

RESEARCH

Open Access



Modules in robust but low-efficiency phyllosphere fungal networks drive saponin accumulation in leaves of different *Panax* species

Guozhuang Zhang¹, Liping Shi¹, Congsheng Liu², Renjun Mao³, Bing Xia¹, Zhixin Huang², Xiuye Wei¹, Lixuan Wu², Yuqing Zheng², Guangfei Wei¹, Jia Xu¹, Shuangrong Gao¹, Shilin Chen^{1,4} and Linlin Dong^{1*}

Abstract

Background The phyllosphere mycobiome plays a crucial role in plant fitness and ecosystem functions. The complex microbial ecological networks (MEN) formed by these fungi remain poorly understood, particularly with regard to their organization strategy and their contributions to plant secondary metabolites such as saponin.

Results In this study, we constructed six MENs from leaf epiphytic and endophytic mycobiomes of three *Panax* species distributed in the northeast and southwest ends of mainland China. Hub nodes were absent in these MENs, which were significantly more complex, robust, and less efficient compared to random graphs ($P < 0.05$), indicating a hub-independent high-robustness strategy to maintain structural homeostasis. The important roles of specific MEN modules in shaping leaf saponin profiles of each *Panax* species were proved by multiple machine learning algorithms. Positive regulation modules (PRMs) of total saponin content were further identified, which exhibited more deterministic ecological assembly and comprised of highly connected nodes as well as higher proportion of plant-associated fungal guilds compared to other network members, indicating their tight links with host plant. The significant and direct effects ($P < 0.05$) of PRMs on total saponin accumulation were validated by well-fitted structural equation models ($\chi^2 < 0.3$, $P > 0.5$). Taxonomic analysis revealed that Pleosporales and Chaetothyriales were significantly overrepresented by positive regulation taxa (PRT) of total saponin content (FDR < 0.05). Across PRT identified in three *Panax* species, *Epicoccum* and *Coniothyrium* were conservatively present, while species-specific taxa such as Agaricales were also found, indicating the conservatism and specificity of plant-fungi interactions associated with leaf saponin accumulation in *Panax* genus.

Conclusions These findings provide a foundation for understanding mechanisms maintaining the steady state of phyllosphere mycobiome in healthy plant, and offer clues for engineering phyllosphere mycobiome to improve the accumulation of bioactive secondary metabolites on the basis of network modules.

Keywords Microbial ecological network, Network robustness, Community assembly, Plant-fungi interaction, Plant secondary metabolite

*Correspondence:

Linlin Dong

lldong@cmm.ac.cn

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Background

The phyllosphere is the aerial parts of plants, and accounts for approximately 60% of the total biomass on Earth, providing habitats for diverse microorganisms such as bacteria and fungi [5, 71]. Among microbial lives in phyllosphere, leaf-associated microbiomes, including those residing on leaf surface (i.e., epiphytes) or in leaf tissues (i.e., endophytes), plays crucial roles in plant health, growth, and biogeochemical cycles [58]. While leaf bacterial communities have been extensively studied in term of mechanisms underlying community assembly and homeostasis, as well as their effects on plant immune response and carbon emission [1, 13, 44], knowledge about leaf mycobiomes is still lacking, particularly with regard to the maintenance strategy of community structure and their contributions to plant secondary metabolism [58, 75]. This knowledge is important for a more comprehensive understanding of phyllosphere biosphere and its role in plant health and ecology.

Intensive exchanges of substances, energy, and information among microorganisms in the phyllosphere shape complex interspecies interactions that can be represented by microbial ecological networks (MEN) [81]. Revealing the organization strategy of MEN can help us understand the potential mechanisms underlying the maintenance of community homeostasis [74]. Two strategies may be employed by microbial communities at the MEN level to adapt to phyllosphere of healthy plant and maintain their structure. The first strategy is high-efficiency. Efficiency is an important property measuring how fast information spreads across network [42]. A high-efficiency network relies on hub nodes which can shorten path length between any nodes, enabling external perturbation to distribute rapidly throughout the entire graph [32, 74]. Regarding the highly metabolic plasticity of fungi, the fast perception of disturbance such as species loss by fungi may allow them change gene expression profiles to adapt to abiotic or biotic changes timely [20, 54, 64]. A dynamic homeostasis can thus be achieved by high-efficiency fungal MENs. The second strategy is high-robustness, which maintains network structure by avoiding the spread of perturbation (e.g., node loss) across network. For instance, nodes can group separately, isolating the adverse effects caused by perturbation in a local area. The high-robustness network may also abandon hub nodes to prevent the destructive effects of hub loss on network structure, at the cost of reducing network efficiency [9]. It is unclear which strategy is preferred by phyllosphere mycobiomes. Given the harsh habitat of the phyllosphere with limited nutrients which may be insufficient to support efficient communication and rapid metabolic response of fungi, we hypothesize that phyllosphere fungal networks apply the hub-independent high-robustness

strategy to organize themselves and maintain structure homeostasis (H1).

Phyllosphere microorganisms have close interactions with host plants [6], [58]. However, our knowledge of the effects of phyllosphere fungi on plant secondary metabolites remains limited. In the rhizosphere and root compartments, researches have shown the significant effects of microbiomes on the accumulation of plant secondary metabolites, especially saponins such as ginsenoside and glycyrrhizin [38, 66], [78]. Saponins are inducible phytoanticipins that play a crucial role in plant defense against bio-invasion, and the biosynthesis of which can be triggered by plant immune responses [23, 52]. As the plant immune response is also effective in the phyllosphere, we expect that certain members of the phyllosphere mycobiome can promote the accumulation of specific and/or total saponin in leaves. MENs typically exhibit modular structure in which nodes are closely connected by sharing environmental or resource preferences [18]. Network modules are thought to represent important ecological units, which may have significant implications for biological or ecological functions [37, 55, 67]. We thus hypothesize that certain MEN modules in the phyllosphere contribute to the formation of saponin profiles and the accumulation of total saponin in plant leaves (H2).

In this study, we aimed to test the above two hypotheses using a total of 162 fungal communities collected from the leaf endosphere (LE) and phylloplane (LP) of three *Panax* species: *P. ginseng* (PG), *P. quinquefolium* (PQ), and *P. notoginseng* (PN) [78]. Although *Panax* genus consists of about 20 species, the three species chosen in this study are especially famous for their wide consumption as dietary supplements, functional food, and medicinal materials for thousands of years, thus becoming the representative and best-selling ginseng species [28]. The three *Panax* species are closely related but distributed distantly, containing diverse ginsenosides as the major bioactive secondary metabolites [28]. By exploring the effects of phyllosphere fungi on saponin metabolism, we hope to improve the production of these pharmacologically important saponins through utilizing foliar microorganisms. This type of exploration has been carried out in several plant species [62]. For instance, the inoculation of *Epichloe festucae* raised the expression level of phenylpropanoid biosynthetic pathway in ryegrass, favoring the biosynthesis of secondary metabolites like flavonoids and anthocyanins [21]. Aly et al. [2] proposed the potential of fungal endophyte acting as biological trigger to increase the production of specific secondary metabolites of host plant. Combined with metadata including the contents of diverse types of saponins in leaves and edaphic properties, this study aims to: (i) estimate the

strategies employed by fungal MENs to maintain community homeostasis, i.e., hub-dependent high-efficiency strategy or hub-independent high-robustness strategy, (ii) identify modules in MENs and their links to saponin profiles and total saponin accumulation in *Panax* leaves; and (iii) explore the relationships among taxa that potentially promoting saponin accumulation across the three *Panax* species to gain a better understanding of the evolutionary characteristics of plant-fungi interactions associated with plant chemical response.

Methods

Collection of samples and metadata

The sample collection has been detailedly described in a previous study [78]. Specifically, *P. ginseng* (PG) and *P. quinquefolium* (PQ) plants were sampled from Jilin Province in the middle of September of 2019, with the mean annual temperature (MAT) ranging from 2.9 to 3.2 °C. While *P. notoginseng* (PN) plants were collected from Yunnan Province in late October of 2019, with the MAT ranging from 16.3 to 16.7 °C (Additional file 1: Fig. S1a). All plants were planted in fields through seed rather than transplanting, and all fields were managed following the Good Agricultural Practice [76], [79]. For each *Panax* species, ninety healthy plants were collected from each of three plantations cultivated with 2-year-old, 3-year-old, and 4-year-old plants, respectively. In each plantation, ten plants representing one sample were sampled from one of nine 2.0 × 1.5 m² plots. A total of 27 samples representing 270 plants were obtained for each *Panax* species.

Fungi residing on leaf surface (LP) were collected according to Yao et al. [72]. Simply, plant leaves were placed into sterile 50-ml centrifuge tubes and immersed in sterile ultrapure water. After alternating sonication and vortex, the suspensions were centrifuged at 10,000 × g for 10 min, and the sediments were collected for DNA extraction and LP mycobiome profiling. After surface sterilized, the treated leaves were stored at −80 °C for profiling of mycobiome in leaf endosphere (LE) and saponin quantification. Additionally, the bulk soil samples (BS) were collected from each plot for measuring edaphic factors using standard methods, including pH, organic carbon (OC), total nitrogen (TN), nitrate and ammonium nitrogen (NIN & AMN), available phosphorus (AP), available potassium (AK), exchangeable calcium (ECa) and magnesium (EMg), available sulfur (AS), and moisture content (MC) (Zhang et al., 2022). As the mean annual temperature (MAT) is similar among three plantations of each plant species, we did not account for temperature in subsequent analysis (Additional file 1: Fig. S1a). In addition, the soil MC rather than mean annual precipitation was used for estimating water conditions

of plant, as all plants were cultivated in sunshade (Additional file 1: Fig. S1b).

