Supporting Information for

Comparative multi-omics of host genome, metabolome and associated rhizosphere microbiota across karst-endemic and widespread *Rhododendron* species

Sulin Wen et al.

This PDF file includes:

Supplementary Notes S1–S9 and Figures S1–S14. Supplementary Tables S1–S12 and Data S1–S5 are provided as separate files.

Supplementary Notes S1-S9

Note S1. Comparative genomics workflow

High-quality reference genomes of *R. bailiense* and *R. delavayi* were obtained from the National Genomics Data Center (NGDC; https://ngdc.cncb.ac.cn/) under accessions GWHERFP00000000 and GWHBKHI00000000, respectively. Whole-genome sequences were aligned with minimap2 (Li, 2018). Syntenic regions, inversions, translocations and duplications were identified with SyRI v1.6.3 (Goel et al., 2019) and visualized with Plotsr v1.1.0 (Goel and Schneeberger, 2022). Regions not assigned to the whole-genome alignment were retained as candidate lineage-differentiated regions; their presence alone was not treated as evidence of adaptive function.

Orthogroups and a single-copy-gene phylogeny were constructed across closely related *Rhododendron* species using OrthoFinder v3.0.1 (Emms and Kelly, 2019), and gene-family expansions and contractions were estimated with CAFE v5.0.0 (Mendes et al., 2020). Expanded families and genes in non-aligned regions were mapped to GO and KEGG, and enrichment was tested with the functional-enrichment module of TBtools v2.467 (Chen et al., 2023), using a hypergeometric test and P_adj < 0.05. Enrichment identifies over-represented annotations and does not establish altered expression, biochemical activity or adaptive benefit.

Note S2. ITS sequencing and processing of rhizosphere fungi

The fungal ITS1 region was amplified with primers ITS1F (5'-CTTGGTCATTTAGAGGAAGTAA-3') and ITS2R (5'-GCTGCGTTCTTCATCGATGC-3'). Amplicons were purified with the AxyPrep DNA Gel Extraction Kit, and libraries were prepared with the TruSeq DNA PCR-Free Sample Preparation Kit before 150-bp sequencing on the Illumina NovaSeq 6000 platform. Reads were demultiplexed by barcode and denoised with the DADA2 plugin in QIIME 2 v2022.2, including error correction and chimera removal, to obtain amplicon sequence variants (ASVs). Representative sequences were classified against UNITE v9.0 with the classify-sklearn naive-Bayes classifier at a confidence threshold of 70% (Wang et al., 2007), and non-fungal sequences were removed.

ASV tables were rarefied to a common sequencing depth before alpha-diversity analysis. Community composition was visualized by principal coordinate analysis, and a phylogenetic tree was constructed from aligned representative sequences with IQ-TREE v3.0.1 (Minh et al., 2020). Ecological guilds were assigned with FUNGuild; only 'Highly Probable' and 'Probable' assignments were retained (Tanunchai et al., 2023). Guild and enzyme-related annotations were interpreted as community-level functional potential rather than

measured enzyme activity or nutrient transfer.

Note S3. Rhizosphere metagenomic sequencing and functional annotation

Rhizosphere-soil DNA from *R. bailiense*, *R. delavayi* and *R. ambiguum* (n = 3 habitat samples per group; nine metagenomes) was used to construct PE150 libraries. The insert sizes reported for these libraries ranged from 447 to 503 bp (Table S4), and each sample yielded at least 6.0 Gb of raw sequence. Reads were quality-filtered with fastp v0.20.0 (Chen et al., 2018), and host-associated reads were removed by mapping against the available *Rhododendron* reference genomes with Bowtie2. Clean reads were assembled with MEGAHIT v1.2.9 (Li et al., 2015), retaining contigs of at least 500 bp. ORFs of at least 100 bp were predicted with Prodigal v2.6.3 (Hyatt et al., 2010) and clustered with CD-HIT v4.6.1 at 95% identity and 90% coverage (Fu et al., 2012). The assemblies contained 3,076,709 predicted ORFs, which were reduced to a non-redundant catalogue of 1,620,158 genes (Tables S7 and S8). Reads were mapped to this catalogue with SOAPaligner v2.21 (Li et al., 2009), and per-sample abundance was expressed as TPM. These DNA-derived TPM values represent relative gene abundance rather than gene expression.

For the taxonomic profiles reported in Data S2, predicted proteins were queried against NR with DIAMOND v2.0.13 (Buchfink et al., 2015), and best-hit assignments were summarized at standard taxonomic ranks. KEGG orthologue annotations and TPM-normalized abundances are reported in Data S3. Targeted P-acquisition comparisons included *gcd* (K00117), *pqqC* (K06137), *phoD* (K01113), *phoA* (K01077), a putative PhoX-family annotation (K07093), *pstSCAB* (K02040, K02037, K02038 and K02036), phosphonate-transport genes *phnCDE* (K02041, K02044 and K02042), and C-P-lyase components *phnMJI* (K06162, K06163 and K06164). Other targeted annotations included N and S transformations, *cheA* and *cheY*, *corA* (K03284; Mg^{2+} transport), *mntH* (K03322; Mn^{2+}/Fe^{2+} transport), and a generic iron-complex outer-membrane receptor (K02014). KO annotations shared between ammonia and methane oxidation, or between nitrate reduction and nitrite oxidation, were treated as functionally ambiguous rather than assigned to a single process. None of these annotations was treated as a Ca^{2+}-specific marker or as evidence of transcription, process rate or nutrient transfer to plants.

Note S4. Non-targeted leaf metabolomics and cross-omics integration

For each plant group, leaf material from six independent biological replicates was analysed by LC-MS (n = 6 per group; 24 profiles in total). Approximately 50 mg of liquid-nitrogen-ground leaf powder was extracted in pre-chilled methanol:water (4:1, v/v) containing an internal standard. After ultrasonication and centrifugation at 12,000 rpm and 4 °C for 15 min, the supernatant was analysed by LC-MS. The column temperature was 40 °C, injection volume 3 μL, capillary temperature 325 °C, and spray voltages +3.5 kV and -3.5 kV. Peaks were picked, aligned and normalized in Progenesis QI and annotated against the Majorbio in-house library, METLIN and HMDB. Annotation confidence was recorded as level B(i) for experimental MS/MS-library matches and level B(ii) for theoretical MS/MS-library matches. Differential metabolite features were defined by VIP > 1.0 and nominal $P < 0.05$.

The full feature table contained 5,776 detected LC-MS features, of which 5,769 had module assignments in the WGCNA output (Data S4). Weighted correlation network analysis was applied to the metabolite-abundance matrix, and dynamic tree cutting defined co-abundance modules summarized by module eigengenes. Module-group correlations and module membership versus metabolite sig*nif*icance were used to prioritize representative features. Because plant identity and habitat were not independently crossed, the modules are described as plant-group-associated rather than species-specific.

For cross-omics integration, only sample-identifier-matched profiles were used: Rb_1-Rb_3, Rd_1-Rd_3 and Ram_1-Ram_3 (n = 9 paired observations). *R. decorum* was excluded because no matched rhizosphere metagenomes were generated. Spearman correlations were evaluated between 258 selected leaf metabolite features and 33 genus-level rhizosphere taxonomic profiles. Data S5 contains 1,972 retained edges satisfying $|\rho| > 0.6$ and nominal two-sided $P < 0.05$; the supplied P values are nominal and were not corrected for multiple testing. These networks are exploratory association maps. The metabolites were measured in leaves, not roots or exudates, and an edge does not demonstrate direct exposure, microbial recruitment, synergy, antagonism, functional complementation or causation.

Note S5. Interpreting genomic divergence in contrasting habitat contexts

The *R. bailiense* and *R. delavayi* genomes differed in non-aligned regions and gene-family histories (Figures S2, S3). *R. bailiense* candidate regions were enriched in broad secondary-metabolism categories, cytochrome P450, pentose and glucuronate interconversions, DNA-replication proteins, lysosome-related annotations and glycosyltransferases. These annotations are compatible with differences in specialized metabolism, detoxification, repair and macromolecular turnover (Nelson and Werck-Reichhart, 2011; Marshall and Vierstra, 2018; Pang et al., 2021; Zhang et al., 2022). Together, they point to a candidate genomic repertoire in which chemical diversification, P450-mediated modification, glycosylation and cellular maintenance could jointly support persistence in thin karst humus soils. Metabolic diversification may broaden chemical defence and microbial signalling, whereas repair and recycling functions may help sustain tissue function under recurrent drought, high exchangeable Ca^{2+} and limited P availability. The accompanying enrichment of plant–pathogen-interaction-related genes further suggests enhanced environmental perception and immune signalling, which may help *R. bailiense* balance beneficial microbial partnerships with defence against opportunists in a microbially active humus layer.

R. delavayi showed a different enrichment profile that included phenylpropanoid- and flavonoid-related annotations. This contrast is consistent with lineage-associated genomic differentiation across distinct edaphic and stress contexts. The soil measurements show that neither habitat is represented by a simple fertility axis: *R. bailiense* soil was rich in organic matter and N but low in P and high in exchangeable Ca^{2+}, whereas *R. delavayi* soil was low in P, organic matter, N and available S. Against this background, the *R. bailiense* repertoire may couple tolerance of P-poor, Ca^{2+}-rich physical constraints with use of a carbon- and N-rich humus horizon, whereas the phenylpropanoid and flavonoid bias of *R. delavayi* may support chemical defence, redox balance and microbial communication under lower resource availability. These associations provide testable host-side hypotheses for how genomic differentiation may contribute to persistence in the two habitats.

Note S6. Fungal community differentiation across contrasting soils

Fungal beta diversity separated the three habitat groups, whereas the measured richness indices did not differ significantly (Figures 2, 3 and S4-S6). *Piloderma* dominated the *R. bailiense* rhizosphere, while *Mortierella* was more prominent in the *R. delavayi* and *R. ambiguum* rhizospheres. The *R. bailiense* profile also contained the largest generalist and smallest specialist fraction; dominance by one associated genus should therefore not be described as community-wide specialist enrichment.

In the low-P, high-Ca^{2+} *R. bailiense* habitat, the co-occurrence of *Piloderma*, ectomycorrhizal guild assignments and acid-phosphatase-related potential is consistent with possible access to poorly available P pools (Finlay et al., 2020; King et al., 2025; Richardson and Simpson, 2011). Chitinase and beta-glucosidase potential may contribute to turnover of the habitat's abundant organic matter. This differs from the earlier carbon-scarcity interpretation: the soil measurements show that *R. bailiense* was not deficient in organic matter or total N.

The relatively saprotroph-rich *R. delavayi* profile is consistent with nutrient recycling in a habitat with low organic matter, N, P and available S. *Mortierella* was also prominent in the relatively P-, alkali-N- and S-rich *R. ambiguum* habitat, indicating that its distribution cannot be attributed to nutrient poverty alone (Ozimek and Hanaka, 2021; Vandepol et al., 2020). These patterns prioritize direct measurements of weathering, P transfer, Ca^{2+} handling and root colonization across paired habitat and host controls.

Note S7. *Bradyrhizobium* abundance and nitrogen-cycling potential

Bradyrhizobium was abundant in the *R. bailiense* rhizosphere. In Data S3, nifH (K02588) was not detected in any *R. bailiense* or *R. delavayi* metagenome and occurred only at trace abundance in *R. ambiguum*, with no supported group difference. This pattern suggests that the prominence of *Bradyrhizobium* in *R. bailiense* may reflect functions other than, or in addition to, nitrogen fixation, including oligotrophic persistence, organic-substrate use or interaction with Ericaceae roots. The genus therefore provides a focused target for strain-resolved genomics and physiological testing.

Because total N was highest in the *R. bailiense* habitat, the ecological relevance of *Bradyrhizobium* is more plausibly linked to plant association and resource use than to compensation for severe habitat-level N deficiency. Strain-resolved genomics, nif-gene screening and physiological assays can distinguish among these

possibilities, while paired bulk-soil source-pool controls can resolve the relative roles of habitat filtering and plant-associated enrichment (Xu et al., 2024).

