The composition characteristics of endophytic communities and their relationship with metabolites profile in Ephedra sinica under wild and cultivated conditions

hui zhang
Shanxi University
Ye Xia
The Ohio State University
Jin-Long Cui (cjl717@163.com)
Shanxi University
Xin Ji
Shanxi University
Xi Liu
Shanxi University
Shuang-Man Miao
Shanxi University
Meng-Liang Wang
Shanxi University
Jun-Hong Wang
Shanxi University

Research Article

Keywords: Medicinal plant, Plant-microbe interaction, Metabolic profile, Environment, Metabolism

Posted Date: February 14th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2572678/v1

Version of Record: A version of this preprint was published at Environmental Science and Pollution Research on August 9th, 2023. See the published version at https://doi.org/10.1007/s11356-023-29145-w.

Abstract

Ephedra sinica is one of the most famous Chinese medicinal plants. The insufficient supply of wild resources has led to the increased use of cultivated products. However, the related medicinal quality differs significantly. Although the influence of external environment on the quality of *E. sinica* has been studied, the impact of endophytic microbes on it remains vague. This study characterized differential metabolites and microbial community compositions in the wild and cultivated *E. sinica* by combining metabolomics with microbiomics, and explored the effect of endophytes on the formation of differential metabolites further. The results showed that the difference in quality between wild and cultivated *E. sinica* was mainly in the productions of alkaloids, flavonoids, and terpenoids. The associated endophytes had special compositional characteristics. For instance, the distribution and abundance of dominant endophytes varied between the wild and cultivated *E. sinica*. Several endophytes had significant, or highly significant correlations with the formations of ephedrine, pseudoephedrine, D-cathinone, methcathinone, coumarin, kaempferol, rhamnetin, or phenylacetic acid. This study will deepen our understanding of the plant-endophyte interactions and provide a strategy for the quality control of *E. sinica* products.

Introduction

Ephedra sinica is a precious medicinal plant for several reasons: ( ) it belongs to a scarce herbaceous gymnosperm; ( ) it has no leaves, therefore, it is also called “leafless herb” (Zhang et al. 2020a); ( ) both its stems and roots are well-known traditional Chinese medicines but with opposite pharmacological activities, which are recorded in the Chinese Pharmacopoeia as “Mahuang” and “Mahuanggen.” According to the conventional experience and research, Mahuang is used for inducing diaphoresis and elevating blood pressure, which is mainly attributed to ephedrine, pseudoephedrine, methylephedrine, and tetramethylpyrazine. Mahuanggen is used for antiperspirant and antihypertensive, which is ascribed to the presence of feruloylhistamine, mahuannin, ephedrannin, and ephedradine A, B, C, and D (Miao et al. 2020). In modern medicine, *E. sinica* is widely used to treat bronchial asthma, angiocardiopathy, immune system disease, and so on (Miao et al. 2022). The extracts of *E. sinica* are also used as dietary supplements and diet products in Occident (Zhang et al. 2018). Surprisingly, Mahuang has been frequently used to treat novel coronavirus pneumonia and has played an essential role in preventing and controlling the disease development, which is the fundamental component of “Maxing Shigan decoction” and “Lianhua Qingwen capsule/granule” in China (Zhou et al. 2020). Overall, the market’s demands for *E. sinica* are increasing. In addition to wild *E. sinica*, the cultivated resources are receiving more and more attention, but their qualities are uneven. In recent years, it has been discovered that metabolomics could uncover more fine-grained differences in diverse plant tissues, including the active ingredients, their precursors, and the trace but essential elements, all of which will provide opportunities to gain insight into the detailed formation of different plant chemical components (Cao et al. 2020).

More research has substantiated that most wild and cultivated medicinal materials differed significantly in their chemical compositions. In addition to the external factors such as light, humidity, and altitude, there were also discrepancies between wild and cultivated medicinal materials even under the same phenological conditions, which is a big puzzle for scientists (Inngjerdingen et al. 2012). Recently, it has been noticed that plants’ internal environment is also a significant factor affecting plant quality, among which the endophytes are the most important. Endophytes are those microorganisms that live inside plant tissues but do not cause apparent diseases (Strobel et al. 2003). After a long period of co-evolution, they developed into a “community of shared destiny.” Meanwhile, endophytes are also dynamic, including “old species” that have been transmitted from their ancestors and “new species” that have been horizontally introduced from the environment (Frank et al. 2017). Their

compositions are influenced not only by the genetics of the species and the environmental factors, but also by the anthropogenic factors of both irrigation and cultivation (Rasmussen et al. 2008). Endophytes, like the “organs” of plants, are the integral parts of the plants. They directly produce metabolites into plant tissues (Cook et al. 2014) through gene exchange with the hosts (Motoyama et al. 2021), signal induction (Zhai et al. 2017), and biotransformation (Liu et al. 2021), which have substantial impacts on the formation of plant metabolites.

There are questions that the existing research has not yet answered. For instance, besides the traditional known active ingredients, what are the other differential metabolites that exist between wild and cultivated *E. sinica*? How do these differential metabolites synthesize? Are the endophytic compositions similar between the corresponding tissues of wild and cultivated *E. sinica*? What are the relationships among different endophytes in the communities? Do endophytes impact the formation of differential metabolites? Are there any endophytes that influence the formation of different compounds? There are many other questions as such. These issues have remained unclear up to now, but they will be explored in this study. Our study will be very helpful to understand the cause of metabolic differences and provide scientific references for the cultivation of *E. sinica* and their applications.

Materials And Methods

Plant Material Collection and Sample Preparation

In July 2019, healthy wild and cultivated *E. sinica* were collected from Ordos (38° 9' 38" N, 107° 26' 7" E, ≈ 1305 m), Inner Mongolia of China. These samples were labeled as CS (Stems of cultivated *E. sinica*), NS (Stems of wild *E. sinica*), CR (Roots of cultivated *E. sinica*), and NR (Roots of wild *E. sinica*). Half of each sample was used for endophytes’ composition investigation and the other half for metabolite investigation. Voucher specimens were numbered as CS2019, CR2019, NS2019, and NR2019, then deposited at the Institute of Applied Chemistry, Shanxi University.

Investigation of the Metabolite Profiles

Plant samples for metabolomics were dried in the shade at 25°C with a humidity of less than 40% and ground into powder with an 80-mesh sieve. Each 0.2 g aliquot was extracted through the supplementation of 15 mL of methanol-water solvent (4:1, *v*: *v*). After ultrasound treatment (210 W, 25°C) for 40 minutes, the mixture was centrifuged at 4°C at 12,000 r/min for 10 minutes. The supernatant was filtered through PTFE (0.22 μm) membrane for UPLC-QTOF/MS analysis. Ten replicates were created for each sample. Equal proportions of wild and cultivated samples were mixed separately to produce quality control samples.

Chromatographic analysis was performed on the Agilent 1290 UPLC system (Agilent Technologies, USA), equipped with a ZORBAX StableBond-C18 column (1.8 μm particle size, 4.6 * 50 mm; Agilent, USA) maintained at 40°C. The mobile phase consisted of A (0.1% formic acid in water) and B (acetonitrile) at a flow rate of 0.25 mL/min, which were applied in the following linear gradient elution: 10–14% B at 0–5 min, 14–20% B at 5–8 min, 20–30% B at 8–12 min, 30–50% B at 12–16 min, 50–70% B, 16–20 min, 70–95% B at 20–24 min, and 95% at 24–30 min. Then the separated components were detected by the mass spectrometer, which was performed with LC-Q/TOF-MS/MS equipped with an ESI source with Jet Stream technology. The instrumental parameters were as follows: the nebulizer pressure was 45 psi; the capillary voltage was 4000 V; the nozzle voltage was 500 V; the sheath gas temperature and flow rates were 400°C and 12 L/min; the drying gas temperature and flow rate were 350°C and 9 L/min; the fragmentation voltage was 175 V; the collision energy, slope = 6.66 and offset = 1. The positive ionization

mode mass spectra were acquired simultaneously in a full-scan operation with a mass range of 80-1200 m/z. All the operations and acquisitions were controlled by Agilent MassHunter Acquisition software (version A.05.01, Agilent Technologies, USA), and data analysis was processed with MassHunter Qualitative Analysis software (version B.07.00, Agilent Technologies, USA).