For saponin quantification, 0.2 g well-mixed dried powder of leaves was weighed accurately and extracted using 15 ml pure methanol. After centrifuging and filtering, the supernatant was injected into HPLC system (Waters Corporation., Milford, MA, USA). The elution gradient was set as follows: 0–12 min, 19% Acetonitrile (A) and 81% ultrapure water (B); 12–70 min, 19–36% A; 70–71 min, 36–19% A; 71–76 min, 19%A (Zhang et al., 2022). The retention times and contents of major saponins, including ginsenosides Rb1, Rb2, Rc, Rd, Re, Rg1, and F1 and notoginsenosides R1, were determined using corresponding standard substances (Shanghai Yuanye Bio-Technology Co., Ltd., Shanghai, China).

DNA extraction and amplicon analysis

Total DNA was extracted from phylloplane precipitates and leaf tissues using FastDNA SPIN Kit for soil (MoBio Laboratories, Inc., Carlsbad, CA, USA). The ITS1F/ITS2R primer pairs were used to amplify fungal sequences, which were then sequenced by Illumina MiSeq PE 300 platform (Shanghai Biozeron Co., Ltd., Shanghai, China). Low quality reads and adaptors in raw sequences were removed using the fastp software (default parameters) [12]. The clean data were then submitted to the cutadapt software to remove the primes with the following parameters: “−errors 0.13 || −overlap 5” [45]. The obtained sequences were imported to QIIME2 using the import plugin [8]. The dada2 denoise-paired command was applied to denoise the sequences to generate amplicon sequence variants with the following parameters: “−p-trunc-len-f 220 || −p-trunc-len-r 220” [11]. The classify-sklearn method in QIIME2 was applied to perform taxonomy annotation using bayes-trained classifier on the basis of UNITE database (V16.10.2022) with the default parameter [48]. Unclassified and singleton ASVs were removed from our data, and remaining ASVs were inspected for plant DNA contamination using blast analysis. The obtained ASV table was then resampled to 12,412 sequences per sample to eliminate the effect of sequence depth difference.

Statistical analysis

All statistical analyses were conducted in *R* (v4.2.2). *R* codes used for statistical analyses are available at <https://github.com/githubzgz>. All *P* values in multiple comparisons were adjusted according to Benjamini and Hochberg [7], and the adjusted *P* values were displayed as FDR.

MENs construction

Fungal MENs were constructed based on Spearman correlations among ASVs, following a random matrix theory

(RMT)-based method to determine the correlation cut-off threshold in a non-arbitrary way [74]. This strategy has been widely applied in microbial ecology researches, greatly promoting our understanding of interaction structure, environmental response, and homeostasis maintenance of microbial communities in diverse habitats [22, 55, 70, 80]. Simply, the RMT-based method estimates nonrandom properties of complex systems like fungal communities from random noise, using the nearest-neighbor spacing distribution of eigenvalues [19], [82]. The transition of eigenvector distribution from GOE (Gaussian orthogonal ensemble) to Poisson distributions is typically used as a reference point to identify system-specific and nonrandom properties embedded in high-dimensional community data. The RMT-based method keeps us away from arbitrary selection of correlation cut-off, which is commonly used in microbial network studies. In the present study, fungal ASVs with relative abundances higher than 0.01% were selected to calculate Spearman correlations to reduce spurious associations caused by rare taxa [31]. After removing non-significant correlations ($FDR > 0.05$), the eigenvalue distribution of correlation matrix was analysed using the *RMThreshold* package to determine the transition point [39]. The matrix was then filtered based on transition point and the *igraph* package was used to construct network [16]. A total of six networks were constructed (three plant species \times two compartments) (Additional file 1: Table S1).

Drivers of MENs

In addition to biotic interaction, environmental filtering and dispersal limitation are also considered as drivers of species co-occurrence or co-exclusion in MENs [36]. A recently developed method, named LTED (Link Test for Environmental filtering or Dispersal limitation), was applied to estimate the potential contributions of environment and space factors to observed links in MENs [68], [74]. In brief, if a link between two ASVs (i.e., Spearman correlations stronger than transition point) results from their covariations with one or more environmental factors, it is considered to be indirect and represent the influence of environmental filtering. For strength of dispersal limitation, all nodes in MENs were tested for their Mantel correlations with spatial distance based on Bray–Curtis dissimilarities [74]. If two linked nodes were both strongly affected by dispersal limitation ($r \geq 0.6$, $P < 0.05$), their link is considered to result from dispersal limitation. Links were identified as the result of potential bio-interactions if they did not meet the above criteria.

MEN efficiency and robustness

To estimate the organization strategy employed by fungal MENs to maintain homeostasis, we estimated a series

of topological properties for observed MENs and random graphs which had the same number of nodes and links with the observed graph. Specifically, the average shortest path length (avgL) and global efficiency were calculated to represent network efficiency. Smaller avgL and higher global efficiency indicated a potentially more efficient network. The natural connectivity was calculated for complete graphs and graphs with 20%, 30%, and 40% of nodes being randomly removed to measure structural robustness [4], [33]. Mean value and standard deviation of the above topological properties were calculated according to 100 random graphs generated based on Erdős–Rényi model using the *sample_gnm* function in the *igraph* package [56]. Deviations greater than 1.96 standard deviations from mean value of random graphs were considered as significant differences between observed and random graphs ($P < 0.05$). Other network properties, including degree centralization, clustering coefficient, and modularity (fast greedy algorithm), were also calculated for true and random graphs to estimate network complexity [41], [56]. To estimate whether fungal MENs were dependent on hub nodes, the z_i (within module degree) and P_i (participation coefficient) scores were calculated for each node according to Guimera and Nunes Amaral [26]. Nodes were then categorized as kinless hubs ($z_i > 2.5$, $P_i > 0.62$), provincial hubs ($z_i > 2.5$, $P_i \leq 0.62$), connectors ($z_i \leq 2.5$, $P_i > 0.62$), or peripherals ($z_i \leq 2.5$, $P_i \leq 0.62$) (Shi et al., 2020). Kinless and provincial hubs were considered as hub nodes.

Contribution of modules to leaf saponin profiles

For each *Panax* species, the differences of saponin profiles in leaves were represented by the first principal components (PC1), as they accounted for the major variations in contents of various saponins (PG: 59.07%; PQ: 73.91%; PN: 94.53%). The principal components analysis (PCA) of saponins were conducted using the *rda* function in the *vegan* package [50]. As the saponin contents were obtained using the absolute quantification method based on standard substances rather the relative ion intensity, we did not perform scale transformation in this analysis. Modules in MENs were detected using the *cluster_fast_greedy* function in the *igraph* package [15]. The fast greedy algorithm has been widely used in MEN analysis [19, 42]. Module eigengene was calculated for modules with 5 or more nodes in all fungal MENs, representing the variations in relative abundance of all ASVs within certain module [82]. Spearman correlation analysis was firstly performed between module eigengenes and saponin variations to identify candidate modules ($FDR < 0.05$) that potentially drove accumulation of different types of saponins. Candidate environmental factors were also selected using the same method. Before

the analysis, environmental variables besides soil pH were log-transformed according to needs to improve normality, and variables with strong collinearity were removed (Spearman $\rho^2 > 0.6$) [29]. As our previous study has shown the minor role of plant growth year in shaping mycobiome structures and saponin variations [78], and *Panax* leaves fell and regrew annually, we did not account for growth year here. To estimate that whether the incorporation of candidate fungal modules could improve our ability to predict saponin profiles, a machine learning framework was applied [51]. Simply, two machine learning algorithms were used for mutual authentication, including random forest analysis and *K*-nearest neighbors regression. Model performances were evaluated by mean squared error (MSE) and Spearman ρ between true observed response and predicted response. Lower MSE or higher Spearman ρ indicated a better model. Performance measures were calculated based on fivefold resampling, and the aggregated scores were extracted. Four models were fitted for each machine learning algorithm, in which saponin PC1 was used as response, and candidate environmental factors (Env), Env and candidate module eigengenes in LE networks (EnvLE), Env and candidate module eigengenes in LP networks (EnvLP), as well as Env and all candidate module eigengenes (EnvALL) were used as features, respectively. Machine learning-based analysis was conducted using the mlr3 package [34]. The `as_task_regr` function was used to construct machine learning tasks. The learners, resampling strategies, and model evaluator were created using the `lrn`, `rsmp`, and `msrs` functions, respectively. The `resample` function was applied to conduct the resampling procedures.

Identification of positive regulation modules and nodes

In modules significantly correlated to saponin variations (i.e., the candidate modules), those containing half of the number of nodes that exhibited significant and positive links to total saponin content (i.e., the sum of different types of saponins) were identified as positive regulation modules (PRMs) of total saponin. Node degree were compared between PRMs and other nodes in the same network to estimate the topological importance of PRM members. For the comparison of functional composition between PRMs and other nodes in the same network, we predicted the potentially functional guilds of MEN nodes using FUNGuild [47]. Only probable and highly probable predictions were retained in this analysis [18]. The community assembly processes of subcommunities comprising of PRM members and other nodes were examined using a taxonomic null model called Normalized Stochasticity Ratio (NST), respectively [49]. The NST measures the relative contribution of stochastic processes

(e.g., dispersal process and ecological drift) to community assembly, which was calculated using the tNST package [49, 59]. A lower NST value indicates that the community is assembled by more deterministic processes such as selection [49, 59].

A series of structural equation modules (SEMs) were then fitted to test the direct effects of PRMs on total saponin accumulation. The candidate environmental variables and PRM eigengenes were incorporated into SEMs. The fit-of-goodness was estimated based on χ^2 test, comparative fit index (CFI), and root mean square error of approximation (RMSEA). The SEM analysis was performed using the lavaan package [53]. In SEM-validated PRMs, fungal nodes significantly and positively correlated to total saponin content were identified as positive regulation taxa (PRT). An order-level enrichment analysis was performed to assess which order was significantly overrepresented by PRT using the `phyper` function in `stats` package. Simply, the ratio of PRT in each fungal order (i.e., the number of PRT in each fungal order/the number of all ASVs in MENs belonging to that order) was compared with the ratio of that order in MENs (i.e., the number of all ASVs in MENs belonging to that order / the number of all ASVs in MENs). Fungal phylogenetic tree was constructed based on taxonomic information using the `taxonomy_to_tree.perl` script developed by Tedersoo et al. [61]. The specificity and conservatism of these PRT across three *Panax* species were further visualized based on the tree using the `ggtree` package [73].

