

# Structure and function of rhizosphere and root endophyte microbial communities associated with healthy and root rot diseased *Panax notoginseng*

Pan Pan Wang (✉ [wangpanpankmust@163.com](mailto:wangpanpankmust@163.com))

Kunming University of Science and Technology <https://orcid.org/0000-0003-0851-5028>

Li Fang Yang

Kunming University of Science and Technology

Jia Ling Sun

Kunming University of Science and Technology

Ye Yang

Kunming University of Science and Technology

Yuan Qu

Kunming University of Science and Technology

Cheng Xiao Wang

Kunming University of Science and Technology

Di Qiu Liu

Kunming University of Science and Technology

Lu Qi Huang

China Academy of Chinese Medical Sciences

Xiu Ming Cui

Kunming University of Science and Technology

Yuan Liu

Kunming University of Science and Technology

---

## Research Article

**Keywords:** *Panax notoginseng* root rot, rhizosphere soil, root endophytes, metagenomic sequencing

**Posted Date:** May 3rd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-478704/v1>

**License:**   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

# Abstract

**Background:** *Panax notoginseng* (Burkill) F. H. Chen is a Chinese medicinal plant of the Araliaceae family commonly used in the treatment of cardiovascular and cerebrovascular diseases in Asia and elsewhere. To meet an increase in Chinese herbal medicine market demand, most *P. notoginseng* is planted artificially, and is vulnerable to various plant diseases. Root rot disease, in particular, causes substantial *P. notoginseng* yield reduction and economic losses. High-depth next-generation sequencing technology was used to analyze the rhizosphere and root endophyte microbial communities of *P. notoginseng* to compare the characteristics of these two communities between healthy and root rot diseased *P. notoginseng* plants, and to clarify the relationship between these microbial communities and root rot disease.

**Results:** The *P. notoginseng* rhizosphere microbial community was more diverse than the root endophyte community, and the difference in functional pathways between healthy and diseased *P. notoginseng* plants was greater in the root endophyte than in the rhizosphere communities. Multi-database annotation results showed that the highest number of endophytic bacteria occurred in the roots of diseased plants. The number of carbohydrate-active enzymes database families was also higher in diseased roots. The RND antibiotic efflux function was higher in the healthy samples. A high abundance of *Variovorax paradoxus* and *Pseudomonas fluorescens* occurred in the healthy and diseased root endophyte communities, respectively. *Ilyonectria mors-panacis* and *Pseudopyrenochaeta lycopersici* were most abundant in the diseased samples. In addition, the complete genome of two unknown *Flavobacteriaceae* species and one unknown *Bacteroides* species were obtained based on binning analysis.

**Conclusions:** The rhizosphere and root endophyte microbial communities of healthy and root rot diseased *P. notoginseng* showed marked differences in diversity and functional pathways. The higher mapping values obtained for the diseased samples reflected the occurrence of root rot disease at the molecular level. *Variovorax paradoxus* and *Pseudomonas fluorescens* may be antagonistic bacteria of root rot in *P. notoginseng*, whereas *Ilyonectria mors-panacis* and *Pseudopyrenochaeta lycopersici* appear to be *P. notoginseng* root rot pathogens. Our study provides a theoretical basis for understanding the occurrence of root rot in *P. notoginseng* and for further research on potential biological control agents.

## Background

*Panax notoginseng* (Burk) F. H. Chen is a plant belonging to the *Panax* genus of the Araliaceae family. It is a valuable Chinese herbal medicine that is mainly grown in the Yunnan and Guangxi Provinces in China, and commonly used for the treatment of coronary heart and cardiovascular diseases [1, 2]. In addition, *P. notoginseng* has hemostatic, anti-thrombotic, anti-atherosclerotic, and neuroprotection effects as well as reducing blood pressure and relieving pain [3, 4]. *Panax notoginseng* is used as a therapeutic drug or health care product in China and other Asian countries, and the root, particularly, is used as a raw material in more than 400 products manufactured by 1,300 companies in China [5, 6]. The World Health Organization reports that approximately one-third of all deaths worldwide are attributed to cardiovascular

diseases [7]. With the increase in the occurrence of cardiovascular diseases, the market demand for *P. notoginseng* root material is increasing. However, *P. notoginseng* is vulnerable to a variety of plant diseases of which root rot disease is the most serious problem limiting *P. notoginseng* production. Root rot in *P. notoginseng*, caused by this disease, can result in a reduction in yield of 5% up to 20% or even above 70% during the entire growing period. Researchers have found that *Cylindrocarpon destructans* var. *destructans* causes root rot disease in *P. notoginseng* [8], and using ITS sequencing technology, the dominant flora in root rot disease of *P. notoginseng* has been identified as *Ilyonectria mors-panacis* [9]. Some natural plant ingredients are known to have a strong inhibitory effect on the pathogenic bacteria *Fusarium oxysporum* and *F. solani* associated with root rot disease in *P. notoginseng* [10]. Previously, researchers analyzed the fungal diversity in the rhizosphere soil of two-year-old healthy *P. notoginseng* plants using ITS sequencing technology, and molecular and morphological characteristics, and found that there are a variety of potential pathogenic fungi in the rhizosphere soil [11]. *Panax notoginseng*, along with most other medicinal plants, is susceptible to continuous cropping obstacles which can cause serious diseases. Approximately 70% of medicinal plants have different degrees of pathogen infection related to continuous cropping obstacles [12, 13]. Under continuous cropping and replanting systems, the *P. notoginseng* mortality rate is reported to be 2.0%–39.4% and 2.6%–81.2%, respectively, with a higher mortality rate under replanting than under continuous cropping [14]. Owing to the influence of various plant diseases and continuous cropping obstacles, and the Daodi (medicinal materials with a long history, suitable origin, excellent varieties, great yield, sophisticated processing, outstanding curative effect, and regional characteristics) of *P. notoginseng*, the available planting area of *P. notoginseng* is decreasing year by year. Therefore, to increase the production of *P. notoginseng*, it is necessary to identify the key pathogenic bacteria associated with root rot in this species. This knowledge will provide a theoretical basis for the development of antagonistic antibacterial or biological control agents.

The rhizosphere microbial community plays an important role in nutrient absorption and resistance to environmental pressures. Similarly, the rhizosphere microbial community is also very important for plant health and disease suppression in rhizosphere soil, and the diversity of the microbial community in the rhizosphere soil is considered to be closely related to soil-borne diseases. Previous studies have shown that the rhizosphere soil microbial diversity of healthy plants is greater than that of diseased plants [15, 16]. Recent studies have also revealed that plant stress responses are linked to the composition of beneficial plant microbial communities [17]. The relationship between different soybean varieties and rhizosphere microbial communities shows that resistant soybean varieties are rich in genes encoding phenazines, phospholipids, and other antifungal-related genes, and more rhizosphere bacteria resistant to the *Fusarium* genotype can be synthesized [18]. Studies have analyzed the microbial communities in the rhizosphere soil and roots of healthy and diseased *P. notoginseng* through bacterial 16S rRNA and fungal 18S rRNA genes. The results show that, compared to healthy *P. notoginseng*, the alpha diversity of the microbial communities was lower in the rhizosphere soil and root of diseased *P. notoginseng*. In the diseased rhizosphere soil, the differences in microbial communities increased, whereas fungal community differences decreased. However, the bacterial and fungal communities in the roots were not markedly different between healthy and diseased *P. notoginseng* [19]. Studies have shown that after root

rot infection of *P. notoginseng*, the content of phenolic acid in the soil is lower than that associated with healthy *P. notoginseng*. It has been found that phenolic acid can inhibit the growth of pathogenic bacteria while also stimulating the production of ferric acid, thus, adjusting the content of phenolic acid can effectively prevent root rot disease occurrence in *P. notoginseng* [20].

Metagenomics refers to the DNA of the entire microbial community under study at one point in time, and evidence of all microorganisms in the environment can be obtained, especially of those that are difficult to cultivate [21]. Recently, an increasing number of researchers have applied metagenomics technology to help solve disease-related problems in plants, such as potato [22], peas [23, 24], citrus [25], and tomato [26]. Researchers use metagenomics technology to detect and discover pathogenic bacteria, and such technology provides an effective approach to the prevention and control of plant diseases. Researchers have used metagenomics to investigate the changes in soil microorganisms under continuous cropping and to effectively alleviate the problems caused by continuous cropping obstacles [27].

Metagenomics has played an important role in research on medicinal plant diseases. Study of the continuous cropping obstacles affecting *P. ginseng* has been a focus of research but has been challenging. With the rapid development of high-throughput sequencing technology, researchers have carried out research on root rot disease in *P. notoginseng* using amplicon technologies such as 16S rRNA and ITS, but mostly focused on fungi. Fungi, endophytes, and rhizosphere microorganisms of other medicinal plants of the *Araliaceae*, such as *P. ginseng* and *P. quinquefolius*, have also been recorded [28–32]; however, there are few reports on the endophytes and rhizosphere microorganisms of *P. notoginseng*. In this study, high-throughput sequencing technology was used to study the rhizosphere soil and root endophytes of healthy and root rot diseased *P. notoginseng* plants, and to analyze and compare the microbial community structure and function of healthy and diseased plants. We expected a study of the correlation between the disease and the microbial community characteristics to provide a theoretical basis for the prevention and treatment of root rot disease in *P. notoginseng*. Consequently, the findings of this study should have important theoretical and production application implications for ensuring the sustainable development of the *P. notoginseng* industry.

## Results

### Metagenomic sequencing data quality control and saturation verification

A total of 333.90 Gb of data were produced from the 12 root endophyte samples, with an average of 27.83 Gb of data per sample. Analysis of the 12 rhizosphere soil samples by Illumina NovaSeq 6000 sequencing produced a total of 173.80 Gb of data with an average of 14.48 Gb of data per sample. After assessing the quality of the original data, the Q30 values of all samples were above 93%. Quality controls were performed on the raw data to remove host contamination and obtain clean data. After removal of host-related sequences of *P. notoginseng*, each rhizosphere soil sample and each root endophyte sample had an average of 13.90 Gb and 12.66 Gb of clean data, respectively (Table 1 and Supplementary Table 1). Being dependent on the alpha diversity chao1 and observed dilution curves (Supplementary Fig. 1) of

rhizosphere soil and root endophyte samples, Supplementary Fig. 1A and B show that the data are saturated at 500,000 sequences (0.075 Gb). Similarly, Supplementary Fig. 1C and D show saturation at 1,000,000 sequences (0.15 Gb). The results showed that the sequencing depth of each sample reached saturation, enabling us to proceed to the next step of the study.