Data Preprocessing and Annotation

Raw data were converted to mzData format using the MassHunter Qualitative Analysis software and uploaded to XCMS through a secure SSL connection. For the feature detection, the XCMS *centWave* algorithm was used with the following parameters: signal/noise threshold = 6, ppm = 30, peak width = (10, 60), and prefilter = (3, 500). Their feature alignment was performed with the default parameters in XCMS with bw = 5 and mzwid = 0.025. The retention time correction was performed with the standard *obiwarp* algorithm in XCMS with prfostep = 0.5. The putative identities of each ion were first given within XCMS by matching features in the METLIN database with the following parameters: sample biosource = PLANT; ppm = 10; adducts = $[M + H]^+$, $[M + Na]^+$, $[M + H - H_2O]^+$, $[M + H - 2H_2O]^+$, $[M + 2Na - H]^+$, $[M + H + Na]^{2+}$and $[M + 2H + Na]^{3+}$ in positive ion mode. The SIMCA-P (version 14.1, Umetrics, Umeε, Sweden) software was used for multivariate statistical analysis of the data obtained via UPLC-MS analysis. A principal component analysis (PCA) was utilized to obtain an overview of the sample distribution and identify outliers. The partial least-square-discrimination analysis (PLS-DA) was used to determine the significant components of the models and thus minimize the overfitting. S-plot modeling and variable importance of projection (VIP) were utilized to explore the chemometric markers. Finally, potential biomarkers were identified according to their exact m/z value, MS/MS spectrum, and retention times. METLIN (http://metlin.scripps.edu/), HMDB (https://hmdb.ca/), MetFrag (https://msbi.ipb-halle.de/Met FragBeta/), and the Agilent MassHunter Pesticides Personal Compound Database and Library (PCDL) were used for metabolites searching. Additionally, to further explore the biological significance of the different metabolites between wild and cultivated *E. Sinica*, the metabolic pathways of different metabolites were identified and constructed by searching the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

DNA Extraction and Sequencing

According to our previous methods, plant samples for microbiomics analysis were surface sterilized within eight hours of collection (Cui et al. 2018). Endophytic bacterial and fungal DNA extractions were carried out with the *E.Z.N.A.* Plant DNA Mini Kit (Omega Bio-Tek, Doraville, GA, USA). The bacterial 16S rRNA genes with V3-V4 regions were amplified with two primers 341F: (5'-CCTACGGGNGGCWGCAG-3') and 801R (5'-GACTACHVGGGTATCTAATCC-3') (Klindworth et al. 2013). The fungal ITS rDNA (ITS3-4) genes with regions were amplified with primer pair of ITS4F (5'-TCCTCCGCTTATTGATATGC-3') and ITS3R (5'-GCATCGATGAAGAACGCAGC-3') (Cui et al. 2019). The polymerase chain reaction (PCR) reactions were implemented as Zhang (ITS) (Zhang et al. 2020a) and Miao (16S rRNA) (Miao et al. 2022). Subsequently, all the amplified products were investigated with the bacteria and fungi through 16S rRNA and ITS sequencing using the Illumina Miseq platform (USA) at Sangon Biotech (Shanghai) Co., Ltd. (Shanghai, China), respectively.

Characteristic Analysis of the Microbial Communities of Associated Endophytes

The bacterial and fungal communities associated with *E. sinica* were characterized using Illumina MiSeq sequencing to explore the alpha diversity, operational taxonomic units (OTU) distribution, community composition, and species classification (Zhang et al. 2020a; Miao et al. 2022). To predict the function of endophytes through

16S rRNA and ITS sequences in KEGG using Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt2) (Douglas et al. 2020). Enzyme associated genes related to the main metabolic pathways of *E. sinica* were screened and displayed with GraphPad Prism 9 (San Diego, California USA). Additionally, to determine the contribution of endophytes in different environmental conditions to these enzyme-associated genes, non-parametric Spearman's correlation (two-tailed) and the R 3.2.2 package “pheatmap” were used to show the graphical data. Likewise, microbial intra-kingdom network analysis was examined by the Spearman’s correlation test using the R package “psych” based on the bacterial and fungal OTUs. Only those elements were conserved with statistically significant positive or negative correlations (Spearman’s $r > 0.6$ or $r < -0.6$; *P*-value < 0.05). The network was visualized, and its topological parameters were computed by Gephi 0.9.2 (Bastian, 2007). The experiment was established with three replicates.

Integrative Analysis of the Marker Metabolites and Specific Endophytes

In order to accurately predict and explain the relationships between microorganisms (including predominant and shared endophytes) and metabolites in *E. sinica*, a correlation analysis (the Spearman algorithm combined with a two-tailed Student’s t-test) was performed to evaluate the relative influence of the microorganic community on chemical constituents of *E. sinica* using SPSS 16.0 software (Chicago, IL, USA), and then it was imagined with the R package “pheatmap”. In the matrices, the correlation coefficients were marked in red and blue for positive and negative correlations, respectively. The spearman’s $r > 0.8$ or $r < -0.8$, $P < 0.05$ (*) or $P < 0.01$ (**) indicated significant or extremely significant correlations.

Results

Screening of Differential Metabolites

The results of the PCA score plot showed considerable differences in metabolic profiles between CS (Stems of cultivated *E. sinica*) and NS (Stems of wild *E. sinica*) (Fig. 1a) as well as CR (Roots of cultivated *E. sinica*) and NR (Roots of wild *E. sinica*) (Fig. 1d). The model had high reliability and good fit based on the PLS-DA (Fig. 1b, e). The S-VIP plot was constructed from OPLS-DA (Fig. 1c, f). Then, based on the screening values VIP > 1.0 and $P < 0.05$, a total of 654 and 288 different metabolites were found in the stems and roots of wild and cultivated *E. sinica*, respectively. However, only 47 and 23 different marker metabolites were ultimately identified in the stems and roots, correspondingly (Supplementary data Table 1).

Table 1
Diversity of endophytic bacteria and fungi in each sample of Ephedra sinica. (Mean ± SD) (n = 3) a

Sample_ID	Seq_num	OTU_num	Shannon_index	ACE_index	Chao1_index	Coverage	Simpson
NSB	74.00 ± 47.51	9.00 ± 4.93	1.30 ± 0.31	34.13 ± 39.43	16.17 ± 14.58	0.93 ± 0.026	0.448 ± 0.115
CSB	152.00 ± 14.80	12.00 ± 3.51	1.14 ± 0.19	162.46 ± 219.36	32.17 ± 16.28	0.94 ± 0.020	0.462 ± 0.040
NRB	6527.67 ± 1524.99	579.00 ± 91.99	4.67 ± 0.13	1074.70 ± 95.07	852.55 ± 88.09	0.96 ± 0.005	0.031 ± 0.008
CRB	6124.00 ± 1772.74	551.00 ± 92.32	4.29 ± 0.20	935.67 ± 122.59	805.45 ± 79.10	0.96 ± 0.011	0.049 ± 0.011
NSF	41495.00 ± 3942.64	1562.00 ± 625.05	2.60 ± 0.49	3403.89 ± 1445.72	2614.12 ± 936.07	0.98 ± 0.0057	0.26 ± 0.52
CSF	43783.67 ± 1717.88	2037.00 ± 244.58	1.92 ± 0.94	5068.21 ± 705.71	3591.16 ± 475.72	0.98 ± 0.0032	0.57 ± 0.29
NRF	39712.67 ± 3029.86	1357.00 ± 96.50	2.05 ± 0.17	3129.93 ± 421.83	2414.68 ± 259.86	0.98 ± 0.0006	0.49 ± 0.05
CRF	42393.00 ± 7380.31	1240.00 ± 352.81	1.69 ± 0.85	2776.90 ± 1033.91	2122.29 ± 674.62	0.99 ± 0.0019	0.57 ± 0.29

[a]OTU richness was observed in the endophytes of wild and cultivated *Ephedra sinica*. Richness = observed number of OTUs; Shannon = Shannon index estimator; ACE = ACE richness estimator; Chao1 = Chao1 richness estimator; coverage = Good's coverage estimator; and Simpson = Simpson index estimator. CSB and NSB are endophytic bacteria in stems of cultivated and wild *E. sinica*, respectively; CRB and NRB are endophytic bacteria in roots of cultivated and wild *E. sinica*, respectively; CSF and NSF are endophytic fungi in stems of cultivated and wild *E. sinica*, respectively; CRF and NRF are endophytic fungi in roots of cultivated and wild *E. sinica*, respectively.

Differential Metabolites and Associated Metabolic Pathways

The selected differential metabolites could be deemed responsible for distinguishing chemical compositions between the wild and cultivated *E. sinica*, which were identified by their mass spectra and mass fragmentation information. Their identification was described as an example of M148T6 (Fig. 2). The differential metabolite M148T6 was identified as pseudoephedrine based on the spectral data of MS/MS match in the HMDB and METLIN databases, combined with m/z values of 148.1116 and its fragmentation pattern (m/z 91.0538 for C_7H_7, m/z 95.0473 for C_7H_{11}, m/z 107.0709 for C_7H_7O, m/z 117.0691 for C_7H_9, m/z 119.0816 for C_9H_{11}, m/z 132.0802 for $C_9H_{10}N$, m/z 133.0877 for C_9H_9O, m/z 135.0920 for $C_9H_{11}O$, m/z 148.1121 for $C_{10}H_{14}N$ and m/z 150.1182 for $C_9H_{12}NO$). The results of other differential compounds, including names, m/z values, retention times, adduct ions, and fragment ions, are summarized in Supplementary data Table 1, and their structures were shown in Supplementary Fig. 1.