Results

Potential bio-interactions were main driver of links in fungal MENs

A total of six networks were constructed for three plant species and two compartments, with node number ranging from 63 to 153 and link number ranging from 199 to 530 (Fig. 1a, Additional file 1: Table S1). All MENs were dominated by positive links, which accounted for 60.80–85.35% of all edges (Fig. 1a). The distribution of node degree in all MENs followed a power law (Kolmogorov–Smirnov test, $P > 0.05$) in upper tail of degree higher than median, indicating the weakest scale-free according to the classification system developed by Broido and Clauset [10]. In which, only the network recovered from the LE of PQ exhibited higher than 50 nodes covered by power law and could be categorized as weak scale-free (Additional file 1: Fig. S2). Compared to phylloplane networks, the endosphere MENs exhibited lower graph density in PG and PQ, while the opposite result was observed in PN (Additional file 1: Table S1).

The potential drivers of network links were further estimated by LTED. In all MENs, only 4.53% to 34.27% of all edges could be assigned to environmental filtering

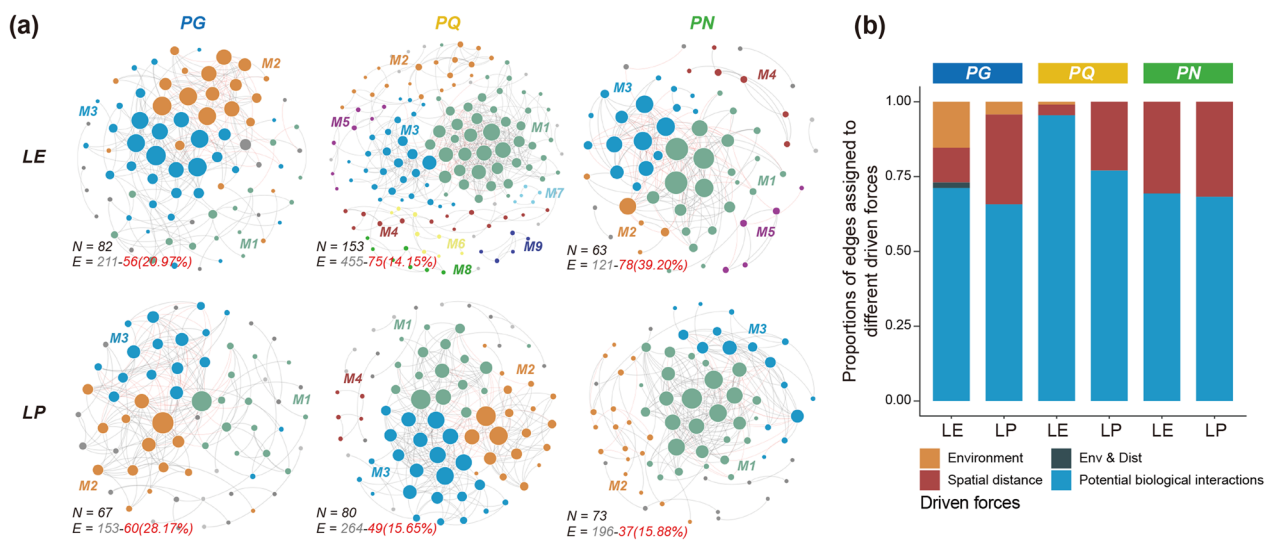


Fig. 1 Fungal MENs and the relative contribution of different drivers to network links. **a** Six fungal networks constructed for two compartments of three *Panax* species. Node colors represent modules with 5 or more nodes. Gray and red links represent positive and negative correlations stronger than threshold, respectively. Digits displayed on left bottom of networks represent the number of nodes, positive and negative links, respectively; **b** The proportion of edges assigned to different driven forces calculated according to LTED. Environment represents edges which can be attributed to covariations of ASVs to environment variables. Spatial distance represents edges that can be attributed to dispersal limitation. Env & Dist represents edges driven by both environmental filtering and dispersal limitation. Remaining edges are assigned to the force of potential biological interactions

and/or dispersal limitations, implying the dominated role of biological interactions in shaping networks (Fig. 1b). However, given the limitations of the environmental factors we measured and the nature of the correlation-based network approach, we still treated the links in MENs as putative bio-interactions. In summary, these results indicated that fungal MENs in *Panax* phyllosphere were weak scale-free networks, in which fungal interactions potentially contributed the most to observed links.

Networks in *Panax* phyllosphere were robust but low-efficiency

Several complexity indices, including degree centralization, clustering coefficient, and modularity, were calculated for observed MENs and corresponding 100 random graphs. All these indices were significantly higher for true networks than random ones (higher than 1.96 standard deviations, $P < 0.05$), indicating the complex organization of phyllosphere mycobiomes (Additional file 1: Fig. S3). The homeostasis maintenance strategies of these MENs were further estimated. Compared with random graphs, the observed networks exhibited significantly lower global efficiency and higher avgL, which both indicated the low exchange efficiency among fungal nodes in *Panax* phyllosphere (Figs. 2a–b). In contrast, the comparison of natural connectivity between complete true and random networks indicated that the observed MENs were significantly more robust than random graphs

($P < 0.05$) (Fig. 2c). The remaining natural connectivity of MENs with 20%, 30%, and 40% of nodes being removed randomly were also significantly higher than those of random graphs (Fig. 2d, Additional file 1: Fig. S4). Additionally, a strong correlation of differences between observed and random natural connectivity with those between observed and random global efficiency was observed (Least square regression, $R^2 = 0.64$, $P = 0.05$), which might represent a trade-off between robustness and efficiency in phyllosphere networks of *Panax* plants (Additional file 1: Fig. S5).

The z_i - P_i framework was applied to identify hub nodes. Only few nodes, however, were categorized as connectors ($z_i \leq 2.5$, $P_i > 0.62$) in all MENs, and no fungal taxa were identified as provincial hubs ($z_i > 2.5$, $P_i < 0.62$), let alone kinless hubs ($z_i > 2.5$, $P_i > 0.62$) (Fig. 2e). In together, these results strongly support our hypothesis **H1**, meaning that the fungal MENs in *Panax* phyllosphere tend to apply the hub-independent high-robustness strategy to maintain homeostasis at the cost of network efficiency.

Network modules contribute to shaping saponin profiles in *Panax* leaves

In the six MENs, module eigengenes were calculated for 27 modules with 5 or more nodes to represent the overall variation pattern of all module members (Fig. 1a). Spearman correlation analysis indicated that the M1 and M2 in LE networks exhibited significant correlations

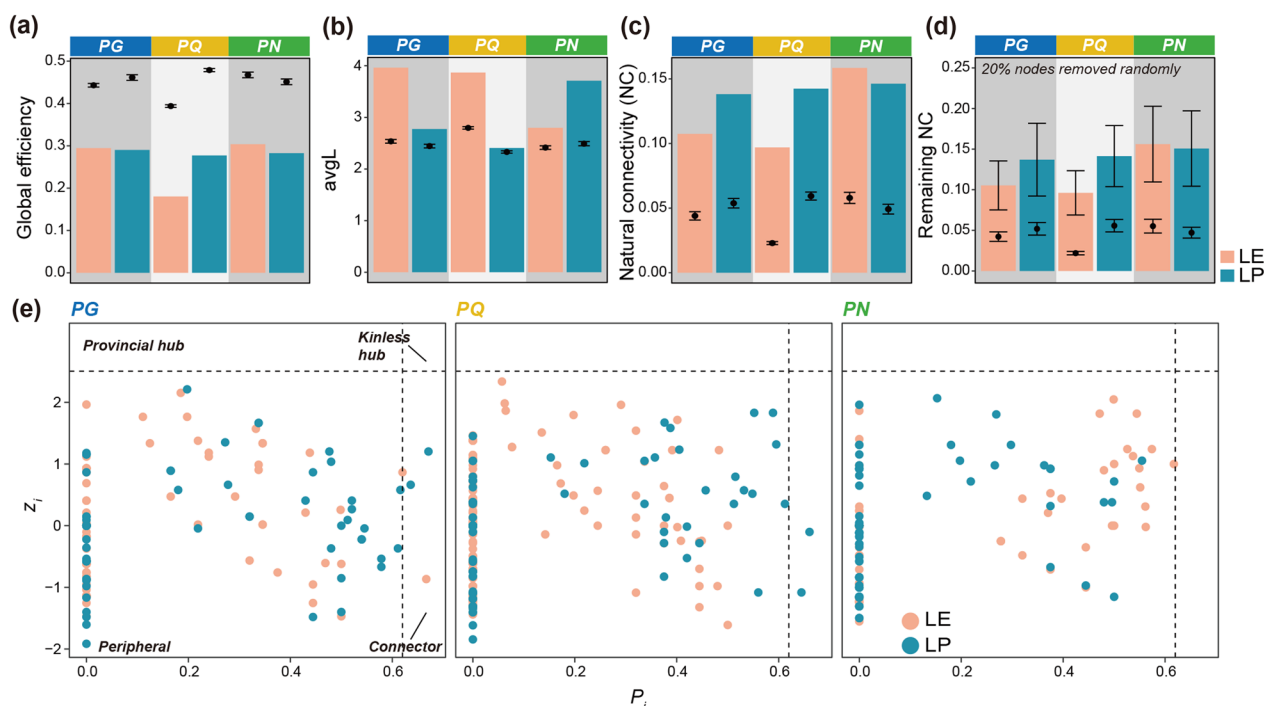


Fig. 2 Comparison of network properties between observed and random graphs and the z - P distribution of nodes. Global efficiency **a**, avgL (average shortest path length) **b**, natural connectivity of complete graph **c**, and remaining natural connectivity after 20% nodes removed randomly **d**. The bar plots represent the results of observed networks, while the error bar plots represent the mean and 1.96-fold of standard deviation. In **d**, the error bar of bar plot represents the standard deviation of 100 times random removal from observed network. **e** The z - P distribution of nodes in the six MENs. The vertical lines were 0.62 and the horizontal lines were 2.5, which represent the classification threshold of node roles

(FDR < 0.05) with saponin variations in PG, while the candidate modules in LP network were M1, M2, and M3 (Fig. 3a). In PQ, LE modules 1 and 3, as well as LP modules 1 and 3 were significantly correlated with PC1 of leaf saponin profiles (Fig. 3b). In PN, the candidate modules were identified as M1 and M5 in LE, as well as M1 in LP (Fig. 3c). Correlations between saponin PC1 and environmental factors were also determined. Soil OC, AK, NIN, MC, and Ca were candidates (FDR < 0.05) potentially contributing to saponin variations in PG (Fig. 3a). In PQ, soil pH, AMN, and MC exhibited significant correlations with leaf saponin PC1 (Fig. 3b). Soil pH, OC, AP, and AK were the selected environmental factors in PN (Fig. 3c).

