

FunEndo: A Comprehensive Global-Scale Fungal Endophyte Repository

Mahdich S. Hosseyni Moghaddam^{1*}, Saleh Rahimlou^{2*}, Nasim Sedighian³, Hajar Soleimani⁴, Yi-Hong Ke⁵, Khadijeh Marzani⁶, Mohammad Bouazar⁴, Hadi Ghasemi⁶, Elahe Bakhshi⁶, Sowlat Khodadadi Manesh⁶, Mahin Izadi⁶, Khalil Kariman⁷, Naser Safaie⁶, Kessy Abarenkov^{8,11}, Mohammad Bahram^{1,9,10}, Leho Tedersoo¹¹

1. Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia
2. Department of Plant Pathology and Environmental Microbiology, Pennsylvania State University, University Park, PA 16802, U.S.A.
3. Centre Armand-Frappier Santé Biotechnologie, Institut National de la Recherche Scientifique (INRS), Laval, Québec, Canada.
4. Department of Plant Protection, School of Agriculture, Shiraz University, Shiraz, 7144113131, Iran.
5. Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, MI 48109, U.S.A.
6. Department of Plant Pathology, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran.
7. UWA School of Agriculture and Environment, The University of Western Australia, Perth, WA 6009, Australia
8. Natural History Museum, University of Tartu, Vanemuise 46, 51003, Tartu, Estonia
9. Department of Agroecology, Aarhus University, Slagelse, Denmark.
10. Department of Ecology, Swedish University of Agricultural Sciences, Uppsala, Sweden.
11. Mycology and Microbiology Center, University of Tartu, Liivi 2, Tartu 50400, Estonia

Authors indicated with an asterisk contributed equally to this work.

Corresponding author:

Mahdich S. Hosseyni Moghaddam,
Institute of Ecology and Earth Sciences
University of Tartu, Tartu, Estonia.
E-mail: mahdiehalsadat.hosseynimoghaddam@ut.ee,
Phone: +1 (734) 353 8121.

Co-corresponding author:

Mohammad Bahram
Department of Ecology Swedish University of Agricultural Sciences
Ulls Väg 16, Uppsala
E-mail: mohammad.bahram@slu.se

Table S1. Phylogenetic structure of host composition within bipartite network modules. Mean pairwise phylogenetic distance (MPD), standardized effect size (Z), and permutation-based P-values (999 permutations) are shown for each detected module. Negative Z-scores indicate phylogenetic clustering, while positive values indicate phylogenetic overdispersion.

Module	MPD	Z-score	P-value
1	377.63	2.97	0.093
2	412.88	53.54	1.000
3	387.50	-5.05	0.242
4	369.94	12.38	0.611
5	398.75	9.90	0.109
6	354.71	-657.10	0.996
7	347.95	-8.85	0.467
8	423.08	-972.70	0.974
9	354.29	NA	1.000
10	353.71	134.01	0.679