There were 43 unique differential compounds in stems between CS and NS (Supplementary data Table 1). Among them, ten compounds were significantly enriched in NS, consisting of terpenoids (including compounds 27, 28, 29, 30 and 31), flavonoids (compounds 18 and 19), organic acids (compounds 32 and 40), and phenylpropanoids (compound 45), with the content variations ranging from 1.78 to 71.07 times. In contrast, the amounts of the other 32 compounds were abundant with flavonoids (including compounds 9, 10, 11, 12, 13, 14, 15, 16, 17, 20, 21, 22, 23, 24, 25 and 26), organic acids (including compounds 33, 34, 35, 36, 37, 38, 39 and 41), alkaloids (including

compounds 4, 5, 6, 7 and 8), and phenylpropanoids (including compounds 42, 43, 44 and 46), which were 1.01 to 20.60 times higher in CS compared to NS. There were 18 unique differential compounds between CR and NR. The results showed that phenylpropanoids (including compounds 17, 18, 19 and 20), alkaloids (compounds 4, 5 and 6), organic acids (compound 13), fatty aldehydes (compounds 21 and 22), and alkyl phenyl ketones (compound 22) had a higher abundance in NR than in CR. However, the relative abundances of flavonoids (including compounds 7, 8, 9, 10, 11, 12) and organic acids (compound 14) were significantly higher in CR than in NR. In addition, five compounds ephedrine, pseudoephedrine, DL-norephedrine, L-phenylalanine, and aminobenzoic acid shared both stems and roots. However, the abundances of ephedrine, pseudoephedrine, and DL-norephedrine in CS and NS were significantly higher than in CR or NR.

The visual network indicated that those differential metabolites involved at least eleven metabolic pathways (Fig. 3): map00360 (Phenylalanine metabolism), map00940 (Phenylpropanoid biosynthesis), map00944 (Flavone and flavonol biosynthesis), map00941 (Flavonoid biosynthesis), map00900 (Terpenoid backbone biosynthesis), map00960 (Tropane, piperidine, and pyridine alkaloids biosynthesis), map01062 (Biosynthesis of terpenoids and steroids), map00400 (Phenylalanine, tyrosine and tryptophan biosynthesis), map00350 (Tyrosine metabolism), map00380 (Tryptophan metabolism), and map00790 (Folate biosynthesis). L-phenylalanine was dehydrogenated by phenylalanine ammonia-lyase to obtain *trans*-cinnamic acid. Then L-phenylalanine and *trans*-cinnamic acid underwent continuous hydrogenation and methylation to produce phenolic acid metabolites, such as phenylacetic acid, dihydro – 3-coumaric acid, ferulic acid, and protocatechualdehyde. The existence of organic acids was a unique feature of plant metabolism. Flavonoids are secondary metabolites derived from phenylpropanoids. In total, eighteen metabolites in stems and six metabolites in roots participated in flavone and flavonol biosynthesis. In addition to flavonoids, coumarin was another benzopyrone of the phenylalanine metabolism. The biosynthetic pathway of terpenoids is involved in the Deoxyxylulose phosphate pathway. This pathway forms 1-deoxy-D-Xylulose 5-phosphate (DXP) from pyruvate and D-aldehyde 3-phosphate as starting materials. After several transformation steps, they reacted to isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). IPP and DMAPP were condensed head to tail under the action of the enzyme to form Geranyl diphosphate (GPP), a critical synthetic precursor of monomers, which further synthesizes limonene, *β*-myrcene, and *α*-pinene. Anthranilic acid involved in phenylalanine, tyrosine and tryptophan biosynthesis may be transformed from chorismate and L-kynurenine from tropane piperidine and pyridine alkaloid biosynthesis tryptophan metabolism, respectively.

Endophytic Community Characteristics in Wild and Cultivated E. sinica

Microbial diversity based on high-throughput sequencing revealed that the number of endophytic bacteria OTUs was the same order of magnitude between NS and CS, which was also similar between NR and CR. The number of endophytic bacteria OTUs in NR or CR was two orders of magnitude higher than in NS or CS. Nevertheless, it was entirely different for endophytic fungi. The number of endophytic fungi in NS, CS, NR and CR was the same order of magnitude (Table 1). The results of the Venn diagram showed that the number of species of endophytes in each tissue was different. The shared OTUs between CSB and NSB, CRB and NRB (CSB, NSB, CRB and NRB are endophytic bacteria in CS, NS, CR, and NR), CSF and NSF, and CRF and NRF (which are endophytic fungi in CS, NS, CR and NR) were 8, 500, 1452, and 1019 species, respectively. The similarity was 18.60%, 33.62%, 22.47%, and 21.82%, respectively (Supplementary Fig. 2). The coverage of bacteria and fungi was 0.93–0.96 and 0.98–0.99, respectively, and the ecological indexes of these endophytic OTUs were shown in Table 1.

Undoubtedly, Proteobacteria was the most predominant endophytic bacterial phylum across all samples. For instance, *Pseudomonas* exhibited high relative abundance in NS (84.23%) and CS (91.01%) (Fig. 4), while

Streptococcus (2.25%) was unique in NS (Fig.S3a). In addition, the shared bacteria between NR and CR (relative abundance > 1%) were mainly *Steroidobacter*, *Pseudomonas*, *Sandaracinus*, *Povalibacter*, *Pseudonocardia*, *Devosia*, and Acidobacteria sp. (Fig. 4). Nevertheless, eleven endophytic bacteria were significantly enriched in NR, including *Phytohabitans*, *Nocardioides*, *Actinophytocola*, *Streptomyces*, *Ohtaekwangia*, *Asticcacaulis*, *Mesorhizobium*, *Phyllobacterium*, *Rhizobium*, *Sphingomonas, Variovorax*, Acidobacteria sp., and *Kofleria* were significantly enriched in CR (Supplementary Fig. 3b). For endophytic fungi, the majority of OTUs were allocated to Ascomycota, except for the unknown species (Fig. 4). Both NS and CS contained *Alternaria* (0.2% vs. 0.07%) and *Passalora* (0.17% vs. 0.09%), but the exclusive fungus of the former was Didymellaceae sp. (0.02%), and the unique fungi of the latter were *Fusarium* (0.02%), *Malbranchea* (0.01%), and *Colletotrichum* (0.01%) (Fig. 4). The shared fungi of NR and CR were species from *Fusarium*, Onygenales sp., *Ilyonectria*, *Paraphoma*, Pleosporales sp., *Alternaria*, and some unknown species of Ascomycota. However, their abundance fluctuated. For example, *Fusarium* (0.61%) in NR was 20 times higher than that in CR (0.03%). In addition, *Aporospora* (0.2%), *Entodesmium* (0.06%), Auriculariales sp. (0.01%), and *Gibberella* (0.01%) were only present in NR, and *Arthrographis* (0.02%) was just detected in CR (Fig. 4).

Interrelation Among Endophytic Microorganisms

The interaction differences of intra-kingdom (bacteria-bacteria; fungi-fungi) were assessed through a co-occurrence network under the wild and cultivated conditions (Fig. 5). Specifically, the topological properties of nodes were utilized to characterize complex patterns of microbiota. Overall, the correlations among microorganisms occurring within co-occurrence networks were predominantly positive (ranging from 60.98–100%), indicating that commensalism or mutualism dominated the interactions among microorganisms (Fig. 5 and Supplementary data Table 2). Remarkably, the microbial network exhibited much more complex and intricate co-occurrence network patterns in the stems or roots of the wild *E. sinica* compared to the cultivated species network (more nodes, edges, average clustering coefficient (ACC), and average degree) (except CSF vs. NSF). The modularity value (an index indicating that OTUs are well-connected in their modules rather than the others) of each network in wild *E. sinica* (both stems and roots) was greater than 0.4, which suggested that the synergistic effect of these colonies was more robust than that in cultivated *E. sinica* (Supplementary data Table 2).

Effect of Endophytes on Differential Metabolism of E. sinica

The Spearman's correlation with the two-tailed Student's t-test demonstrated that eight alkaloids showed positive correlations with endophytes in stems (Fig. 6a, b). For instance, ephedrine was correlated with six endophytic fungi OTUs: OTU137, OTU329, OTU141, OTU744, OTU337, and OTU707. Additionally, both pseudoephedrine and DL-norephedrine were associated with the endophytic bacterium OTU243 (*Escherichia/ Shigella*) (Fig. 6a), the endophytic fungus OTU147, and so on (Fig. 6b). The *Pseudomonas* (OTU18/OTU19) also had significant positive correlations with coumarin, kaempferol, rhamnetin, and phenylacetic acid, respectively (Fig. 6a). Additionally, the endophytic fungus OTU53653 (*Alternaria*) shared in the wild and cultivated stems of *E. sinica*, was significantly positively associated with three terpenoids including *β*-myrcene, farnesal, and thymol. However, *Alternaria* was significantly negatively correlated with the majority of metabolites, such as D-cathinone, triacetin, 6-hydroxykaempferol, (+)-gallocatechin, 3-*O*-methylquercetin, anthranilic acid, *trans*-beta-D-glucosyl-2-hydroxycinnamate, phenylacetic acid, and aminobenzoic acid (Fig. 6b). The unique endophytic bacterium OTU440 (*Streptococcus*) from the stems of wild *E. sinica* was the largest contributor to the production of isovitexin 2'''-*O*-arabinoside (Fig. 6a).