Note S8. Leaf metabolites and candidate cross-kingdom associations

***R. bailiense*-associated modules.**

Safynol, luteolinidin, artonin K and long-chain fatty-acid derivatives were prominent in *R. bailiense*-associated leaf modules. Polyacetylenes and flavonoids can alter microbial growth or signalling in experimental systems (Greger, 2025; Huang et al., 2019; Korenblum et al., 2020). A shoot-to-root connection is biologically plausible because flavonoid intermediates can move from shoots towards root tips (Buer et al., 2007), while comparisons across Arabidopsis, Brachypodium and Medicago show that root exudates contain both a conserved core and species-specific compounds that overlap with tissue metabolomes (McLaughlin et al., 2023). Microbiome-dependent systemic signalling can also alter root exudation of acylsugars, fatty acids, hydroxycinnamate conjugates and other metabolites (Korenblum et al., 2020). In Data S5, luteolinidin and artonin K were negatively associated with Ktedonobacter, Ktedonosporobacter and several other genera. Together with the higher *cheA*, *cheY* and aromatic-compound-processing potential observed at the *R. bailiense* group level, these patterns suggest chemical filtering in which root-accessible analogues may favour microbes able to sense, tolerate or transform specialized metabolites while restricting more sensitive taxa.

***R. ambiguum*-associated modules.**

Capillarisin- and curcumin-related features were prominent in the *R. ambiguum*-associated module. Its metagenome also showed relatively high abundance of genes assigned to the KEGG pathway labelled 'biofilm formation-Pseudomonas aeruginosa'. This label represents a conserved biofilm-related gene set, so its ecological implication here is a candidate capacity for surface attachment and matrix formation. Curcumin I and demethoxycurcumin were positively associated in Data S5 with *Streptomyces*, *Candidatus Angelobacter*, Reyranella, Candidatus Acidiferrum and uncultured Dormiibacterota and Binataceae lineages. Because phenolics can act as selective inhibitors, signals or transformation substrates, these edges suggest that tolerant or biotransforming taxa may be favoured in the *R. ambiguum* rhizosphere, with biofilm capacity potentially stabilizing root-surface colonization and competition with opportunistic pathogens (Koprivova and Kopriva, 2022). Aboveground pathogen exposure can systemically reprogramme root exudation and generate disease-suppressive soil feedbacks (Yuan et al., 2018), providing a precedent for a shoot-to-rhizosphere chemical connection and making modulation of quorum sensing, biofilm behaviour and pathogen competition testable mechanisms here.

***R. delavayi*, *R. decorum* and shared modules.**

The *R. delavayi*-associated red module contained flavonoid-, phenylpropanoid- and benzoxazinoid-related features, whereas the *R. decorum*-associated yellow module showed a distinct enrichment profile. The proximity of *R. delavayi* and *R. decorum* in metabolite ordination, despite their species difference, is consistent with contributions from their shared yellow-soil habitat, conserved plant chemistry or both. Root-accessible flavonoid and phenylpropanoid analogues could serve as microbial signals, substrates or selective filters, while benzoxazinoid-related chemistry provides an additional route through which plant metabolic state may shape microbial assembly (Pang et al., 2021; Koprivova and Kopriva, 2022). Contrasting positive and negative genus-metabolite edges in Data S5 therefore identify candidate microbial responses involving utilization, biotransformation or tolerance. Targeted root and exudate profiling can test whether the shared aboveground chemical signatures extend belowground and distinguish habitat-associated convergence from lineage-dependent chemical filtering.

Note S9. Network-associated hub taxa and candidate functions

The cross-omics network linked leaf metabolite features with genus-level rhizosphere profiles in nine sample-identifier-matched observations (Figure 8; Data S5). The retained network contains exploratory edges meeting $|\rho| > 0.6$ and nominal $P < 0.05$. Because between-group separation can generate parallel correlations, network degree is used here to prioritize a compact set of key microorganisms for functional follow-up (Banerjee et al., 2018).

Specialized-metabolite-associated candidates.

Streptomyces was positively associated with Curcumin I, demethoxycurcumin and numerous other

specialized-metabolite features. Members of this genus can produce or transform bioactive compounds and influence plant defence. Together, these associations suggest a reciprocal chemical interaction in which root-accessible metabolites may favour tolerant or transformative *Streptomyces* strains, whose biotransformation and antimicrobial capacities could in turn modify the local chemical environment and support plant defence.

Biosynthetic and carbon-processing candidates.

Pseudorhodoplanes and *Candidatus Angelobacter* occupied high-degree network positions and were positively associated with Curcumin I; *Candidatus Angelobacter* was also linked to demethoxycurcumin. Genomic surveys of *Candidatus Angelobacter*-related Acidobacteriota report biosynthetic gene clusters (Crits-Christoph et al., 2022). Their high-degree positions, together with this biosynthetic potential, suggest a candidate link between plant-derived carbon, microbial secondary metabolism and competitive interactions in the rhizosphere.

Resource- and mineral-associated candidates.

High-degree nodes also included an unclassified Candidatus Dormiibacterota lineage, an unclassified Candidatus Binataceae lineage, Reyranella and Candidatus Acidiferrum, several of which were positively associated with curcuminoid features, whereas Ktedonobacter and Ktedonosporobacter showed predominantly negative associations with luteolinidin, artonin K and other retained metabolites. Metagenome-assembled genomes of Binatota indicate broad methylotrophic and alkane-degradation potential (Murphy et al., 2021), suggesting C1- and hydrocarbon-processing capacity for the Binatota-related node. Genomes of Ktedonobacteria from volcanic deposits encode CO oxidation and diverse respiratory functions (Hernández et al., 2020), supporting a respiratory-versatility hypothesis specifically for Ktedonobacter and possible sensitivity to the *R. bailiense*-associated chemical profile. In another plant-soil system, Candidatus Acidiferrum-enriched microbiomes were associated with Fe reduction, weathering and mineral-P mobilization (Epihov et al., 2021), raising the possibility that the Acidiferrum-related node contributes to mineral nutrient mobilization in these P-limited soils. These taxa can now be prioritized for MAG- or strain-resolved functional tests of chemical transformation, respiratory versatility and access to mineral-associated nutrients.

Collectively, the network supports a working model of resource-function matching among contrasting soil contexts, plant-group-associated chemistry and co-occurring rhizosphere taxa. Habitat filtering and plant-associated chemical selection may act jointly: the soil source pool constrains which microorganisms are available, while mobile metabolites, precursors or systemically regulated root exudates may favour particular members near roots. The network therefore provides prioritized taxon–metabolite pairs for targeted root and exudate metabolomics, MAG or strain sequencing, compound-response assays and manipulative inoculation tests that can determine which key microorganisms provide functional complementarity to *Rhododendron* hosts.

References

Banerjee S., Schlaeppi K. & Van Der Heijden M.G.A. 2018. Keystone taxa as drivers of microbiome structure and functioning. Nature Reviews Microbiology, 16, 567-576.

Buer C.S., Muday G.K. & Djordjevic M.A. 2007. Flavonoids are differentially taken up and transported long distances in Arabidopsis. Plant Physiology, 145, 478–490.

Brewer T.E., Aronson E.L., Arogyaswamy K., Billings S.A., Botthoff J.K., Campbell A.N., Dove N.C., Fairbanks D., Gallery R.E., Hart S.C., Kaye J., King G., Logan G., Lohse K.A., Maltz M.R., Mayorga E., O'neill C., Owens S.M., Packman A., Pett-Ridge J., Plante A.F., Richter D.D., Silver W.L., Yang W.H. & Fierer N. 2019. Ecological and Genomic Attributes of Novel Bacterial Taxa That Thrive in Subsurface Soil Horizons. Mbio, 10.

Buchfink B., Xie C. & Huson D.H. 2015. Fast and sensitive protein alignment using DIAMOND. Nat Methods, 12, 59-60.

Chen C., Wu Y., Li J., Wang X., Zeng Z., Xu J., Liu Y., Feng J., Chen H., He Y. & Xia R. 2023. TBtools-II: A "one for all, all for one"bioinformatics platform for biological big-data mining. Molecular Plant, 16, 1733-1742.

Chen S., Zhou Y., Chen Y. & Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics, 34, i884-i890.

Crits-Christoph A., Diamond S., Al-Shayeb B., Valentin-Alvarado L. & Banfield J.F. 2022. A widely distributed genus of soil Acidobacteria genomically enriched in biosynthetic gene clusters. Isme Communications, 2.

Emms D.M. & Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biology, 20.

Epihov D.Z., Saltonstall K., Batterman S.A., Hedin L.O., Hall J.S., Van Breugel M., Leake J.R. & Beerling D.J. 2021. Legume-microbiome interactions unlock mineral nutrients in regrowing tropical forests. Proceedings of the National Academy of Sciences of the United States of America, 118, e2022241118.

Finlay R.D., Mahmood S., Rosenstock N., Bolou-Bi E.B., Kohler S.J., Fahad Z., Rosling A., Wallander H., Belyazid S., Bishop K. & Lian B. 2020. Reviews and syntheses: Biological weathering and its consequences at different spatial levels - from nanoscale to global scale. Biogeosciences, 17, 1507-1533.

Fu L., Niu B., Zhu Z., Wu S. & Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28, 3150-2.

Goel M. & Schneeberger K. 2022. plotsr: visualizing structural similarities and rearrangements between multiple genomes. Bioinformatics, 38, 2922-2926.

Goel M., Sun H.Q., Jiao W.B. & Schneeberger K. 2019. SyRI: finding genomic rearrangements and local sequence differences from whole-genome assemblies. Genome Biology, 20, 277.

Greger H. 2025. Comparative Phytochemistry of Polyacetylenes of the Genus Artemisia (Asteraceae): Compounds with High Biological Activities and Chemotaxonomic Significance. Molecules, 30, 537.

Hernández M., Vera-Gargallo B., Calabi-Floody M., King G.M., Conrad R. & Tebbe C.C. 2020. Reconstructing genomes of carbon monoxide oxidisers in volcanic deposits including members of the class *Ktedonobacter*ia. Microorganisms, 8, 1880.

Huang A.C., Jiang T., Liu Y.-X., Bai Y.-C., Reed J., Qu B., Goossens A., Nützmann H.-W., Bai Y. & Osbourn A. 2019. A specialized metabolic network selectively modulates Arabidopsis root microbiota. Science, 364, eaau6389.

Hyatt D., Chen G.L., Locascio P.F., Land M.L., Larimer F.W. & Hauser L.J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics, 11, 119.

King K., Fransson P., Finlay R. & Sanchez-Garcia M. 2025. Transporter gene family evolution in ectomycorrhizal fungi in relation to mineral weathering capabilities. Biogeosciences, 22, 8047-8064.

Koprivova A. & Kopriva S. 2022. Plant secondary metabolites altering root microbiome composition and function. Current Opinion In Plant Biology, 67.

Korenblum E., Dong Y., Szymanski J., Panda S., Jozwiak A., Massalha H., Meir S., Rogachev I. & Aharoni A. 2020. Rhizosphere microbiome mediates systemic root metabolite exudation by root-to-root signaling. Proceedings Of The National Academy Of Sciences Of The United States Of America, 117, 3874-3883.

Li D., Liu C.M., Luo R., Sadakane K. & Lam T.W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. Bioinformatics, 31, 1674-6.

Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics, 34, 3094-3100.

Li R., Yu C., Li Y., Lam T.W., Yiu S.M., Kristiansen K. & Wang J. 2009. SOAP2: an improved ultrafast tool for short read alignment. Bioinformatics, 25, 1966-7.

Marshall R.S. & Vierstra R.D. 2018. Autophagy: The Master of Bulk and Selective Recycling. Annual Review of Plant Biology, 69, 173-208.