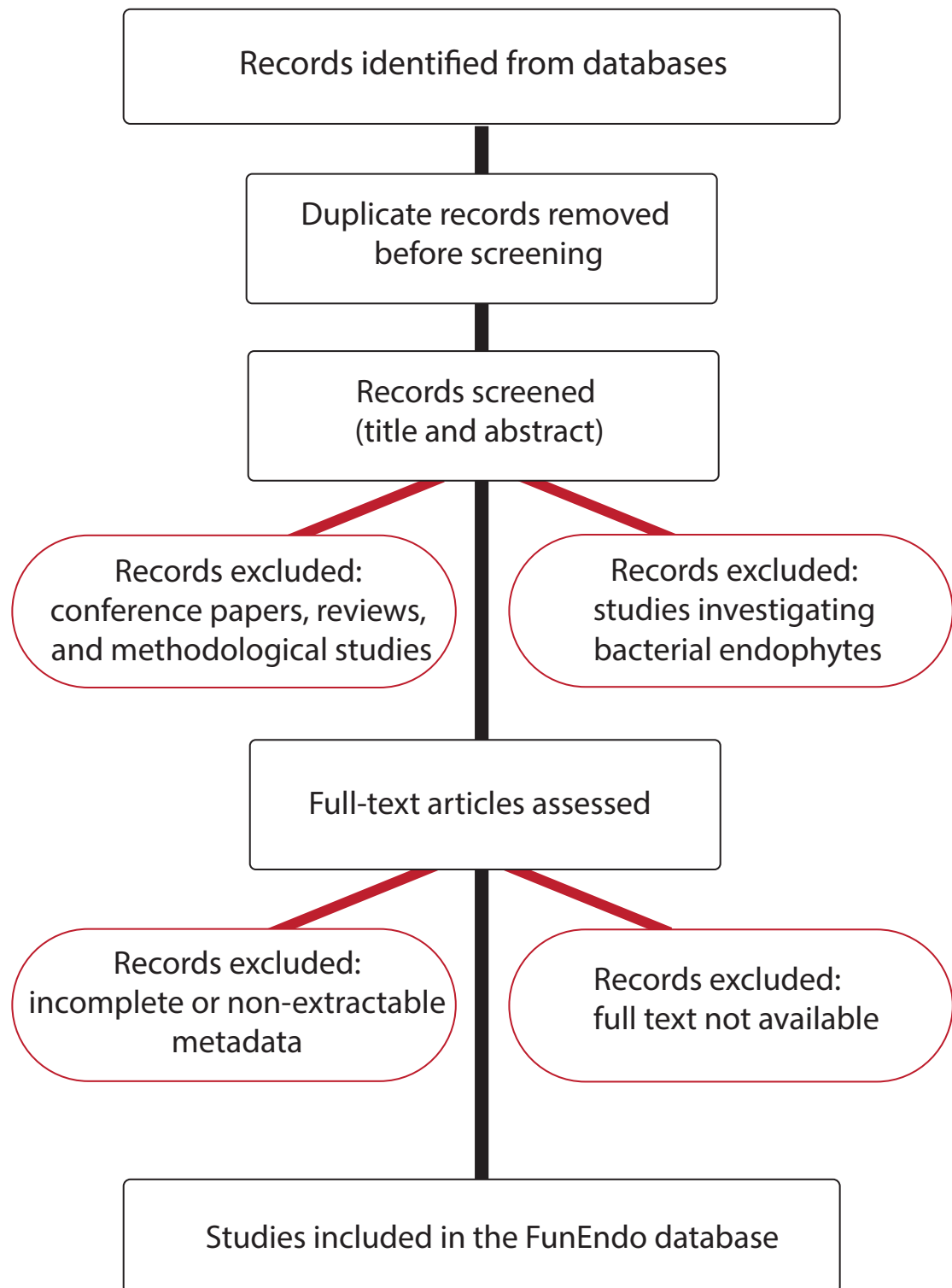


Fig S1. Workflow for literature screening and curation of the FunEndo database.

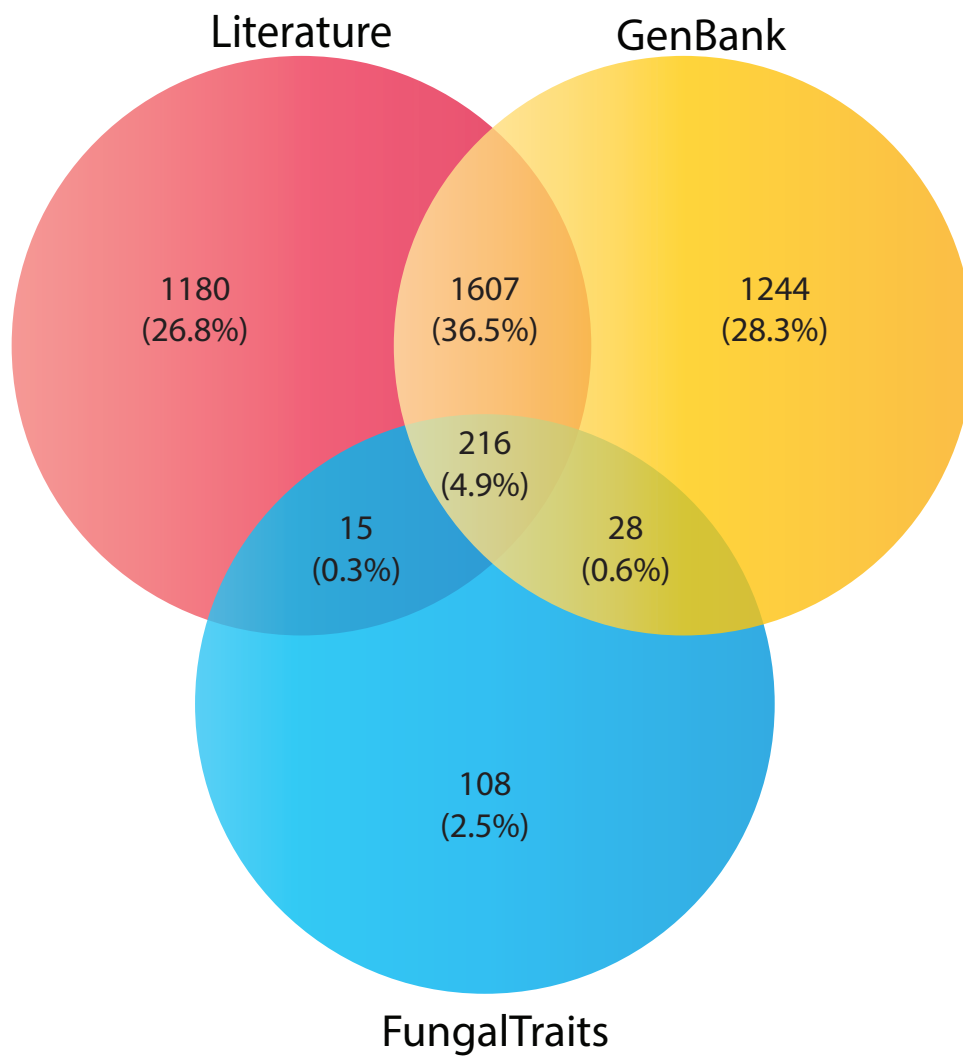


Fig S2. Venn diagram illustrates the unique and overlapping number of fungal species across three datasets: Scientific literature, GenBank, and FungalTraits.