A machine learning-based strategy was applied to evaluate the contribution of fungal network modules to saponin variations in *Panax* leaves. In K-nearest neighbors regression, the mean squared error (MSE) and Spearman ρ both indicated that models incorporated with candidate fungal modules, no matter from LE or LP, performed better than models only containing candidate environmental features in all three *Panax* species (Fig. 3d–e). The random forest regression also supported the above observation (Fig. 3d–e). These results supported the critical roles of specific phyllosphere fungal modules in shaping

saponin profiles of *Panax* leaves, providing evidence for our hypothesis **H2**.

3.4 Positive regulation modules of total saponin accumulation were assembled by more deterministic ecological processes, comprising highly connected nodes and more plant-associated guilds

Fungal ASVs in candidate modules identified above were further analysed for their correlations with contents of individual and total saponin (Additional file 1: Fig. S6). Modules with more than half of the number of positively correlated ASVs with total saponin content were identified as positive regulation modules (PRMs) of saponin accumulation, including M2 in LE and M3 in LP of PG, M3 in LE, M1 and M3 in LP of PQ, as well as M1 in LE and M1 in LP of PN (Additional file 1: Fig. S6). Members in these PRMs exhibited significantly (Wilcoxon rank sum test, $P < 0.05$) or nearly significantly ($P = 0.08$ for LE of PQ) higher degree than other nodes in the same MEN (Fig. 4a), indicating a more central position of these nodes in community structure. Besides, the proportions of plant-associated fungal guilds such as plant pathogen, endophyte, and epiphyte, were generally higher in PRMs compared to the proportions in other nodes,

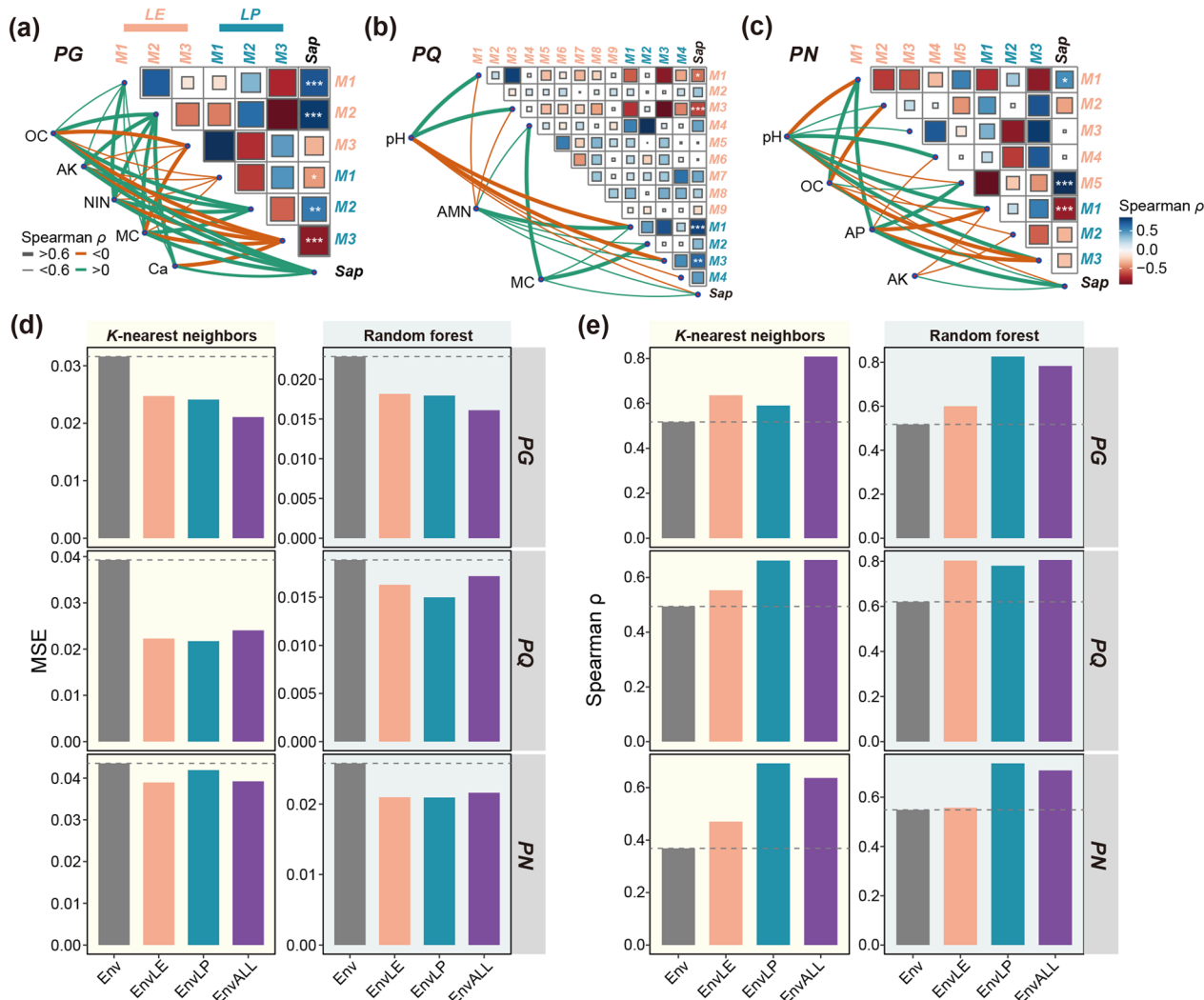


Fig. 3 Contributions of fungal network modules to leaf saponin profiles. Correlations between eigengenes of modules, saponin profiles (Sap, the PC1 of leaf saponins), and environmental variables in PG **a**, PQ **b**, and PN **c**. Lines represent significant Spearman correlations (FDR < 0.05) between environmental factors and biotic factors. Heatmaps represent Spearman correlations among fungal modules and saponin profiles. *, FDR < 0.05; **, FDR < 0.01; ***, FDR < 0.001. MSE **d** and Spearman ρ **e** measures of four models constructed using two machine learning algorithms, respectively. Env represent models only containing environmental factors as features. Module eigengenes of LE and LP networks were further incorporated in EnvLE and EnvLP models, respectively. EnvALL models containing all environmental and module factors as features

suggesting the tighter links between PRMs and *Panax* plant (Fig. 4b). The examination of ecological assembly mechanisms (NST) also supported the potentially tighter links between PRMs and host plant compared with other nodes, as the relative contributions of stochastic processes to the subcommunities assembly were higher for other nodes, indicating a more deterministic assembly of PRM members (Fig. 4c).

The SEM was fitted to estimate whether there were significant and direct effects of PRMs on total saponin accumulation based on the prior model (Fig. 5a). The χ^2 test, CFI and RMSEA statistics all supported the good fitting

of models of three *Panax* species (Fig. 5b–d). In PG, the M3 in LP (Effect size=0.585, $P < 0.05$) showed significant and direct effect on total saponin content (Fig. 5a). The M3 in LE (-0.998, $P < 0.001$) and M1 in LP (0.480, $P < 0.001$) were direct regulators on leaf total saponin in PQ (Fig. 5c). In PN, the M1 in LE (-1.372, $P < 0.05$) and M1 in LP (-1.714, $P < 0.01$) exhibited significant and direct effects on total saponin accumulation (Fig. 5d). Notably, the direct effects of environmental factors on total saponin content were rare and weak, highlighting the importance of fungal network modules in promoting leaf saponin accumulation (Fig. 5b–d). In general, these