McLaughlin S., Zhalnina K., Kosina S., Northen T.R. & Sasse J. 2023. The core metabolome and root exudation dynamics of three phylogenetically distinct plant species. Nature Communications, 14, 1649.

Mendes F.K., Vanderpool D., Fulton B. & Hahn M.W. 2020. CAFE 5 models variation in evolutionary rates among gene families. Bioinformatics, 36, 5516-5518.

Minh B.Q., Schmidt H.A., Chernomor O., Schrempf D., Woodhams M.D., Von Haeseler A. & Lanfear R. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Molecular Biology and Evolution, 37, 1530-1534.

Murphy C.L., Sheremet A., Dunfield P.F., Spear J.R., Stepanauskas R., Woyke T., Elshahed M.S. & Youssef N.H. 2021. Genomic analysis of the yet-uncultured Binatota reveals broad methylotrophic, alkane-degradation, and pigment production capacities. mBio, 12, e00985-21.

Nelson D. & Werck-Reichhart D. 2011. A P450-centric view of plant evolution. Plant Journal, 66, 194-211.

Ozimek E. & Hanaka A. 2021. Mortierella Species as the Plant Growth-Promoting Fungi Present in the Agricultural Soils. Agriculture, 11, 7.

Pang Z., Chen J., Wang T., Gao C., Li Z., Guo L., Xu J. & Cheng Y. 2021. Linking Plant Secondary Metabolites and Plant Microbiomes: A Review. Frontiers In Plant Science, 12.

Richardson A.E. & Simpson R.J. 2011. Soil microorganisms mediating phosphorus availability update on microbial phosphorus. Plant Physiol, 156, 989-96.

Tanunchai B., Ji L., Schroeter S.A., Wahdan S.F.M., Hossen S., Delelegn Y., Buscot F., Lehnert A.-S., Alves E.G., Hilke I., Gleixner G., Schulze E.-D., Noll M. & Purahong W. 2023. FungalTraits vs. FUNGuild: Comparison of Ecological Functional Assignments of Leaf- and Needle-Associated Fungi Across 12 Temperate Tree Species. Microbial Ecology, 85, 411-428.

Vandepol N., Liber J., Desiro A., Na H., Kennedy M., Barry K., Grigoriev I.V., Miller A.N., O'donnell K., Stajich J.E. & Bonito G. 2020. Resolving the Mortierellaceae phylogeny through synthesis of multi-gene phylogenetics and phylogenomics. Fungal Diversity, 104, 267-289.

Wang Q., Garrity G.M., Tiedje J.M. & Cole J.R. 2007. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol, 73, 5261-7.

Xu Y.L., Li J.J., Qiao C., Yang J.C., Li J., Zheng X.A., Wang C., Cao P.J., Li Y. & Chen Q.S. 2024. Rhizosphere bacterial community is mainly determined by soil environmental factors, but the active bacterial diversity is mainly shaped by plant selection. Bmc Microbiology, 24.

Yuan J., Zhao J., Wen T., Zhao M., Li R., Goossens P., Huang Q., Bai Y., Vivanco J.M., Kowalchuk G.A., Berendsen R.L. & Shen Q. 2018. Root exudates drive the soil-borne legacy of aboveground pathogen infection. Microbiome, 6.

Zhang Y., Guo R., Wang S.-S., Jiang X.-Y., Cui H.-Y., Guo Y., Song X.-Y., Guo Q.-Q. & Cao L. 2022. Autophagy-related Proteins in Genome Stability : Autophagy-Dependent and Independent Actions. International Journal Of Biological Sciences, 18, 5329-5344.

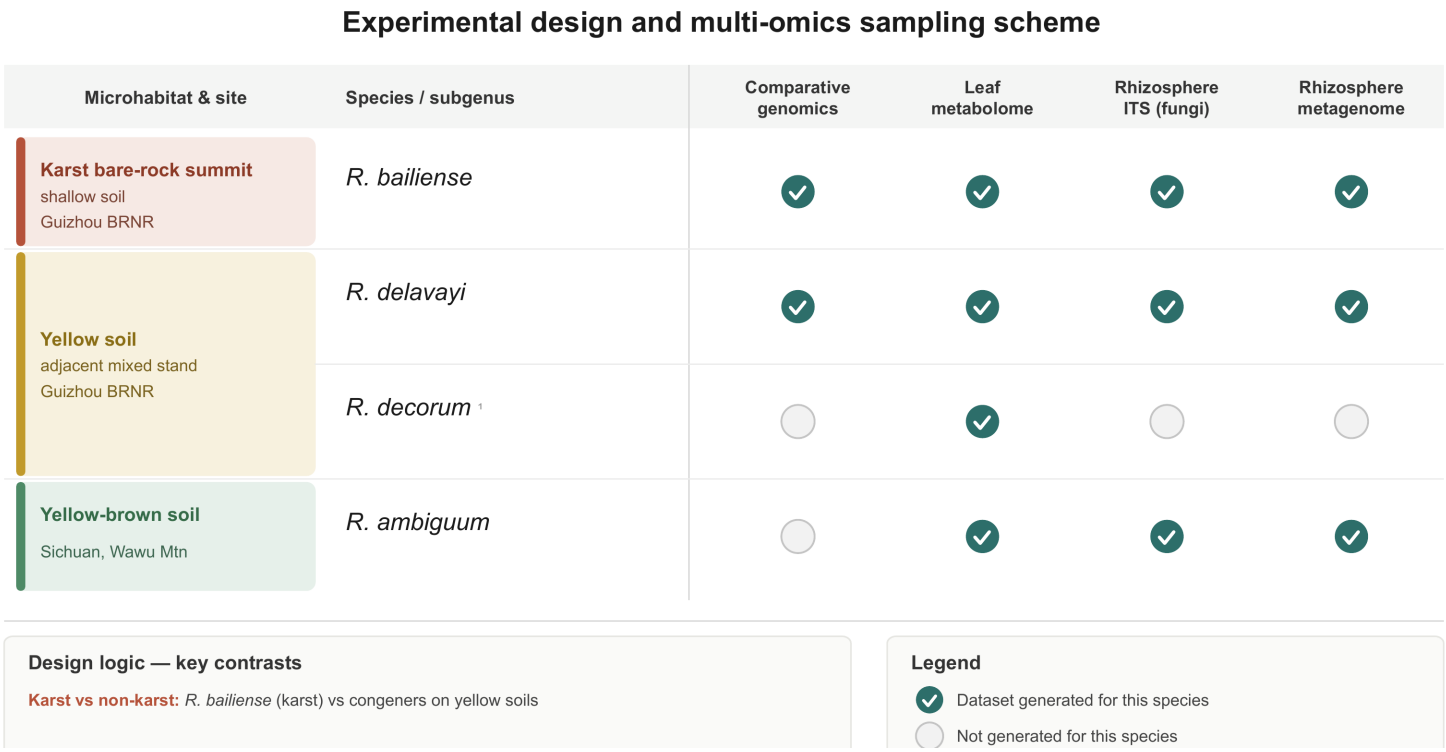


Figure S1. Experimental design and multi-omics sampling scheme. Four *Rhododendron* species were sampled across three habitat settings. Filled circles indicate datasets generated for a plant group, whereas open circles indicate datasets not generated. Comparative genomics focused on *R. bailiense* and *R. delavayi*; gene-family analyses additionally incorporated published *Rhododendron* genomes.

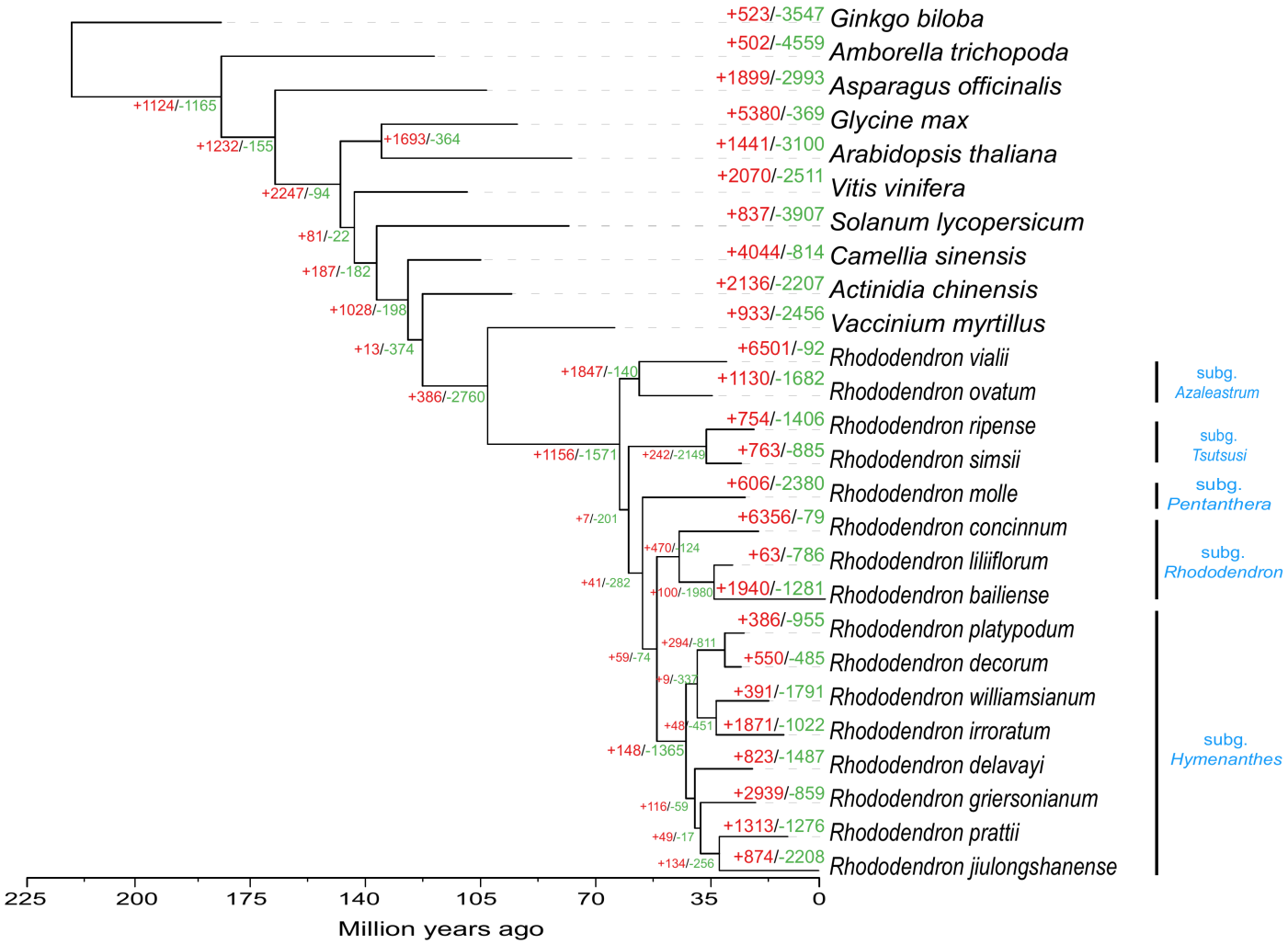


Figure S2. Phylogenetic tree and divergence chronogram depicting evolutionary relationships and gene-family dynamics among *Rhododendron* species. Branches are annotated with estimated divergence times and gene-family expansion and contraction counts.