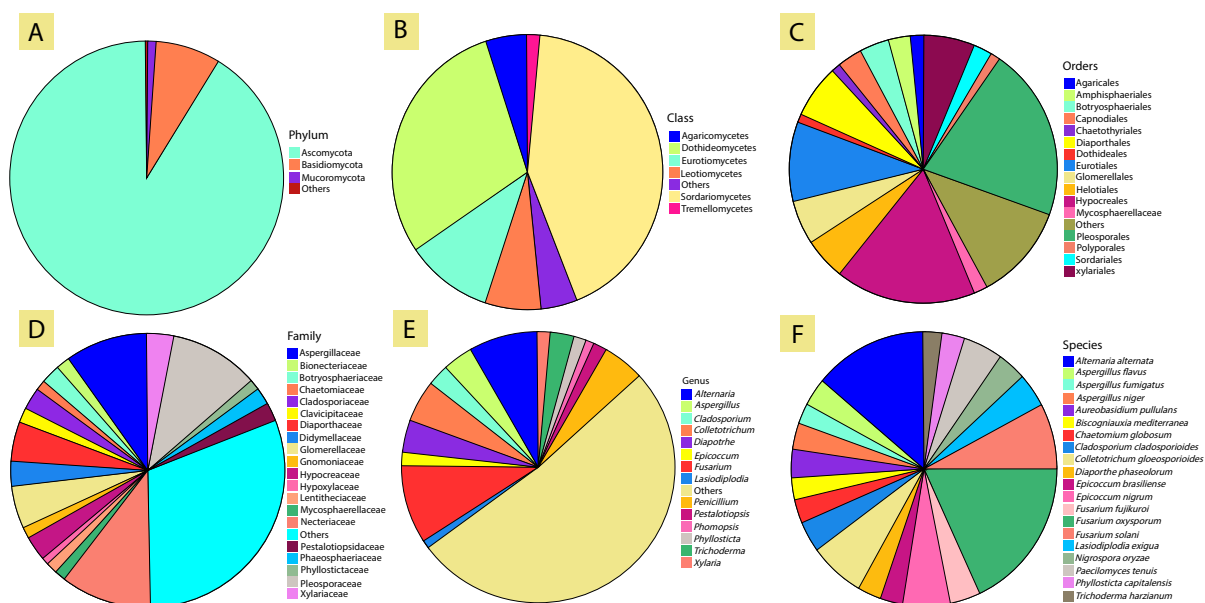


Fig S3. The proportion of dominant fungal endophyte taxa is listed in the scientific literature. A-E show the proportion of fungal endophyte records from the phylum to genus levels. F illustrates the top 20 dominant fungal endophyte species. The data not assigned to any taxonomic level are not included.