It is worth noting that some common fungi and bacteria in roots were remarkably correlated with the bioactive components (Fig. 6c, d). For example, OTU5034 (*Nocardia*) and OTU6296 (Rhizobiales sp.) exhibited a positive correlation with pseudoephedrine but a negative correlation with ephedrine (Fig. 6d). Endophytic fungi OTU54294 *(Fusarium)* (Fig. 6c) and endophytic bacteria OTU6230 (*Sphingomonas*), OTU2116 (*Ohtaekwangia*), OTU45 (Acidobacteria sp.), OTU6266 (*Caulobacter*), and OUT5887 (*Bosea*) (Fig. 6d) were positively correlated with ferulic acid and cinnamyl alcohol, while negatively correlated with afzelechin and epicatechin-(4beta->8)-ent-epicatechin. Furthermore, the unique endophytic fungus OTU51523 (*Aporospora*) in the roots of wild *E. sinica* had the mostly positive relationship with propanoylagmatine, cinnamaldehyde, chavicol, (+)-syringaresinol, cinnamyl alcohol, and (*S*)-2,3,4,5-tetrahydropyridine-2-carboxylate (Fig. 6c).

By predicting the function of interacting genes several functional genes of microorganisms involved in the regulation of the metabolism of secondary metabolites of *E. sinica* were found (Fig. 3, Fig. 7 and Supplementary Fig. 4). For example, in the metabolic pathway of phenylalanine metabolism, a total of four enzymes were involved in L-phenylalanine to phenylacetic acid (Fig. 3 and Fig. 7): aromatic-L-amino-acid/L-tryptophan decarboxylase (K01593 DDC), monoamine oxidase (K00274 MAO), primary-amine oxidase (K00276 AOC3), and aldehyde dehydrogenase (NAD (P)+) (K00129 ALDH3). Of these, the monoamine oxidase (K00274 MAO) coming from OTU243 (*Escherichia*/*Shigella*) was observed to be considerably up-regulated in CSB (Fig. 7 and Supplementary Fig. 4a), which indicated that OTU243 (*Escherichia*/*Shigella*) might contribute to the production of phenylacetic acid in CS (r = 0.621) (Fig. 3 and Fig. 6). Notably, several microorganisms regulated some enzymes involved in metabolite formation simultaneously. For instance, the enzyme para-aminobenzoate synthetase component II (K01664 pabA), showed significant positive correlations with fifteen endophytic bacteria, including the following OTU5023 (*Actinophytocola*), OTU4889 (*Streptomyces*), OTU6239 (*Phyllobacterium*), OTU4890 (*Phytohabitans*), *Steroidobacter* (OTU60 and OTU226), OTU5026 (*Promicromonospora*), OTU6242 (*Rhizomicrobium*), OTU4774 (*Micromonospora*), OTU6252 (*Roseomonas*), OTU201 (*Pseudoxanthomonas*), OTU2526 (*Flavitalea*), OTU5039 (*Kribbella*), OTU225, and OTU766 (*Opitutus*). Such correlations could convert chorismate to aminobenzoic acid in the folate biosynthesis. Moreover, 4-coumarate–CoA ligase (K01904 4CL) was found in many bacteria, including OTU6239, OTU201, OTU5026, OTU2526, OTU6242, OTU766, and OTU4891 (*Nocardioides*), with direct or indirect effects on cinnamaldehyde production (Supplementary Fig. 4b) (Fig. 3). The other enzyme-encoding genes, such as kynureninase (K01556 KYNU), tryptophan 2,3-dioxygenase (K00453 TDO2), and aldehyde dehydrogenase (NAD+) (K00128 ALDH) were widespread in bacteria OTU192 (*Ralstonia*) and fungus OTU52748 (*Fusarium*) which contributed significantly to the tryptophan metabolism, etc. (Fig. 3 and Supplementary Fig. 4a, d).

Discussion

The quality of Chinese medicinal materials depends on the variable content and quantity of chemical compositions (Zhang et al. 2022). Due to technical constraints, the limited and fragmented knowledge of the disparity in quality between wild and cultivated *E. sinica* was mainly reflected in alkaloids with high relative contents (Sun et al. 2012; Miao et al. 2022). In our study, 47 and 23 differential metabolites were found through holistic analyses of the metabolome in the stems and roots of the wild and cultivated varieties. Among them, the content of alkaloids (ephedrine, pseudoephedrine, and DL-norephedrine), flavonoids (such as luteolin, kaempferol, quercetin, tricin, and vitexin), and organic acids (such as anthranilic acid, quinaldic acid, phenylacetic acid, and aminobenzoic acid) in CS were significantly higher than those in NS. In comparison, the terpenoids, including limonene, *β*-myrcene, farnesal, thymol, and *α*-pinene, were produced higher in NS than in CS. In addition, a lot of phenylpropanoids (chavicol, (+)-syringaresinol, cinnamaldehyde and cinnamyl alcohol) and alkaloids (norpseudoephedrine,

propanoylagmatine and piperettine) were accumulated in NR, whereas higher contents of flavonoids (such as afzelechin, biochanin A, proanthocyanidin A2, and Delphinidin 3-glucosylglucoside) and organic acids (eg. terephthalic acid) were accumulated in CR (Table A1). The accumulation had an essential contribution to the pharmaceutical effect of *E. sinica* (Oshima et al. 2019).

The effects of endophytes on the differentiation of chemical constituents between the wild and cultivated *E. sinica* were studied under the same phenological conditions. It was found that there were discrepancies in the number of OTUs and community distribution of endophytes between the wild and cultivated *E. sinica*, which might be caused by the agricultural activities and the natural selection of host plants. Furthermore, both wild and cultivated varieties had specific endophytic communities, indicating the selective effects on particular endophytes. For example, *Streptococcus* was only presented in NS, while the unique endophytic fungi in NS, CS, NR, and CR were Didymellaceae spp., *Fusarium*, *Aporospora* and *Arthrographis*, respectively. However, Suleiman et al. (2017) reported that plants might have robust core microbiome compositions that were less prone to the alteration due to variation in land use, soil type, or edaphic factors. This means that there was a highly conserved and co-evolved microbiome preserved during the domestication of *E. sinica*. These observations are consistent with the other studies (Abdullaeva et al. 2020; Guo et al. 2021). *Pseudomonas* were the dominant bacterial genus in both NS and CS, which were correlated both significantly and positively with the most differential metabolites, such as D-cathinone, methcathinone, coumarin, kaempferol, and rhamnetin. Hence, *Pseudomonas* most likely played an essential role in metabolites forming in the stems of *E. sinica*. Common predominant bacteria of NR and CR included *Steroidobacter*, *Pseudomonas*, *Sandaracinus*, *Povalibacter*, and *Pseudonocardia*, all of which could promote growth and protect plants from host plant diseases as previously reported (Kinkel et al. 2012; Sah et al. 2021). Moreover, a high proportion of *Sphingomonas* in the roots of wild *E. sinica* enhanced the drought resistance of host plants by releasing volatile organic compounds (Luo et al. 2020). The co-occurrence networks also clarified that endophytes improved the stress resistance of wild *E. sinica* by modulating its phenotypic plasticity, which was essential for its healthy growth.

The gene functional prediction results showed that there were microbial genes involved in differential metabolite accumulation in the wild or cultivated *E. sinica*. Therefore, endophytes are very critical as the hosts' endo-environment for their growth, development, and metabolite formation of medicinal plants. For example, the endophytic bacteria OTU243 (*Escherichia/Shigella*) produced monoamine oxidase (K00274 MAO) (Ramesh et al. 2016), which was highly expressed in CSB and had a close correlation with phenylacetic acid ($r = 0.621$). Therefore, OTU243 (*Escherichia/Shigella*) might be involved in forming phenylacetic acid, a metabolite of *E. sinica* (Zhai et al. 2017; Motoyama et al. 2021; Liu et al. 2021). However, the more elaborate mechanisms of action are not well understood, and more evidence is required. Such reports have sprung up in recent years including several studies revealed that *Salvia miltiorrhiza*–microbe interactions might enhance the biomass production of *S. miltiorrhiza*, which could affect the metabolic pathway for tanshinone production (Huang et al. 2018). In summary, further exploration of the effects of endophytes on the metabolism is of great significance for the precise cultivation of *E. sinica*.

Declarations

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Jin-Long Cui, Xin Ji, Xi Liu and Shuang-Man Miao. The first draft of the manuscript was written by Hui Zhang and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This work was financially supported by the Basic Research Program of Shanxi Province (20210302123428), Research Project Supported by Shanxi Scholarship Council of China (2020-013), and Fund Program for the Scientific Activities of Selected Returned Overseas Professionals in Shanxi Province (20220004).