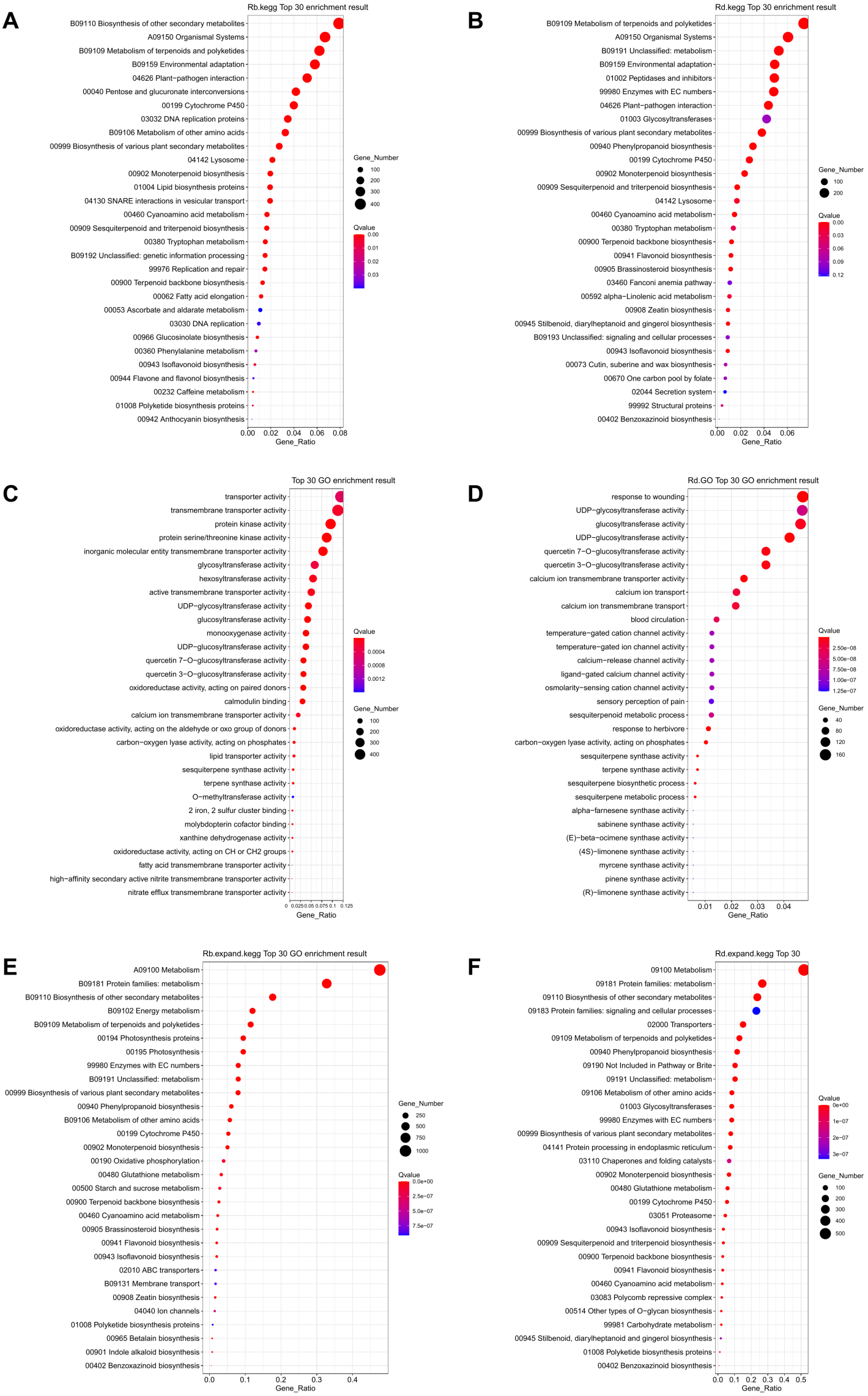


Figure S3. Comparative enrichment of non-aligned and expanded gene sets in *R. bailiense* and *R. delavayi*. (A, B) KEGG enrichment analyses of genes in non-aligned genomic regions of *R. bailiense* relative to *R. delavayi* (A) and *R. delavayi* relative to *R. bailiense* (B). (C, D) Corresponding GO enrichment analyses of genes in non-aligned regions of *R. bailiense* (C) and *R. delavayi* (D). (E, F) KEGG enrichment analyses of expanded gene families at the *R. bailiense* (E) and *R. delavayi* (F) divergence nodes, corresponding to 1,940 and 823 expanded gene families, respectively, in Figure S2.

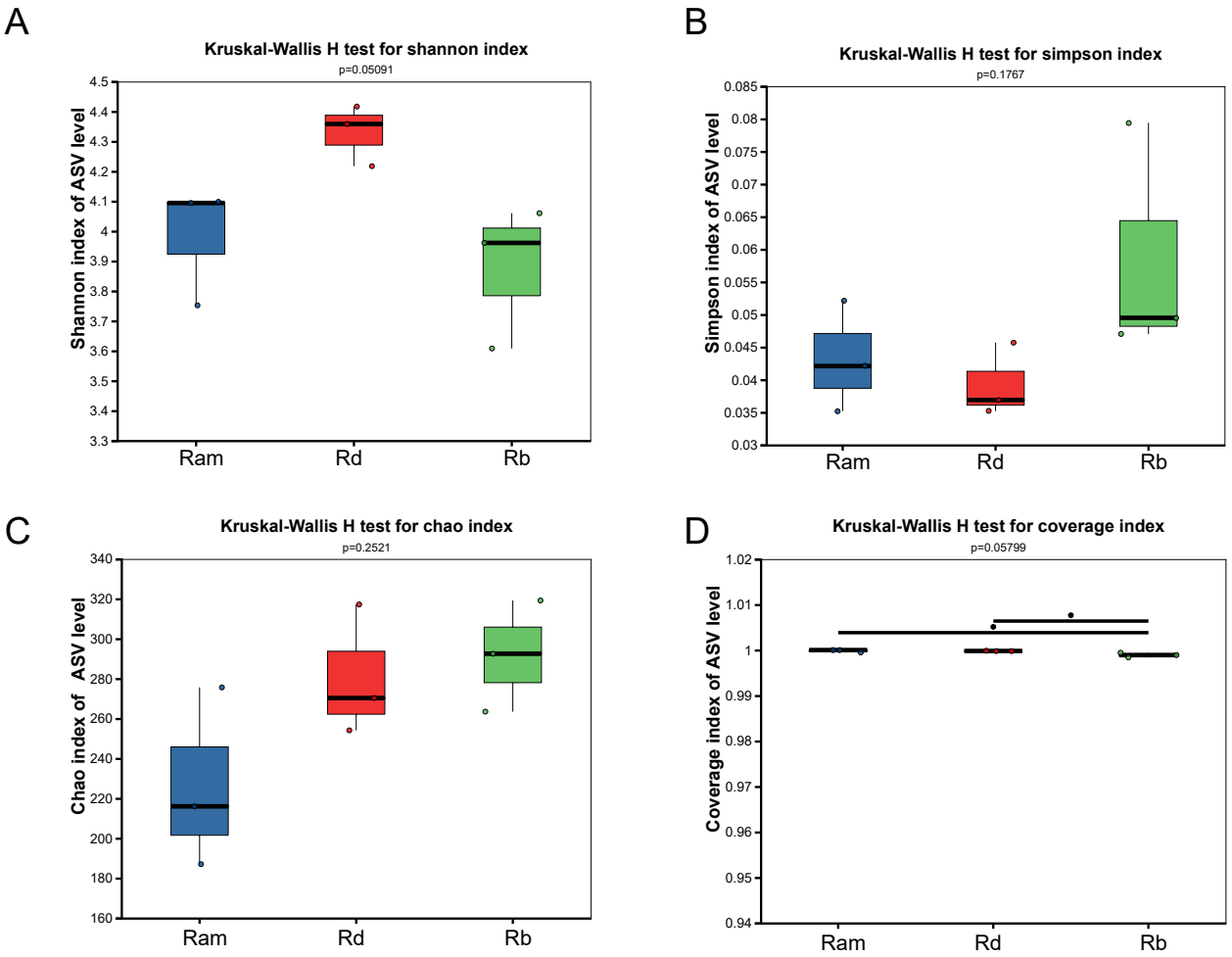


Figure S4. The alpha diversity indices (four indices) for each species in the fungal ITS sequencing analysis. (A − D) Shannon index, Simpson index, Chao index, and coverage index data for fungal alpha diversity in soil samples from the habitats of three rhododendron species (*R. bailiense*, *R. delavayi*, *R. ambiguum*). Significance was assessed using the Kruskal-Wallis H test (*, P < 0 .05). The Shannon index is a measure of community diversity (more sensitive to rare species); a higher index value indicates greater community diversity. The Simpson index also reflects community diversity (more sensitive to dominant species); a higher value of this index actually indicates lower diversity. The Chao index estimates the total number of species in a community; a higher index value indicates a greater total number of species. The coverage index reflects whether the sequencing coverage is sufficient; typically, high coverage (99% or higher) indicates that the vast majority of species in the sample have been detected, and the current sequencing coverage is sufficient to accurately reflect the community structure.

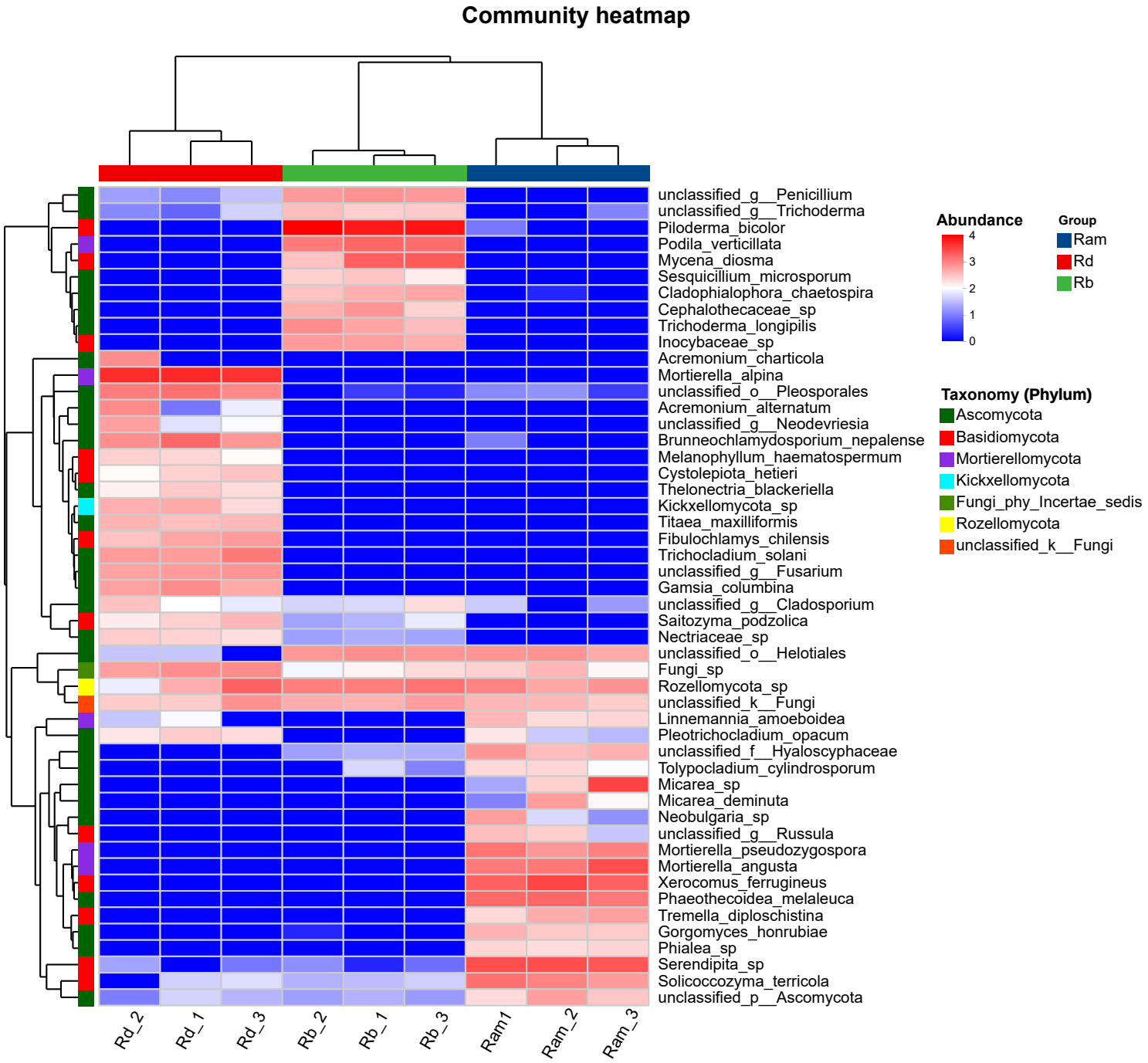
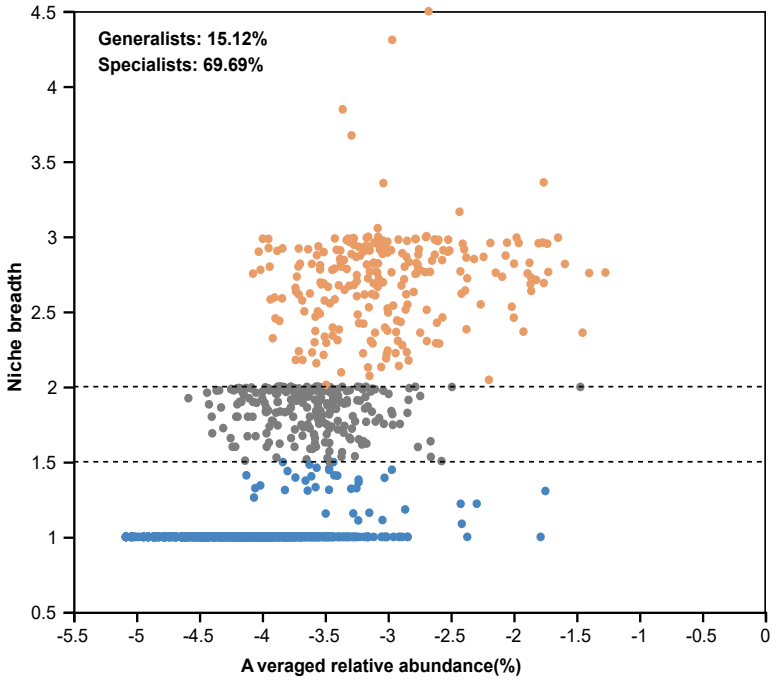