### **Microbial community characteristics between healthy and diseased *Panax notoginseng***

Kraken2 is a high-precision metagenomic sequence classification software based on the k-mer algorithm, which can quickly classify sequencing data by species [33]. In this study, the species diversity of each group was analyzed, and the alpha diversity boxes plotted (Fig. 1B). The other diversity index boxes are shown in Supplementary Fig. 2 and 3. Box plots of the top five species in each group at the genus level are shown in Fig. 1C. The overall richness of the rhizosphere soil was higher than that of the root endophytes. The number of operational taxonomic units (OTUs) in the rhizosphere soil of diseased plants was higher than that associated with healthy plants, whereas in the root endophyte samples, the number of OTUs of healthy plants was higher than that of diseased plants (Fig. 1B). The abundance of *Bradyrhizobium* in healthy rhizosphere soil was higher than in diseased rhizosphere soil (Fig. 1C). In the healthy rhizosphere soil, the top three species were *Bradyrhizobium*, *Micromonospora*, and *Streptomyces*. The genus-level differences in the root endophytes between healthy and diseased plants were more obvious. The three most abundant species in the healthy root endophytes were *Lysobacter*, *Pseudomonas*, and *Paraburkholderia*. The top three species in diseased root endophytes were *Pseudomonas*, *Lelliottia*, and *Variovorax*. *Lysobacter* and *Paraburkholderia* were more abundant in healthy samples than in root rot diseased samples, whereas *Pseudomonas* and *Lelliottia* were more abundant in diseased root endophyte samples (Fig. 1C).

We used the MetaPhlan2 [34] in the HUMAnN2 [35] pipeline to analyze the composition of the microbial communities (including bacteria, archaea, eukaryotes, and viruses), and obtained a species composition table for the rhizosphere soil and root endophyte samples (Supplementary Table 2). Based on the species composition table, we used the Graphlan [36] software to construct species composition circle diagrams for the rhizosphere soil and root endophyte samples (Fig. 2). We screened the top 100 most abundant species in the species composition table and classified them by phylum. As indicated by the colors in the species circle diagrams, there were four main phyla: Actinobacteria, Bacteroides, Firmicutes, and Proteobacteria. Species that did not classify into one of these four phyla were grouped as *Others* in the circle diagrams. The circle diagrams show that Proteobacteria occurred in the highest proportion in both the rhizosphere soil and root endophyte samples. The relative occurrence of the phyla in the rhizosphere soil and root endophyte samples, respectively, was: Actinobacteria: 14% and 7%, Bacteroides: 10% and 7%, Firmicutes: 20% and 6%, Proteobacteria: 50% and 77%, and Others: 6% and 3% (Fig. 2). The species composition heat maps together with the histograms show the differences at the genus level between the samples with different health statuses (Fig. 3). A total of 87 genera were detected in the rhizosphere soil samples and the top seven most abundant genera were *Rhodopseudomonas*, *Actinoplanes*, *Burkholderia*, *Caulobacter*, *Ktedonobacter*, *Mesorhizobium*, and *Granulicella* (Fig. 3A, C). The rhizosphere soil bacterial flora in the diseased and healthy rhizosphere soil samples from the sampling sites in Lijiang (LJ) and

Qiubei (QB) showed opposite results. The abundance of *Rhodopseudomonas* was higher in the healthy rhizosphere soil at the LJ site (LJTC.2), whereas it was more abundant in the root rot diseased rhizosphere soil at the QB site (QBTB.1). *Ktedonobacter* was found only in the rhizosphere soil from QB, and was more abundant in the root rot diseased samples. In total, 108 genera were detected in the root endophyte samples (Fig. 3B, D). The eight most abundant genera were *Pseudomonas*, *Burkholderia*, *Variovorax*, *Afipia*, *Agrobacterium*, *Sphingobium*, *Rhodanobacter*, and *Janthinobacterium*. *Pseudomonas* showed a high abundance in the root rot diseased root endophyte samples from LJ, and a high abundance in the healthy root endophyte samples from QB, except for one pair of samples (QBGB.3 and QBGC.3). The abundance of *Burkholderia* in the healthy root samples was higher than that in the root rot diseased samples. *Variovorax* and *Afipia* showed high abundance in the healthy root endophyte samples from LJ. Overall, the abundance of *Agrobacterium* in the diseased root endophyte samples was higher than that in the healthy samples, and *Sphingobium* showed the opposite trend. The LEfSe software was used to analyze the abundance differences among species in each group, and LDA was used to estimate the effect of each difference feature [37]. There were eight species with significant differences in the root endophyte and rhizosphere soil samples (LDA > 4 for significant differences; Fig. 4).

Regarding eukaryotes identified in the samples, a total of 12 eukaryote species were detected in the rhizosphere soil, and 29 species in the root endophyte samples (Supplementary Table 3). *Acaulopage tetraceros*, *Acrobelloides nanus*, *Arthrobotrys oligospora* ATCC 24927, *Dactylellina haptotyla* CBS 200.50, and *Folsomia candida* were unique to root endophytes, and *Acaulopage tetraceros* and *Chaetomium cochliodes* were more abundant in the healthy samples. *Acrobelloides nanus*, *Arthrobotrys oligospora* ATCC 24927, *Folsomia candida*, and *Pseudopyrenochaeta lycopersici* were more abundant in the diseased samples, with *Ilyonectria mors-panacis*, the most common species (Fig. 3E).

### Functional composition and functional pathways

We used HUMAnN2 [35] to determine the functional composition and functional pathways, as well as to stratify the functional pathways and produce a functional composition table. We used STAMP [38] to screen the significantly different pathways in the functional composition table (at the 95% confidence interval and  $P < 0.05$ ; Fig. 5). There were 16 significantly different pathways in the rhizosphere soil. The pathways: CDP-diacylglycerol biosynthesis I/II, superpathway of L-threonine biosynthesis, and superpathway of L-isoleucine biosynthesis I were significantly up-regulated in the diseased rhizosphere soil samples, whereas the L-lysine biosynthesis III and Pentose phosphate pathway were significantly up-regulated in the healthy rhizosphere soil (Fig. 5A). There were 12 significantly different functional pathways in the root endophyte samples. The pathways: superpathway of L-phenylalanine biosynthesis, 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis, and Pyridoxal 5'-phosphate biosynthesis I were significantly up-regulated in the root endophyte samples, whereas in the healthy root endophyte samples, the pathways: L-valine biosynthesis, L-isoleucine biosynthesis I, and Pyruvate fermentation to isobutanol were significantly up-regulated (Fig. 5B).

### Function annotation results of healthy and root rot diseased *Panax notoginseng* plants

We de novo assembled the clean data, evaluated the quality of the contig, and then performed de-redundancy, quantification, and obtained pure contig for subsequent functional annotation. We obtained 11.22 Gb contigs with an average contig N50 of 1,287 bp (Supplementary Table 4). This we compared to the eggNOG [39] database, integrated the annotation results, and counted the annotated clusters of orthologous genes (COG) orthologous proteins in each group. The COG classified 24 functional descriptions (Supplementary Table 5), and the annotation results showed that there were no significant differences among the 24 functional descriptions of the diseased and healthy rhizosphere samples ( $P > 0.05$ ). The three dominant functions with higher abundance in the rhizosphere soil samples were S: function unknown (TC = 20.61%, TB = 20.85%), E: amino acid transport and metabolism (TC = 9.82%, TB = 9.54%), and C: energy production and conversion (TC = 8.69%, TB = 8.22%; Supplementary Table 5). Among the 24 functional descriptions of the root endophytes, four of them were significantly different between the diseased and healthy samples ( $P < 0.05$ ), and all of them showed a higher abundance in the diseased samples at the following  $P$  values: E: amino acid transport and metabolism ( $P = 0.011$ ), N: cell motility ( $P = 0.028$ ), O: posttranslational modification, protein turnover, chaperones ( $P = 0.027$ ), and S: function unknown ( $P = 0.027$ ), respectively (Supplementary Fig. 4). The three dominant functions with a higher abundance in the root endophyte samples were S: function unknown (GC = 20.73%, GB = 19.38%), E: amino acid transport and metabolism (GC = 8.88%, GB = 9.60%), and P: inorganic ion transport and metabolism (GC = 8.29%, GB = 8.63%), the functional pathway of which was relatively abundant in the root rot diseased samples (Supplementary Table 5).

The Kegg orthologs (KO) function results for each group were counted (Supplementary Table 6). Among them, the TC group had a total of 7,898 KOs, the TB group had 8,078 KOs, the GC group had 9,886 KOs, and the GB group had 10,736 KOs. It can be observed that the number of KOs in the root endophytes was higher than in the rhizosphere soil, and the number of KOs in the diseased group was higher than in the healthy group. There were 6,469 core KOs in the four groups, accounting for 54.4% (Supplementary Fig. 5). In addition, each group had unique KOs, which were 106 (0.9%), 124 (1.0%), 435 (3.7%), 921 (7.7%), sorted by TC, TB, GC, and GB, respectively. The GO term annotation results are mainly divided into three categories: molecular function (MF), cellular component (CC), and biological process (BP). The top 50 GO terms of each group were classified and enriched, and the identified GO terms were found to belong to the BP category (Supplementary Fig. 6). The enrichment degree of the GO terms in the root endophytes (Supplementary Fig. 6C and D) was higher than that in the rhizosphere soil (Supplementary Fig. 6A and B). The pathways of the top three enrichments in the rhizosphere soil were the biological process, metabolic process, and organic substance metabolic process (Supplementary Fig. 6A and B). The carboxylic acid metabolic process and the oxoacid metabolic process were enrichments in the root rot diseased soil group that did not appear in the top 20 processes associated with the healthy rhizosphere soil (Supplementary Fig. 6A and B). The macromolecule metabolic process, cellular aromatic compound metabolic process, and organic cyclic compound metabolic process were significantly enriched in healthy root endophytes, and did not occur in the top 20 processes of the diseased root endophytes (Supplementary Fig. 6C and D).

