped conidia which are typically distinctive between species (19, 38). Although conidia have not been observed for the strain isolated as an endophyte of wheat (Sweden, 2007), *T. sp.* 82A210, this strain falls within the species *T. furcatum*, and is highly similar to strain F11883 (Czech Republic, 1983) which was isolated from foam as a typical aquatic hyphomycete (Table S1). This result supports the idea that *T. furcatum* is multifunctional, as opposed to having different ecologies for morphologically similar species. *Tetracladium furcatum* was first reported as a root endophyte in 1996, based on the morphological identification of conidia (81). It has also been associated with the endophytic lifestyle in metabarcoding studies, including studies of terrestrial plants from the high Arctic (35) and submerged aquatic plants in Norway (15). *Tetracladium nainitalense* which was isolated as endophyte of *Eupatorium adenophorum* (11), is most closely related to species isolated as aquatic hyphomycetes and has itself been isolated from foam in a stream (morphological identification (82)). It should be noted that *Tetracladium* species are found as endophytes even in non-stream, non-riparian habitats. This study also confirms the phylogenetic position of T. sp. F10008, collected as an aquatic hyphomycete from Malaysia in 2008, as the sister species of *T. apiense* (19).

In the majority of cases when more than one strain per species was sequenced the strains are resolved together as expected. Most strains in Group A were isolated by experts in aquatic

hyphomycetes (Table S1), who will have relied on spore morphology for initial identification. This suggests that morphological identification of most Group A species from field samples can be reliable. However, two species are potentially problematic. Strains historically identified as *T. marchalianum*, F12812 (called *T. sp.* F12812 herein) and F26399, are not conspecific. The same is true of two strains of *T. breve* (F12505 and F10501). Rather, the *T. marchalianum*-like strain F12812 is the sister species of the *T. breve*-like F12505, and *T. marchalianum*-like F26399 is sister to *T. breve*-like F10501, and the two pairs are divergent. This finding is consistent with previous studies based on one or few genes (19, 83). The *T. marchalianum* and *T. breve*-like strains also differ in predicted secondary metabolite profiles (Table S10), suggesting that secondary metabolite profiles might be valuable tools for species identification in the same way that protein fingerprinting is being developed (51). *Tetracladium marchalianum*-like strains can readily be catagorizied as F12812 or JA005/F26399-like using beta-tubulin sequence clustering (SI Fig. X), based on a preliminary analysis of strains identified as *T. marchalianum* from a population genetics study (84). Note, the strains from that study all cluster with JA005/F26399(Fig. S5). Both *T. marchalianum* and *T. breve* require further study and taxonomic revision. Taxonomic revision is beyond the scope of the work presented here, but is ongoing

As more fungi related to Groups B and C are discovered it is probable that these groups will be described as separate genera. In addition to being divergent from Group A in the phylogeny, the fungi in Group B have not been observed to produce the conidia typical of the genus. *Tetracladium elipsoideum* and *T. globosum* are named for their ellipsoid and globose conidia. *T. psychrophilum*, although described without conidia, does produce floating multiseptate elongated allantoid-lunate conidia when submerged (Anderson personal observation). Sporulation in the two endobryophytic strains (*T. sp*. AFCN889 and AFCN900) has not been observed. Based on strains included here, including the ITS phylogeny below, it appears that Group B may predominantly contain psychrophilic or psychrotolerant fungi. The sole representative of Group C, *T. sp*. T11Df, is particularly unusual in that it grows as a yeast, producing short pseudohyphae in culture. All other species in the genus are known in filamentous forms. Spores have not yet been observed in this species.

The family level phylogenetic relationships resolved in this study are in line with the results from previous studies (85-87). Within the Leotiomycetes, all clades identified by Johnston and colleagues (88) and were represented here, are recovered. The family containing *Tetracladium*, Vandijckellaceae (88, 89), was included in an analysis based on 15 concatenated sequences (88). The tree herein differs from that result in the branching order of the Vandijckellaceae and helotioid taxa. However, the corresponding nodes are not well supported in the Johnston tree where they receive

maximum 93% UFBoot, but are fully supported herein (100% UFBoot). The minimum support considered reliable in ultrafast bootstrapping (UFBoot) in IQ-TREE is 95% (90, 91).

Broad ecological and geographical diversity in *Tetracladium*

Tetracladium-like fungi are increasingly being reported from aquatic and terrestrial sources, including from soils and as endobionts of plants, from around the world. Identification as "*Tetracladium*-like" is frequently based only on ITS sequence data and ITS data can be useful to identify species of *Tetracladium* in at least some cases (19). Here, from analysis of the *Tetracladium*-like ITS sequences available in GenBank, it is clear that both ecological and geographical diversity are wide-spread within the genus (Fig. 3). These sequences represent diversity in *Tetracladium* from 32 countries, regions, and territories, from sea level to high elevation, and from the Arctic to the Antarctic (Table S5). It is striking how many *Tetracladium* sequences are coming from polar and alpine regions, high elevations, or are glacier associated (Fig. 3). Frequently, fungi from these extreme environments that were sequenced from soil, water or non-orchid plants are similar in ITS to those from orchids from more temperate climates. Overall, patterns of strains from different sources or lifestyles being distinct from each other are not observed. Rather, in most cases, the ecologies and lifestyles are mixed among the groups of sequences most similar based on ITS. These results suggest that multifuncationality is widespread across the genus. Further, *Tetracladium* fungi are found in association with a broad diversity of plants (Table S5).

Ecological clues from CAZymes and secondary metabolites

The CAZyme and secondary metabolite profiles of *Tetracladium* species are more similar to each other than to other taxa, regardless of ecology (Fig. 4). This suggests a strong role for phylogeny in shaping the CAZyme and SM content of genomes within the genus. In comparison to genomes of other Leotiomycetes and more distantly related ascomycetes, *Tetracladium* genomes are enriched for PL domains which are associated with degradation of pectin and pectate. *Tetracladium* genomes contain around 3-6 times more PL domains than the other fungi in the study which may be related to their ecological role as aquatic hyphomycetes. Aquatic hypyomycetes are major decomposers of leaves in streams and pectin and pectate are complex polysaccharides that are abundant in leaves. In contrast, the genome of *Xylona heveae*, a horizontally transmitted endophyte of sapwood in rubber trees, contains no predicted PL (92).

Tetracladium genomes contain a particularly large number of pectate lyase PL3-2 (EC 4.2.2.2) genes, which is a feature they share with *A. tetracladia*, the only non-*Tetracladium* aquatic hyphomycete in the study (Table S6). The PL repertories of *Tetracladium* species and *A. tetracladia* are very similar

514 overall, however they are also qualitatively similar to the two *Cadophora* species in the study (Table
515 S6). The sequenced strain of *C. malorum*, was isolated from a deep sea shrimp from a depth of 2300
516 m below sea level at a hydrothermal vent along the Mid Atlantic Ridge (93), the other *Cadophora sp.*
517 was isolated as an endophyte of *Salix rosmarinifolia*; neither is expected to decompose leaves. Thus,
518 determining relative importance of ecology and phylogeny in shaping the PL content of these
519 genomes requires genomes from additional taxonomically and ecologically diverse fungi.
520
521 Fungal saprotrophs (of plants), endophytes, and plant pathogens are typically able to degrade plant
522 cell walls in order to enter a plant host or grow through a plant substrate and to obtain nutrition
523 from plants. Thus, some overlap in the CAZyme and SM content in the genomes of fungi that exhibit
524 these ecologies should be expected especially among taxa that are multifunctional (94). PL3-2, the
525 most abundant PL in *Tetracladium* genomes (above), is best known as part of the molecular arsenal
526 of plant pathogens including *Botrytis cinerea* and is highly expressed in developing infections (*e.g.*
527 tomato (95)). Also, pectic enzymes can elicit defense responses in plants (96). In *Tetracladium*,
528 expression of PL3-2, and other PLs, may require tight context dependent control, with high
529 expression during saprotrophy to degrade leaves, but no-to-low expression when living as
530 endobionts to avoid triggering plant defenses. The Dimethylcoprogens, nonribosomal peptide
531 synthetases involved in synthesis of siderophores (97) which have iron uptake and storage functions,
532 have also been reported as common among taxa within the Pleosporales, Dothideomycetes (98) an
533 order that includes many plant pathogens. Dimethylcoprogen, and siderophores generally, have
534 been associated with pathogenicity for some fungi including the corn pathogen *Cochliobolus*
535 *heterostrophus* and wheat pathogen *Fusarium graminearum* (99). However, iron homeostasis and
536 storage have other important roles in fungi, including resistance to reactive oxygen species (97, 99)
537 and the maintenance of mutualism in endophytic fungus–plant interactions (100). Going forward,
538 comparisons expression patterns associated with these PLs and SM clusters can help dissect how
539 organisms with similar molecular "toolboxes" selectively wield these tools appropriately for
540 saprotrophs, endobionts, and pathogens.
541
542 *Tetracladium* genomes are also rich in GH71 domains which are α-1,3-glucanases and their
543 associated non-catalytic CBM24 α-1,3-glucan/mutan binding modules, suggesting that these fungi
544 are well-suited to breaking down fungal cell walls; is an important component of the cell walls of
545 filamentous fungi and dimorphic yeast. As was the case for *S. sclerotiorum* and *B. cinerea* when
546 enrichment for GH71 in these fungi was first reported, it is unknown whether the abundant α-1,3-
547 glucanases in these genomes is associated degradation of the fungi's own cell walls or those of