Figure S5. This heatmap illustrates the differences in abundance among 50 representative fungal genera in the soil of three *Rhododendron* habitats, based on fungal ITS sequencing analysis. The x-axis lists the sample names, the y-axis lists the genus names, and a color gradient (red for high abundance, blue for low abundance) indicates the relative abundance of each genus.

A



B

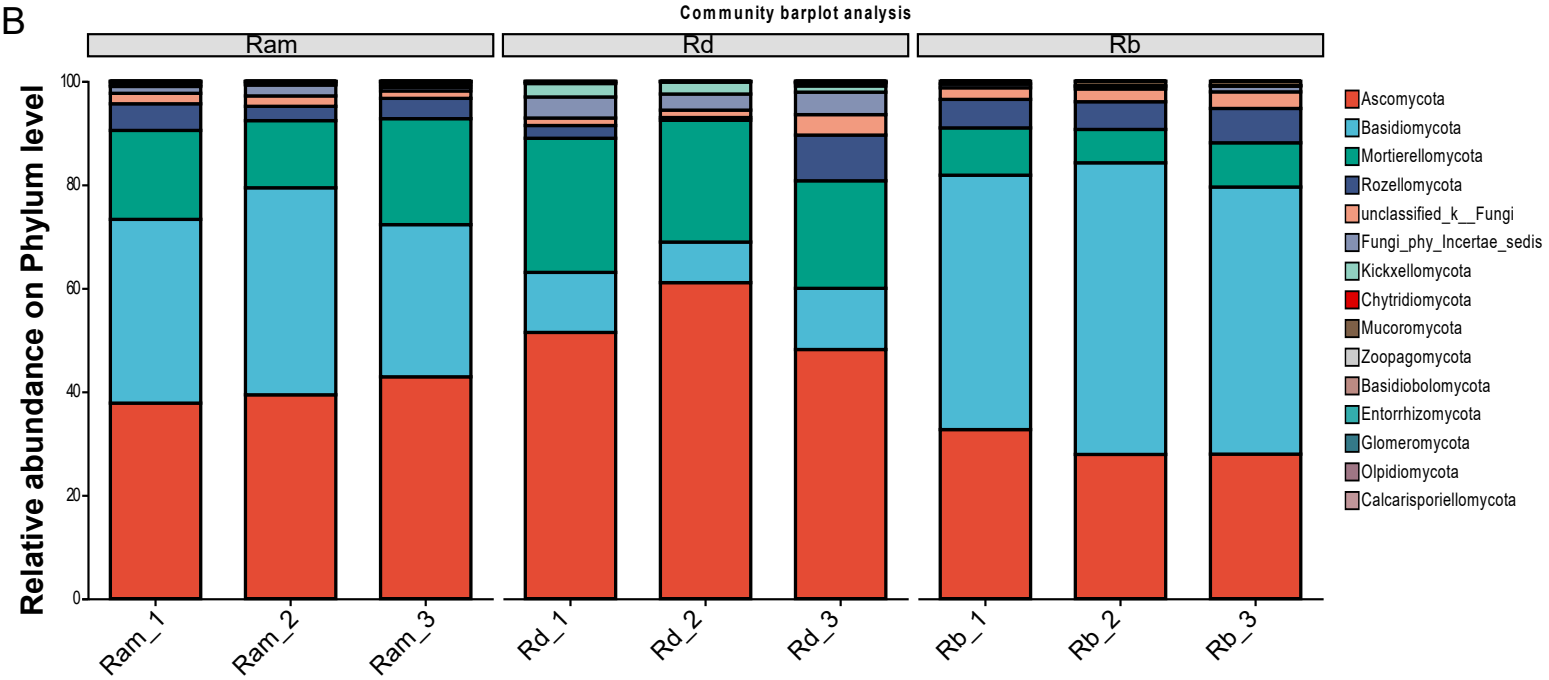


Figure S6. Species community abundance and phylum distribution characteristics based on fungal ITS sequencing analysis. (A) Scatter plot of species community abundance versus niche breadth. The x-axis represents the transformed average relative abundance of species (Averaged relative abundance (%)) (i.e., log10 transformation), and the y-axis represents niche breadth. Each point in the figure represents a species: yellow-brown points indicate generalists, while gray points indicate non-key species and specialists. (B) Statistical data on the distribution of fungal phyla in the soil of the habitats of three rhododendron species (*R. bailiense*, *R. delavayi*, *R. ambiguum*) based on microbial diversity analysis.

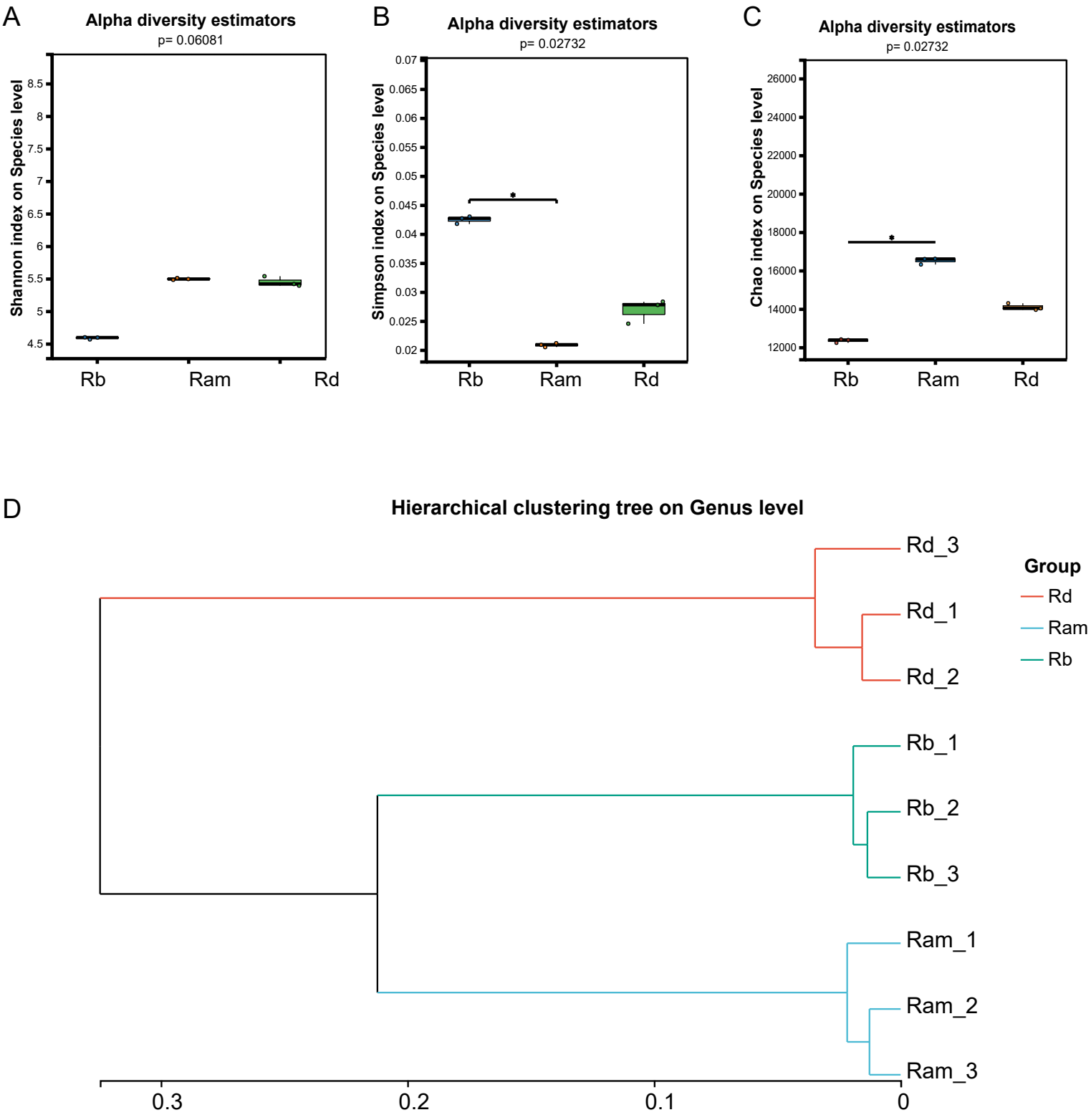


Figure S7. The alpha diversity indices (three indices) and hierarchical clustering for each species in the microbial metagenomic analysis. (A − C) Shannon, Simpson, and Chao indices for fungal diversity analysis of soil samples from the habitats of three rhododendron species (*R. bailiense*, *R. delavayi*, *R. ambiguum*). Significance was assessed using the Kruskal-Wallis H test (*, $P < 0.05$). For an interpretation of the diversity indices, see the legend in Figure S4. (D) Hierarchical clustering tree at the genus level for the nine soil samples from rhododendron habitats analyzed via metagenomics.