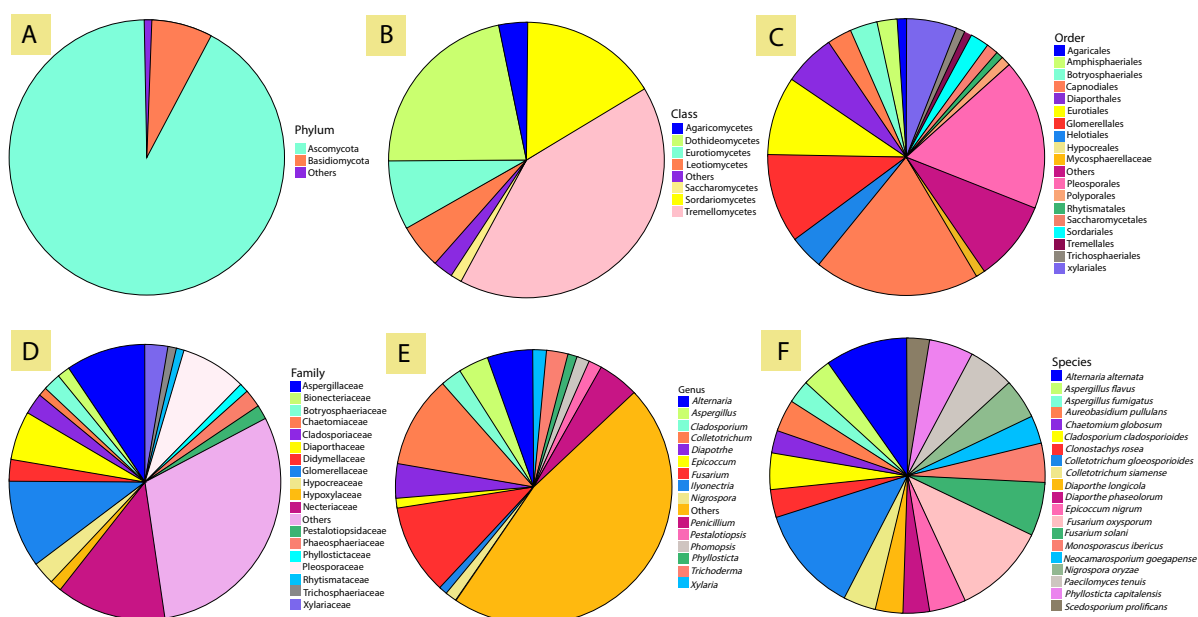


Fig S4. The proportion of dominant fungal endophyte taxa is listed from GenBank. A-E show the proportion of fungal endophyte records from the phylum to genus levels. F illustrates the top 20 dominant fungal endophyte species. The data not assigned to any taxonomic level are not included.

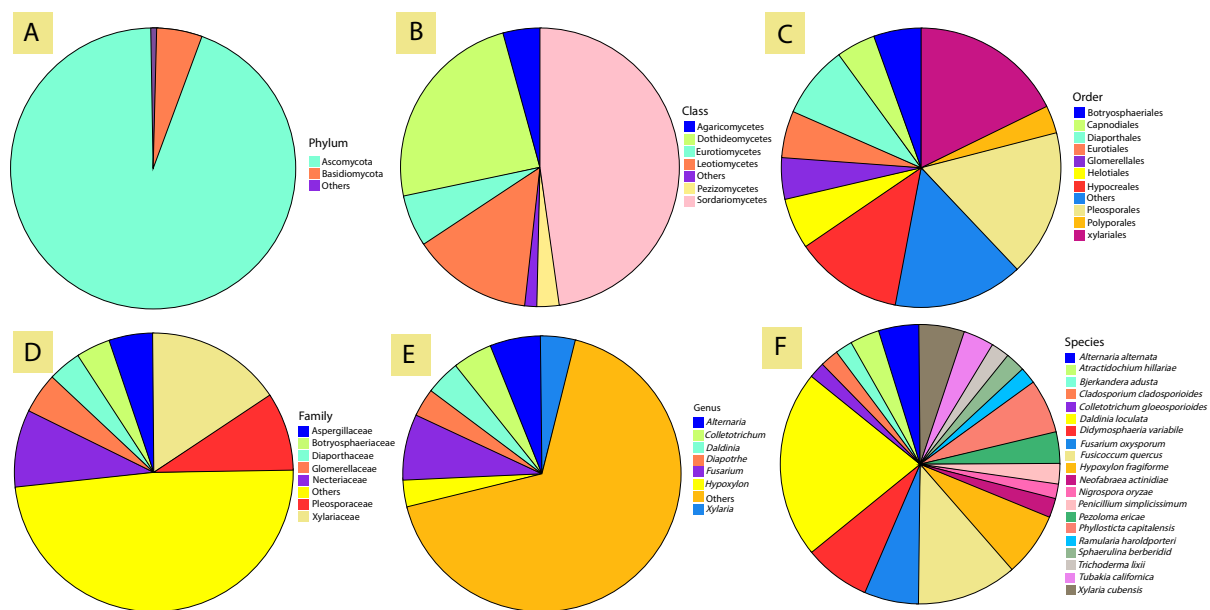


Fig S5. The proportion of dominant fungal endophyte taxa is listed from FungalTraits. A-E shows the proportion of fungal endophyte records from the phylum to genus levels. F illustrates the top 20 dominant fungal endophyte species. The data not assigned to any taxonomic level are not included.