antagonistic fungi(101), or whether these enzymes play other roles entirely. Interestingly, α-1,3-glucan is important in the matrix of fungal and bacterial biofilms (102) which can be disrupted by glucanases (103, 104). Decomposing leaves in streams and the roots of plants are covered by microbial biofilms and it is possible that enrichment for *α*-1,3-glucanase in *Tetracladium* genomes is associated with interactions with those biofilms. Further, there is some evidence that α-1,3-glucanases prevent plant detection of *β*-1,3-glucan in in invading pathogens, negatively impacting plant defensive responses. Transgenic rice plants expressing bacterial (105) or fungal (106) α-1,3-glucanases demonstrate protection against fungal pathogens. Thus, it is also possible that these enzymes are important in plant-fungal mutualism and beneficial to plants (107, 108).

Tetracladium genomes also contain more GH32 catalytic domains and associated CBM38 modules than other taxon groups studied here. GH32 enzymes can function as invertases and also act on inulin and fructose (77). Given the inulin-binding function of CBM28 the inulinase function may be most enriched in *Tetracladium*. Inulin is a reserve carbohydrate in some plants that is stored in roots, taproots, and bulbs. Some strains/species of *Tetracladium* were found to have these components in a GH32+CBM38+CBM38+CBM38+GH32 conformation; unique among the taxa in this study. These observations suggest that *Tetracladium* species obtain nutrition from inulin when living as endobionts of plants. Studies of *Tetracladium* species as root endophytes are needed to test this hypothesis.

CONCLUSIONS

The genomes of 24 *Tetracladium* and *Tetracladium*-like fungi, including representatives of all described species, were sequenced and used to resolve relationships among the taxa and to improve our understanding of ecological and genomic diversity in this group of ecologically important, multifunctional fungi. All genome-sequenced *Tetracladium* and *Tetracladium*-like fungi in this study form a monophyletic group, which may in time be subdivided into separate genera. From analysis of ITS sequences from water, soil, and plants from around the world, it emerges that multifunctionality may be widespread throughout the genus, that many species have multifunctional lifestyles. Further, *Tetracladium* is frequently sampled from extreme and cold environments, suggesting that these fungi may have important roles in those ecosystems and also may produce secondary metabolites or enzymes of interest for industrial applications. Studies are needed to investigate the terrestrial and endiobiont roles of *Tetracladium* fungi, including of where in plant roots these fungi are found, whether they utilize inulin as a source of nutrition as endobionts, and how PL expression is controlled in saprotroph and endiobiont contexts. Lastly, these fungi are more similar to each

582 other in genome content for SMs and CAZymes than they are to other taxa, regardless of variation in
583 the ecology of the fungi, suggesting that within *Tetracladium*, broad ecological diversity and
584 multifunctionality can be achieved among taxa using highly similar genomic toolkits.
585

586 **ACKNOWLEDGEMENTS**

587 We thank Iker Irisarri and Dan Vanderpool for advice on phylogenomics and annotations. We thank
588 Anna Rosling for reagents and Doug Scofield, Ioana Onut Brännström, and Diem Nguyen for analysis
589 support. We are grateful to Monika Laichmanová and the CCM, Marcelo Baeza, Dai Hirose, and
590 Magdalena Grudzinska-Sterno for cultures. Sequencing was performed by the SNP&SEQ Technology
591 Platform in Uppsala. This research, and J. Anderson, are funded by grant #2016-03595 from
592 Vetenskapsrådet, The Swedish Research Council.
593

594 **LITERATURE CITED**

595 1. Grossart H-P, Van den Wyngaert S, Kagami M, Wurzbacher C, Cunliffe M, Rojas-Jimenez K. Fungi in
596 aquatic ecosystems. Nat. Rev. Microbiol. 2019.
597 2. Gulis V, Su R, Kuehn KA. Fungal decomposers in freshwater environments. In: Hurst CJ, editor. The
598 Structure and Function of Aquatic Microbial Communities. Cham, Switzerland: Springer;
599 2019. p. 121-55.
600 3. Bärlocher F, Kendrick B. Dynamics of Fungal Population on Leaves in a Stream. J Ecol.
601 1974;62(3):761-91.
602 4. Fisher SG, Likens GE. Energy flow in Bear Brook, New Hampshire: An integrative approach to
603 stream ecosystem metabolism. Ecological Monographs. 1973;43(4):421-39.
604 5. Bärlocher F. Reproduction and dispersal in aquatic hyphomycetes. Mycoscience. 2009;50(1):3-8.
605 6. Chauvet E, Cornut J, Sridhar KR, Selosse M-A, Bärlocher F. Beyond the water column: Aquatic
606 hyphomycetes outside their preferred habitat. Fungal Ecology. 2016;19:112-27.
607 7. Dang CK, Gessner MO, Chauvet E. Influence of conidial traits and leaf structure on attachment
608 success of aquatic hyphomycetes on leaf litter. Mycologia. 2007;99(1):24-32.
609 8. Fabre E. Changes in concentration of aquatic hyphomycete conidia in water passing through a
610 concrete pipe. Mycol Res. 1997;101:908-10.
611 9. Selosse MA, Schneider-Maunoury L, Martos F. Time to re-think fungal ecology? Fungal ecological
612 niches are often prejudged. New Phytol. 2018;217(3):968-72.
613 10. Brundrett MC. Understanding the Roles of Multifunctional Mycorrhizal and Endophytic Fungi. In:
614 Schulz BJE, Boyle CJC, Sieber TN, editors. Microbial Root Endophytes. Soil Biology, vol 9.
615 Berlin, Heidelberg: Springer; 2006. p. 281-98.
616 11. Sati SC, Arya P, Belwal M. *Tetracladium nainitalense sp nov.*, a root endophyte from Kumaun
617 Himalaya, India. Mycologia. 2009;101(5):692-5.
618 12. Selosse MA, Vohnik M, Chauvet E. Out of the rivers: are some aquatic hyphomycetes plant
619 endophytes? New Phytol. 2008;178(1):3-7.
620 13. Ghate SD, Sridhar KR. Endophytic aquatic hyphomycetes in roots of riparian tree species of two
621 Western Ghat streams. Symbiosis. 2017;71(3):233-40.
622 14. Sridhar KR, Barlocher F. Endophytic Aquatic Hyphomycetes of Roots of Spruce, Birch and Maple.
623 Mycol Res. 1992;96:305-8.
624 15. Kohout P, Sykorova Z, Ctvrtlikova M, Rydlova J, Suda J, Vohnik M, *et al.* Surprising spectra of root-
625 associated fungi in submerged aquatic plants. Fems Microbiology Ecology. 2012;80(1):216-
626 35.