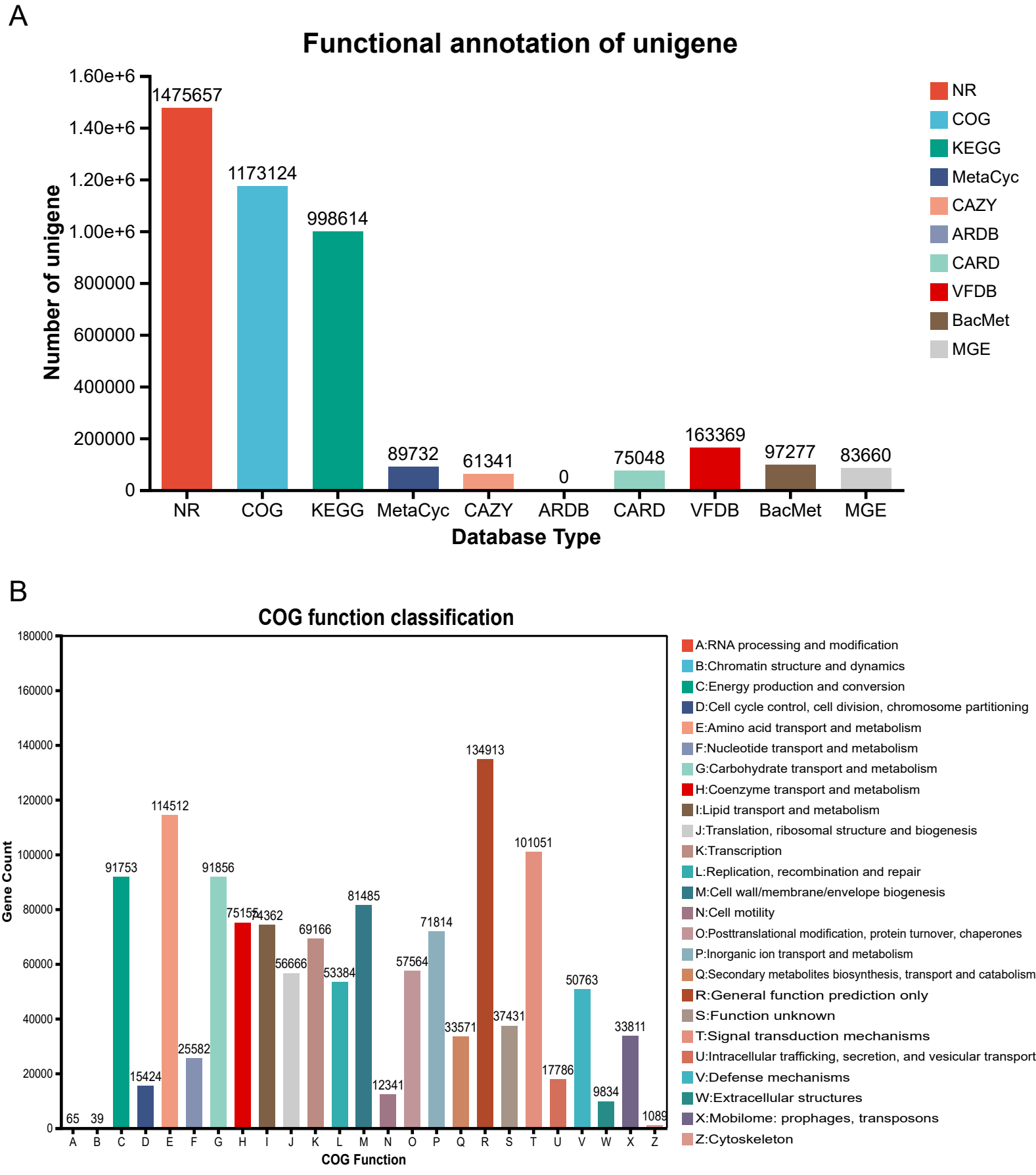


Figure S8. Overview of species and functional annotation in metagenomic analysis. (A) Statistical summary of metagenomic annotation based on multiple databases. (B) Bar chart showing the distribution of COG annotations. The x-axis represents different COG functional categories, and the y-axis represents the number of genes (Gene_Count) annotated to each functional category.

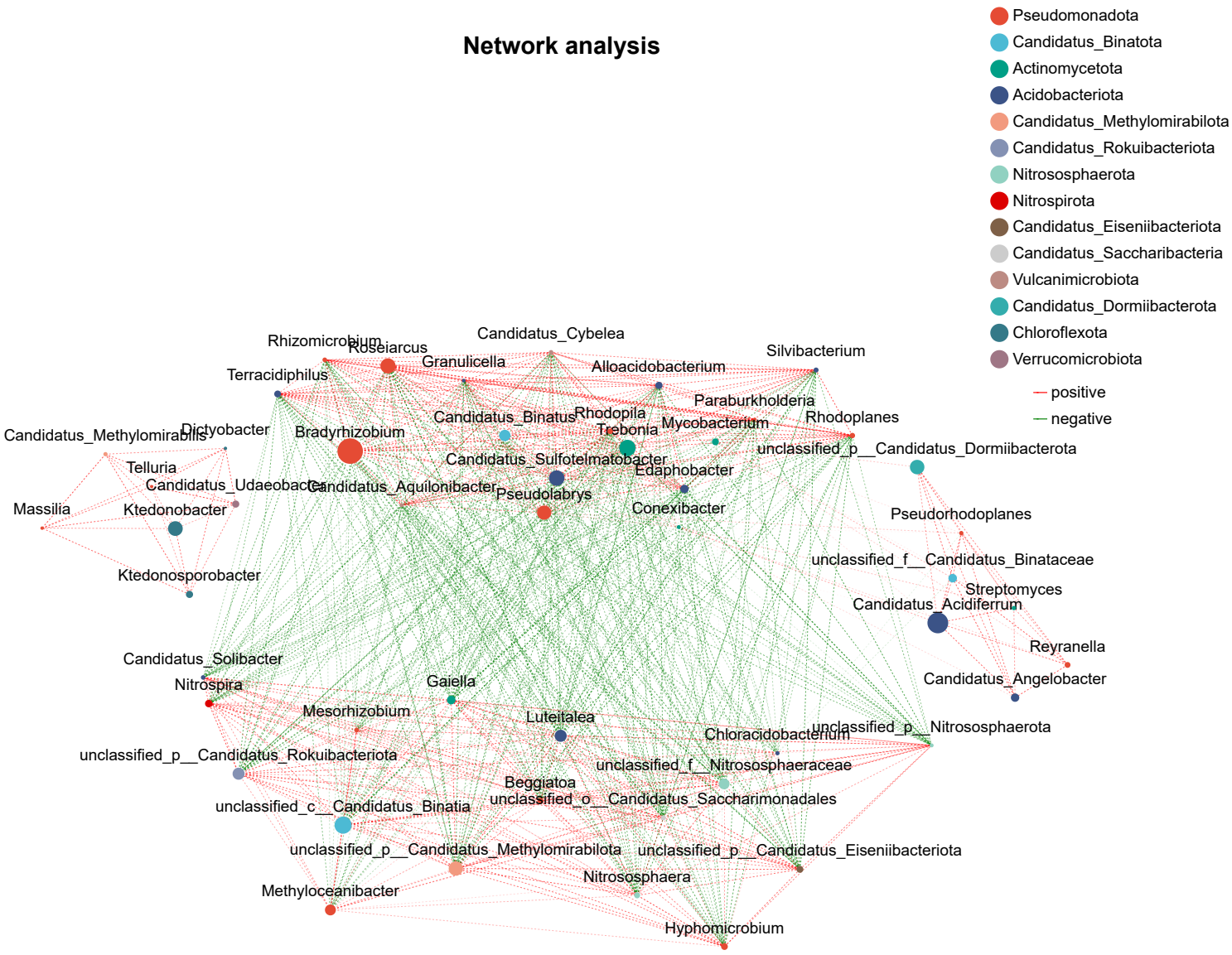


Figure S9. A single-factor correlation network diagram reveals the correlations among different microbial genera. The size of the nodes in the figure represents species abundance; larger nodes correspond to higher abundance values. Node colors represent different phyla (see legend in the upper right corner). The color of the connecting lines indicates positive or negative correlation: red indicates a positive correlation between species, while green indicates a negative correlation. The thickness of the lines indicates the magnitude of the correlation coefficient; thicker lines indicate a stronger correlation between species. A greater number of lines indicates a closer connection between a species and other species. Clearly, the genera in the upper half of the figure generally exhibit a negative correlation with those in the lower half.