Data Availability The raw sequence data generated as part of this study has been deposited in NCBI (https://submit.ncbi.nlm.nih.gov) and was publicly available with Short Read Archive accession: PRJNA735567.

Competing Interest The authors declare no competing interests.

References

1. Abdullaeva Y, Ambika Manirajan B, Honermeier B, Schnell S, Cardinale M (2020) Domestication affects the composition, diversity, and co-occurrence of the cereal seed microbiota. J Adv Res 31: 75-86 doi: https://doi.org/10.1016/j.jare.2020.12.008
2. Bastian M, Heymann S, Jacomy M (2009) Gephi: an open source software for exploring and manipulating networks. International AAAhI Conference on Weblogs and Social Media 3: 361-362 doi: https://doi.org/10.13140/2.1.1341.1520
3. Cao M, Liu Y, Jiang W, Meng X, Zhang, W, Chen, W, Peng D, Xing, S (2020) UPLC/MS-based untargeted metabolomics reveals the changes of metabolites profile of *Salvia miltiorrhiza* bunge during sweating processing. Sci Rep 10: 19524 doi: https://doi.org/10.1038/s41598-020-76650-w
4. Cook D, Gardner DR, Pfister JA (2014) Swainsonine-containing plants and their relationship to endophytic fungi. J Agric Food Chem 62: 7326-7334 doi: https://doi.org/10.1021/jf501674r
5. Cui JL, Gong Y, Vijayakumar V, Zhang G, Wang ML, Wang JH, Xue XZ (2019) Correlation in chemical metabolome and endophytic mycobiome in *Cynomorium songaricum* from different desert locations in China. J Agric Food Chem 67: 3554-3564 doi: https://doi.org/10.1021/acs.jafc.9b00467
6. Cui JL, VijayakumarV, Zhang G (2018) Partitioning of fungal endophyte assemblages in root-parasitic plant *Cynomorium songaricum* and its host *Nitraria tangutorum*. Front Microbiol 9: 666 doi: https://doi.org/10.3389/fmicb.2018.00666
7. Douglas GM, Maffei VJ, Zaneveld JR, Yurgel SN, Brown JR, Taylor CM, Huttenhower C, Langille M (2020) PICRUSt2 for prediction of metagenome functions. Nat Biotechnol 38: 685-688 doi: https://doi.org/10.1038/s41587-020-0548-6
8. Frank AC, Saldierna Guzmán JP, Shay JE (2017) Transmission of bacterial endophytes. Microorganisms 5: 70 doi: https://doi.org/10.3390/microorganisms5040070
9. Guo J, Ling N, Li Y, Li K, Ning H, Shen Q, Guo S, Vandenkoornhuyse P (2021) Seed-borne, endospheric and rhizospheric core microbiota as predictors of plant functional traits across rice cultivars are dominated by deterministic processes. The New phytologist 230: 2047-2060 doi: https://doi.org/10.1111/nph.17297
10. Huang W, Long C, Lam E (2018) Roles of plant-associated microbiota in traditional herbal medicine. Trends Plant Sci 23: 559-562 doi: https://doi.org/10.1016/j.tplants.2018.05.003
11. Inngjerdingen KT, Meskini S, Austarheim I, Ballo N, Inngjerdingen M, Michaelsen TE, Diallo D, Paulsen BS (2012) Chemical and biological characterization of polysaccharides from wild and cultivated roots of *Vernonia kotschyana*. J Ethnopharmacol 139: 350-358 doi: https://doi.org/10.1016/j.jep.2011.10.044
12. Kinkel LL, Schlatter DC, Bakker MG, Arenz BE (2012) *Streptomyces* competition and co-evolution in relation to plant disease suppression. Res Microbiol 163: 490-499 doi: https://doi.org/10.1016/j.resmic.2012.07.005

13. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, Glockner FO (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. Nucleic Acids Res 41: e1 doi: https://doi.org/10.1093/nar/gks808
14. Luo Y, Zhou M, Zhao Q, Wang F, Gao J, Sheng H, An L (2020) Complete genome sequence of *Sphingomonas* sp. Cra20, a drought resistant and plant growth promoting rhizobacteria. Genomics 112: 3648-3657 doi: https://doi.org/10.1016/j.ygeno.2020.04.013
15. Liu X, Zhou ZY, Cui JL, Wang ML, Wang JH (2021) Biotransformation ability of endophytic fungi: from species evolution to industrial applications. Appl Microbiol Biotechnol 105: 7095-7113 doi: https://doi.org/10.1007/s00253-021-11554-x
16. Miao SM, Zhang Q, Bi XB, Cui JL, Wang ML (2020) A review of the phytochemistry and pharmacological activities of *Ephedra* herb. Chin J Nat Med 18: 321-344 doi: https://doi.org/10.1016/S1875-5364(20)30040-6
17. Miao SM, Xia Y, Cui JL, Wang JH, Wang ML (2022) Correlation analysis between differential metabolites and bacterial endophytes of *Ephedra sinica* in different years. Ind Crops Prod 114250: 175 doi: https://doi.org/10.1016/S1875-5364(20)30040-6
18. Motoyama T, Yun CS, Osada H (2021) Biosynthesis and biological function of secondary metabolites of the rice blast fungus *Pyricularia oryzae*. J Ind Microbiol Biotechnol 48: kuab058 doi: https://doi.org/10.1093/jimb/kuab058
19. Oshima N, Yamashita T, Uchiyama N, Hyuga S, Hyuga M, Yang J (2019) Non-alkaloidal composition of *Ephedra* Herb is influenced by differences in habitats. J Nat Med 73: 303-311 doi: https://doi.org/10.1007/s11418-018-1265-z
20. Rasmussen S, Parsons AJ, Fraser K, Xue H, Newman JA (2008) Metabolic profiles of *Lolium perenne* are differentially affected by nitrogen supply, carbohydrate content, and fungal endophyte infection. Plant Physiol 146: 1440-1453 doi: https://doi.org/10.1104/pp.107.111898
21. Sah S, Krishnani S, Singh R (2021) *Pseudomonas* mediated nutritional and growth promotional activities for sustainable food security. Curr Res Microb Sci 2: 100084 doi: https://doi.org/10.1016/j.crmicr.2021.100084
22. Suleiman AKA, Pylro VS, Roesch LFW (2017) Replacement of native vegetation alters the soil microbial structure in the Pampa. Sci agric 74: 77-84 doi: https://doi.org/10.1590/1678-992X-2015-0494
23. Sun H, Zhang AH, Wang XJ (2012) Potential role of metabolomic approaches for Chinese medicine syndromes and herbal medicine. Phytother Res 26: 1466-1471 doi: https://doi.org/10.1002/ptr.4613
24. Strobel G, Daisy B (2003) Bioprospecting for microbial endophytes and their natural products. Microbiol Mol Biol Rev 67: 491-502 doi: https://doi.org/10.1128/MMBR.67.4.491-502.2003
25. Zhang BM, Wang ZB, Xin P, Wang QH, Bu H, Kuang HX (2018) Phytochemistry and pharmacology of genus *Ephedra*. Chin J Nat Med 16: 811-828 doi: https://doi.org/10.1016/S1875-5364(18)30123-7
26. Zhang GF, Huang QL, Bi XQ, Liu YL, Yuan ZS (2020b) Analysis of endophytic bacterial community diversity and metabolic correlation in *Cinnamomum camphora*. Arch Microbiol202:181-189 doi: https://doi.org/10.1007/s00203-019-01733-w
27. Zhang H, Zhang Y, Zhang T, Liu C (2022) Research progress on quality markers of traditional Chinese medicine. J Pharm Biomed Anal 211: 114588 doi: https://doi.org/10.1016/j.jpba.2022.114588
28. Zhang Q, Xue XZ, Miao SM, Cui JL, Qin XM (2020a) Differential relationship of fungal endophytic communities and metabolic profiling in the stems and roots of *Ephedra sinica* based on metagenomics and metabolomics. Symbiosis 81: 115-125 doi: https://doi.org/10.1007/s13199-020-00685-w

29. Zhou Z, Gao N, Wang Y, Chang P, Tong Y, Fu S (2020) Clinical studies on the treatment of novel coronavirus pneumonia with traditional Chinese medicine-A literature analysis. Front Pharmacol 11: 560448 doi: https://doi.org/10.3389/fphar.2020.560448
30. Zhai X, Jia M, Chen L, Zheng CJ, Rahman K, Han T, Qin LP (2017) The regulatory mechanism of fungal elicitor-induced secondary metabolite biosynthesis in medical plants. Crit Rev Microbiol 43: 238-261 doi: https://doi.org/10.1080/1040841X.2016.1201041