16. Bessey EA. *Varicosporium elodeae* Kegel, an uncommon soil fungus. Papers of the Michigan Academy of Science Arts and Letters. 1939;35:15-7.
17. de Wildeman É. Notes mycologiques IV. Annales de Société Belge de Microscopie. 1893;17(2):35-40.
18. Webster JR, Descals E. Morphology, distribution, and ecology of conidial fungi in freshwater habitats. In: Cole GT, Kendrick WB, editors. Biology of conidial fungi. 1. Cambridge, UK: Cambridge University Press; 1981. p. 295-348.
19. Letourneau A, Seena S, Marvanova L, Barlocher F. Potential use of barcoding to identify aquatic hyphomycetes. Fungal Diversity. 2010;40(1):51-64.
20. Wang M, Jiang X, Wu W, Hao Y, Su Y, Cai L, *et al.* Psychrophilic fungi from the world's roof. Persoonia. 2015;34:100-12.
21. Nemec S. Fungi associated with strawberry root rot in Illinois. Mycopathologia et mycologia applicata. 1970;41(3):331-46.
22. Louarn S, Nawrocki A, Thorup-Kristensen K, Lund OS, Jensen ON, Collinge DB, *et al.* Proteomic changes and endophytic micromycota during storage of organically and conventionally grown carrots. Postharvest Biol Tec. 2013;76:26-33.
23. Gkarmiri K, Mahmood S, Ekblad A, Alstrom S, Hogberg N, Finlay R. Identifying the active microbiome associated with roots and rhizosphere soil of oilseed rape. Appl Environ Microb. 2017;83(22).
24. Wei XM, Wang XY, Cao P, Gao ZT, Chen AJ, Han JP. Microbial community changes in the rhizosphere soil of healthy and rusty *Panax ginseng* and discovery of pivotal fungal genera associated with rusty roots. Biomed Res Int. 2020;2020.
25. Scherwinski K, Grosch R, Berg G. Effect of bacterial antagonists on lettuce: active biocontrol of *Rhizoctonia solani* and negligible, short-term effects on nontarget microorganisms. FEMS Microbiol Ecol. 2008;64(1):106-16.
26. Friberg H, Persson P, Jensen DF, Bergkvist G. Preceding crop and tillage system affect winter survival of wheat and the fungal communities on young wheat roots and in soil. Fems Microbiology Letters. 2019;366(15).
27. Grudzinska-Sterno M, Yuen J, Stenlid J, Djurle A. Fungal communities in organically grown winter wheat affected by plant organ and development stage. Eur J Plant Pathol. 2016;146(2):401-17.
28. Park MS, Eimes JA, Oh SH, Suh HJ, Oh SY, Lee S, *et al.* Diversity of fungi associated with roots of *Calanthe* orchid species in Korea. J Microbiol. 2018;56(1):49-55.
29. Hirose D, Hobara S, Matsuoka S, Kato K, Tanabe Y, Uchida M, *et al.* Diversity and community assembly of moss-associated fungi in ice-free coastal outcrops of continental Antarctica. Fungal Ecology. 2016;24:94-101.
30. Liang HB, Wang XW, Yan JW, Luo LX. Characterizing the intra-vineyard variation of soil bacterial and fungal communities (vol 10, 1239, 2019). Front Microbiol. 2019;10.
31. Dresch P, Falbesoner J, Ennemoser C, Hittorf M, Kuhnert R, Peintner U. Emerging from the ice-fungal communities are diverse and dynamic in earliest soil developmental stages of a receding glacier. Environ Microbiol. 2019;21(5):1864-80.
32. Perini L, Gostincar C, Gunde-Cimerman N. Fungal and bacterial diversity of Svalbard subglacial ice. Sci. Rep. 2019;9:20230.
33. Carrasco M, Rozas JM, Barahona S, Alcaino J, Cifuentes V, Baeza M. Diversity and extracellular enzymatic activities of yeasts isolated from King George Island, the sub-Antarctic region. Bmc Microbiology. 2012;12:251.
34. Praeg N, Pauli H, Illmer P. tMicrobial diversity in bulk and rhizosphere soil of *Ranunculus glacialis* along a high-Alpine altitudinal gradien. Front Microbiol. 2019;10:1429.
35. Zhang T, Yao YF. Endophytic fungal communities associated with vascular plants in the High Arctic Zone are highly diverse and host-plant specific. Plos One. 2015;10(6).

677 36. Durán P, Barra PJ, Jorquera MA, Viscardi S, Fernandez C, Paz C, *et al.* Occurrence of soil fungi in
678 Antarctic pristine environments. Front Bioeng Biotech. 2019;7:28.
679 37. Tiquia-Arashiro SM, Grube M, editors. Fungi in Extreme Environments: Ecological Role and
680 Biotechnological Significance. Cham: Springer International Publishing; 2019.
681 38. Roldán A, Descals E, Honrubia M. Pure Culture Studies on T*etracladium*. Mycol Res. 1989;93:452-
682 65.
683 39. Abdel-Raheem AM. Laccase activity of lignicolous aquatic hyphomycetes isolated from the River
684 Nile in Egypt. Mycopathologia. 1997;139(3):145-50.
685 40. Carrasco M, Rozas JM, Alcaino J, Cifuentes V, Baeza M. Pectinase secreted by psychrotolerant
686 fungi: identification, molecular characterization and heterologous expression of a cold-
687 active polygalacturonase from *Tetracladium sp.* Microb Cell Fact. 2019;18.
688 41. Carrasco M, Alcaino J, Cifuentes V, Baeza M. Purification and characterization of a novel cold
689 adapted fungal glucoamylase. Microb Cell Fact. 2017;16:75.
690 42. Sati SC, Pant P. Evaluation of phosphate Solubilization by root endophytic aquatic Hyphomycete
691 *Tetracladium setigerum*. Symbiosis. 2019;77(2):141-5.
692 43. Arya P, Sati SC. Evaluation of endophytic aquatic hyphomycetes for their antagonistic activity
693 against pathogenic bacteria. Internat. Res. Journal of Microbiology. 2011;2(9):343-347.
694 44. Gulis VI, Stephanovich AI. Antibiotic effects of some aquatic hyphomycetes. Mycol Res.
695 1999;103:111-115.
696 45. Sharma VK, Li X-Y, Wu G-l, Bai W-X, Parmar S, White Jr JF, *et al.* Endophytic community of Pb-Zn
697 hyperaccumulator *Arabis alpina* and its role in host plants metal tolerance. 2019;437:397-
698 411.
699 46. Franke-Whittle IH, Manici LM, Insam H, Stres B. Rhizosphere bacteria and fungi associated with
700 plant growth in soils of three replanted apple orchards. Plant Soil. 2015;395(1-2):317-333.
701 47. Yim B, Nitt H, Wrede A, Jacquiod S, Sorensen SJ, Winkelmann T, *et al.* Effects of soil pre-
702 treatment with Basamid (R) granules, *Brassica juncea*, *Raphanus sativus*, and *Tagetes patula*
703 on bacterial and fungal communities at two apple replant disease sites. Front Microbiol.
704 2017;8:1604.
705 48. Spiteller P. Chemical ecology of fungi. Natural Product Reports. 2015;32(7):971-993.
706 49. Knapp DG, Nemeth JB, Barry K, Hainaut M, Henrissat B, Johnson J, *et al.* Comparative genomics
707 provides insights into the lifestyle and reveals functional heterogeneity of dark septate
708 endophytic fungi. Sci Rep. 2018;8:6321.
709 50. Anderson JL, Beever J, Shearer CA. Eight polymorphic microsatellite loci for the aquatic fungus
710 *Tetracladium marchalianum*. Mol Ecol Notes. 2006;6(3):703-705.
711 51. Cornut J, De Respinis S, Tonolla M, Petrini O, Bärlocher F, Chauvet E, *et al.* Rapid characterization
712 of aquatic hyphomycetes by matrix-assisted laser desorption/ionization time-of-flight mass
713 spectrometry. Mycologia. 2019;111(1):177-89.
714 52. Baschien C, Manz W, Neu TR, Marvanová L, Szewzyk U. In situ detection of freshwater fungi in an
715 alpine stream by new taxon-specific fluorescence in situ hybridization probes. Appl Environ
716 Microb. 2008;74(20):6427-36.
717 53. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data.
718 Bioinformatics. 2014;30(15):2114-20.
719 54. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, *et al.*, editors.
720 Assembling Genomes and Mini-metagenomes from Highly Chimeric Reads2013; Berlin,
721 Heidelberg: Springer Berlin Heidelberg.
722 55. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome
723 assemblies. Bioinformatics. 2013;29(8):1072-5.
724 56. Simao FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome
725 assembly and annotation completeness with single-copy orthologs. Bioinformatics.
726 2015;31(19):3210-2.