We used Metascape [40] to perform a functional enrichment analysis of these genes and displayed them as a network diagram to facilitate understanding of the relationship between pathways and biological processes (Supplementary Fig. 7). Most of the terms in each group showed a strong correlation (Supplementary Fig. 6). Although some terms were highly enriched, they were not strongly associated with other terms, such as anatomical structure homeostasis (Supplementary Fig. 6A). Independent terms included: diseases of metabolism, glycoside compound metabolic process, and cellular amino acid biosynthetic process (Supplementary Fig. 6C).

The CAZy database consists of five categories and one related module. Each category is subdivided into families. Among them, the glycoside hydrolases (GHs), which mainly hydrolyze and rearrange glycosidic bonds, number 167 families in total [41]. We found 123 GH families in the TC group, 127 in the TB group, 126 in the GC group, and 131 in the GB group (Supplementary Table 7). The number of GH families in the diseased samples was slightly higher than that in the healthy group. Similarly, there are 110 families of glycosyltransferases (GTs), mainly involved in the formation of glycosidic bonds. There were 70 and 75 GT families in TC and GC, respectively, and 71 and 83 GT families in the TB and GB groups, respectively. In addition, polysaccharide lyases (PLs) have 40 families, carbohydrate esterases (CEs) have 17 families, auxiliary activities (AAs) have 16 families, and carbohydrate-binding modules (CBMs) have 86 families. These families all occurred in slightly higher numbers in the diseased group than in the healthy group (Supplementary Table 7). Sequences compared to the CAZy database in each group also showed an upward trend. The TC group had the lowest number of reads matched to the CAZy database, the GB group had the most reads matched to the CAZy database, and the GT family matched ratio was the highest, followed by that of the GH family (Supplementary Fig. 8).

The Resfam database analysis results showed that there were 15 classifications of mechanisms with 94 functional descriptions (Supplementary Fig. 9). Among them, the TB group had the most resistance genes (106,554), with an average of 17,759 resistance genes per sample. The resistance genes of the other groups were TC (89,250), GB (84,135), and GC (57,122) (Supplementary Table 8). Overall, there were more resistance genes in the rhizosphere soil than in the root endophytes, and the number of resistance genes in the diseased group was higher than in the healthy group. The top three functions of the resistance genes in the TC, TB GC, and GB groups, respectively, were consistently: ABC transporter (33.77%, 34.27%, 31.50%, and 36.23%), Gene modulating resistance (20.24%, 20.56%, 21.60%, and 20.12%), and RND antibiotic efflux (17.96%, 16.79%, 16.73%, and 14.92%). The ABC transporter, RND antibiotic efflux, and Gene modulating resistance functions in the root rot disease group were higher than those in the healthy group (TC and GC) (Supplementary Table 8).

## Metagenomic data binning

After the clean data were assembled, de-redundant contigs were binned to mine the single bacterial genome. Ultimately, nine bins obtained in the TC group, 24 bins in the TB group, and two bins in the TB group had > 99% completeness. *Thaumarchaeota* had 99.51% completeness and the genome length was 1.40 Mb, and *Clostridium* had 99.01% completeness, and the genome length was 6.31Mb in the TB group

(Supplementary Fig. 10 and Supplementary Table 9). A total of 41 bins in the GC group, two of which were 100% complete, namely *Flavobacteriaceae* (genome length = 5.00 Mb) and *Bacteroidetes* (genome length = 6.34 Mb). There were 91 bins in the GB group, one of which was 100% complete (*Flavobacteriaceae*, genome length = 4.3 Mb). We found that these two 100% complete *Flavobacteriaceae* genomes were two different species under the *Flavobacteriaceae* genus through similarity comparison. The GC content results showed that the abundance of orders in the rhizosphere soil was greater than that in the root endophytes (Fig. 6). These results are consistent with the findings of a higher microbial diversity in the rhizosphere soil than that in the root endophytes (Fig. 1B).

The abundance of the purified single bacterial genome in each sample was determined by bin quantification, and the quantitative results were displayed in a heat map (Supplementary Fig. 10). Bacteria (bin 2) were rich in the QBTC.3 sample and *Sphingomonadaceae* (bin 13) was rich in the QBGC.1 sample. *Alphaproteobacteria* (bin 3) were rich in the LJ diseased root endophyte samples (Supplementary Fig. 10 and Supplementary Table 9). We also used the Krona software for visualization (Supplementary Fig. 11).

## Discussion

### Microbial communities in different ecological niches of root rot diseased *Panax notoginseng*

We found that the shape of the dilution curves for the rhizosphere soil and root endophyte samples were significantly different (Supplementary Fig. 1). Among the rhizosphere soil samples (Supplementary Fig. 1A and B), only the healthy rhizosphere soil dilution curve shape changed noticeably, and the dilution curves of the diseased rhizosphere soil samples were uniform, whereas the shape of the root endophyte dilution curves was more distinct (Supplementary Fig. 1C and D). Research on poplar shows that the obvious change in the OTU dilution curve of root endophytes may be caused by the uneven distribution of the microbial community associated with the roots and other plant parts [42]. In addition, when comparing the alpha diversity of rhizosphere soil and root endophyte samples (Fig. 1B), we found that the microbial diversity of the rhizosphere soil was higher than that of the root endophytes. Studies have shown that the microbial community of rhizosphere soil is more diverse than that of root endophytes [43, 44], and that there is a clear compartmental separation between the rhizosphere soil and root endophytes of plants [45–51]. Our results reflect this uniqueness of the microbial communities (Fig. 3A and B). Previous studies have shown that the bacteria of root endophytes are recruited from the rhizosphere soil [52, 53], which may account for why the microbial diversity in the rhizosphere soil is higher than that in the root endophyte samples. Our findings support this view.

### Bacterial community differences between healthy and diseased *Panax notoginseng* plants

We compared the differences in bacterial communities between healthy and diseased samples and found that *Rhodopseudomonas* was most abundant in the rhizosphere soil (Fig. 3A and C), especially in the healthy samples from the LJ site. It is reported that a compound produced by *Rhodopseudomonas* can promote plant growth, and studies have shown that *Rhodopseudomonas* can induce resistance in plants

[54]. In tobacco, *Rhodopseudomonas* induces tobacco mosaic resistance to protect against tobacco mosaic immunity [55]. Studies have also evaluated the effect of KTSSR54, an active compound produced by *R. palustris*, which proves that KTSSR54 can help rice inhibit pathogens [56]. Our results are in agreement with previous studies showing that *Rhodopseudomonas* can protect *P. notoginseng* from root rot diseases. Among the bacteria studied, *Actinoplanes* have attracted increasing attention because they can inhibit plant pathogens and produce more than 100,000 compounds such as antifungal compounds, siderophores, hydrocyanic acid, hydrolytic enzymes, and ammonia. *Actinoplanes* can accelerate the absorption of nutrients and have a beneficial effect on plant growth [57–59]. *Actinoplanes* produce many antibiotics, and account for about 45% of the antibiotics on the market [60]; most of these antibiotics have been shown to inhibit plant diseases [61–64]. In our study, *Actinoplanes* were more abundant in LJ (LJTC.3/ LJTB.3) and QB rhizosphere soil samples (Fig. 3A), suggesting that the diseased *P. notoginseng* plants may have recruited *Actinoplanes* to resist root rot pathogens, resulting in an increase in the abundance of *Actinoplanes* in the soil of diseased plants. *Burkholderia* was relatively abundant in most samples. Many studies have shown that *Burkholderia* has antibacterial properties and promotes plant growth and is a potential and effective biological control agent [65, 66]. Researchers have found that *Burkholderia* is an antagonist of black spot disease in cherry [67]. It is also known that *Burkholderia* secretes beneficial compounds to protect bananas from Fusarium wilt [68]. We infer that *Burkholderia* can protect *P. notoginseng* from plant diseases.

*Pseudomonas*, the most abundant dominant bacterial species in the root endophyte samples in our study, was more abundant in the diseased samples from LJ (Fig. 3B and D). In previous studies on *P. notoginseng*, *Pseudomonas* was predicted to be the pathogen causing root rot disease [69]. In previous studies, *Pseudomonas* was only identified at the genus level. We identified *Pseudomonas* to the species level, and found that the species with highest abundance was *P. fluorescens*. *Pseudomonas fluorescens* is known to protect many crop plants against soil-borne diseases caused by phytopathogens [70, 71]. Studies have shown that *P. fluorescens*, as a biological control agent, can effectively reduce root knot nematode infestation of cucumbers, and replace chemical measures for nematode control, resulting in an increase in cucumber yield that has the added benefit of reduced cultivation costs [72]. Screened under in vitro conditions, *P. fluorescens*, as a biological control agent, has been shown to have antagonistic activity against pea root rot, with high biological control activity [73]. A recent report found that *Pseudomonas* is recruited from roots to resist plant pathogens, and *Pseudomonas* was used to verify the disease incidence of plants, where it was found to reduce the disease incidence of *Arabidopsis* [74]. We suggest that, in the root rot diseased samples, *Pseudomonas* may have been recruited by rhizosphere exudates to resist plant pathogens, causing a higher abundance of *Pseudomonas* in the diseased samples. In our study, the abundance of *Burkholderia* in the healthy root endophyte samples was higher than that in the diseased samples. This finding is consistent with previous studies [65–68], suggesting that *Burkholderia* is a beneficial bacterium for *P. notoginseng*. *Variovorax* has been shown to be harmful to the growth of *P. notoginseng* and to reduce the yield of *P. notoginseng* [75]. Our results showed that *Variovorax* had a higher abundance in most of the healthy root endophytic samples. We identified *Variovorax* to the species level, and found that the species with the highest abundance was *V. paradoxus*.