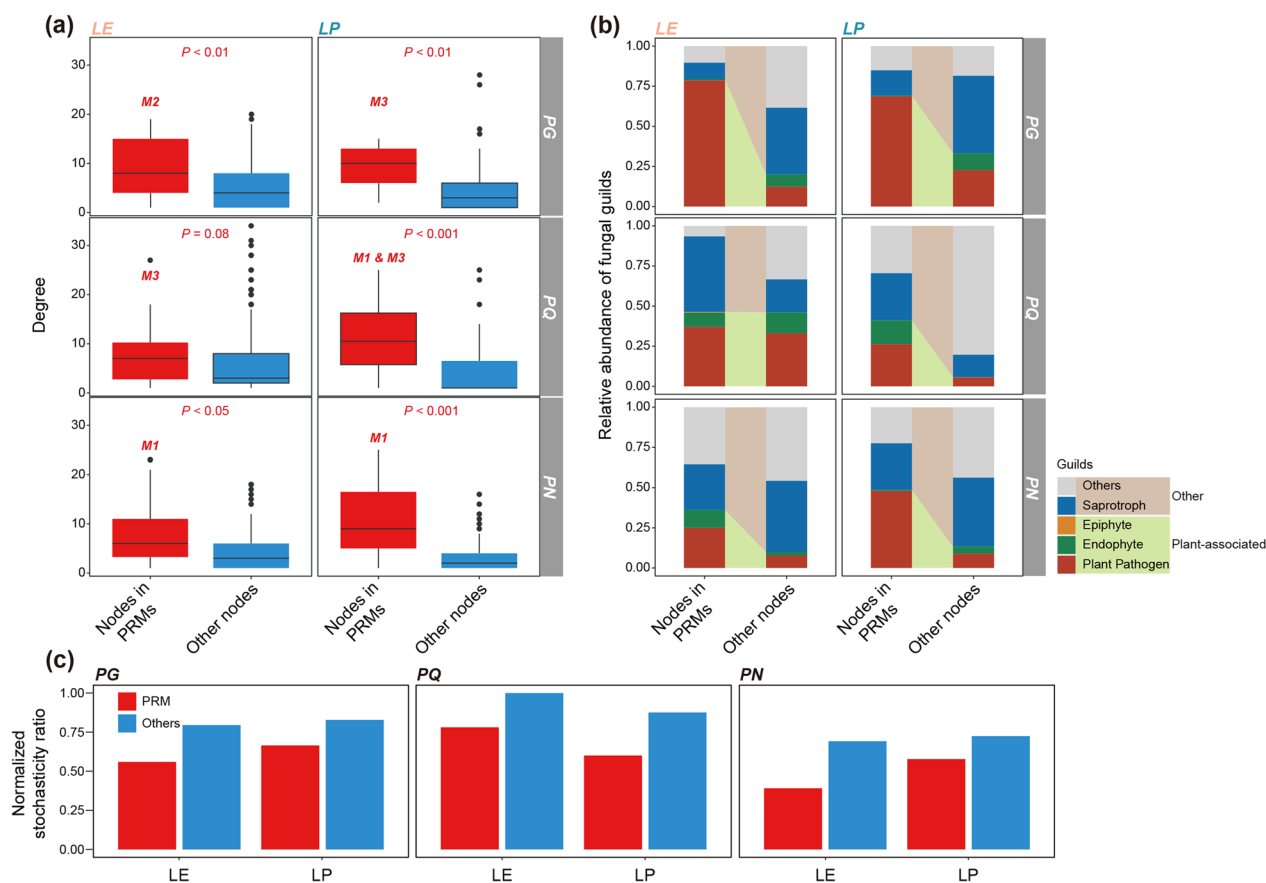


Fig. 4 Comparison of positive regulation modules with other network members in terms of topological importance, guilds, and ecological assembly. **a** Comparison of degree between nodes in positive regulation modules (PRMs) and other nodes in the same network. P values were calculated using Wilcoxon rank-sum test. **b** Relative abundances of different functional guilds in PRMs and other nodes in the same network. Fungal guilds plant pathogen, endophyte, and epiphyte are considered as plant-associated guilds. **c** The relative contribution of stochastic processes to the assembly of PRM subcommunity and other network members. Red and blue bars represent RPM subcommunities and other nodes, respectively

results indicated the topological importance of positive regulation modules, and emphasized their potentially tight interactions with plant host, as well as their direct effects on total saponin accumulation in *Panax* leaves.

Specific and conservative positive regulation taxa of total saponin accumulation across *Panax* species

In SEM-validated positive regulation modules, nodes with significant and positive correlations with total saponin content were identified as positive regulation taxa (PRT) of leaf saponin accumulation (Additional file 1: Table S2). As there were strong collinearity between M2 in LE and M3 in LP of PG (Spearman $\rho = -0.92$, $P < 0.001$; Additional file 1: Fig. S7), which might make the former be excluded from SEM, we also bring the M2 in LE into PRT analysis. Order-level enrichment analysis of these taxa from all six MENs showed that only two fungal orders were significantly overrepresented by PRT (Fisher's exact test, $FDR < 0.05$), including Pleosporales

and Chaetothyriales (Fig. 6a). In which, the Pleosporales was observed in networks of all three *Panax* species, while the Chaetothyriales only exhibited in MENs of PN (Fig. 6a). Additionally, all PRT were mapped on a taxonomy-based phylogenetic tree to estimate whether they were specific or conservative across *Panax* species. Almost all PRT present in PG networks exhibited relatives in PRT of PQ and PN (Fig. 6b). Obvious taxonomic specificity, however, could be observed for PQ and PN, respectively. For instance, PRT belonging to Chaetothyriales were specific to PN, and PRT from Agaricales only exhibited in PQ networks (Fig. 6b). Taxa of Helotiales presented in networks of both PQ and PN, while taxa of Microbotryomycetes were found in networks of both PG and PN. At genus level, only PRT belonging to *Epicoccum* and *Coniothyrium* could be found in MENs all three *Panax* species, indicating the conservatism of these taxa in improving leaf saponin accumulation across *Panax* species (Fig. 6b). In summary, these results revealed the

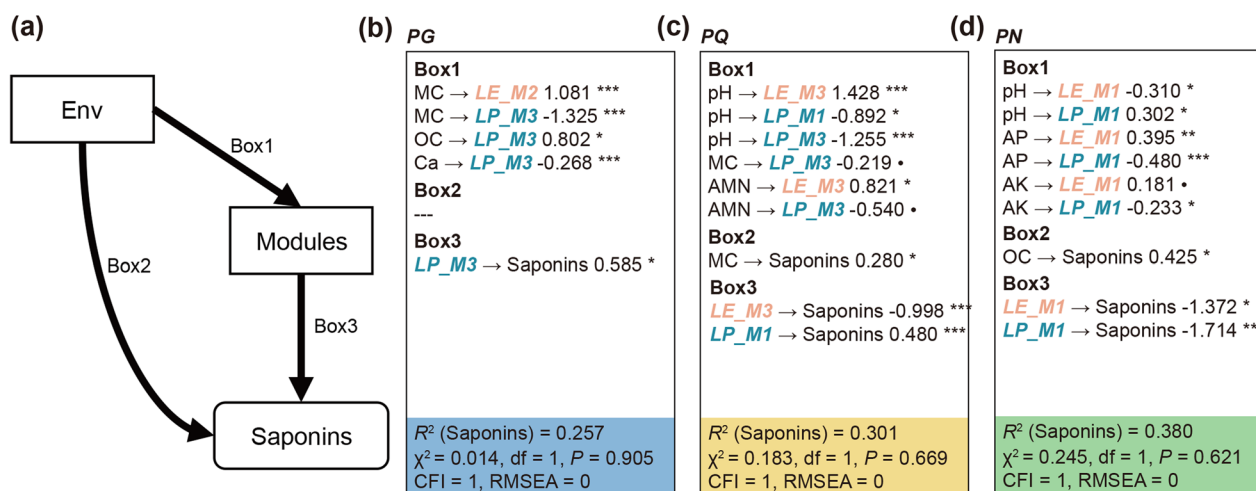


Fig. 5 Structural equation model describing the direct effects of positive regulation modules on total saponin accumulation. **a** The prior model, on the basis of which the structural equation models were fitted. In the prior model, “Modules” indicated the eigengenes of positive regulation modules identified in each *Panax* species, and “Saponins” indicated the sum of contents of different saponins, representing total saponin accumulation in *Panax* leaves. The significant effect sizes calculated by structural equation models in PG **b**, PQ **c**, and PN **d** were displayed. Model statistics were shown in bottom of each box

specificity and conservatism of fungal taxa which potentially contributed to the accumulation of leaf total saponin across three *Panax* species.

Discussion

In this study, we utilized microbial ecology network analysis to test two hypotheses regarding MEN organization strategy and the effects of network members on saponin metabolism in leaf-associated niches of three *Panax* species. Our findings revealed that the phyllosphere mycobiome employed a hub-independent structural robustness strategy at the cost of low efficiency. Moreover, certain modules in fungal networks contributed to the formation of leaf saponin profiles and the accumulation of leaf total saponin. The positive regulation modules of total saponin content located in more central positions of networks and displayed potentially tighter interactions with plant host than other nodes. Additionally, the positive regulation taxa of total saponin demonstrated specificity and conservatism across three *Panax* species.

Highly robust but low-efficiency fungal networks in phyllosphere

Understanding how MEN was organized can help us reveal the maintenance strategies of mycobiome structural homeostasis, which can enhance our ability to manipulate microbial communities and is important for improving ecosystem services and plant performance [30, 69]. Many studies have revealed the small-world property (i.e., short path length between nodes) of ecological networks in bacterial communities, with highly

connected hub nodes that provide a structural basis for efficient exchange of substances or information among community members [74, 77]. However, the organization and characteristics of phyllosphere mycobiome network is rarely studied. The present study showed that phyllosphere fungal MENs were highly robust but less efficiency than random graphs, implying an extreme homeostasis maintenance strategy of phyllosphere mycobiome. A previous study also found that fungal network structure was more stable than bacterial networks in response to drought stress [17]. The resource-poor environment in phyllosphere may be the driver of the extreme strategy employed by fungal communities. Compared to rhizosphere which is rich in minerals and plant-derived nutrients, the barren phyllosphere cannot afford fungi to quickly response to environmental disturbance in terms of energy demands [75]. In addition, stress like nutrition deficiency could drive the replacement of active taxa by slow-growing and stress-tolerant species which would also decrease the overall efficiency of networks [27].