727 57. Okonechnikov K, Conesa A, García-Alcalde F. Qualimap 2: advanced multi-sample quality control
728 for high- throughput sequencing data. Bioinformatics. 2015;32(2):292-4.
729 58. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform.
730 Bioinformatics. 2009;25(14):1754-60.
731 59. Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, *et al.* MAKER: An easy-to-use annotation
732 pipeline designed for emerging model organism genomes. Genome Research.
733 2008;18(1):188-96.
734 60. Smit AFA, Hubley R, Green P. RepeatMasker. Available from: http://repeatmasker.org.
735 61. Ter-Hovhannisyan V, Lomsadze A, Chernoff YO, Borodovsky M. Gene prediction in novel fungal
736 genomes using an *ab initio* algorithm with unsupervised training. Genome Research.
737 2008;18(12):1979-90.
738 62. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, *et al.* BLAST plus :
739 architecture and applications. Bmc Bioinformatics. 2009;10.
740 63. Bateman A, Martin MJ, Orchard S, Magrane M, Alpi E, Bely B, *et al.* UniProt: a worldwide hub of
741 protein knowledge. Nucleic Acids Res. 2019;47(D1):D506-D15.
742 64. Jones P, Binns D, Chang HY, Fraser M, Li WZ, McAnulla C, *et al.* InterProScan 5: genome-scale
743 protein function classification. Bioinformatics. 2014;30(9):1236-40.
744 65. Li L, Stoeckert CJ, Roos DS. OrthoMCL: Identification of ortholog groups for eukaryotic
745 genomes. Genome Research. 2003;13(9):2178-89.
746 66. Whelan S, Irisarri I, Burki F. PREQUAL: detecting non-homologous characters in sets of unaligned
747 homologous sequences. Bioinformatics. 2018;34(22):3929-30.
748 67. Katoh K, Standley DM. A simple method to control over-alignment in the MAFFT multiple
749 sequence alignment program. Bioinformatics. 2016;32(13):1933-42.
750 68. Criscuolo A, Gribaldo S. BMGE (Block Mapping and Gathering with Entropy): a new software for
751 selection of phylogenetic informative regions from multiple sequence alignments. Bmc
752 Evolutionary Biology. 2010;10:210.
753 69. Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: A fast and effective stochastic
754 algorithm for estimating maximum-likelihood phylogenies. MBE. 2015;32(1):268-74.
755 70. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast model
756 selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-589.
757 71. Wang HC, Minh Q, Susko E, Roger AJ. Modeling site heterogeneity with posterior mean site
758 frequency profiles accelerates accurate phylogenomic estimation. Syst Biol. 2018;67(2):216-
759 35.
760 72. Hoang DT, Chernomor O, von Haeseler A, Minh BQ, Vinh LS. UFBoot2: Improving the Ultrafast
761 Bootstrap approximation. MBE. 2018;35(2):518-22.
762 73. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments.
763 Nucleic Acids Res. 2019;47(W1):W256-W9.
764 74. Zhang H, Yohe T, Huang L, Entwistle S, Wu PZ, Yang ZL, *et al.* dbCAN2: a meta server for
765 automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2018;46(W1):W95-
766 W101.
767 75. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, Lee SY, *et al.* antiSMASH 5.0: updates to the
768 secondary metabolite genome mining pipeline. Nucleic Acids Res. 2019;47(W1):W81-W7.
769 76. Team RC. R: A language and environment for statistical computing Vienna, Austria: R Foundation
770 for Statistical Computing; 2020 [Available from: https://www.R-project.org.
771 77. Lombard V, Ramulu HG, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes
772 database (CAZy) in 2013. Nucleic Acids Res. 2014;42(D1):D490-D5.
773 78. Kautsar SA, Blin K, Shaw S, Navarro-Muñoz JC, Terlouw BR, van der Hooft JJJ, *et al.* MIBiG 2.0: a
774 repository for biosynthetic gene clusters of known function. Nucleic Acids Res.
775 2019;48(D1):D454-D8.
776 79. Nikolcheva L, Bärlocher F. Phylogeny of *Tetracladium* based on 18S rDNA. Czech Mycol.
777 2002;53(4):285-95.

778 80. Baschien C, Tsui CK-M, Gulis V, Szewzyk U, Marvanová L. The molecular phylogeny of aquatic
779 hyphomycetes with affinity to the Leotiomycetes. Fungal biology. 2013;117(9):660-72.
780 81. Seetharam R, Sridhar K, Bärlocher F. Endophytic aquatic hyphomycetes of roots of plantation
781 crops and ferms from India. Sydowia. 1996;48:152-60.
782 82. Fiuza PO, Gusmao LFP. Ingoldian fungi from the semi-arid Caatinga biome of Brazil. Mycosphere.
783 2013;4(6):1133-50.
784 83. Baschien C, Marvanová L, Szewzyk U. Phylogeny of selected aquatic hyphomycetes based on
785 morphological and molecular data. Nova Hedwigia. 2006;83(3-4):311-52.
786 84. Anderson JL, Shearer CA. Population genetics of the aquatic fungus *Tetracladium marchalianum*
787 over space and time. Plos One. 2011;6(1):e15908.
788 85. Carbone I, White JB, Miadlikowska J, Arnold AE, Miller MA, Kauff F, *et al.* T-BAS: Tree-Based
789 alignment selector toolkit for phylogenetic-based placement, alignment downloads and
790 metadata visualization: an example with the Pezizomycotina tree of life. Bioinformatics.
791 2017;33(8):1160-8.
792 86. Gazis R, Miadlikowska J, Lutzoni F, Arnold AE, Chaverri P. Culture-based study of endophytes
793 associated with rubber trees in Peru reveals a new class of Pezizomycotina: Xylonomycetes.
794 Mol Phylogenet Evol. 2012;65(1):294-304.
795 87. Schoch CL, Sung GH, Lopez-Giraldez F, Townsend JP, Miadlikowska J, Hofstetter V, *et al.* The
796 Ascomycota Tree of Life: A phylum-wide phylogeny clarifies the origin and evolution of
797 fundamental reproductive and ecological traits. Syst Biol. 2009;58(2):224-39.
798 88. Johnston PR, Quijada L, Smith CA, Baral HO, Hosoya T, Baschien C, *et al.* A multigene phylogeny
799 toward a new phylogenetic classification of Leotiomycetes. IMA Fungus. 2019;10:1.
800 89. Untereiner WA, Yue Q, Chen L, Li Y, Bills GF, Stepanek V, *et al. Phialophora* section Catenulatae
801 disassembled: New genera, species, and combinations and a new family encompassing taxa
802 with cleistothecial ascomata and phialidic asexual states. Mycologia. 2019;111(6):998-1027.
803 90. Minh BQ, Nguyen MAT, von Haeseler A. Ultrafast Approximation for Phylogenetic Bootstrap.
804 MBE. 2013;30(5):1188-95.
805 91. iqtree.org. How do I interpret ultrafast bootstrap (UFBoot) support values? [updated Nov 27,
806 2018]. Available from: http://www.iqtree.org/.
807 92. Gazis R, Kuo A, Riley R, Labutti K, Lipzen A, Lin JY, *et al.* The genome of *Xylona heveae* provides a
808 window into fungal endophytism. Fungal Biology. 2016;120(1):26-42.
809 93. Burgaud G, Le Calvez T, Arzur D, Vandenkoornhuyse P, Barbier G. Diversity of culturable marine
810 filamentous fungi from deep-sea hydrothermal vents. Environ Microbiol. 2009;11(6):1588-
811 600.
812 94. Martino E, Morin E, Grelet GA, Kuo A, Kohler A, Daghino S, *et al.* Comparative genomics and
813 transcriptomics depict ericoid mycorrhizal fungi as versatile saprotrophs and plant
814 mutualists. New Phytol. 2018;217(3):1213-29.
815 95. Petrasch S, Silva CJ, Mesquida-Pesci SD, Gallegos K, van den Abeele C, Pepin V, *et al.* Infection
816 strategies deployed by *Botrytis cinerea*, *Fusarium acuminatum*, and *Rhizopus stolonifer* as a
817 function of tomato fruit ripening stage. Front Plant Sci. 2019;10:223.
818 96. Yang YK, Zhang Y, Li BB, Yang XF, Dong YJ, Qiu DW. A *Verticillium dahliae* pectate lyase induces
819 plant immune responses and contributes to virulence. Front Plant Sci. 2018;9:1271.
820 97. Chen LH, Lin CH, Chung KR. A nonribosomal peptide synthetase mediates siderophore
821 production and virulence in the citrus fungal pathogen *Alternaria alternata*. Mol Plant
822 Pathol. 2013;14(5):497-505.
823 98. Gluck-Thaler E, Haridas S, Binder M, Grigoriev IV, Crous PW, Spatafora JW, *et al.* The architecture
824 of metabolism maximizes biosynthetic diversity in the largest class of fungi. bioRxiv.
825 2020:2020.01.31.928846.
826 99. Oide S, Moeder W, Krasnoff S, Gibson D, Haas H, Yoshioka K, *et al.* NPS6, encoding a
827 nonribosomal peptide synthetase involved in siderophore-mediated iron metabolism, is a

conserved virulence determinant of plant pathogenic ascomycetes. Plant Cell. 2006;18(10):2836-53.

100. Johnson LJ, Koulman A, Christensen M, Lane GA, Fraser K, Forester N, *et al.* An extracellular siderophore is required to maintain the mutualistic interaction of *Epichloe festucae* with *Lolium perenne*. Plos Pathog. 2013;9(5):e1003332.

101. Amselem J, Cuomo CA, van Kan JAL, Viaud M, Benito EP, Couloux A, *et al.* Genomic analysis of the necrotrophic fungal pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. Plos Genet. 2011;7(8) e1002230.

102. Mitchell KF, Zarnowski R, Andes DR. Fungal super flue: the biofilm matrix and its composition, assembly, and functions. Plos Pathog. 2016;12(9):e1005828.

103. Gulati M, Nobile CJ. Candida albicans biofilms: development, regulation, and molecular mechanisms. Microbes Infect. 2016;18(5):310-21.