Studies have shown that *Variovorax* can participate in the interaction between plants and microorganisms by manipulating plant hormone levels to balance the growth of plant roots. *Variovorax paradoxus* can significantly inhibit potato tuber soft rot, and is considered to be a new type of potato tuber soft rot biological control agent [76]. In our study, the abundance of *V. paradoxus* in healthy root endophyte samples was high. This suggests that *V. paradoxus* may promote the growth of *P. notoginseng* and may be an effective biological control agent for root rot disease.

### **Eukaryotic microbial communities associated with healthy and diseased *Panax notoginseng* plants**

Our results showed that *Ilyonectria mors-panacis* had the highest ratio in the diseased root endophyte samples (Fig. 3E) corroborating the findings of a previous study that *I. mors-panacis* is the main pathogen of *P. notoginseng* root rot [9]. *Ilyonectria mors-panacis* is also the main pathogen of *P. ginseng* root rot [77, 78]. *Pseudopyrenochaeta lycopersici* is also a plant fungal pathogen that causes severe root rot and root knot disease [79]. In our study, *P. lycopersici* had a higher number of reads in the diseased samples, except for QBTC.3. *Pseudopyrenochaeta lycopersici* has not previously been reported in association with root rot diseased *P. notoginseng*, and we infer that *P. lycopersici* may be a pathogen of *P. notoginseng* root rot. *Arthrobotrys oligospora* ATCC 24927, a fungus that feeds on nematodes, is often used as a biological control agent for plant and animal parasitic nematodes [80]. The number of reads of this fungus was relatively high in the diseased samples in our study. *Acroboloides nanus* is a type of nematode that feeds on bacteria and was mainly associated with diseased samples in our study. This could account for the high occurrence of *A. oligospora* ATCC 24927, which feeds on nematodes, in the diseased samples.

### **The interaction of functional genes and pathways helps improve plant resistance**

Amino acids play an important role in plant growth. Studies have shown that amino acids also play a vital role in pathogen infection as they are indispensable nitrogen sources for many nutritional pathogens. After infection with pathogens, changes in the expression of genes are involved in amino acid metabolism and transport [81, 82], so regulating the content of amino acids is critical to the growth and defense of plants. Studies have demonstrated that the increased expression of UMAMIT (usually multiple acids move in and out transporters) amino acid transporters can induce plant resistance to pathogens [83]. In addition, the existence of the amino acid-sensing mechanism in plants indicates that changing the level of amino acids by changing the metabolism or transport of amino acids may trigger a defense response. For example, overexpression of the cationic amino acid transporter CAT1 helps the system to develop resistance to *Pseudomonas*. Furthermore, resistance-related genes are upregulated [84]. Our study found higher amino acid transport and metabolism in healthy rhizosphere soil than in diseased rhizosphere soil. Based on previous research results, it was concluded that the up-regulation of amino acid transport and metabolism in soil and the enhancement of transport activity would help to resist the invasion of pathogens and improve the resistance of *P. notoginseng* (Supplementary Table 5). In a study on the resistance signal mechanism of the related tomato root-knot nematode disease, it was found that the protein with the greatest difference in the differentially expressed protein group belongs to energy

production and conversion, because the defense response is an energy-consuming process, which may play a key role in the plant's disease resistance process [85]. In our study, we found that the energy production and conversion process in diseased roots is three times that of healthy roots and may be closely related to the disease resistance process of *P. notoginseng* (Supplementary Table 5).

### **The antibiotic resistance-related gene can help *Panax notoginseng* resist root rot disease**

The misuse of antibiotics has contributed to the widespread development of antimicrobial resistance among clinically significant bacterial species, so understanding the mechanism of antibiotics in *P. notoginseng* is important for facilitating the healthy growth of this plant species. Studies have shown that Chinese herbal medicines, such as *P. notoginseng*, can produce compounds that interfere with quorum sensing, which can effectively resist pathogens and help delay antibiotic resistance [86]. In our study, annotated results showed that ABC transporter (34.16%), Gene modulating resistance (20.55%), and RND antibiotic efflux (16.62%) accounted for the highest proportions, and the highest number of diseased samples (Supplementary Table 8). Studies have shown that transporters exist in all organisms, among which the ABC transporter reaction can help plants play an important role in disease resistance [87]. The G and B subfamily of the ABC transporter family has the most research on related secondary metabolism, and is also the key transport family to the defensive process [88–90]. The proportion of ABC transporters in the two groups of diseased samples was higher than that in the corresponding two groups of healthy samples, which may be related to the disease resistance of *P. notoginseng*. Many studies have reported that gene modulating resistance helps plants to inhibit pathogenic bacteria and improve plant resistance [91–95]. Studies have shown that the RND antibiotic efflux mechanism has an efflux effect on a variety of antibiotics, and inhibiting these pumps is one of the ways to resolve the problem of antibiotic resistance [96–98]. Our study showed that the RND antibiotic efflux mechanism in diseased samples was greater than that in healthy samples. We speculate that the RND antibiotic efflux mechanism in the diseased samples was active and had an efflux effect on antibiotics. As a result, the reduction of antibiotics in the diseased samples weakens the resistance to pathogens. This phenomenon may induce plant roots to produce antibiotic efflux, making it easier for root rot to invade.

### ***Panax notoginseng* root rot CAZy genes**

According to the CAZy database comparison results, the proportion of GHs in each group was the largest, and the proportion in diseased samples was higher than that of healthy samples. In a study of peanut stem rot caused by fungi, it was found that the number of GHs is largest in the pathogenic secreted protein [99]. In addition, in research on soybean seed rot, 149 plant cell wall-degrading enzymes were detected, most of which are GHs [100]. Our results are in agreement with these previous reports. The number of GHs in the diseased samples was higher than that in the healthy samples, indicating that the root rot-affected parts of *P. notoginseng* secrete more enzymes related to the GH family. Regarding the glycosyltransferases (GTs), in a study of wheat infection with *Fusarium wilt*, it was found that the GTs in wheat infected with *F. graminearum* showed an upward trend [101], which is also consistent with our results, indicating that infection with *P. notoginseng* root rot can cause GTs to increase. In short, the

increase in the number of GH and GT families can predict the tendency of *P. notoginseng* to suffer from root rot (Supplementary Fig. 8 and Supplementary Table 7).

## Conclusions

In this study, we analyzed the rhizosphere soil and root endophyte microbial communities of *P. notoginseng* from different geographical locations using second-generation high-throughput sequencing. We showed that the structural variability of the root endophyte microbial community was higher than that of the soil, but the diversity of the microbial community in the rhizosphere soil was higher than that in the root endophyte samples. Among them, *Rhodopseudomonas*, *Actinoplanes*, *Burkholderia*, *V. paradoxus*, and *P. fluorescens* can help *P. notoginseng* resist the invasion of root rot disease. *Ilyonectria mors-panacis* and *P. lycopersici* are pathogenic bacteria of *P. notoginseng* root rot. The up-regulation of amino acid transport and metabolism in soil would help to resist pathogens and improve the resistance of *P. notoginseng*. The ABC transporter and Gene modulating resistance genes can improve the disease resistance of *P. notoginseng*. In addition, the increase in the number of GT and GH families may be a molecular manifestation of *P. notoginseng* root rot. This study demonstrated the microbial and functional diversity in the rhizosphere microbial community of *P. notoginseng* and provided useful information for understanding the microbial community in *P. notoginseng* root rot. The results of this study provide a basis for a more comprehensive and complete study of the molecular mechanism of *P. notoginseng* root rot and other plant rhizosphere microbial communities, and lays a foundation for the development of biological control agents and related molecular monitoring methods.

## Methods

### Sample collection

Twelve samples were collected from Lijiang City, Yunnan Province (26°49'59" N, 100°3'24" E, altitude: 2,640 m) and Qiubei County, Yunnan Province (24°2'49" N, 103°58'40" E, altitude: 2,020 m), respectively. Twenty-four healthy and diseased rhizosphere soil and root endophyte samples were collected, totally. The samples were divided into four groups: healthy rhizosphere soil (TC), diseased rhizosphere soil (TB), healthy root endophytes (GC), and diseased root endophytes (GB), with three replicates in each group. The samples of diseased root endophytes were collected from *P. notoginseng* plants suffering from root rot disease (the aboveground parts of diseased plants were wilted and yellowed, and the roots were necrotic), whereas the endophyte samples from healthy plants were collected from plants showing no signs of disease (Fig. 1A). Both diseased and healthy plants were gently lifted from the soil and shaken to remove the dirt and other impurities attached to the surface of the roots. Thereafter, the roots were carefully rinsed with sterile water until the root surfaces were free of impurities. A prepared sterile scalpel was used to cut off the cleaned roots into 50 ml sterile centrifuge tubes, which were immediately placed into liquid nitrogen for preservation. When collecting the rhizosphere soil from diseased and healthy plants, each plant was gently removed from the soil and a standard soil ring knife was used to collect the samples. Rhizosphere soil 20 cm under the ground was collected from each plant and placed into 50 ml

sterile centrifuge tubes which were stored in liquid nitrogen for later use. While selecting the diseased samples, another routine sample was selected within a radius of 1 m around diseased sample as a control. Were six samples taken from diseased plants (three rhizosphere and three endophytes) and six from healthy plants (three rhizosphere and three endophytes) at each sampling site.

## **Metagenomic sequencing**

A Qubit 2.0 fluorometer (ThermoFisher Scientific Inc., Waltham, MA, USA) was used to accurately quantify the DNA concentration. After the DNA samples were qualified, they were randomly interrupted using a Covaris Focused-ultrasonicator (model; Covaris, Inc., Woburn, MA, USA), and then the entire library was prepared following the steps of end repair, A-tailing, sequencing adapters, purification, and PCR amplification. After the library was constructed, it was diluted, and the inserted fragments of the library were detected using an Agilent 2100 bioanalyzer (Agilent, Santa Clara, CA, USA). After the size of the inserted fragments met expectations, the effective concentration of the library was accurately quantified using the Q-PCR method to ensure the quality of the library. After the library was qualified, different libraries were pooled to flow cells according to the effective concentration and target data volume requirements. After the cBOT was clustered, a NovaSeq 6000 high-throughput sequencing platform (Illumina, Inc., San Diego, CA, USA) was used for sequencing. Each root endophyte sample had approximately 20–30 Gb of sequencing data, and each rhizosphere soil sample had approximately 10–30 Gb of sequencing data.