Figures

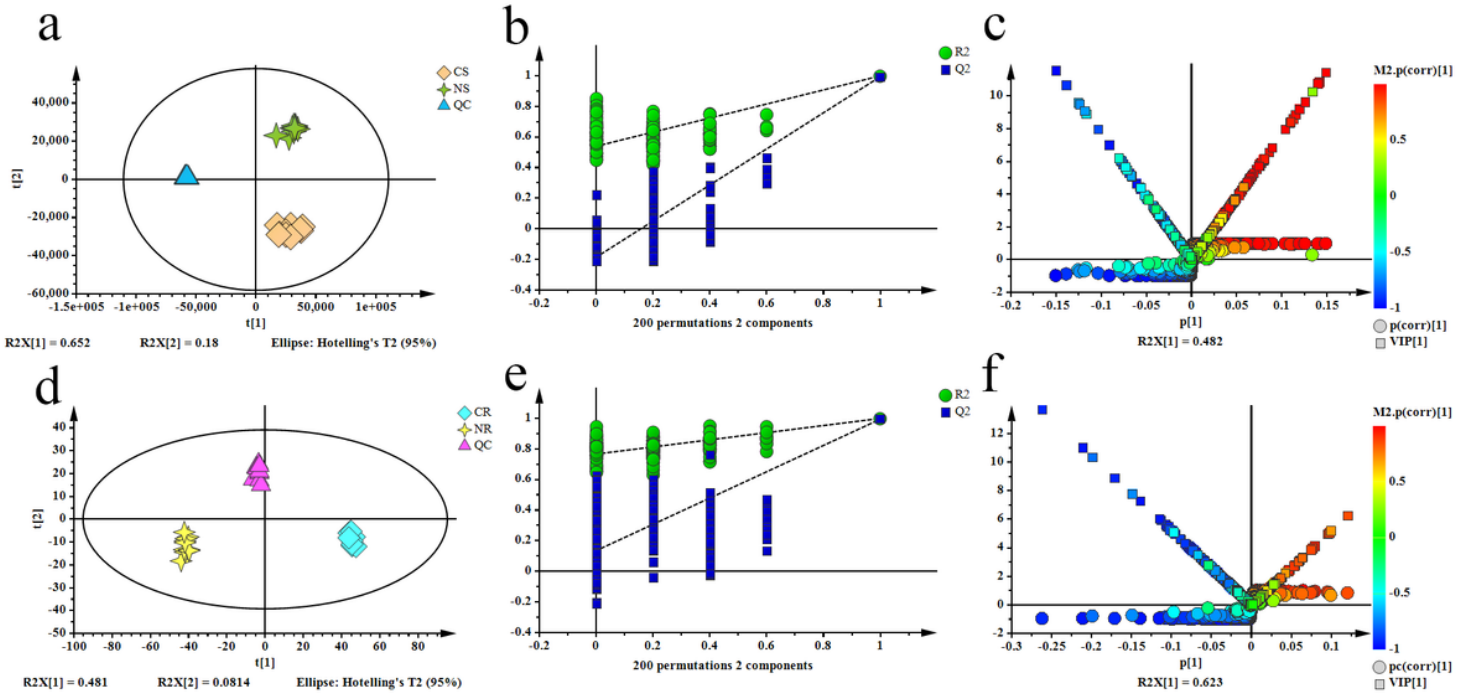
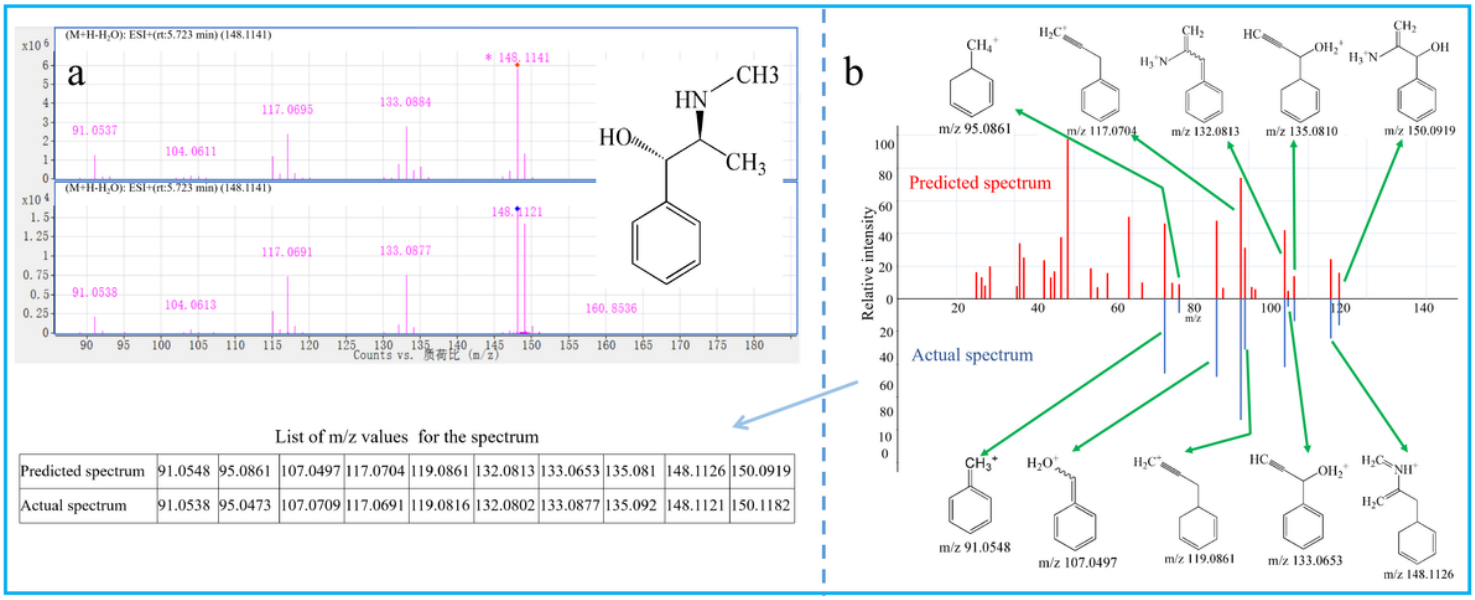


Figure 1

Model of multivariate analysis and its cross-validation based on the LC-Q/TOF-MS data. PCA score plot (a, d), permutation test plots (b, e), and combination of S- and VIP score plots constructed from OPLS-DA (c, f). CS and NS represent stems of cultivated and wild *E. sinica*, respectively; CR and NR represent roots of cultivated and wild *E. sinica*, respectively



Predicted spectrum	91.0548	95.0861	107.0497	117.0704	119.0861	132.0813	133.0653	135.081	148.1126	150.0919
Actual spectrum	91.0538	95.0473	107.0709	117.0691	119.0816	132.0802	133.0877	135.092	148.1121	150.1182

Figure 2

Mass spectrometric identification of the differential marker metabolite M148T6 as Pseudoephedrine. Predicted MS/MS match with Agilent MassHunter Qualitative analysis in the PCDL database and the structure of M148T6 (a). The predicted and actual list of m/z values for the M148T6 spectrum and MS/MS match in the HMDB and METLIN databases (b)