104. Honma K, Ruscitto A, Sharma A. β-glucanase activity of the oral bacterium *Tannerella forsythia* contributes to the frowth of a partner species, *Fusobacterium nucleatum*, in cobiofilms. Appl Environ Microb. 2018;84(1):e01759-17.

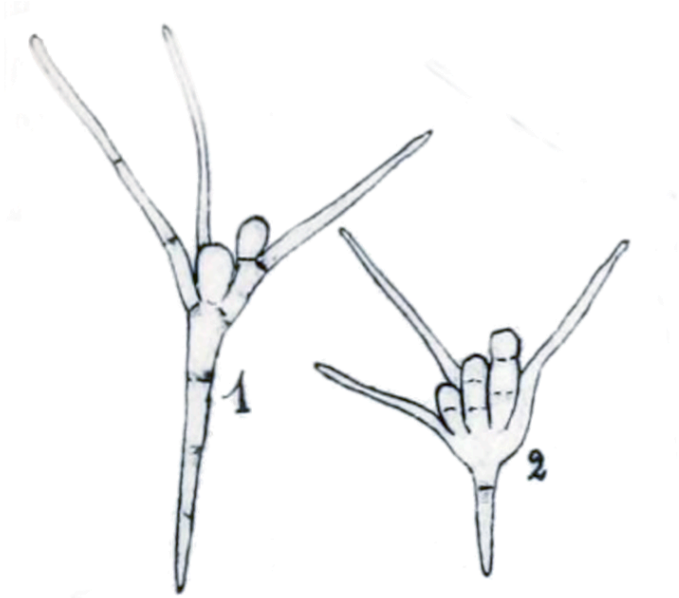
105. Fujikawa T, Sakaguchi A, Nishizawa Y, Kouzai Y, Minami E, Yano S, *et al.* Surface α-1,3-glucan facilitates fungal stealth infection by interfering with innate immunity in plants. Plos Pathog. 2012;8(8)e1002882.

106. Kumar R, Kumari K, Hembram KC, Kandha L, Bindhani BK. Expression of an endo α-1, 3-Glucanase gene from *Trichoderma harzianum* in rice induces resistance against sheath blight. J. of Plant Biochem. and Biotech. 2019;28(1):84-90.

107. Busby PE, Ridout M, Newcombe G. Fungal endophytes: modifiers of plant disease. Plant Mol Biol. 2016;90(6):645-55.

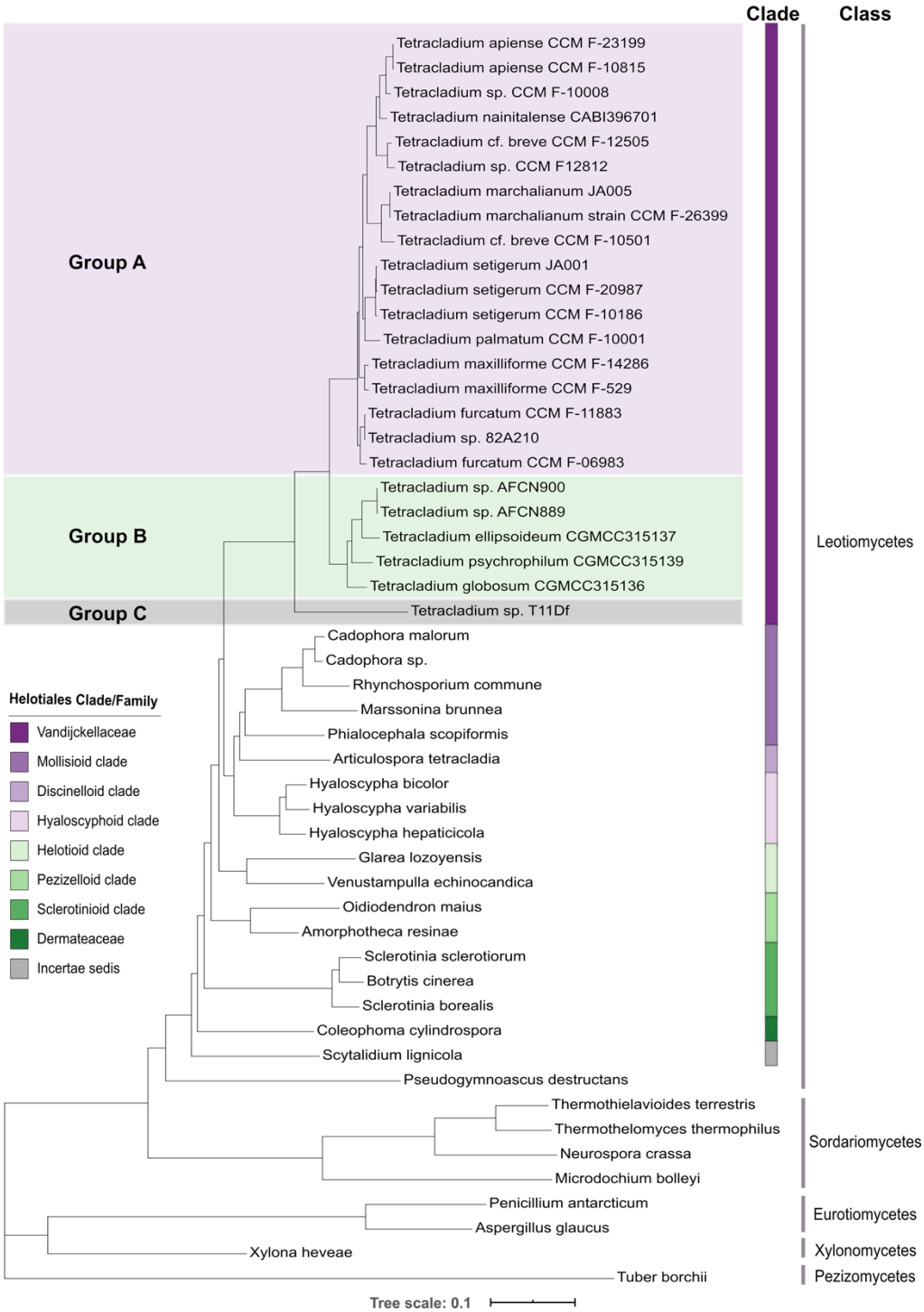
108. Masih EI, Paul B. Secretion of β-1,3-glucanases by the yeast *Pichia membranifaciens* and its possible role in the biocontrol of *Botrytis cinerea* causing grey mold disease of the grapevine. Current Microbiol. 2002;44(6):391-5.

856 **FIGURE 1**
857
858
859



860
861
862
863 **Figure 1**: Conidia of *Tetracladium marchalianum* (1) and *T. cf. breve*(2), both originally described as
864 *T. marchalianum*. Illustrations by de Wildeman (17)(b.1866-d.1947). Modified from plate IV(17). For
865 reference, from the base of conidia 1 to the rounded terminus is around 35 µm.
866

867 **FIGURE 2**



868
869
870 Figure 2: Phylogeny of *Tetracladium* and *Tetracladium*-like taxa (Groups A, B, and C) within the
871 Leotiomycetes. Maximum likelihood phylogram inferred from 1820 single copy orthologs present in
872 all 51 taxa in this analysis using the PMSF model in IQ-Tree. For fungi within the order Helotiales the
873 color bar indicates the clade or family in which they belong, after Johnston *et al.* 2019 (details Table
874 S2). All branches in this tree are fully supported (100% UFBoot and SH-aLRT). The tree scale
875 represents the expected number of nucleotide substitutions per site. Phylogram drawn with *Tuber*
876 *borchii* as the root.