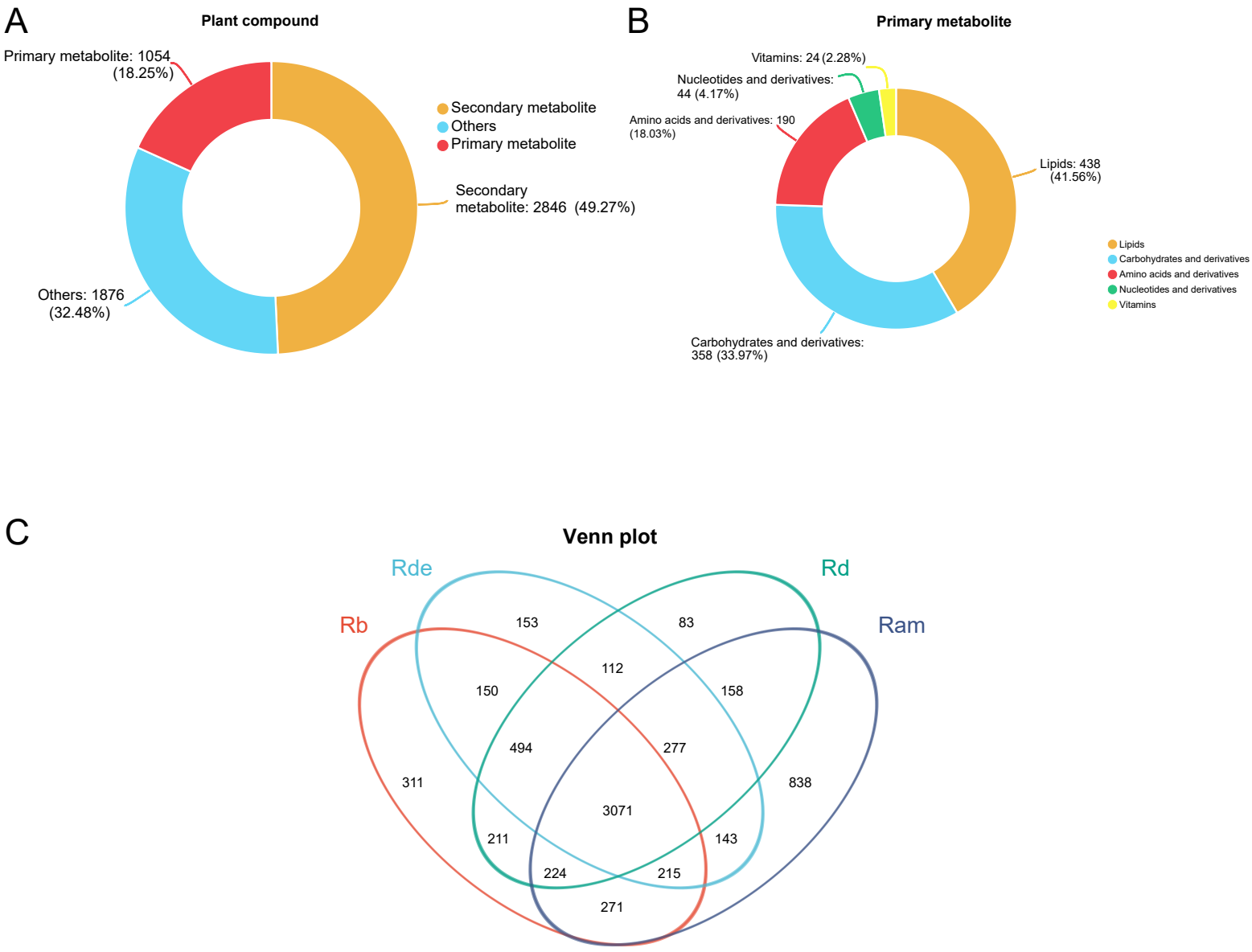
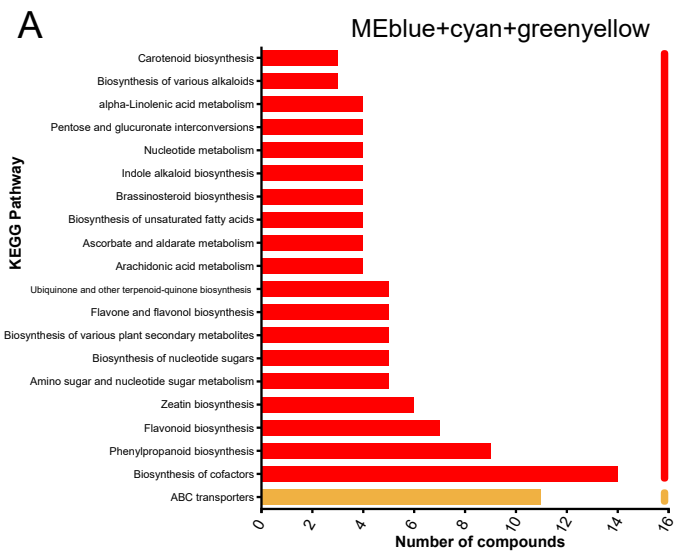
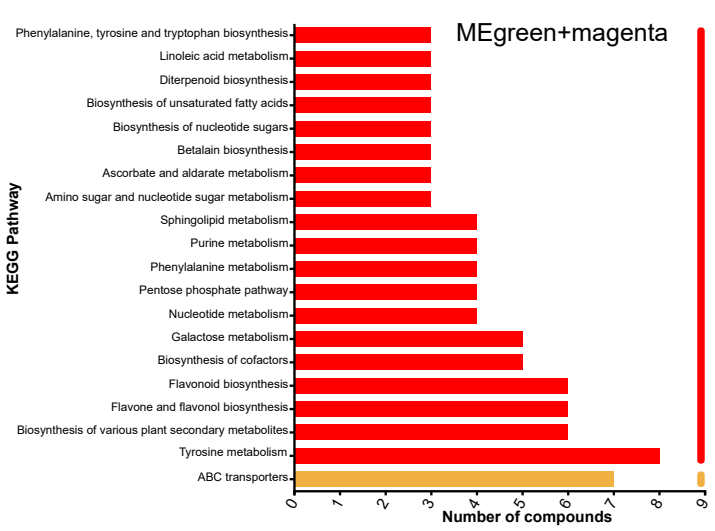
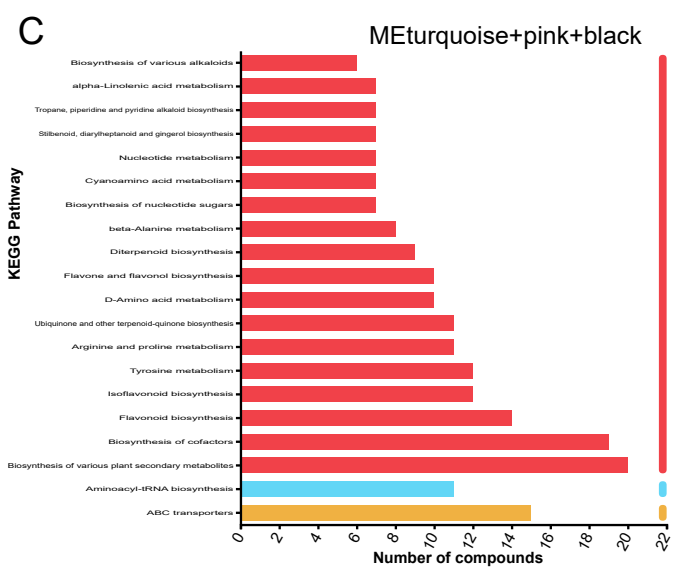
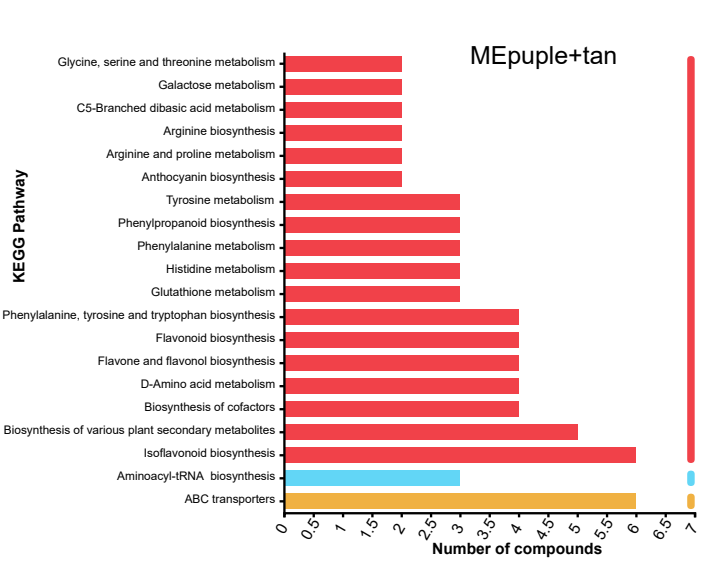


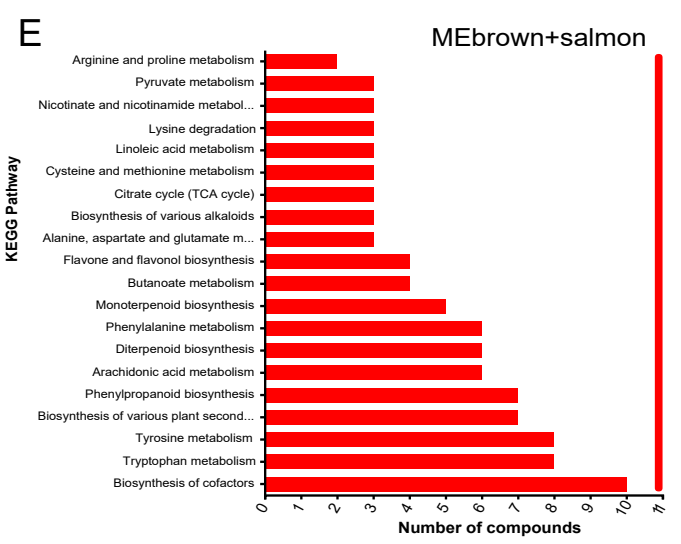
Figure S10. Metabolite annotation profiles in the leaves of four *Rhododendron* species (*R. bailiense*, *R. delavayi*, *R. decorum*, *R. ambiguum*). (A) Distribution and proportion of total plant compounds. (B) Distribution and proportion of primary metabolites. (C) A Venn diagram showing the distribution of shared and unique metabolites among the four rhododendron species. Among them, *R. delavayi* and *R. decorum* are two rhododendron species sharing the same habitat. Meanwhile, *R. bailiense*, *R. delavayi*, and *R. ambiguum* belong to three distinct habitats.

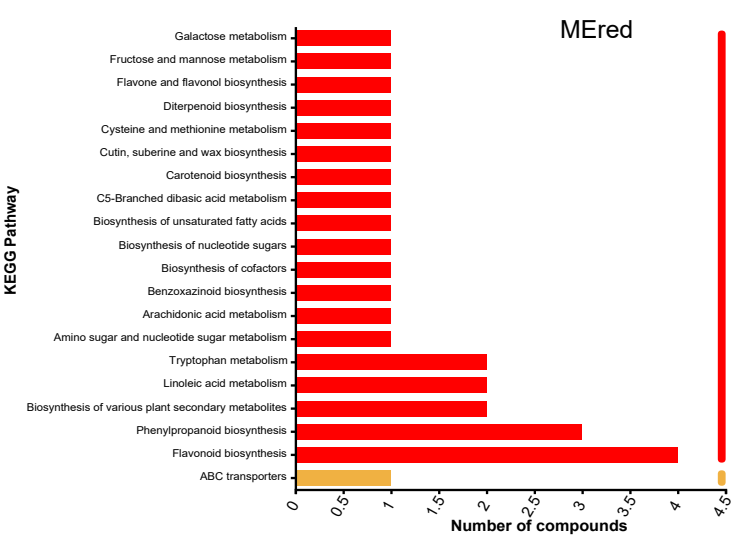


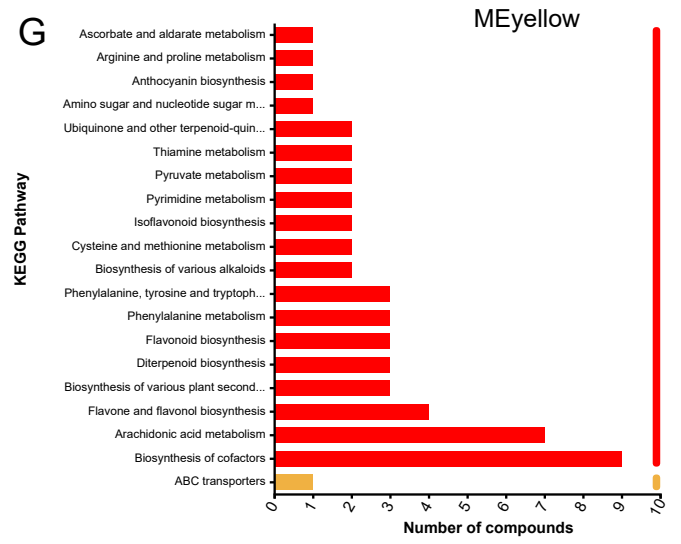












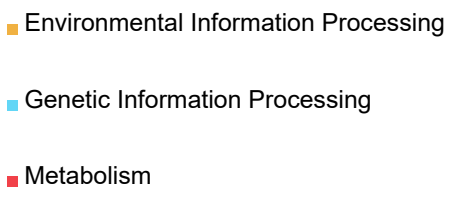


Figure S11. KEGG enrichment analysis of metabolites associated with different modules in the WGCNA analysis. (A) KEGG enrichment analysis of the MEblue + MEcyan + MEgreenyellow modules (corresponding to the major dominant metabolites of *R. bailiense*). (B) KEGG enrichment analysis of the MEgreen + MEmagenta modules (corresponding to some positively associated metabolites of *R. delavayi* and *R. decorum*). (C) KEGG enrichment analysis of the MEturquoise + MEpink + MEblack modules (corresponding to the major dominant metabolites of *R. ambiguum*). (D) KEGG enrichment analysis of the MEpuple + MEtan module (corresponding to some positively associated metabolites of *R. delavayi*, *R. ambiguum*, and *R. decorum*). (E) KEGG enrichment analysis of the MEpuple + MEtan module (corresponding to some positively associated metabolites of *R. bailiense*, *R. delavayi*, and *R. decorum*). (F) KEGG enrichment analysis of the MEred module (corresponding to the major dominant metabolites of *R. delavayi*). (G) KEGG enrichment analysis of the MEyellow module (corresponding to the major dominant metabolites of *R. decorum*).