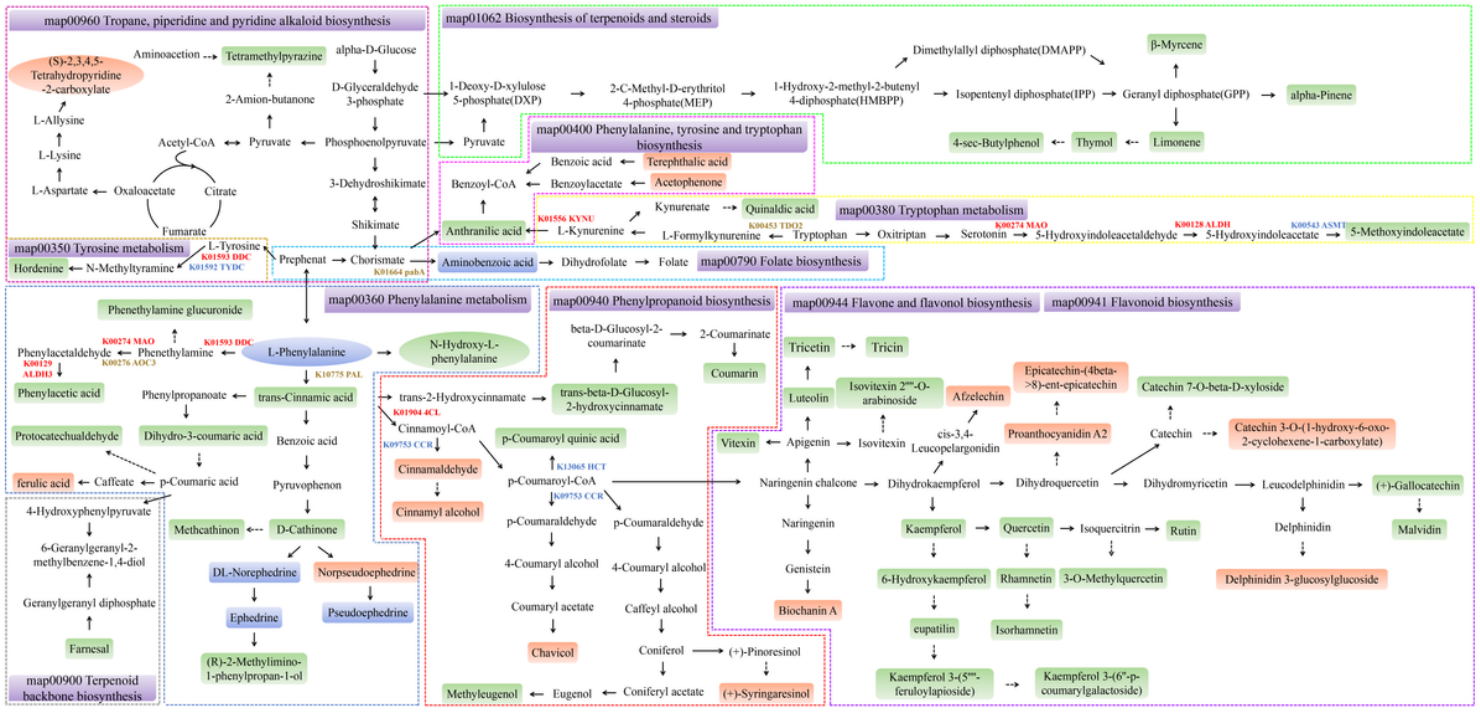


Figure 3

Pathway of differential metabolites between wild and cultivated *Ephedra sinica*. The orange boxes represent unique differential metabolites in roots; the green boxes represent unique differential metabolites in stems; the blue boxes

represent the common differential metabolites in both stems and roots; the Peru KOs represent bacterial genes; the steel blue KOs represent fungal genes; the red KOs represent the common genes in both bacteria and fungi

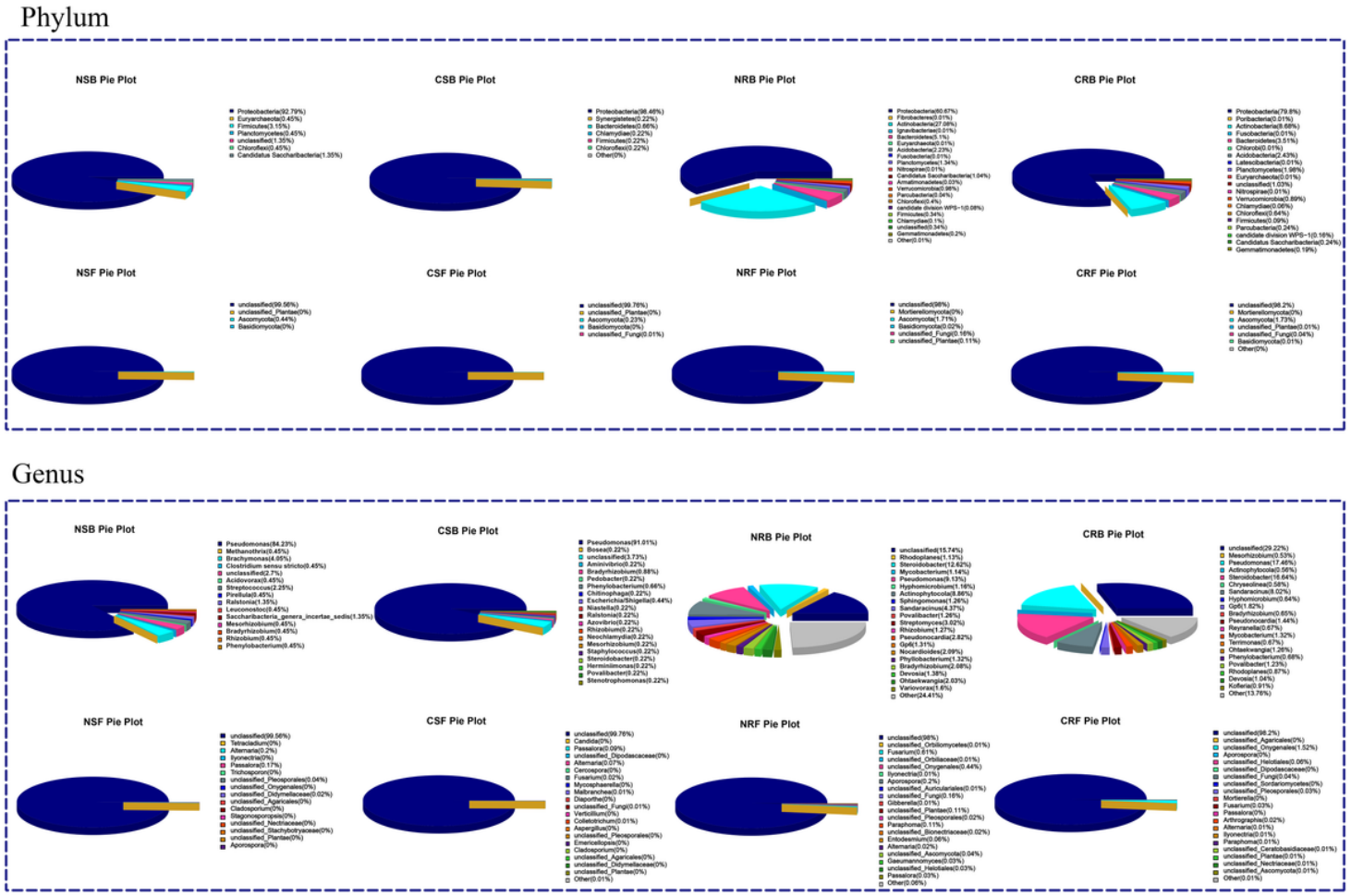


Figure 4

Relative abundance (%) at the Phylum and genus level of endophytes from each sample of *Ephedra sinica*. CSB (CSF) and NSB (NSF) are endophytic bacteria (fungi) in the stems of cultivated and wild *E. sinica*, respectively; CRB (CRF) and NRB (NRF) are endophytic bacteria (fungi) in the roots of cultivated and wild *E. sinica*, respectively