FIGURE 3

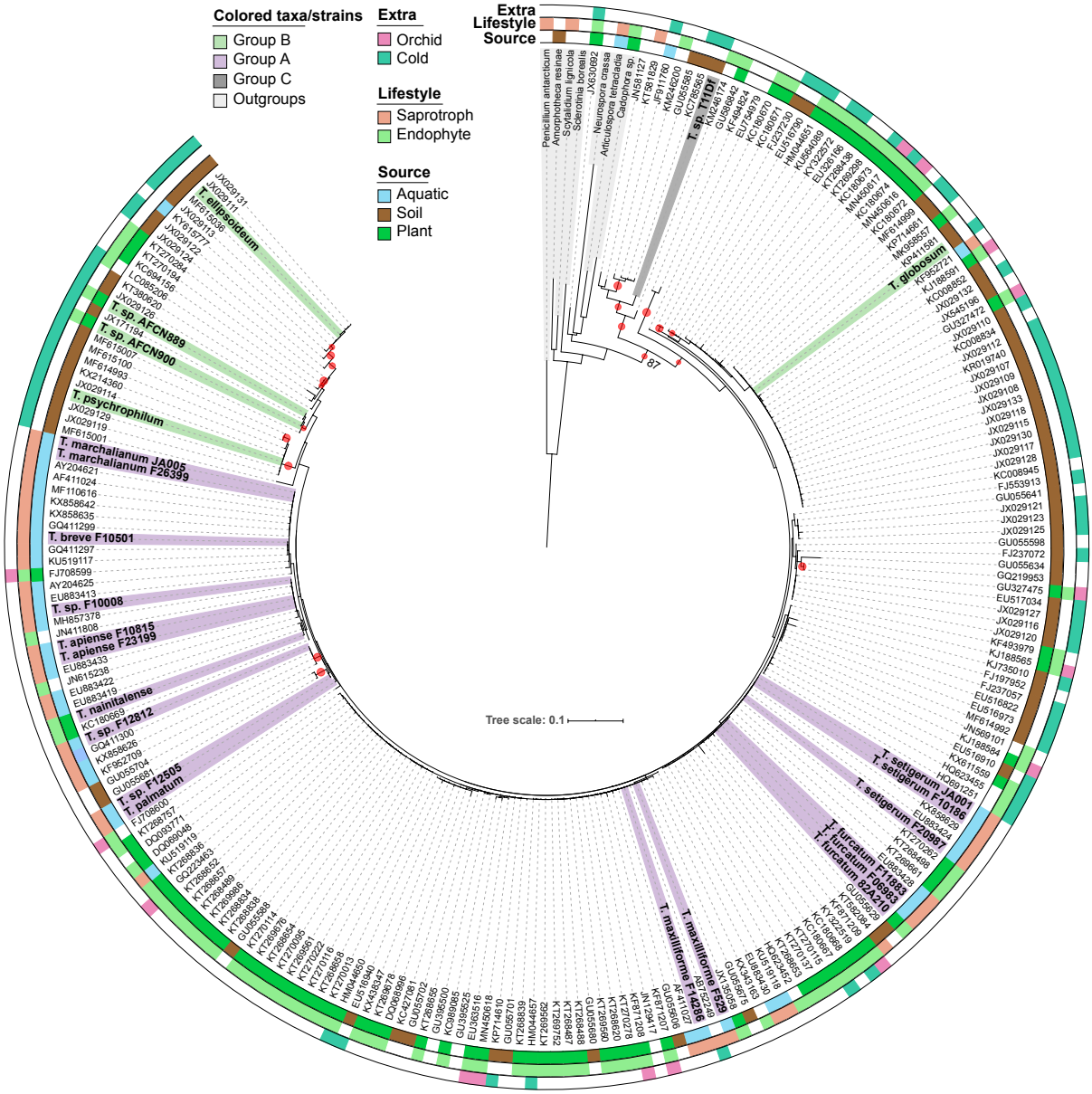


Figure 3: ITS and ecological diversity within the *Tetracladium* and *Tetracladium*-like fungi based on ITS region data and sample information for 198 accessions from GenBank, the 24 newly sequenced genomes herein (Groups A, B, and C colored as in Fig. 2), and outgroups (grey). Maximum likelihood phylogram inferred using the model TIM3e+R3. Note: Only branches receiving non-parametric bootstrap support > 85% are highlighted in this figure. Sequences identified as "cold" in the Extra category originate from high latitude, high altitude, alpine, or glacier associated locales. The tree scale represents the expected number of nucleotide substitutions per site. Phylogram drown with *Penicillium antarcticum* as the root.

FIGURE 4

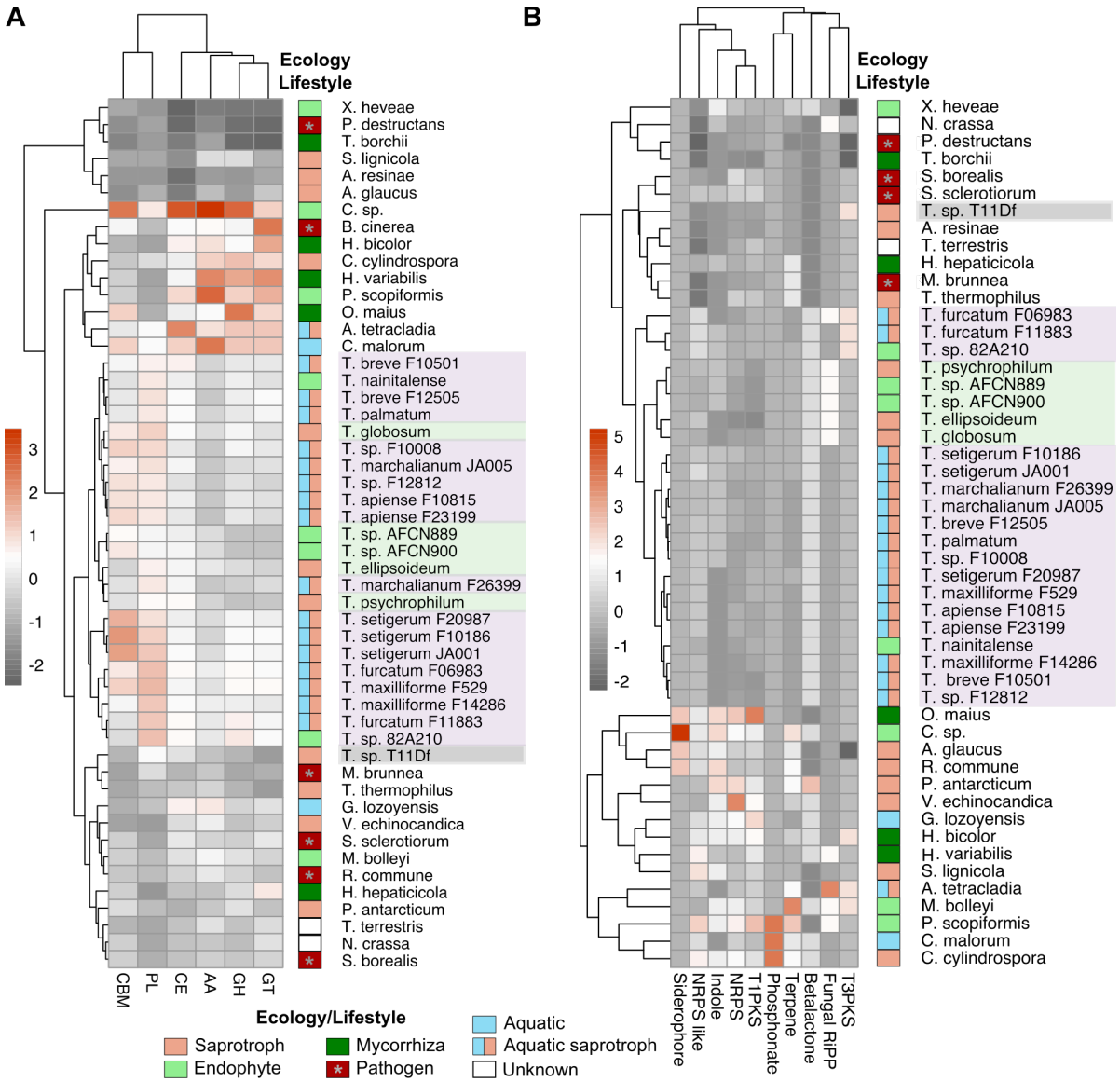


Figure 4: *Tetracladium* and *Tetracladium*-like strains cluster together in analyses of the number of CBM and CAZyme domains (A) and secondary metabolite clusters (B) of each type regardless of ecology. Heat map colors are relative to scaled datasets with orange colors representing high abundances and grey colors representing lower abundances of the domain/cluster in each genome. Note: *P. destructans* is an animal pathogen, all other pathogens infect plants. CAZyme classes and associated modules: Auxiliary Activities (AA), carbohydrate binding modules (CBM), carbohydrate esterase (CE), glycoside hydrolase (GH), glycosyltransferase (GT), and polysaccharide lyase (PL).