## **Data quality control**

The FastQC software was used to estimate the quality of the original sequence of 24 samples to generate a quality evaluation report [102]. The Kneaddata tool (<http://huttenhower.sph.harvard.edu/kneaddata>) was used to perform quality control and de-hosting of the *P. notoginseng* genome sequence. The Kneaddata process relies on Trimmomatic to remove primers [103], linkers, and low-quality sequences (parameters: sequence quality  $\geq 20$ , minimum sequence length  $\geq 50$  bp). Bowtie2 [104] was used to compare the genome of *P. notoginseng* [105] to screen out the sequence of the non-host genome, and finally obtain a clean sequence (Table 1). Cleaned reads were used for downstream analyses.

## **Analysis of the microbial diversity of *Panax notoginseng* root rot**

The HUMAnN2 software was used to calculate the species and functional composition [35]. HUMAnN2 does not consider double-ended information in the analysis, so the double-ended sequence after quality control is combined as one input file, and the other parameters are the default parameters. MetaPhlAn2 was used to calculate the species composition [34], and Bowtie2 [104] was compared to the nucleic acid sequence database that comes with HUMAnN2 and generates a species abundance composition table for each group. We used GraPhlAn software to analyze the top 100 species in the species composition table at the phylum level [36]. To display the species circle map, R 3.4.3 software was used to run the R script to display the heat map of the species at the genus level [106], and to compare the differences in species composition between groups. Finally, LEfSe was used for the species difference analysis [37].

The LEfSe used the Kruskal-Wallis and Wilcoxon signed rank tests, signed linear judgment analysis (LDA) logarithmic score, and related  $P$ -values to identify species with significant differences in each group (LDA  $> 4$  was considered as significantly different species) [107].

Kraken2 is a high-precision metagenomic sequence classification software based on the k-mer algorithm [33]. Kraken2 is used for species annotation at the read level, which can quickly classify sequencing reads for species classification. We compared the data to Kraken2's standard library (human, bacteria, archaea, viruses, and vectors), and the fungi, protozoa, and plasmid databases. The bacteria, archaea, fungi, protozoa, plasmids, and virus libraries were provided by the RefSeq database (<https://www.ncbi.nlm.nih.gov/refseq/>); for humans, the GRCh38 human genome ([https://www.ncbi.nlm.nih.gov/assembly/GCF\\_000001405.26/](https://www.ncbi.nlm.nih.gov/assembly/GCF_000001405.26/)) was used; and for the vectors, the UniVec\_Core vector core library (<https://www.ncbi.nlm.nih.gov/>). The species composition table was used for species diversity analysis, and the species level annotations were extracted and combined to the minimum sequencing amount, and alpha diversity was calculated using R 3.4.3, to draw the alpha diversity box plot, and to draw the top 30 genera of each group to compare the differences.

The EukDetect method [108] was used to detect eukaryotic microorganisms in the clean data. The EukDetect pipeline uses a Snakemake workflow engine [109]. Microbial eukaryotic genomes were downloaded from NCBI GenBank for all species designated as fungi, protists, other non-vertebrate metazoans, and non-streptophyte archaeplastida species. In addition to the GenBank genomes, 314 genomes and transcriptomes comprising 282 protists, 30 *Archaeplastida* species, and 2 metazoans curated by the EukProt project were downloaded [110].

### **Functional composition analysis**

Diamond [111] was used to compare contigs to the UniProt protein library (<https://www.uniprot.org/>), which includes the UniRef90 protein database, and the MetaCyc metabolic pathway database was used to predict gene pathways (<https://metacyc.org/>). MinPath (<https://omics.informatics.indiana.edu/MinPath/>) defines the minimum set of pathways, producing results at the protein, gene and pathway level, and a functional composition table. The STAMP [38] software was used to calculate the functional pathways ( $P < 0.05$ ) with significant functional differences between the groups and a histogram was drawn (Fig. 5).

### **Assembly results of metagenomic data**

MEGAHIT [112] was used for each sample assembly. The contigs of each sample were mixed by group, and the quality of contigs was assessed by QUASt [113]. MetaProdigal [114] was used for gene prediction, and cd-hit-est [115] to cluster and de-redundancy the predicted genes to construct a non-redundant gene set (similarity  $\geq 95\%$ , coverage  $\geq 90\%$ ). The non-redundant gene set obtained was used for subsequent analysis. The nucleic acid sequences were quantified using Salmon [116]. The parameter was set to the *-meta* metagenomic model so that all reads could be binned, and the abundance of genes in each sample was obtained.

## **Multi-database annotation of *P. notoginseng* root rot**

The nucleic acid sequence was translated into a protein sequence and compared to the eggNOG [39] database using the Diamond [111] software. The eggNOG database integrates annotation information, such as GO, COG, and KO. We performed multiple sequence alignment between the query and the eggNOG database, to determine the conserved sites and analyze their evolutionary relationships [39]. We sorted out the possible gene names of each sequence predicted by the eggNOG database and eliminated the duplication results. The eggNOG annotation results were used for subsequent analysis. Metascape [40] was used to perform a functional enrichment analysis of these genes to facilitate understanding of the relationship between pathways and biological processes (Supplementary Fig. 6). Similarly, Diamond [111] was used to align the protein sequence with the carbohydrate-active enzymes database (CAZy) [117]. The CAZy database, providing an understanding of the nature and extent of the complex carbohydrate metabolism focuses on the differences in carbon source metabolism between species, and comparing it to the ResFam database using Diamond [111, 118], we predicted the unknown resistance gene and obtained the mechanism and function of the resistance genes in each sample.

## **Binning analysis of metagenomic data**

We used MetaWRAP [119] to excavate the draft genome of a single strain. This process integrates three popular binning software. We chose the MaxBin2 [120] and MetaBAT2 [121] software for binning. The de-redundant contigs of each group were binned, and the bins were purified. We evaluated and comprehensively analyzed the results to obtain better results. The Checkm database [122] requires a completeness of more than 70% and a pollution rate of less than 5%, which was used for purification. After purification, the Blobology module was used to visualize the GC content and abundance of contigs by comparing the NCBI nt and tax databases. Samlon [116] was used to quantify the bins and calculate the abundance of a single bacterial genome in each sample. Subsequently, each contig in the bin was annotated, and the species of each bin was estimated. Finally, Krona was used to visualize the results of bin annotation [123].

## **Declarations**

### **Ethics approval and consent to participate**

Not applicable

### **Consent for publication**

Not applicable

### **Availability of data and materials**

The sequencing data for all samples have been deposited to China National

GeneBank DataBase (CNGDB) with the accession number CNP0001803. The authors declare that all data necessary for confirming the conclusions presented in the article are represented fully within the article and additional files.

### Competing interests

The authors declare that they have no conflict of interest.

### Funding

YL acknowledges the financial support from the National Science Foundation of China (Grant No. 31960134) and Key project at central government level: The ability establishment of sustainable use for valuable Chinese medicine resources (Grant No. 2060302).

### Authors' contributions

PPW and YL wrote and reviewed the manuscript, PPW, LFY, JLS and YL collected the samples, YY, YQ, CXW, DQL and LQH provide analysis ideas and reviewed the manuscript. PPW and YL reviewed the manuscript and draw the picture, XMC and YL supervised the work. Both authors read and approved the final paper.

### Acknowledgements

We thank the editor and anonymous reviewers for providing valuable comments on the manuscript.

### Author details

<sup>1</sup> Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming 650500, Yunnan, China. <sup>2</sup> Yunnan Provincial Key Laboratory of *Panax notoginseng*, Kunming 650500, China. <sup>3</sup> Key Laboratory of *Panax notoginseng* Resources Sustainable Development and Utilization of State Administration of Traditional Chinese Medicine, Kunming 650500, China. <sup>4</sup> Kunming Key Laboratory of Sustainable Development and Utilization of Famous-Region Drug, Kunming 650500, China. <sup>5</sup> National Resource Center for Chinese Meteria Medica, Chinese Academy of Chinese Medical Sciences, Beijing 100700, China. <sup>6</sup> Sanqi Research Institute of Yunnan Province, Kunming 650500, Yunnan, China.

## References

1. Fan ZY, Miao CP, Qian XG, Zheng YK, Chen HH, Guan HL, et al. Diversity, distribution, and antagonistic activities of rhizobacteria of *Panax notoginseng*. *J Ginseng Res.* 2016;40:97-104.
2. Guo HB, Cui XM, An N, Cai GP. Sanchi ginseng (*Panax notoginseng* (Burkill) F. H. Chen) in China: distribution, cultivation and variations. *Genet ResourCrop Ev.* 2010;57:453-60.
3. Yang M, Yuan Y, Huang HC, Ye C, Guo CW, Zhu SS, et al. Steaming combined with biochar application eliminates negative plant-soil feedback for sanqi cultivation. *Soil and Tillage*