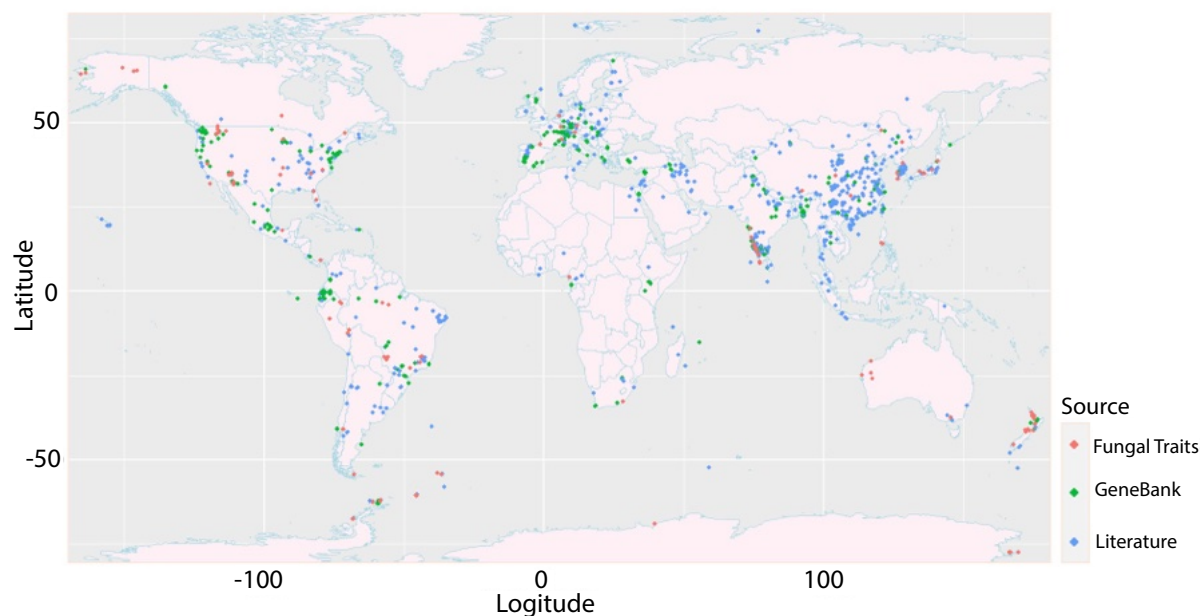


Fig S6. Geographic distribution of fungal endophytes. The map indicates the precise locations of fungal endophyte reports, each marked by a dot. Different dots of different colors indicate data collected from other sources. It's noteworthy that coordinate information was available for less than 25% of the dataset.

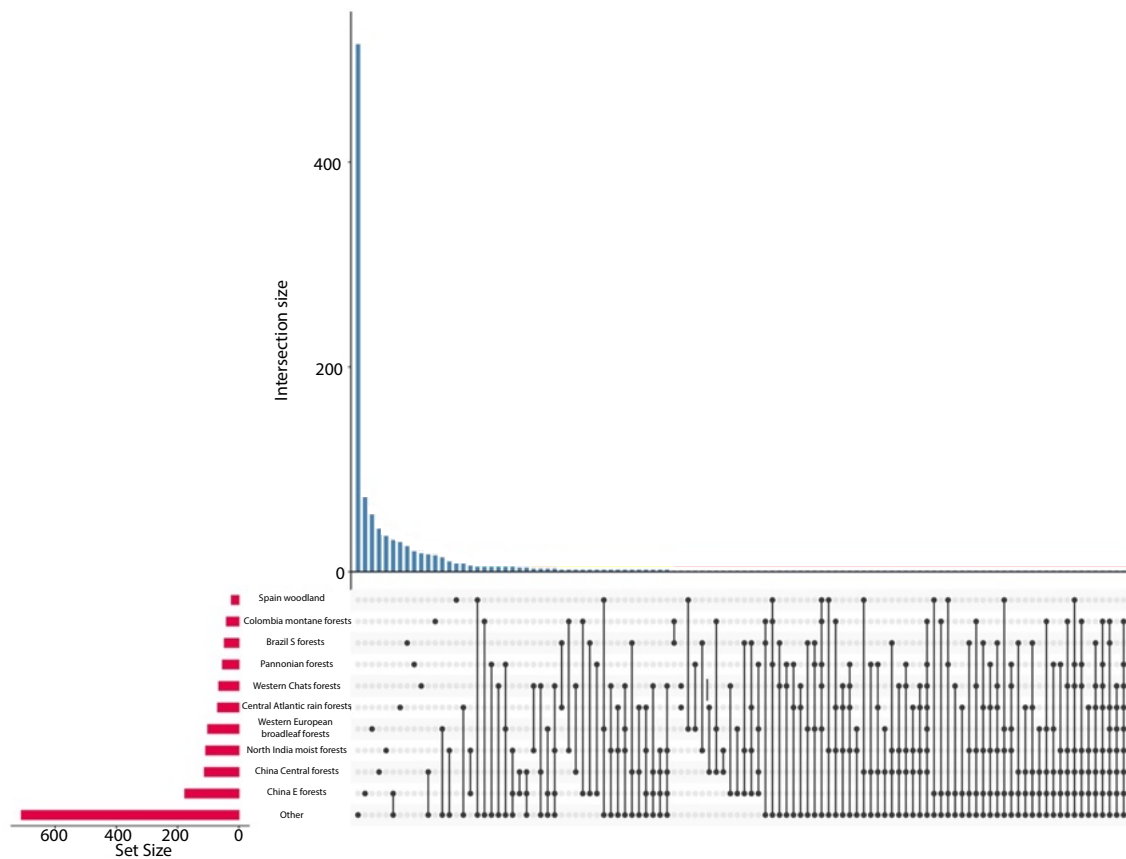


Fig S7. Distribution of unique and shared species among the studied super ecoregion, with set size representing the number of species reported from each ecoregion. Super-ecoregions outside the top 10 most frequent in the dataset are grouped together as 'Other' for clarity.