FIGURE 5

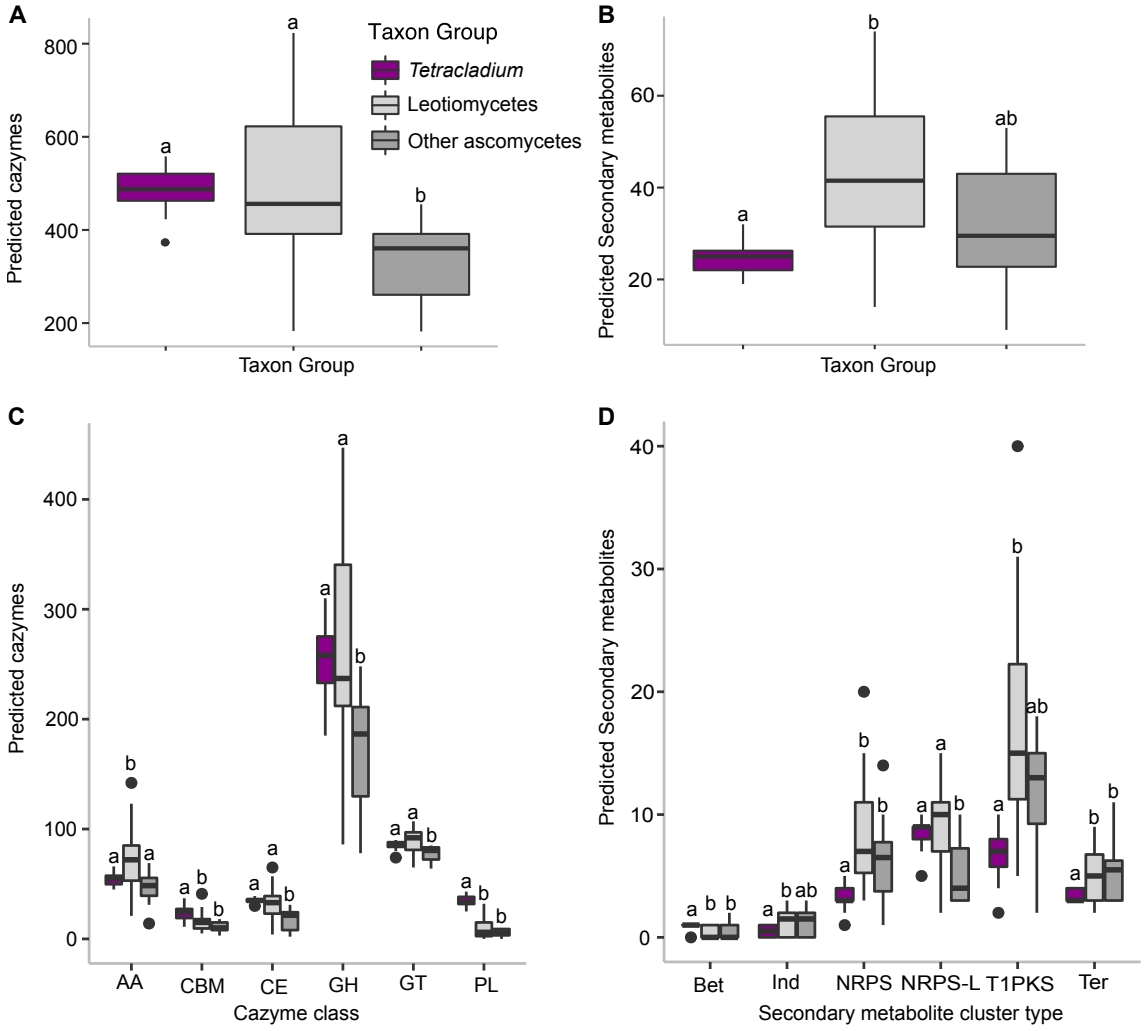


Figure 5: Variation in total number of CAZyme and CBM domains (A), and SM clusters (B) identified in taxon groups containing *Tetracladium* and *Tetracladium*-like genomes (N=24), and the genomes of the other Leotiomycete (N=18) or other ascomycetes (N=8) in the study. Results were analyzed using analysis of variance with Tukey's HSD for comparisons among taxon-groups. In C and D, separate analyses were performed for each CAZyme or SM type. Taxon-groups sharing the same letter are not statistically different ($p > 0.05$). CAZyme classes and associated modules: Auxiliary Activities (AA), carbohydrate binding modules (CBM), carbohydrate esterase (CE), glycoside hydrolase (GH), glycosyltransferase (GT), and polysaccharide lyase (PL). SM cluster types: betalactone (Bet), indole (Ind), nonribosomal peptides (NRPS), NRPS-Like (NRPS-L) and terpene (ter). Other SM categories did not differ by taxon group (see Table S10).

FIGURE S1

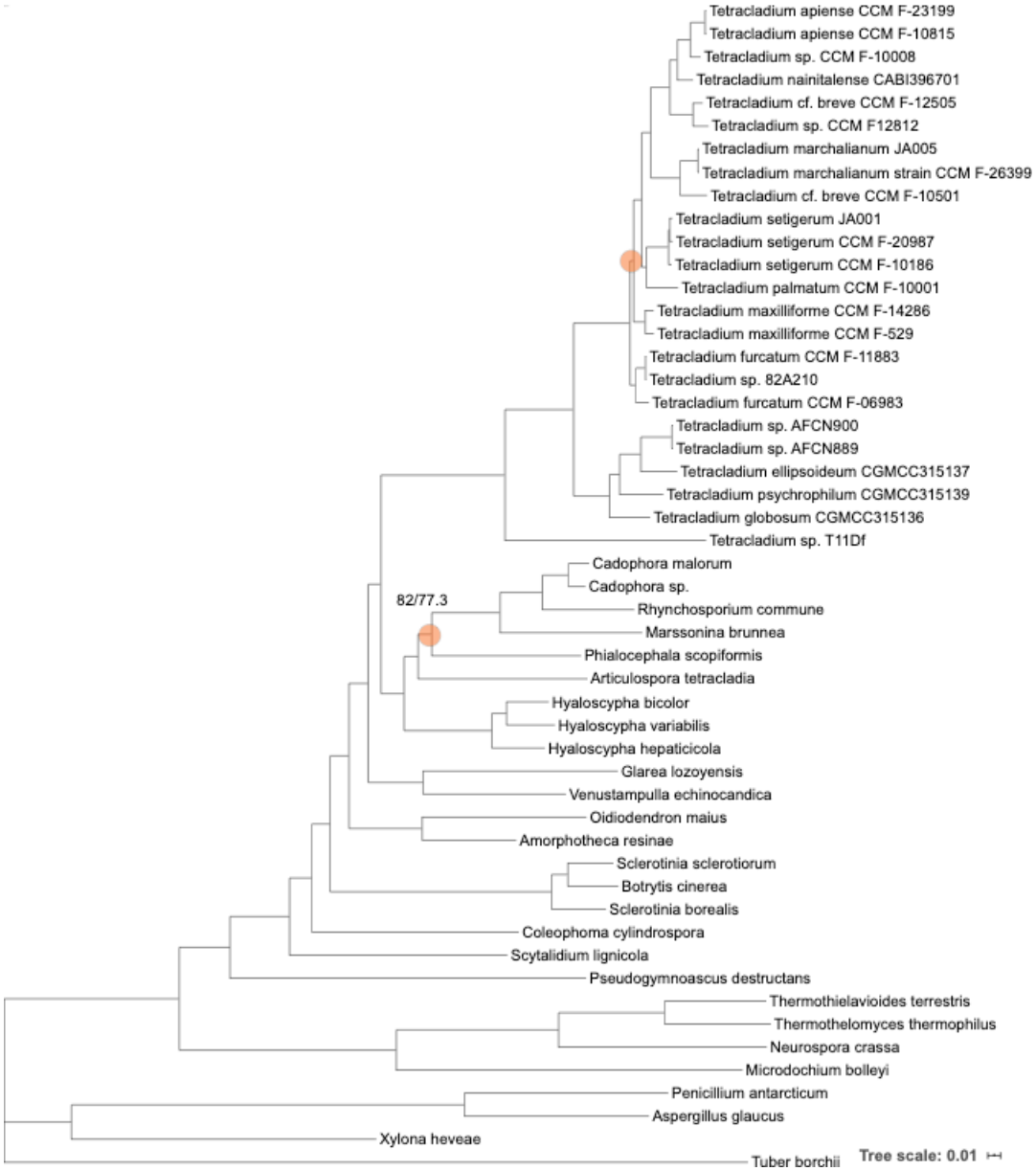


Figure S1. ML tree from analysis of the all-SCO dataset with model LG+F+I+G4. All branches received UFBoot values > 95% and SH-aLRT > 80% unless indicated (UFBoot%/ SH-aLRT%). The tree scale represents the expected number of nucleotide substitutions per site.