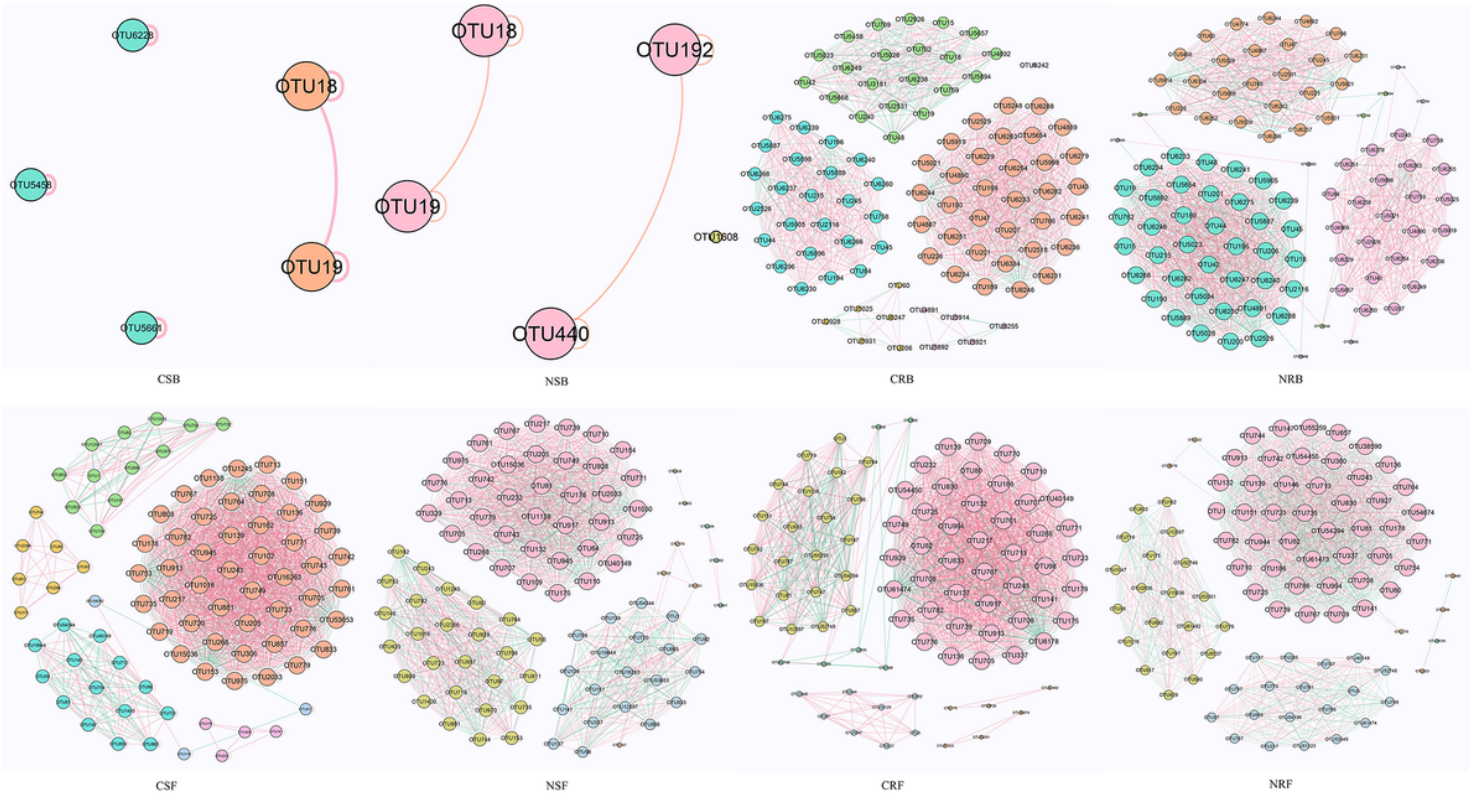


Figure 5

The co-occurrence network of bacteria and fungi from *Ephedra sinica* based on correlation analysis. CSB (CSF) and NSB (NSF) are endophytic bacteria (fungi) in the stems of cultivated and wild *E. sinica*, respectively; CRB (CRF) and NRB (NRF) are endophytic bacteria (fungi) in the roots of cultivated and wild *E. sinica*, respectively

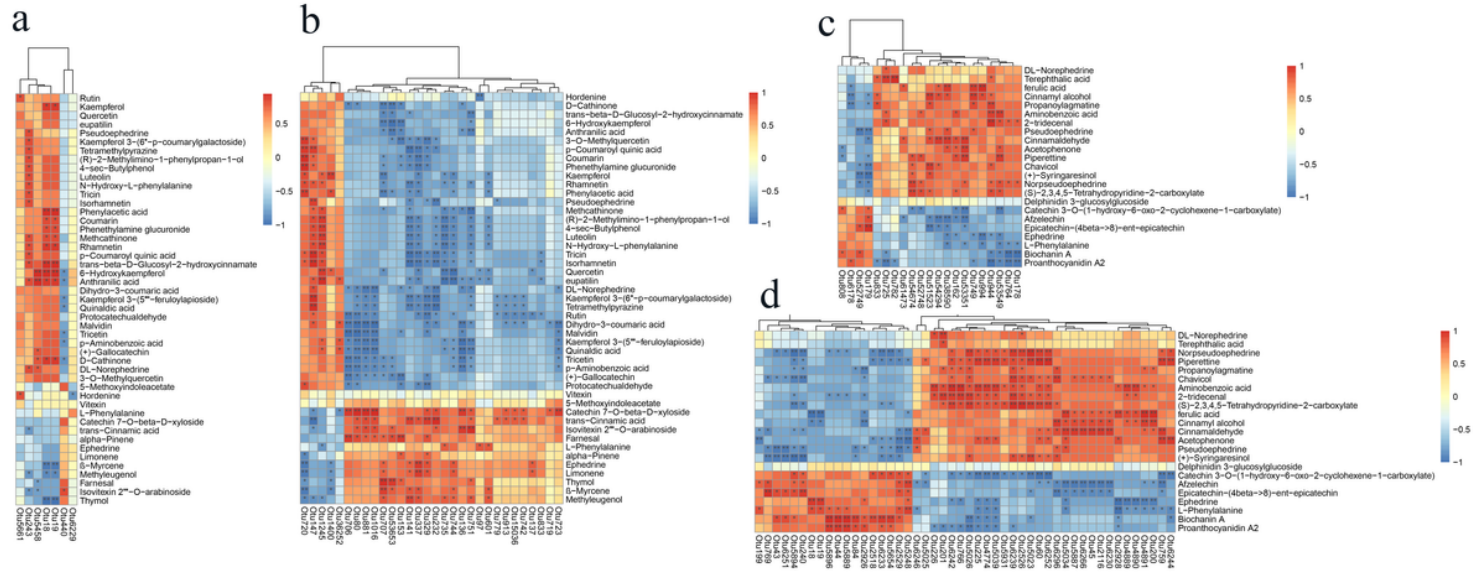


Figure 6

Heat maps of the correlations between bacterial and fungal OTUs and metabolites in each sample of *Ephedra sinica*. Color intensity indicates the degree of correlation: significant (*$P < 0.05$) and extremely significant differences (**$P < 0.01$). (a) and (b) refer correlation coefficient matrix of bacterial and fungal OTUs and

metabolites in stems of *E. sinica*, respectively. (c) and (d) refer correlation coefficient matrix of fungal and bacterial OTUs and metabolites in roots of *E. sinica*, respectively

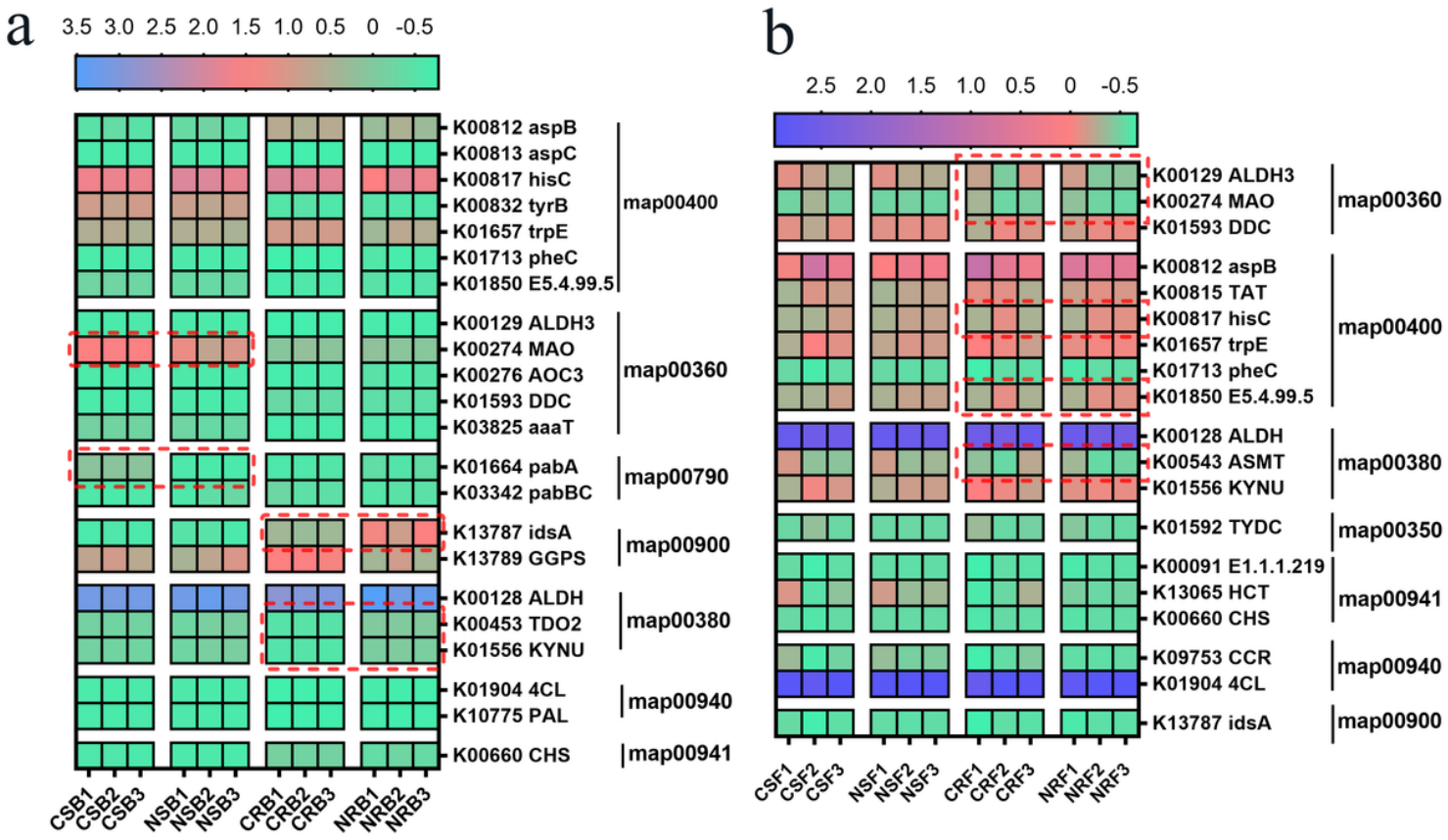


Figure 7

Variation of bacterial (a) and fungal (b) functional profiles from each sample of *Ephedra sinica* analyzed by PICRUSt2. CSB (CSF) and NSB (NSF) are endophytic bacteria (fungi) in the stems of cultivated and wild *E. sinica*, respectively; CRB (CRF) and NRB (NRF) are endophytic bacteria (fungi) in the roots of cultivated and wild *E. sinica*, respectively

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarymaterial.zip