- Research. 2019;189:189-98.
4. Guo WQ, Chen YG, Shi RZ, He K, Wang JF, Shao JH, et al. 20(S)-Protopanaxdiol suppresses the abnormal granule-monocyte differentiation of hematopoietic stem cells in 4T1 breast cancer-bearing mouse. *Evidence-Based Complementary and Alternative Medicine*. 2020;2020:8747023.
  5. Li MZ, Chen ZJ, Qian J, Wei FG, Zhang GZ, Wang Y, et al. Composition and function of rhizosphere microbiome of *Panax notoginseng* with discrepant yields. *Chinese Medicine*. 2020;15:85.
  6. Miao ZQ, Li SD, Liu XZ, Hen YJ, Li YH, Wang Y, et al. The causal microorganisms of *Panax notoginseng* root rot disease. *Scientia Agricultura Sinica*. 2006;39(7):1371-8.
  7. Santhakumar AB, Battino M, Alvarez-Suarez JM. Dietary polyphenols: Structures, bioavailability and protective effects against atherosclerosis. *Food Chem Toxicol*. 2018;113:49-65.
  8. Mao ZS, Long YJ, Zhu YY, Zhu SS, He XH, Chen ZJ. et al. First report of *Cylindrocarpon destructans* var. *destructans* causing black root rot of Sanqi (*Panax notoginseng*) in China. *Plant Dis*. 2014;98:162.
  9. Mi CY, Yang RH, Rao JF, Yang SZ, Wei FG, Hu XF. et al. Unveiling of dominant fungal pathogens associated with rusty root rot of *Panax notoginseng* based on multiple methods. *Plant Dis*. 2017;101:2046-52.
  10. Ma YN, Chen CJ, Li QQ, Xu FR, Cheng YX, Dong X. Monitoring antifungal agents of *Artemisia annua* against *Fusarium oxysporum* and *Fusarium solani*, associated with *Panax notoginseng* root-rot disease. *Molecules*. 2019;24:213.
  11. Miao CP, Mi QL, Qiao XG, Zheng YK, Chen YW, Zhao LX, et al. Rhizospheric fungi of *Panax notoginseng*: diversity and antagonism to host phytopathogens. *J Ginseng Res*. 2016;40:127-34.
  12. Zhao YP, Lin S, Chu LX, Gao JT, Azeem S, Lin WX. et al. Insight into structure dynamics of soil microbiota mediated by the richness of replanted *Pseudostellaria heterophylla*. *Scientific Reports*. 2016;6:26175.
  13. Wu LK, Li ZF, Li J, Huang WM, Zhang ZY, Lin WX, et al. Assessment of shifts in microbial community structure and catabolic diversity in response to *Rehmannia glutinosa* monoculture. *Applied Soil Ecology*. 2013;67:1-9.
  14. Dong LL, Xu J, Feng GQ, Li XW, Chen SL. Soil bacterial and fungal community dynamics in relation to *Panax notoginseng* death rate in a continuous cropping system. *Scientific Reports*. 2016;6:31802.
  15. Bakker MG, Glover JD, Mai JG, Kinkel LL. Plant community effects on the diversity and pathogen suppressive activity of soil streptomycetes. *Applied Soil Ecology*, 2010;46:35-42.
  16. Brussaard L, De Ruiter PC, Brown GG. Soil biodiversity for agricultural sustainability. *Agr Ecosyst Environ*. 2007;121:233-44.
  17. Bakker P, Pieterse C, Jonge RD, Berendsen RL. The soil-borne legacy. *Cell*. 2018;172:1178-80.
  18. Mendes LW, Raaijmakers JM, Hollander MD, Mendes R, Tsai SM. Influence of resistance breeding in common bean on rhizosphere microbiome composition and function. *The ISME Journal*. 2018;12:212-24.

19. Wu ZX, Hao ZP, Zeng Y, Guo LP, Huang LQ, Chen BD. Molecular characterization of microbial communities in the rhizosphere soils and roots of diseased and healthy *Panax notoginseng*. *Antonie Van Leeuwenhoek*. 2015;108:1059-74.
20. Zhang YM, Cheng YX, Ma YN, Chen CJ, Xu FR, Dong X. Role of phenolic acids from the rhizosphere soils of *Panax notoginseng* as a double-edge sword in the occurrence of root-rot disease. *Molecules*. 2018;23:819.
21. Chen K, Pachter L. Bioinformatics for whole-genome shotgun sequencing of microbial communities. *Plos Comput Biol*. 2005;1:106-12.
22. Goss EM, Tabima JF, Cooke DE, Restrepo S, Fry WE, Forbes GA, et al. The Irish potato famine pathogen *Phytophthora infestans* originated in central Mexico rather than the Andes. *Proc Natl Inst Sci USA*. 2014;111:8791-6.
23. Gaulin E, Jacquet C, Bottin A, Dumas B. Root rot disease of legumes caused by *Aphanomyces euteiches*. *Mol Plant Pathol*. 2007;8:539-48.
24. Chatterton S, Gossen B, McLaren D, Esmaeili Taheri A. Metagenomic analysis of oomycete communities from the rhizosphere of field pea on the Canadian prairies. *Can J Microbiol*. 2017;63:758-68.
25. Duan Y, Zhou L, Hall DG, Li W, Lin H, Liu L, et al. Complete genome sequence of citrus huanglongbing bacterium, '*Candidatus Liberibacter asiaticus*' obtained through metagenomics. *Mol PlantMicrobe Interact*. 2009;22:1011-20.
26. Min-Jung Kwak, Hyun Gi Kong, Kihyuck Choi, Soon-Kyeong Kwon, Ju Yeon Song, Jidam Lee, et al. Rhizosphere microbiome structure alters to enable wilt resistance in tomato. *Nat Biotechnol*. 2018;36:1100-9.
27. Punja ZK. Fungal pathogens of American ginseng (*Panax quinquefolium*) in British Columbia. *Can JPlant Pathol*. 1997;9:301-6.
28. Vendan RT, Yu YJ, Lee SH, Rhee YH. Analysis of bacterial community in the Ginseng soil using denaturing gradient gel electrophoresis (DGGE). *Indian J Microbiol*. 2010;48:559-65.
29. Ji WX, Leng X, Jin ZX, Li HL. Plant growth promoting bacteria increases biomass, effective constituent, and modifies rhizosphere bacterial communities of *Panax ginseng*. *Soil Plant Science*. 2018;69:135-46.
30. Wang QX, Sun H, Li MJ, Xu CL, Zhang YY. Different age-induced changes in rhizosphere microbial composition and function of *Panax ginseng* transplantation mode. *Front Plant Sci*. 2020;11:563240.
31. Deng L, Mao X, Liu D, Ning XQ, Shen Y, Chen B, et al. Comparative analysis of physicochemical properties and microbial composition in high-temperature *Daqu* with different colors. *Frontiers in Microbiology*. 2020;11:58817.
32. Xing XK, Guo SX, Fu JG. Biodiversity and distribution of endophytic fungi associated with *Panax quinquefolium*L. cultivated in a forest reserve. *Symbiosis*. 2010; 51:161-6.

33. Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* 2019;20:257.
34. Truong DT, Franzosa EA, Tickle TL, Scholz M, Weingart G, Pasolli E, et al. MetaPhlAn2 fir enhanced metagenomic taxonomic profiling. *Nat Methods.* 2015; 12:902-3.
35. Franzosa EA, Mclver LJ, Rahnavard G, Thompson LR, Schirmer M, Weingart G, et al. Species-level functional profiling of metagenomes and metatranscriptomes. *Nat Methods.* 2018;15:962-8.
36. Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ.*2015;3:e1029.
37. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011;12:R60..
38. Paeks DH, Tyson GW, Hugenoltz P, Beiko RG. STAMP: Statistical analysis of taxonomic and functional profiles. *Bioinformatics.* 2014;30:3123-4.
39. Cepas JH, Kristoffer F, Coelho LP, Szklarczyk D, Jensen LJ, Mering CV, et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Molecular Biology and Evolution.* 2017;34:2115-22.
40. Zhou YY, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun.* 2019;10:1523.
41. Lombard V, Golaconda RE, Drula E, Coutinho PM, Henrissat B. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 2013; 42:D490-5.
42. Gottel NR, Castro HF, Kerley M, Yang Z, Pelletier DA, Podar M, et al. Distinct microbial communities within the endosphere and rhizosphere of *Populus deltoides* roots across contrasting soil types distinct. *Appl Environ Microbiol.* 2011;77:5934-44.
43. Bulgarelli D, Schlaeppi K, Spaepen S, Themaat E, Schulze-Lefert P. Structure and functions of the bacterial microbiota of plants. *Annu Rev Plant Biol.* 2013;64:807-38.
44. Jumpponen A, Johnson LC. Can rDNA analyses of diverse fungal communities in soil and roots detect effects of environmental manipulations: A case study from tallgrass prairie. *Mycologia.* 2005;97:1177-94.
45. Lundberg DS, Lebeis SL, Paredes SH, Yourstone S, Gehring J, Malfatti S, et al. Defining the core *Arabidopsis thaliana* root microbiome. *Nature.* 2012;488:86-90.
46. Bulgarelli D, Rott M, Schlaeppi K, Ver Loren van Themaat E, Ahmadinejad N, Assenza F, et al. Revealing structure and assembly cues for *Arabidopsis* root-inhabiting bacterial microbiota. *Nature.* 2012;488:91-5.
47. Hartman K, van der Heijden MG, Roussely-Provent V, Walser J-C, Schlaeppi K. Deciphering composition and function of the root microbiome of a legume plant. *Microbiome.* 2017;5:2.
48. Peiffer JA, Spor A, Koren O, Jin Z, Tringe SG, Dangl JL, et al. Diversity and heritability of the maize rhizosphere microbiome under field conditions. *Proc Natl Acad Sci.* 2013;110:6548-53.