FIGURE S2

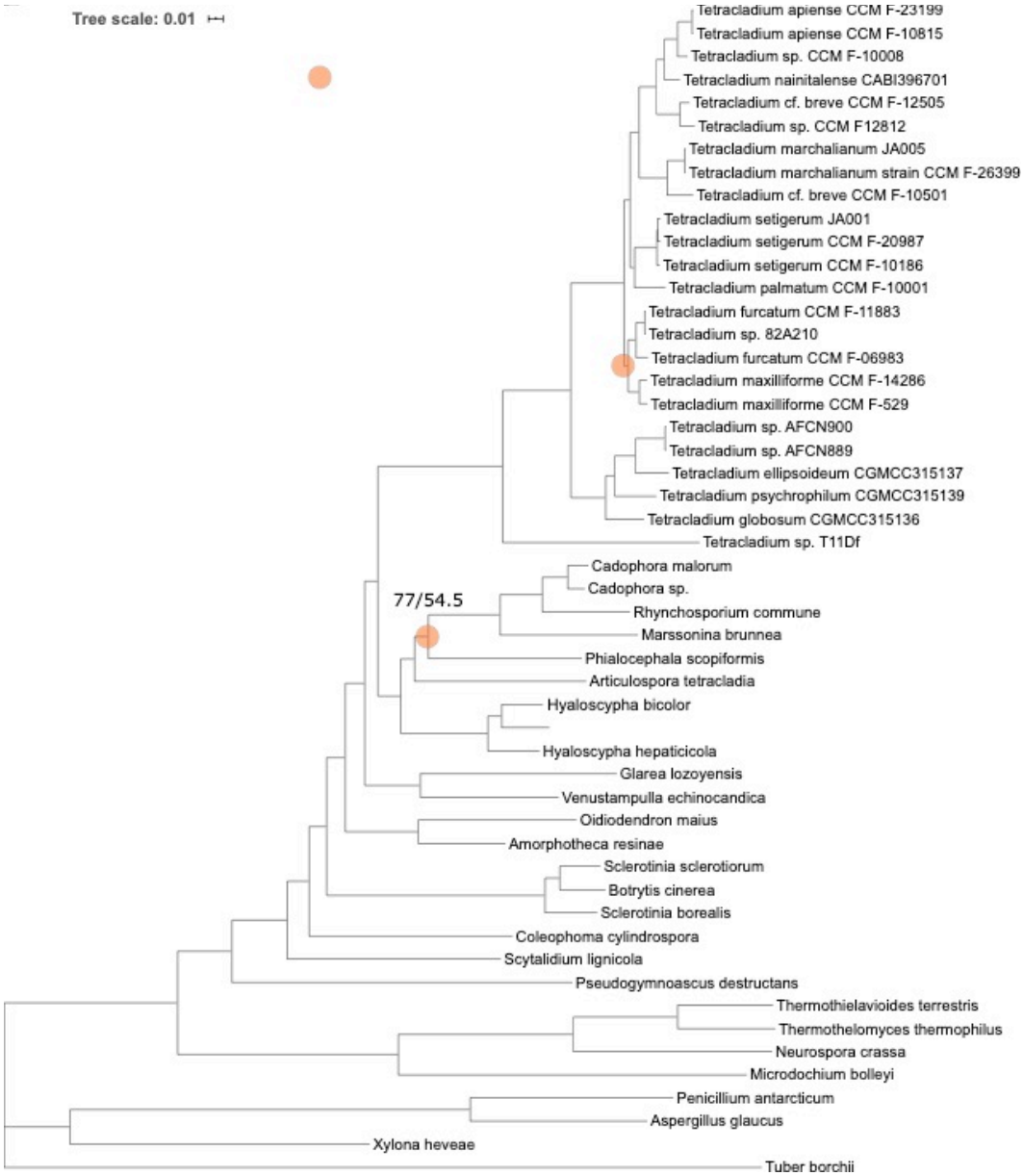


Figure S2. ML tree from analysis of the asco-SCO dataset with model LG+F+R6. All branches received UFBoot values > 95% and SH-aLRT > 80% unless indicated (UFBoot%/ SH-aLRT%). The tree scale represents the expected number of nucleotide substitutions per site.

FIGURE S3

Figure S4. Summary of ML trees from analysis of the fungi-SCO dataset model LG+F+I+G4 and using the FreeRate model. All branches received UFBoot values > 95% and SH-aLRT > 80% unless indicated. Support values are presented as UFBoot%/ SH-aLRT% for LG+F+I+G4 above, the FreeRate model below. Tree scale refers to the FreeRate model analysis. The tree scale represents the expected number of nucleotide substitutions per site.

FIGURE S4

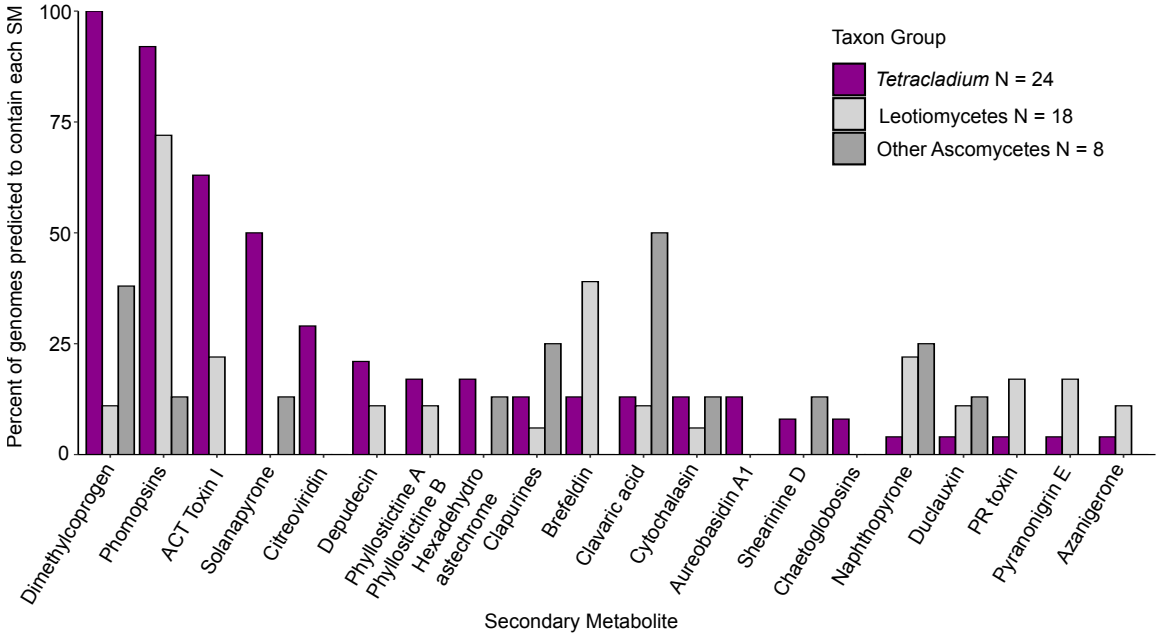


Figure S4. The 18 secondary metabolite clusters in *Tetracladium* genomes that were identifiable to type and the percent of genomes in each taxon-group where each SM type was also identified. Note the difference in sample size among taxon-groups.

FIGURE S5

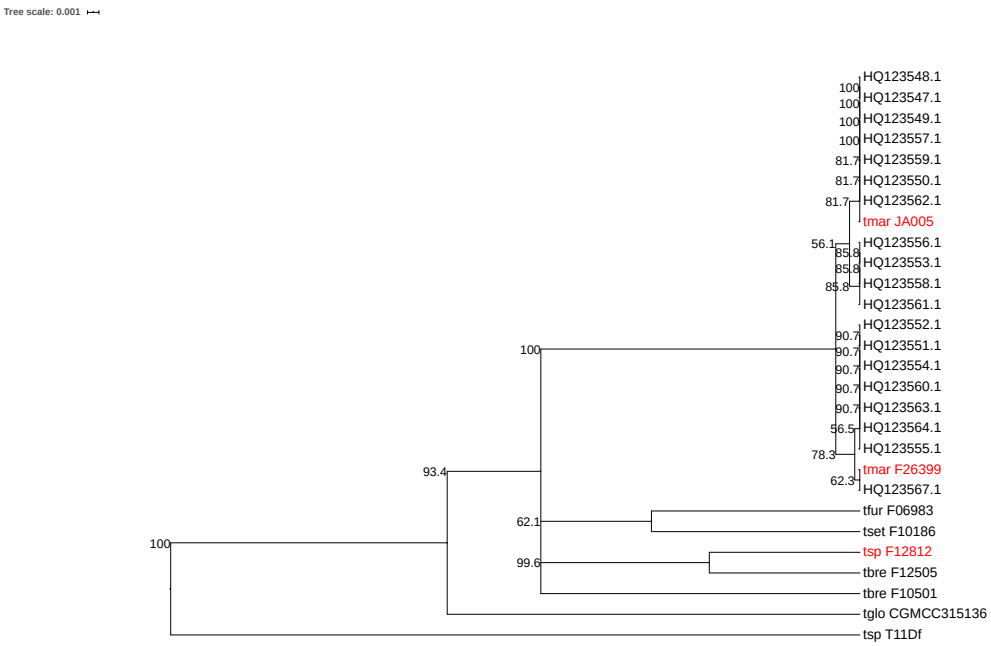


Figure S5. UPGMA clustering of 46 beta-tubulin sequences. Analysis performed in Geneious 10.2.4 using a Jukes-Cantor genetic distance model branch support estimated using 1000 bootstraps with resampling. All sequences with GenBank accessions starting with HQ come from the population genetics study of Anderson and Shearer (84).