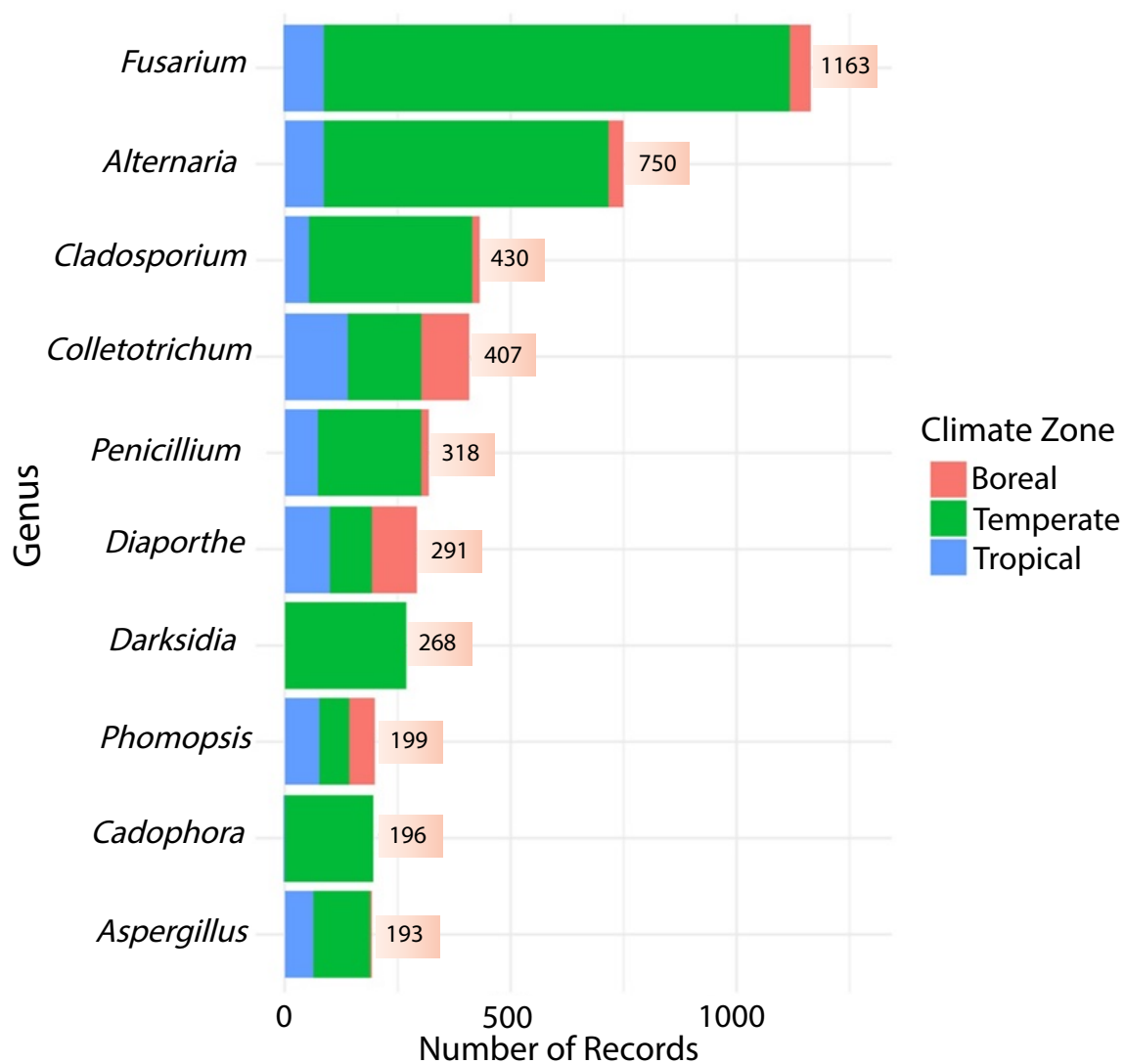


Fig S8. Distribution of the most frequent fungal endophyte genera across boreal, temperate, and tropical climate zones. This subset of data only includes taxa for which the geographic coordinates are available.