Researches in grassland, alpine meadow, and agriculture ecosystems have emphasized the tight links between microbial network complexity and multiple ecosystem functions [79], [30, 74]. The highly complexity of fungal MENs observed in *Panax* phyllosphere thus provides basis for their important biological or ecological functions. Many studies of soil microbiomes have shown that ecosystem functions like biogeochemical cycling were mediated by hub nodes in MENs, which have key functional potential and are

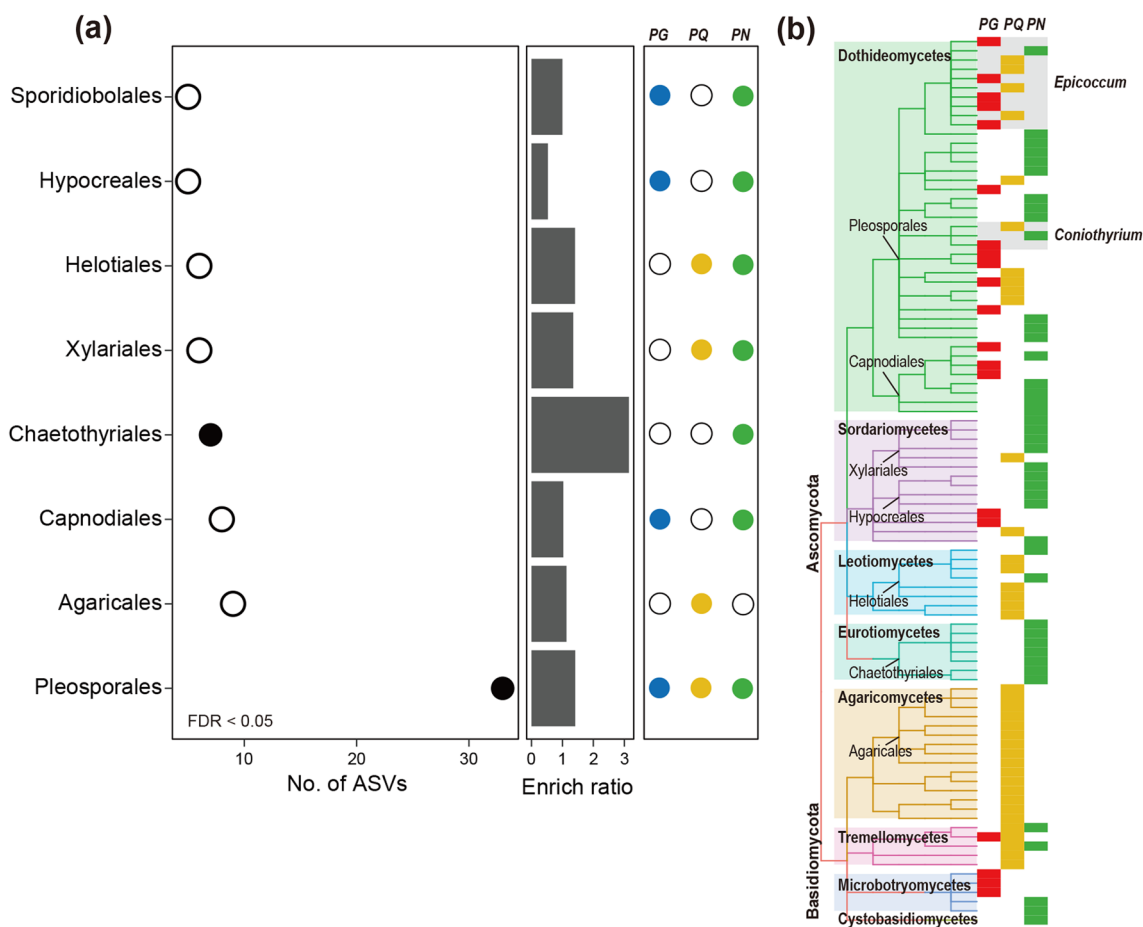


Fig. 6 Order-level enrichment and taxonomic distribution of positive regulation taxa across three *Panax* species. **a** Order-level enrichment analysis of positive regulation taxa (PRT) across all six networks. Solid points represent the corresponding orders were significantly overrepresented by PRT (Fisher’s exact test, FDR < 0.05). **b** Taxonomic distribution of PRT and their absence/presence in the PRT group of each *Panax* species. The phylogenetic tree was constructed based on taxonomy, and clade colors represent different fungal classes

disproportionally important in structuring microbial communities [56]. The present study, nevertheless, indicated that there might not be such strong functional hubs in phyllosphere mycobiomes, in which members tended to tightly connect to each other in subgraphs. These subgraphs, rather than specific one or several nodes, may be the pivotal units for ecological or biological functions of phyllosphere mycobiomes [51]. Additionally, hub microorganisms have been recognized as the key for establishing desired microbiomes for sustainable agriculture [51], [63]. The hub-dependent mycobiome optimization strategy may face challenges in phyllosphere, as communities around a core may not be robust enough to persist in phyllosphere. Constructing and applying stable modules through synthetic community provides an alternative way to reshape phyllosphere mycobiome for improved plant performance and ecosystem services [75].

Phyllosphere fungal modules contribute to saponin accumulation in *Panax* leaves and are tightly linked to host plant

Our analysis of fungal modules in LE and LP provided further validation for their important roles in shaping saponin profiles in *Panax* leaves, indicating that phyllosphere fungi may interact with plant host in the form of tightly connected clusters. Network modules as functional units of microorganisms have been widely reported in various ecosystems. For example, Liu et al. [75] found that modules in soil multitrophic networks could regulate the level of secondary metabolites in licorice root. Other researches in ecosystems such as sponge and human gut also suggested the functional importance of microbial modules [40, 67]. More importantly, positive regulation modules identified in the present study comprised a higher proportion of plant-associated fungi particularly potential plant pathogen, and exhibited more deterministic assembly processes.

This finding is in line with our expectations, as saponins are a type of secondary metabolites closely related to plant-biotic factor interplay, and the plant–microbe war is a key driver of saponin evolution and biosynthesis [3, 24, 52]. In addition, the topological importance of members in PRMs indicated that these taxa not only mediated plant-mycobiome interactions but also within-mycobiome interplays. Actually, the importance in these two types of interactions is intrinsically related, as fungi tightly linked to plants can perceive and spread plant-associated information or materials, such as hormones and VOCs, more quickly than other nodes, thus occupying more central positions in MENs [71]. Previous studies in the model plant *Arabidopsis thaliana* also indicated that topologically important taxa in plant-associated microbial networks mediated interplays between the plant and microbial communities (van der Heijden and Hartmann, 2016). The significant and direct effects of most PRMs on total saponin accumulation were confirmed by SEM, further demonstrating the potential of these functional units to improve the accumulation of desired bioactive compounds in medicinal plant.

The specificity and conservatism of positive regulation taxa across three *Panax* species were identified on the basis of module members. Only fungi belonging to *Epicoccum* and *Coniothyrium* of Pleosporales were identified as PRT in all three *Panax* species. *Epicoccum* is a ubiquitous fungal genus. Many *Epicoccum* members have been found in niches associated with diverse plant species, some of which exhibit an invasion lifestyle [60]. *Coniothyrium* is also reported to be associated with the leaves of many plant species [46]. The conservative chemical response of *Panax* leaves to these fungi suggests that certain molecular patterns present in the two genera might be recognized by receptors shared by *Panax* species, ultimately leading to the up-regulation of saponin accumulation [43]. These fungi are thus useful for exploring conservative molecular mechanisms underlying the regulation of saponin biosynthesis in *Panax* genus. Additionally, many *Panax* species-specific PRT were also found diverse taxonomic groups, which may reflect the differences in the fungal species pool exposed to three *Panax* species caused by environmental filtering and/or dispersal limitation [35]. Furthermore, new types of plant-fungi interactions driven by plant divergence and evolution may also contribute to this specificity [57, 58]. Consequently, the positive regulation taxa and the modules formed by them may serve as the foundation for engineering mycobiome to improving leaf saponin accumulation in *Panax* and help us understand the molecular and chemical mechanisms of plant-fungi interactions.

Conclusions

In summary, this study revealed the high-robustness but low-efficiency network organization strategy of phyllosphere mycobiomes and validated the contributions of network modules to saponin profiles and total saponin content in *Panax* leaves. These findings provide foundations for module-based phyllosphere mycobiomes engineer aiming at the promotion of bioactive secondary metabolite accumulation. The specificity and conservatism of positive regulation fungal taxa across three *Panax* species enhance our understanding of plant-fungi interactions and coevolution associated with secondary metabolism.

Abbreviations

AK	Available potassium
AMN	Ammonium nitrogen
AP	Available phosphorus
AS	Available sulfur
ASV	Amplicon sequence variants
Eca	Exchangeable calcium
EMg	Exchangeable magnesium
GOE	Gaussian orthogonal ensemble
LTED	Link test for environmental filtering or dispersal limitation
MC	Moisture content
MEN	Microbial ecological network
NIN	Nitrate nitrogen
NST	Normalized stochasticity ratio
OC	Organic carbon
PG	<i>Panax ginseng</i>
PQ	<i>Panax quinquefolium</i>
PN	<i>Panax notoginseng</i>
PRM	Positive regulation module of total saponin accumulation
PRT	Positive regulation taxa of total saponin accumulation
TN	Total nitrogen
RMT	Random matrix theory
SEM	Structural equation module

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40793-023-00516-7>.

Additional file 1. Supplementary figures and tables.

Author contributions

LLD, SLC, and GZZ developed the concept. GZZ analysed the data and wrote the manuscript. LPS, CSL, RJM, BX, ZXH, XYW, LXW, and YQZ collected the data and discussed the content. GFW, JX, and SRG revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was funded by the National Key Research and Development Program (2022YFC3501803, 2022YFC3501804), Yunnan Province Biological Medicine Field Major Science and Technology special plan project (202102AA100052; 202102AA310048-1) and Fundamental Research Funds for the Central Public Welfare Research Institutes (ZXKT22001, ZXKT22052, ZXKT22060, ZXKT23018, and ZZ16-XRZ-072).

Availability of data and materials

The datasets generated and/or analysed during the current study are available in the National Genomics Data Center (<https://ngdc.cncb.ac.cn>) under the BioProject PRJCA007643. R codes used for statistical analyses are available at <https://github.com/githubzgz>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interest

The authors declare that they have no competing interests.

Author details

¹Key Laboratory of Beijing for Identification and Safety Evaluation of Chinese Medicine, Institute of Chinese Materia Medica, China Academy of Chinese Medical Sciences, Beijing 100700, China. ²Zhangzhou Pien Tze Huang Pharmaceutical Co., Ltd., Fujian 363000, China. ³School of Life Sciences, Yan'an University, Yan'an 716000, China. ⁴Institute of Herbgonomics, Chengdu University of Traditional Chinese Medicine, Chengdu, Sichuan, China.