49. Yeoh YK, Paungfoo-Lonhienne C, Dennis PG, Robinson N, Ragan MA, Schmidt S, et al. The core root microbiome of sugarcane cultivated under varying nitrogen fertilizer application. *Environ Microbiol.* 2015;18:1338-51.
50. Zarraonaindia I, Owens SM, Weisenhorn P, West K, Hampton-Marcell J, Lax S, et al. The soil microbiome influences grapevine-associated microbiota. *MBio.* 2015;6:e02527-14.
51. Bulgarelli D, Garrido-Oter R, Münch PC, Weiman A, Dröge J, Pan Y, et al. Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe.* 2015;17:392-403.
52. Edwards J, Johnson C, Santos-Medellín C, Lurie E, Podishetty NK, Bhatnagar S, et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc Natl Acad Sci.* 2015;112:E911-20.
53. Chaparro JM, Badri DV, Vivanco JM. Rhizosphere microbiome assemblage is affected by plant development. *ISME J.* 2014;8:790-803.
54. Hsu SH, Shen MW, Chen JC, Lur HS, Liu CT. The Photosynthetic Bacterium *Rhodopseudomonas palustris* Strain PS3 Exerts Plant Growth-Promoting Effects by Stimulating Nitrogen Uptake and Elevating Auxin Levels in Expanding Leaves. *Front Plant Sci.* 2021;12:573634.
55. Su P, Tan X, Li C, Zhang D, Cheng J, Zhang S, et al. Photosynthetic bacterium *Rhodopseudomonas palustris* GJ-22 induces systemic resistance against viruses. *Microb Biotechnol.* 2017;10:612-24.
56. Nookongbut P, Kantachotea D, Khuong NQ, Tantirungkij M. The biocontrol potential of acid-resistant *Rhodopseudomonas palustris* KTSSR54 and its exopolymeric substances against rice fungal pathogens to enhance rice growth and yield. *Biol Control.* 2020;150:104354.
57. Palaniyandi SA, Yang SH, Zhang L, Suh JW. Effects of actinobacteria on plant disease suppression and growth promotion. *Appl Microbiol Biotechnol.* 2013; 97:9621-36.
58. Sreevidya M, Gopalakrishnan S, Kudapa H, Varshney RK. Exploring plant growth-promotion actinomycetes from vermicompost and rhizosphere soil for yield enhancement in chickpea. *Braz J Microbiol.* 2016;47:85-95.
59. Anwar S, Ali B, Sajid I. Screening of rhizospheric actinomycetes for various *in vitro* and *in vivo* plant growth promoting (PGP) traits and for agroactive compounds. *Front Microbiol.* 2016;7:1334.
60. Liu X, Bolla K, Ashforth EJ, Zhuo Y, Gao H, Huang P, et al. Systematics-guided bioprospecting for bioactive microbial natural products. *Anton Leeuw Int J G.* 2012; 101:55-66.
61. Jinhua C, Yang SH, Palaniyandi SA, Han JS, Yoon TM, Kim TJ, et al. Azalomycin F complex is an antifungal substance produced by *Streptomyces malaysiensis* MJM1968 isolated from agricultural soil. *J Korean Soc Appl Bi.* 2010;53:545-52.
62. Palaniyandi SA, Yang SH, Cheng JH, Meng L, Suh JW. Biological control of anthracnose (*Colletotrichum gloeosporioides*) in yam by *Streptomyces* sp. MJM5763. *J Appl Microbiol.* 2011;111:443-55.
63. Samaca DA, Willertc AM, McBrideb MJ, Kinkel LL. Effects of antibiotic-producing *Streptomyces* on nodulation and leaf spot in alfalfa. *Appl Soil Ecol.* 2003;22:55-66.

64. Zucchi TD, Almeida LG, Dossi FCA, Cônsoli FL. Secondary metabolites produced by *Propionicimonas* sp. (ENT-18) induce histological abnormalities in the sclerotia of *Sclerotinia sclerotiorum*. *BioControl*. 2010;55:811-9.
65. Jeong Y, Kim J, Kim S, Kang YS. Toxoflavin produced by *Burkholderia glumae* causing rice grain rot is responsible for inducing bacterial wilt in many field crops. *Plant Dis*. 2003;87:890-5.
66. Kunakom S, Eustáquio AS. *Burkholderia* as a source of natural products. *J Nat Prod*. 2019;82:2018-37.
67. Ding L, Xu LJ, Chu X, Yang L, Zhu HL, Huang JX. Dissimilarity analysis of microbial communities in the rhizosphere and tissues of diseased and healthy cherry trees (*Cerasus pseudocerasus*). *Can J Plant Pathol*. 2021.
68. Xu ZZ, Wang MY, Du JP, Huang T, Liu JF, Dong T, Chen YL. Isolation of *Burkholderia* sp. HQB-1, a promising biocontrol bacteria to protect banana against Fusarium wilt through phenazine-1-carboxylic acid secretion. *Frontiers in Microbiology*. 2020;11:605152.
69. Miao ZQ, Li SD, Liu XZ, Chen YJ, Li YH, Wang Y, et al. The causal microorganisms of *Panax notoginseng* root rot disease. *Scientia Agricultura Sinica*. 2006;39:1371-8.
70. Liang F, Zhang B, Yang QQ, Zhang Y, Zheng DH, Zhang LQ, et al. Cyclic-di-GMP regulates the quorum-sensing system and biocontrol activity of *Pseudomonad fluorescens* 2P24 through the RsmA and RsmE proteins. *Genet Mol Biol*. 2020;86:e02016-20.
71. Silby MW, Cerdano-Tarrage AM, Vernikos GS, Giddens, SR, Jackson RW, Preston GM, et al. Genomic and genetic analyses of diversity and plant interactions of *Pseudomonas fluorescens*. *Genome Biol*. 2009;10:R51.
72. Panpatte DG, Shelat HN, Jhala YK, Vyas RV. Fortified bacterial consortium -A novel approach to control root knot nematode in cucumber (*Cucumis sativum*). *Biol Control*. 2021;155:104528
73. Godebo AT, Germida JJ, Walley FL. Isolation, identification, and assessment of soil bacteria as biocontrol agents of pea root rot caused by *Aphanomyces euteiches*. *Can J Soil Sci*. 2020;100:206-16.
74. Wen T, Zhao ML, Yuan J, A. Kowalchuk G, Shen QR. Root exudates mediate plant defense against foliar pathogens by recruiting beneficial microbes. *SEL*. 2020;3:42-51.
75. Li MZ, Chen ZJ, Qian J, Wei FG, Zhang GZ, Wang Y, et al. Composition and function of rhizosphere microbiome of *Panax notoginseng* with discrepant yields. *Chinese Medicine*. 2020;15:85.
76. Ha NT, Minh TQ, Hoi PX, Thuy NTT, Furuya N, Long HH. Biological control of potato tuber soft rot using N-acyl-L-homoserine lactone-degrading endophytic bacteria. *Curr Sci India*. 2018;115:1921-7.
77. Kim YJ, Abbai R, Singh P, Jung KH, Kim YJ, Yang CD, et al. Pathogenesis strategies and regulation of ginsenosides by two species of *Ilyonectria* in *Panax ginseng*: power of speciation. *J Ginseng Res*. 2020;44:332-40.
78. DesRochers N, Walsh JP, Renaud JB, Seifert KA, Yeung KC, Sumarah MW. Metabolomic profiling of fungal pathogens responsible for root rot in American Ginseng. *Metabolites*. 2020;10:35.

79. Marra R, Crescenzi A, Davino SW, Fanigliulo A, Woo SL, Lorito M, et al. Integrated management strategies of *Meloidogyne incognita* and *Pseudopyrenochaeta lycopersici* on tomato using a *Bacillus firmus*-based product and two synthetic nematicides in two consecutive crop cycles in greenhouse. *Crop Prot.* 2019;122:159-64.
80. Ahman J, Johansson T, Olsson M, Punt PJ, Hondel VD, Tunlid A. Improving the pathogenicity of a nematode-trapping fungus by genetic engineering of a subtilisin with nematotoxic activity. *Appl Environ Microbiol.* 2002;68:3408-18.
81. Douglas AE. The nutritional quality of phloem sap utilized by natural aphid populations. *Ecol Entomol.* 1993;18:31-8.
82. Rico A, Preston GM. *Pseudomonas syringae* pv. *tomato* DC3000 uses constitutive and apoplast-induced nutrient assimilation pathways to catabolize nutrients that are abundant in the tomato apoplast. *Mol Plant Microbe In.* 2008; 21:269-82.
83. Besnard J, Sonawala U, Maharjan B, Collakova E, Finlayson SA, Pilot G, et al. Increased expression of UMAMIT amino acid transporters results in activation of salicylic acid dependent stress response. *Front Plant Sci.* 2021;11:606386.
84. Yang H, Postel S, Kemmerling B, Ludewig U. Altered growth and improved resistance of *Arabidopsis* against *Pseudomonas syringae* by overexpression of the basic amino acid transporter *AtCAT1*. *Plant Cell Environ.* 2014;37:1404-14.
85. Zhao WC, Hao JH, Xing JY, Yang R, Zhao FK, Wang JL, et al. Proteomic analysis of phloem proteins leads to the identification of potential candidates for JA-mediated RKN-resistant elements in *Solanum lycopersicum*. *J Plant Growth Regul.* 2017;36:96-105.
86. Koh KH, Tham FY. Screening of traditional Chinese medicinal plants for quorum-sensing inhibitors activity. *Journal of Microbiology.* 2011;44:144-8.
87. Devanna BN, Jaswal R, Singh PK, Kapoor R, Jain P, Kumar G, et al. Role of transporters in plant disease resistance. *Physiol Plant.* 2021;171:849-67.
88. Hwang JU, Song WY, Hong D, Ko D, Yamaoka Y, Jang S, et al. Plant ABC transporters enable many unique aspects of a terrestrial plant's lifestyle. *Mol Plant.* 2016;9:338-55.
89. Borghi L, Kang J, Francisco R. Filling the gap: Functional clustering of ABC proteins for the investigation of hormonal transport in *planta*. *Front Plant Sci.* 2019;10:422.
90. Dhara A, Raichaudhuri A. ABCG transporter proteins with beneficial activity on plants. *Phytochemistry.* 2021;184:112663.
91. Zia R, Nawaz MS, Yousaf S, Amin I, Hakim S, Imran A. Seed inoculation of desert-plant growth-promoting rhizobacteria induce biochemical alterations and develop resistance against water stress in wheat. *Physiologia Plantarum.* 2021.
92. Gong C, Cheng MZ, Li JF, Chen HY, Zhang ZZ, Qi HN, Zhang Y, et al. The  $\alpha$ -subunit of the chloroplast ATP synthase of tomato reinforces resistance to gray mold and broad-spectrum resistance in transgenic tobacco. *Phytopathology.* 2021;111:485-95.