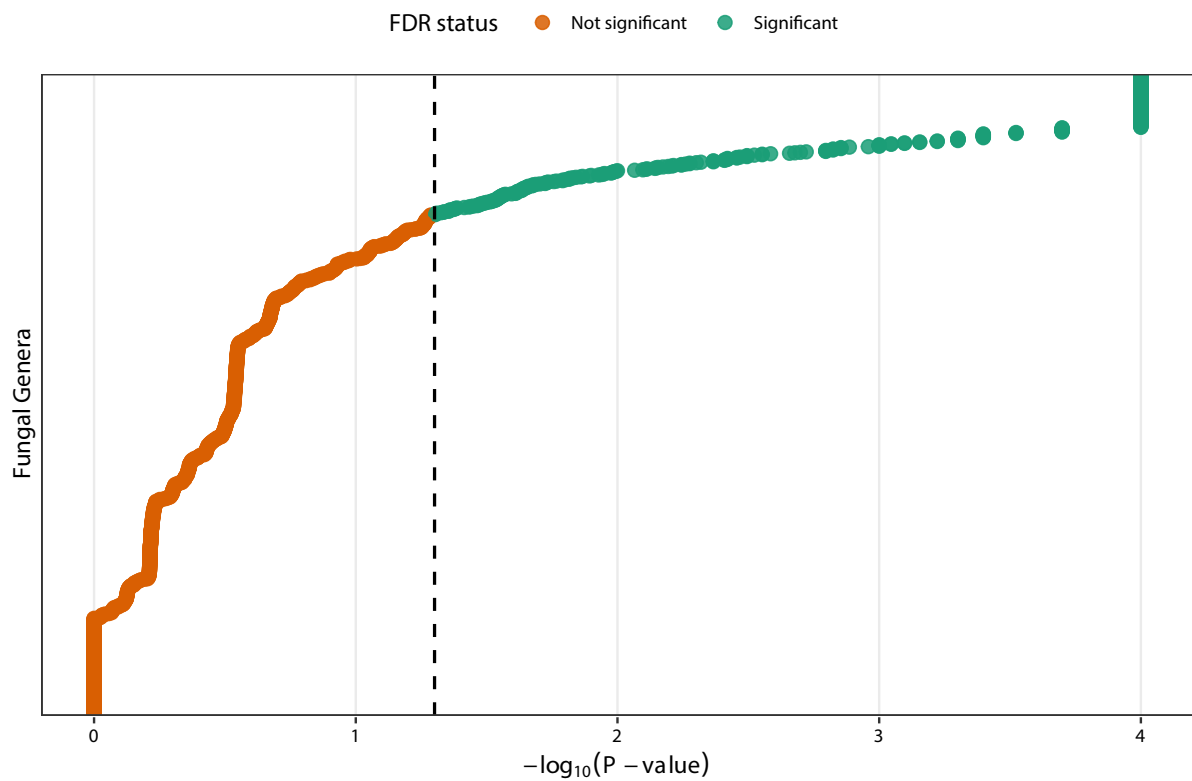


Figure S9. Organ specificity of fungal genera. Fungal genera were tested for organ-specific enrichment using Fisher's exact test. Points represent $-\log_{10}(\text{P-values})$, with the dashed line indicating the FDR-adjusted significance threshold (0.05). Significant genera are shown in green and non-significant genera in orange.

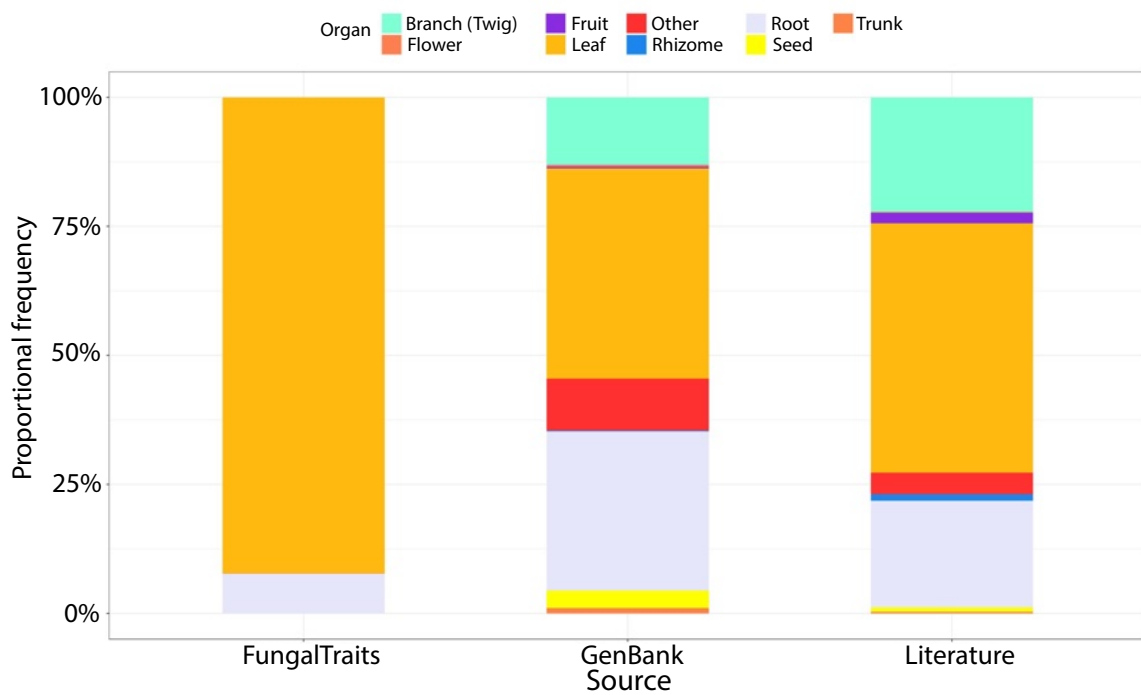


Fig S10. The proportion of fungal endophytes recovered from various organs in different sources.