Received: 23 April 2023 Accepted: 7 July 2023

Published online: 12 July 2023

References

- Abanda-Nkpawt D, Musch M, Tschiersch J, Boettner M, Schwab W. Molecular interaction between *Methylobacterium extorquens* and seedlings: growth promotion, methanol consumption, and localization of the methanol emission site. *J Exp Bot*. 2006;57:4025–32. <https://doi.org/10.1093/jxb/erl173>.
- Aly AH, Debbab A, Proksch P. Fungal endophytes: unique plant inhabitants with great promises. *Appl Microbiol Biotechnol*. 2011;90:1829–45. <https://doi.org/10.1007/s00253-011-3270-y>.
- Augustin JM, Kuzina V, Andersen SB, Bak S. Molecular activities, biosynthesis and evolution of triterpenoid saponins. *Phytochemistry*. 2011;72:435–57. <https://doi.org/10.1016/j.phytochem.2011.01.015>.
- Banerjee S, Walder F, Buchi L, Meyer M, Held AY, Gatteringer A, et al. Agricultural intensification reduces microbial network complexity and the abundance of keystone taxa in roots. *ISME J*. 2019;13:1722–36. <https://doi.org/10.1038/s41396-019-0383-2>.
- Bar-On YM, Phillips R, Milo R. The biomass distribution on Earth. *Proc Natl Acad Sci USA*. 2018;115:6506–11. <https://doi.org/10.1073/pnas.1711842115>.
- Bashir I, War AF, Rafiq I, Reshi ZA, Rashid I, Shouche YS. Phyllosphere microbiome: diversity and functions. *Microbiol Res*. 2021;254:126888. <https://doi.org/10.1016/j.micres.2021.126888>.
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J Roy Stat Soc B*. 1995;57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol*. 2019;37:852–7. <https://doi.org/10.1038/s41587-019-0209-9>.
- Bougheas S. Contagion in networks: Stability and efficiency. *Math Soc Sci*. 2022;115:64–77. <https://doi.org/10.1016/j.mathsocsci.2021.10.006>.
- Broido AD, Clauset A. Scale-free networks are rare. *Nat Commun*. 2019;10:1017. <https://doi.org/10.1038/s41467-019-08746-5>.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13:581–3. <https://doi.org/10.1038/nmeth.3869>.
- Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. *iMeta*. 2023. <https://doi.org/10.1002/imt2.107>.
- Chen T, Nomura K, Wang X, Sohrabi R, Xu J, Yao L, et al. A plant genetic network for preventing dysbiosis in the phyllosphere. *Nature*. 2020;580:653–7. <https://doi.org/10.1038/s41586-020-2185-0>.
- Chen W, Wang J, Chen X, Meng Z, Xu R, Duoqi D, et al. Soil microbial network complexity predicts ecosystem function along elevation gradients on the Tibetan Plateau. *Soil Biol Biochem*. 2022. <https://doi.org/10.1016/j.soilbio.2022.108766>.
- Clauset A, Newman ME, Moore C. Finding community structure in very large networks. *Phys Rev E*. 2004;70:066111. <https://doi.org/10.1103/PhysRevE.70.066111>.
- Csárdi G, Nepusz T, Traag V, Horvát S, Zanini F, Noom D, et al. igraph: Network analysis and visualization in R. R package version 1.5.0.9002 2023.
- de Vries FT, Griffiths RI, Bailey M, Craig H, Girlanda M, Gweon HS, et al. Soil bacterial networks are less stable under drought than fungal networks. *Nat Commun*. 2018;9:3033. <https://doi.org/10.1038/s41467-018-0516-7>.
- Delgado-Baquerizo M, Reich PB, Trivedi C, Eldridge DJ, Abades S, Alfaro FD, et al. Multiple elements of soil biodiversity drive ecosystem functions across biomes. *Nat Ecol Evol*. 2020;4:210–20. <https://doi.org/10.1038/s41559-019-1084-y>.
- Deng Y, Jiang YH, Yang Y, He Z, Luo F, Zhou J. Molecular ecological network analyses. *BMC Bioinformatics*. 2012;13:113. <https://doi.org/10.1186/1471-2105-13-113>.
- Doll K, Chatterjee S, Scheu S, Karlovsky P, Rohlf M. Fungal metabolic plasticity and sexual development mediate induced resistance to arthropod fungivory. *Proc Biol Sci*. 2013;280:20131219. <https://doi.org/10.1098/rspb.2013.1219>.
- Dupont PY, Eaton CJ, Wargent JJ, Fechtner S, Solomon P, Schmid J, et al. Fungal endophyte infection of ryegrass reprograms host metabolism and alters development. *New Phytol*. 2015;208:1227–40. <https://doi.org/10.1111/nph.13614>.
- Feng K, Zhang Z, Cai W, Liu W, Xu M, Yin H, et al. Biodiversity and species competition regulate the resilience of microbial biofilm community. *Mol Ecol*. 2017;26:6170–82. <https://doi.org/10.1111/mec.14356>.
- Fitzpatrick CR, Salas-Gonzalez I, Conway JM, Finkel OM, Gilbert S, Russ D, et al. The plant microbiome: from ecology to reductionism and beyond. *Annu Rev Microbiol*. 2020;74:81–100. <https://doi.org/10.1146/annurev-micro-022620-014327>.
- Frantzeskakis L, Di Pietro A, Rep M, Schirawski J, Wu CH, Panstruga R. Rapid evolution in plant-microbe interactions - a molecular genomics perspective. *New Phytol*. 2020;225:1134–42. <https://doi.org/10.1111/nph.15966>.
- Gao C, Xu L, Montoya L, Madera M, Hollingsworth J, Chen L, et al. Co-occurrence networks reveal more complexity than community composition in resistance and resilience of microbial communities. *Nat Commun*. 2022. <https://doi.org/10.1038/s41467-022-31343-y>.
- Guimera R, Nunes Amaral LA. Functional cartography of complex metabolic networks. *Nature*. 2005;433:895–900. <https://doi.org/10.1038/nature03288>.
- Hernandez DJ, David AS, Menges ES, Searcy CA, Afkhami ME. Environmental stress destabilizes microbial networks. *ISME J*. 2021;15:1722–34. <https://doi.org/10.1038/s41396-020-00882-x>.
- Hou M, Wang R, Zhao S, Wang Z. Ginsenosides in *Panax* genus and their biosynthesis. *Acta Pharm Sin B*. 2021;11:1813–34. <https://doi.org/10.1016/j.apsb.2020.12.017>.
- Jiao S, Lu Y. Soil pH and temperature regulate assembly processes of abundant and rare bacterial communities in agricultural ecosystems. *Environ Microbiol*. 2020;22:1052–65. <https://doi.org/10.1111/1462-2920.14815>.
- Jiao S, Lu Y, Wei G. Soil multitrophic network complexity enhances the link between biodiversity and multifunctionality in agricultural systems. *Glob Chang Biol*. 2022;28:140–53. <https://doi.org/10.1111/gcb.15917>.
- Jiao S, Yang Y, Xu Y, Zhang J, Lu Y. Balance between community assembly processes mediates species coexistence in agricultural soil microbiomes across eastern China. *ISME J*. 2020;14:202–16. <https://doi.org/10.1038/s41396-019-0522-9>.
- Jun T, Kim J-Y. Connectivity, stability and efficiency in a network as an information flow. *Math Soc Sci*. 2007;53:314–31. <https://doi.org/10.1016/j.mathsocsci.2007.02.001>.
- Jun W, Barahona M, Yue-Jin T, Hong-Zhong D. Natural connectivity of complex networks. *Chin Phys Lett*. 2010. <https://doi.org/10.1088/0256-307x/27/7/078902>.
- Lang M, Binder M, Richter J, Schratz P, Pfisterer F, Coors S, Au Q, Casalicchio G, Kotthoff L, Bischl B mlr3: a modern object-oriented machine learning framework in R. *J Open Sour Softw*. 2019;4(44):1903.
- Leibold MA, Holyoak M, Mouquet N, Amarasekare P, Chase JM, Hoopes MF, et al. The metacommunity concept: a framework for multi-scale community ecology. *Ecol Lett*. 2004;7:601–13. <https://doi.org/10.1111/j.1461-0248.2004.00608.x>.