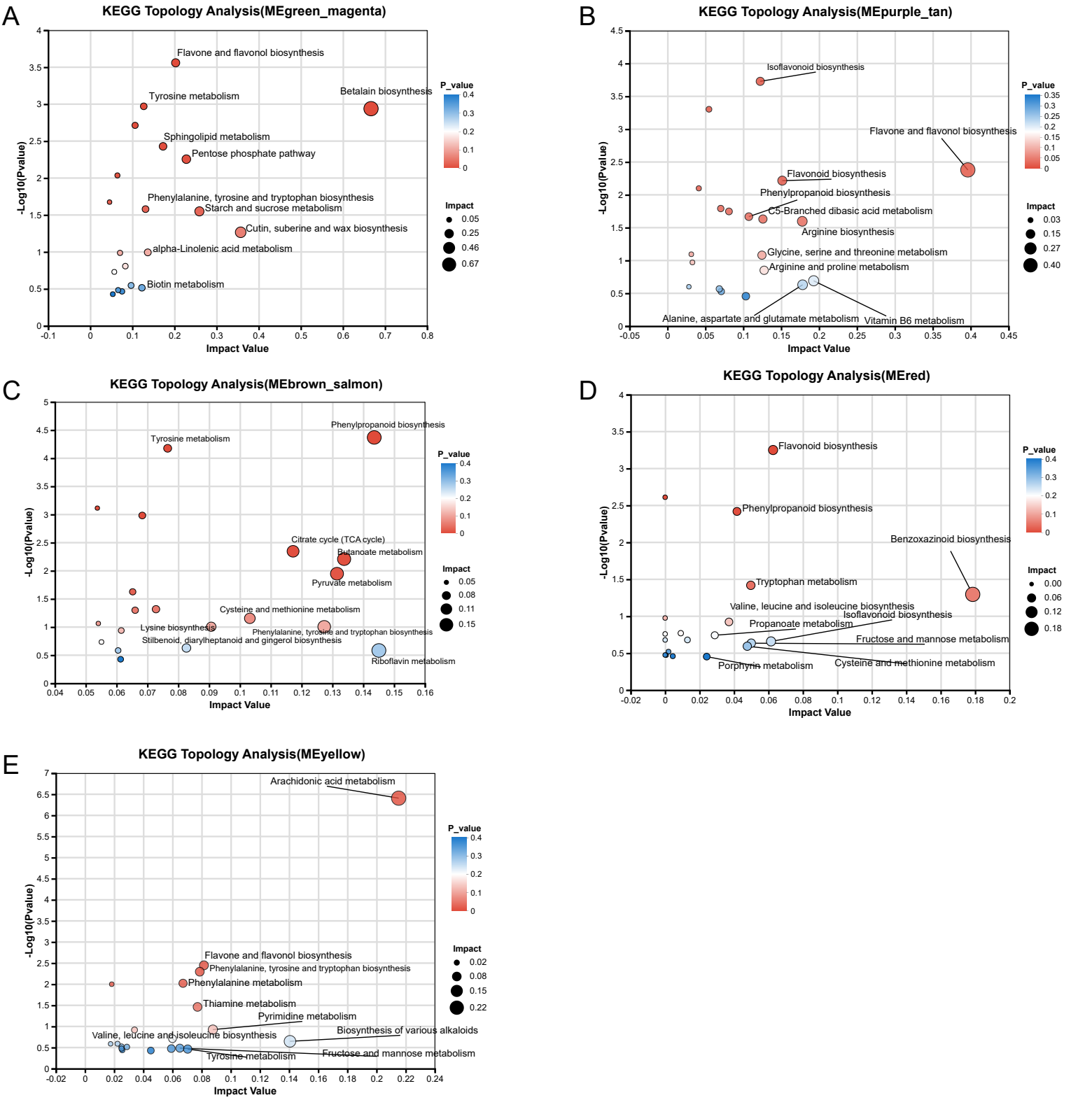


Figure S12. KEGG topology analysis of different modules identified in the metabolite WGCNA analysis. (A) KEGG topology analysis of the MEgreen + MEmagenta modules (corresponding to some positively correlated metabolites in *R. delavayi* and *R. decorum*). (B) KEGG topology analysis of the MEpuple + MEtan module (corresponding to some positively associated metabolites of *R. delavayi*, *R. ambiguum*, and *R. decorum*). (C) KEGG topology analysis of the MEpuple + MEtan module (corresponding to some positively associated metabolites of *R. bailiense*, *R. delavayi*, and *R. decorum*). (D) KEGG enrichment analysis of the MErred module (corresponding to the major dominant metabolites of *R. delavayi*). (E) KEGG topology analysis of the MEyellow module (corresponding to the major dominant metabolites of *R. decorum*).

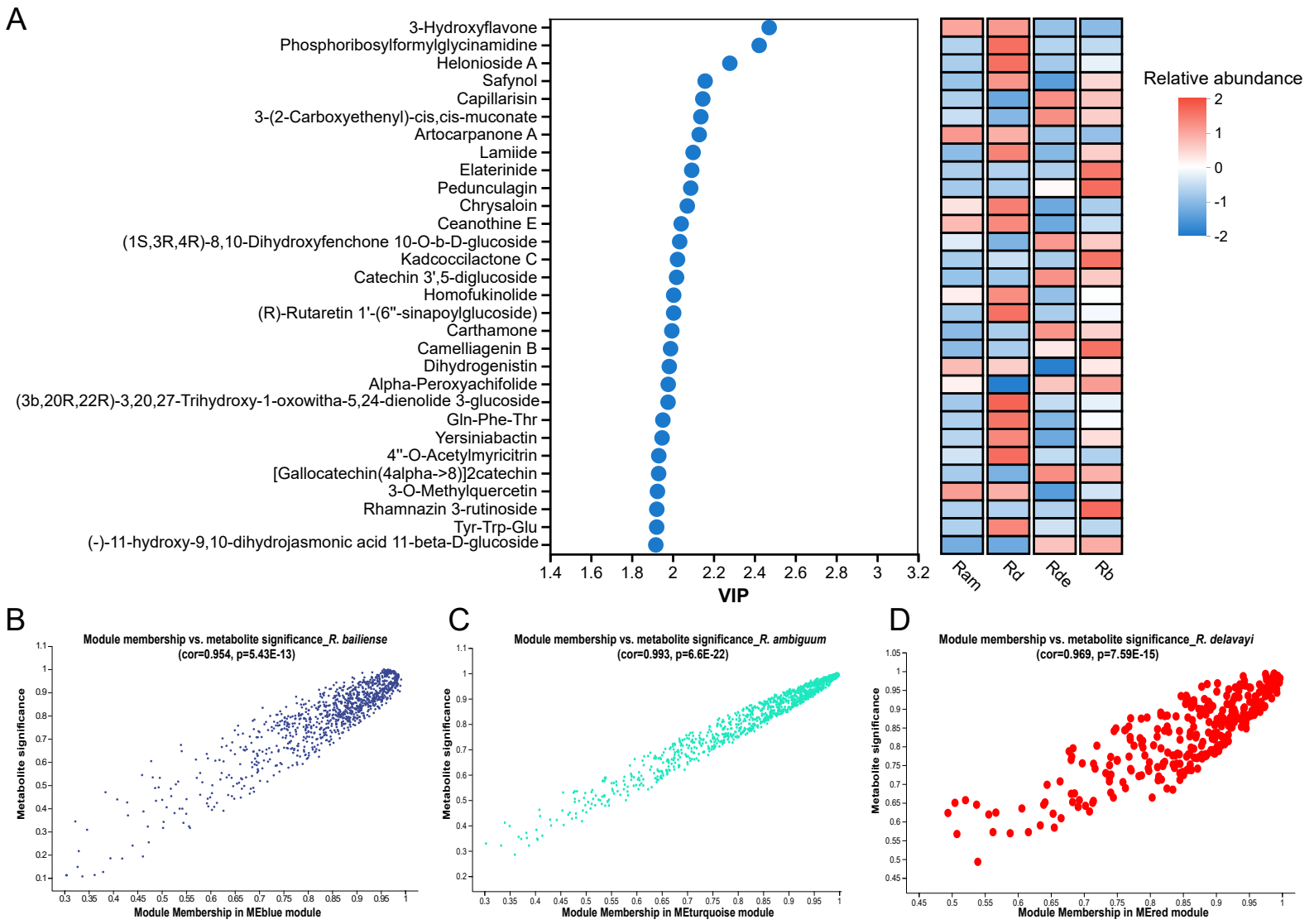


Figure S13. Identification of dominant metabolites interacting with microorganisms. (A) The heatmap displays the top 30 differentially expressed metabolites (DAMs) with the most significant differentiation in accumulation patterns among the four rhododendron species. This figure combines Variable Importance in Projection (VIP) scores with relative abundance (heatmap) for a comprehensive evaluation. The relationship between module membership and gene significance for (B) metabolites in MEblue (corresponding to the major dominant metabolites of *R. bailiense*), (C) MEturquoise (corresponding to the major dominant metabolites of *R. ambiguum*), and (D) MEred (corresponding to the major dominant metabolites of *R. delavayi*) .

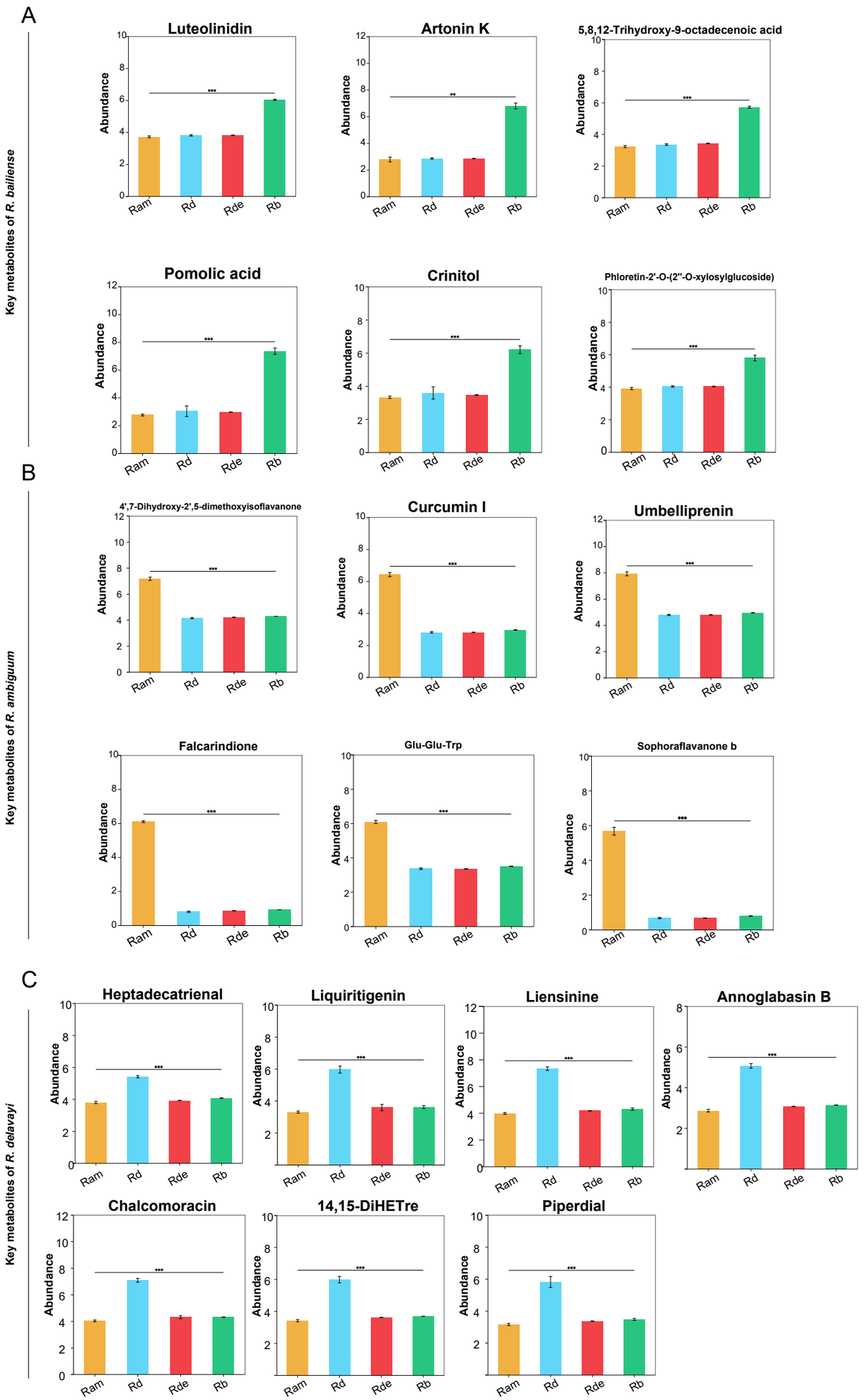


Figure S14. Dominant metabolites identified in each species that have known interactions with microorganisms. (A) Major dominant metabolites identified in *R. bailiense* that interact with microorganisms. (B) Major dominant metabolites identified in *R. ambiguum* that interact with microorganisms. (C) Major dominant metabolites identified in *R. delavayi* that interact with microorganisms. Significance was assessed using the Kruskal-Wallis H test (***, P < 0.001). Raw data on metabolite accumulation and metabolite-module associations are stored in Data S4.