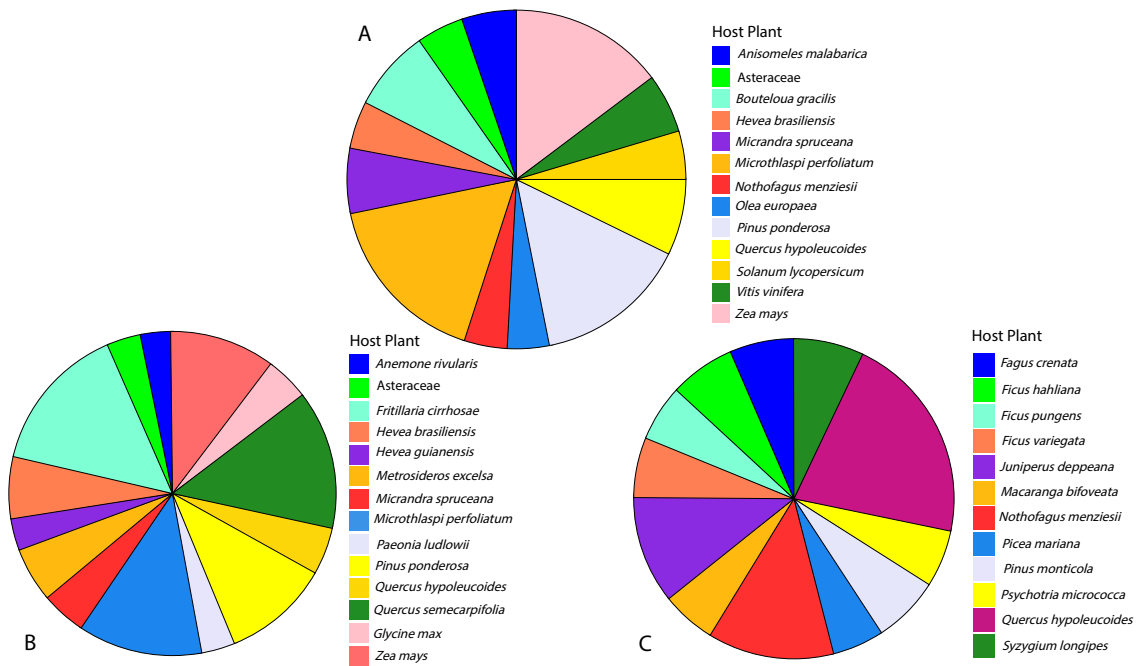


Fig S11. Host plant with the highest number of reported fungal endophytes across three datasets including scientific literature (A), GenBank (B), and the FungalTraits (C).

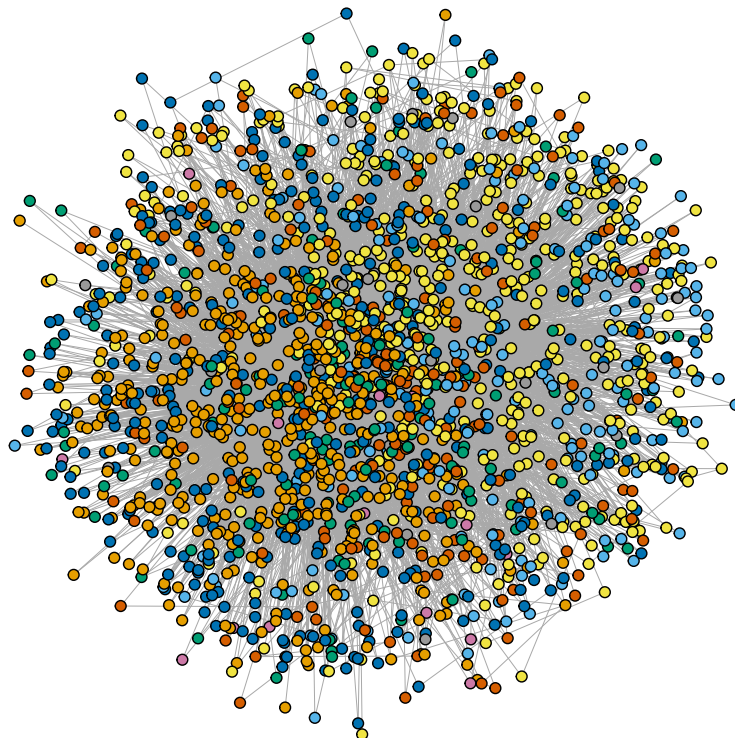


Fig S12. Bipartite host–fungus interaction network and phylogenetic structure within Louvain modules