36. Lima-Mendez G, Faust K, Henry N, Decelle J, Colin S, Carcillo F, et al. Ocean plankton. Determinants of community structure in the global plankton interactome. *Science*. 2015;348:1262073. <https://doi.org/10.1126/science.1262073>.
37. Liu Y, Li D, Gao H, Li Y, Chen W, Jiao S, et al. Regulation of soil micro-foodwebs to root secondary metabolites in cultivated and wild licorice plants. *Sci Total Environ*. 2022;828:154302. <https://doi.org/10.1016/j.scitotenv.2022.154302>.
38. Liu Y, Wang H, Peng Z, Li D, Chen W, Jiao S, et al. Regulation of root secondary metabolites by partial root-associated microbiotas under the shaping of licorice ecotypic differentiation in northwest China. *J Integr Plant Biol*. 2021. <https://doi.org/10.1111/jipb.13179>.
39. Luo F, Yang Y, Zhong J, Gao H, Khan L, Thompson DK, et al. Constructing gene co-expression networks and predicting functions of unknown genes by random matrix theory. *BMC Bioinformatics*. 2007;8:299. <https://doi.org/10.1186/1471-2105-8-299>.
40. Lurgi M, Thomas T, Wemheuer B, Webster NS, Montoya JM. Modularity and predicted functions of the global sponge-microbiome network. *Nat Commun*. 2019;10:992. <https://doi.org/10.1038/s41467-019-08925-4>.
41. Ma B, Wang H, Dsouza M, Lou J, He Y, Dai Z, et al. Geographic patterns of co-occurrence network topological features for soil microbiota at continental scale in eastern China. *ISME J*. 2016;10:1891–901. <https://doi.org/10.1038/ismej.2015.261>.
42. Ma B, Wang Y, Ye S, Liu S, Stirling E, Gilbert JA, et al. Earth microbial co-occurrence network reveals interconnection pattern across microbiomes. *Microbiome*. 2020;8:82. <https://doi.org/10.1186/s40168-020-00857-2>.
43. Ma KW, Niu Y, Jia Y, Ordon J, Copeland C, Emonet A, et al. Coordination of microbe-host homeostasis by crosstalk with plant innate immunity. *Nat Plants*. 2021;7:814–25. <https://doi.org/10.1038/s41477-021-00920-2>.
44. Maier BA, Kiefer P, Field CM, Hemmerle L, Bortfeld-Miller M, Emmenegger B, et al. A general non-self response as part of plant immunity. *Nat Plants*. 2021;7:696–705. <https://doi.org/10.1038/s41477-021-00913-1>.
45. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J*. 2011;17(1):10–2. <https://doi.org/10.14806/EJ.17.1.200>.
46. Michalak I, Aliman J, Hadziabulic A, Komlen V. Novel trends in crop bioprotection. In: Chojnacka K, Saeid A, editors. *Smart agrochemicals for sustainable agriculture*. Academic Press, 2022.
47. Nguyen NH, Song Z, Bates ST, Branco S, Tedersoo L, Menke J, et al. FUN-Guild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol*. 2016;20:241–8. <https://doi.org/10.1016/j.funeco.2015.06.006>.
48. Nilsson RH, Larsson K-H, Taylor AFS, Bengtsson-Palme J, Jeppesen TS, Schigel D, et al. The UNITE database for molecular identification of fungi: handling dark taxa and parallel taxonomic classifications. *Nucleic Acids Res*. 2019;47:D259–64. <https://doi.org/10.1093/nar/gky1022>.
49. Ning D, Deng Y, Tiedje JM, Zhou J. A general framework for quantitatively assessing ecological stochasticity. *Proc Natl Acad Sci*. 2019;116:16892–8. <https://doi.org/10.1073/pnas.1904623116>.
50. Oksanen J, Blanchet G, Friendly M, Kindt R, Legendre P, McGlenn D, et al. *vegan: community ecology package*. R package version 2.5–6 2019.
51. Peng Z, Liang C, Gao M, Qiu Y, Pan Y, Gao H, et al. The neglected role of micronutrients in predicting soil microbial structure. *NPJ Biofilms Microbiomes*. 2022;8:103. <https://doi.org/10.1038/s41522-022-00363-3>.
52. Piasecka A, Jedzejczak-Rey N, Bednarek P. Secondary metabolites in plant innate immunity: conserved function of divergent chemicals. *New Phytol*. 2015;206:948–64. <https://doi.org/10.1111/nph.13325>.
53. Rosseel Y. lavaan: an R package for structural equation modeling. *J Stat Softw*. 2012;48:1–36. <https://doi.org/10.18637/jss.v048.i02>.
54. Shen YQ, Burger G. Plasticity of a key metabolic pathway in fungi. *Funct Integr Genomics*. 2009;9:145–51. <https://doi.org/10.1007/s10142-008-0095-6>.
55. Shi S, Nuccio EE, Shi ZJ, He Z, Zhou J, Firestone MK. The interconnected rhizosphere: high network complexity dominates rhizosphere assemblages. *Ecol Lett*. 2016;19:926–36. <https://doi.org/10.1111/ele.12630>.
56. Shi Y, Delgado-Baquerizo M, Li Y, Yang Y, Zhu YG, Penuelas J, et al. Abundance of kinless hubs within soil microbial networks are associated with high functional potential in agricultural ecosystems. *Environ Int*. 2020;142:105869. <https://doi.org/10.1016/j.envint.2020.105869>.
57. Shim H, Waminal NE, Kim HH, Yang TJ. Dynamic evolution of *Panax* species. *Genes Genomics*. 2021;43:209–15. <https://doi.org/10.1007/s13258-021-01047-6>.
58. Sohrabi R, Paasch BC, Liber JA, He SY. Phyllosphere microbiome. *Annu Rev Plant Biol*. 2023. <https://doi.org/10.1146/annurev-arplant-102820-032704>.
59. Stegen JC, Lin X, Fredrickson JK, Konopka AE. Estimating and mapping ecological processes influencing microbial community assembly. *Front Microbiol*. 2015;6:370. <https://doi.org/10.3389/fmicb.2015.00370>.
60. Taguam JD, Evallo E, Balendres MA. Epicoccum species: ubiquitous plant pathogens and effective biological control agents. *Eur J Plant Pathol*. 2021;159:713–25. <https://doi.org/10.1007/s10658-021-02207-w>.
61. Tedersoo L, Sánchez-Ramírez S, Kõljalg U, Bahram M, Döring M, Schigel D, et al. High-level classification of the Fungi and a tool for evolutionary ecological analyses. *Fungal Divers*. 2018;90:135–59. <https://doi.org/10.1007/s13225-018-0401-0>.
62. Thakur M, Bhattacharya S, Khosla PK, Puri S. Improving production of plant secondary metabolites through biotic and abiotic elicitation. *J Appl Res Med Aromat Plants*. 2019;12:1–12. <https://doi.org/10.1016/j.jarmap.2018.11.004>.
63. Toju H, Peay KG, Yamamichi M, Narisawa K, Hiruma K, Naito K, et al. Core microbiomes for sustainable agroecosystems. *Nat Plants*. 2018;4:247–57. <https://doi.org/10.1038/s41477-018-0139-4>.
64. Trivedi P, Leach JE, Tringe SG, Sa T, Singh BK. Plant-microbiome interactions: from community assembly to plant health. *Nat Rev Microbiol*. 2020;18:607–21. <https://doi.org/10.1038/s41579-020-0412-1>.
65. van der Heijden MG, Hartmann M. Networking in the plant microbiome. *PLoS Biol*. 2016;14:e1002378. <https://doi.org/10.1371/journal.pbio.1002378>.
66. Wei G, Zhang G, Li M, Liu C, Wei F, Wang Y, et al. Core rhizosphere microbiome of *Panax notoginseng* and its associations with belowground biomass and saponin contents. *Environ Microbiol*. 2022. <https://doi.org/10.1111/1462-2920.16245>.
67. Wu G, Zhao N, Zhang C, Lam YY, Zhao L. Guild-based analysis for understanding gut microbiome in human health and diseases. *Genome Med*. 2021;13:22. <https://doi.org/10.1186/s13073-021-00840-y>.
68. Xiao N, Zhou A, Kempfer ML, Zhou BY, Shi ZJ, Yuan M, et al. Disentangling direct from indirect relationships in association networks. *Proc Natl Acad Sci USA*. 2022. <https://doi.org/10.1073/pnas.2109995119>.
69. Xiong C, Zhu YG, Wang JT, Singh B, Han LL, Shen JP, et al. Host selection shapes crop microbiome assembly and network complexity. *New Phytol*. 2021;229:1091–104. <https://doi.org/10.1111/nph.16890>.
70. Xiong W, Li R, Ren Y, Liu C, Zhao Q, Wu H, et al. Distinct roles for soil fungal and bacterial communities associated with the suppression of vanilla Fusarium wilt disease. *Soil Biol Biochem*. 2017;107:198–207. <https://doi.org/10.1016/j.soilbio.2017.01.010>.
71. Xu N, Zhao Q, Zhang Z, Zhang Q, Wang Y, Qin G, et al. Phyllosphere microorganisms: sources, drivers, and their interactions with plant hosts. *J Agric Food Chem*. 2022;70:4860–70. <https://doi.org/10.1021/acs.jafc.2c01113>.
72. Yao H, Sun X, He C, Maitra P, Li XC, Guo LD. Phyllosphere epiphytic and endophytic fungal community and network structures differ in a tropical mangrove ecosystem. *Microbiome*. 2019;7:57. <https://doi.org/10.1186/s40168-019-0671-0>.
73. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY, McInerney G. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol Evol*. 2016;8:28–36. <https://doi.org/10.1111/2041-210x.12628>.
74. Yuan MM, Guo X, Wu L, Zhang Y, Xiao N, Ning D, et al. Climate warming enhances microbial network complexity and stability. *Nat Clim Chang*. 2021;11:343–8. <https://doi.org/10.1038/s41558-021-00989-9>.
75. Zhan C, Matsumoto H, Liu Y, Wang M. Pathways to engineering the phyllosphere microbiome for sustainable crop production. *Nat Food*. 2022;3:997–1004. <https://doi.org/10.1038/s43016-022-00636-2>.
76. Zhang B, Peng Y, Zhang Z, Liu H, Qi Y, Liu S, et al. GAP production of TCM herbs in China. *Planta Med*. 2010;76:1948–55. <https://doi.org/10.1055/s-0030-1250527>.
77. Zhang B, Zhang J, Liu Y, Shi P, Wei G. Co-occurrence patterns of soybean rhizosphere microbiome at a continental scale. *Soil Biol Biochem*. 2018;118:178–86. <https://doi.org/10.1016/j.soilbio.2017.12.011>.
78. Zhang G, Wei F, Chen Z, Wang Y, Jiao S, Yang J, et al. Evidence for saponin diversity-mycobiome links and conservatism of plant-fungi interaction

patterns across Holarctic disjunct *Panax* species. *Sci Total Environ.* 2022;830:154583. <https://doi.org/10.1016/j.scitotenv.2022.154583>.

79. Zhang G, Wei G, Wei F, Chen Z, He M, Jiao S, et al. Dispersal limitation plays stronger role in the community assembly of fungi relative to bacteria in rhizosphere across the arable area of medicinal plant. *Front Microbiol.* 2021;12:713523. <https://doi.org/10.3389/fmicb.2021.713523>.
80. Zhao ZB, He JZ, Geisen S, Han LL, Wang JT, Shen JP, et al. Protist communities are more sensitive to nitrogen fertilization than other microorganisms in diverse agricultural soils. *Microbiome.* 2019;7:33. <https://doi.org/10.1186/s40168-019-0647-0>.
81. Zhou J, Deng Y, Luo F, He Z, Tu Q, Zhi X. Functional molecular ecological networks. *MBio.* 2010. <https://doi.org/10.1128/mBio.00169-10>.
82. Zhou J, Deng Y, Luo F, He Z, Yang Y. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO₂. *MBio.* 2011. <https://doi.org/10.1128/mBio.00122-11>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