93. [Corredor-Moreno P](#), [Minter F](#), [Davey PE](#), [Wegel E](#), [Kular B](#), [Brett P](#), et al. The branched-chain amino acid aminotransferase TaBCAT1 modulates amino acid metabolism and positively regulates wheat rust susceptibility. *Plant Cell*. 2021;1-20.
94. Liang YB, Li Z, Zhang Y, Meng FL, Qiu DW, Zeng HM, et al. Nbnrp1 mediates *Verticillium dahliae* effector PevD1-triggered defense responses by regulating sesquiterpenoid phytoalexins biosynthesis pathway in *Nicotiana benthamiana*. *Gene*. 2021;768:145280.
95. Wang JY, Wang RY, Fang H, Zhang CY, Zhang F, Hao ZY, et al. Two VOZ transcription factors link an E3 ligase and an NLR immune receptor to modulate immunity in rice. *Mol Plant*. 2021;14:253-66.
96. [Ichinose YK](#), [Nishimura T](#), [Harada M](#), [Kashiwagi R](#), [Yamamoto M](#), [Noutoshi Y](#), et al. Role of two sets of RND-type multidrug efflux pump transporter genes, *mexAB-oprM* and *mexEF-oprN*, in virulence of *Pseudomonas syringae* pv. *tabaci* 6605. *Plant Pathol J*. 2020;36:148-56.
97. Liang ZB, Huang LH, He F, Zhou XF, Shi ZR, Zhou JN, et al. A substrate-activated efflux pump, DesABC, confers zeamine resistance to *Dickeya zea*. *mBio*. 2019;10:e00713-19.
98. Siriyong T, Srimanote P, Chusri S, Yingyongnarongkul B, Suaismom C, Tipmanee V, et al. Conessine as a novel inhibitor of multidrug efflux pump systems in *Pseudomonas aeruginosa*. *BMC Complement Altern M*. 2017;17:405.
99. Iquebal MA, Tomar RS, Parakhia MV, Singla D, Jaiswal S, Kumar N, et al. Draft whole genome sequence of groundnut stem rot fungus *Athelia rolfsii* revealing genetic architect of its pathogenicity and virulence. *Scientific Reports*. 2017;7:5299.
100. Li SX, Musungu B, Lightfoot D, Ji PS. The Interactomic analysis reveals pathogenic protein networks in *Phomopsis longicolla* underlying seed decay of soybean. *Front Genet*. 2018;9:104.
101. [Ha X](#), [Koopmann B](#), [Tiedemann AV](#). Wheat blast and fusarium head blight display contrasting interaction patterns on ears of wheat genotypes differing in resistance. *Phytopathology*. 2016;106:270-81.
102. Andrews, S. FastQC: a quality control tool for high throughput sequence data. *AJPS*. 2010;11.
103. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30:2114-20.
104. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357-9.
105. Zhang D, Li W, Xia EH, Zhang QJ, Liu Y, Niu YC, et al. The medicinal herb *panax notoginseng* genome provides insights into ginsenoside biosynthesis and genome evolution. *Mol Plant*. 2017;10:903-7.
106. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. 2011.
107. Zuo JJ, Zu MT, Liu L, Song XM, Yuan YD. Composition and diversity of bacterial communities in the rhizosphere of the Chinese medicinal herb *Dendrobium*. *BMC Plant Biol*. 2021;21:127.
108. Lind, AL., Pollard, KS. Accurate and sensitive detection of microbial eukaryotes from whole metagenome shotgun sequencing. *Microbiome*. 2021;9:58.

109. Köster J, Rahmann S. Snakemake—a scalable bioinformatics workflow engine. *Bioinformatics*. 2012;28:2520-2.
110. Richter DJ, Berney C, Strasser JFH, Burki F, De Vargas C. EukProt: a database of genome-scale predicted proteins across the diversity of eukaryotic life. Cold Spring Harbor Laboratory. 2020;2020:180687.
111. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods*. 2015;12:59-60.
112. Li DH, Liu CM, Luo RB, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de *Bruijn* graph. *Bioinformatics*. 2015;31:1674-6
113. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29:1072-5.
114. Doug Hyatt, LoCasio1 PF, Hauser LJ, Uberbacher EC. Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics*. 2012;28: 2223-30.
115. Fu LM, Niu BF, Zhu ZW, Wu ST, Li WZ. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics*. 2012;28:3150-2.
116. [Patro R](#), [Duggal G](#), [Love MI](#), [Irizarry RA](#), [Kingsford C](#). Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017;14:417-9.
117. Lombard V, Golaconda RE, Drula E, Coutinho PM, Henrissat B. The Carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res*. 2013; 42:D490-5.
118. Gibson MK, Forsberg KJ, Dantas G. Improved annotation of antibiotic resistance determinants reveals microbial resistomes cluster by ecology. *ISME J*. 2015; 9:207-16.
119. V.Uritskiy G, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome*. 2018;6:158.
120. Wu YW, Simmons BA, Singer SW. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics*. 2016;32:605-7.
121. Li F, An H, [Kirton E](#), [Thomas A](#), [Egan R](#), Wang Z, et al. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ*. 2019;7:e7359.
122. Paeks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2014;25:1043-55.
123. Ondov DB, Bergman NH, Phillippy AM. Interactive metagenomic visualization in a Web browser. *BMC Bioinformatics*. 2011;12:385.

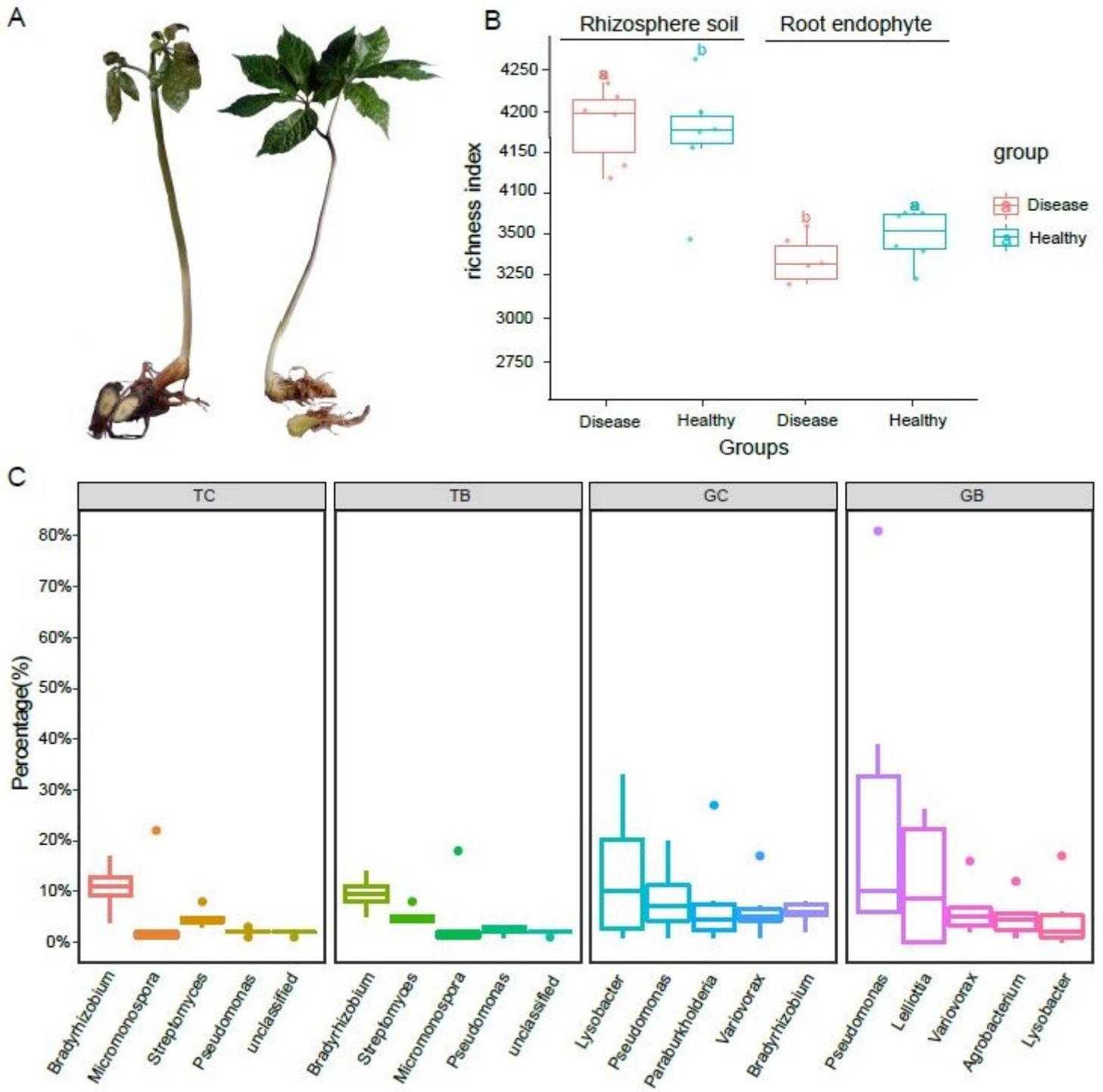
## Table

**Table 1 Summary table of samples and sequencing data**

| Sample Type    | Plant status | Number           | Raw Reads (Gb) | Clean reads (Gb) |      |      |
|----------------|--------------|------------------|----------------|------------------|------|------|
| Root endophyte | Disease      | LJGB-1           | 25.7           | 10.9             |      |      |
|                |              | LJGB-2           | 32.1           | 24.9             |      |      |
|                |              | LJGB-3           | 25.5           | 4.16             |      |      |
|                |              | QBGB-1           | 33.3           | 17.4             |      |      |
|                |              | QBGB-2           | 30.4           | 22.8             |      |      |
|                |              | QBGB-3           | 31.8           | 28.1             |      |      |
|                | Healthy      | LJGC-1           | 12.8           | 1.7              |      |      |
|                |              | LJGC-2           | 27.9           | 4.0              |      |      |
|                |              | LJGC-3           | 28.6           | 6.5              |      |      |
|                |              | QBGC-1           | 28.5           | 8.6              |      |      |
|                |              | QBGC-2           | 27.0           | 13.6             |      |      |
|                |              | QBGC-3           | 30.3           | 9.2              |      |      |
|                |              | Rhizosphere soil | Disease        | LJTB-1           | 13.5 | 12.8 |
|                |              |                  |                | LJTB-2           | 13.2 | 12.8 |
|                |              |                  |                | LJTB-3           | 13.0 | 12.5 |
| QBTB-1         | 13.0         |                  |                | 12.4             |      |      |
| QBTB-2         | 12.7         |                  |                | 12.3             |      |      |
| QBTB-3         | 13.1         |                  |                | 12.6             |      |      |
| Healthy        | LJTC-1       |                  | 33.9           | 32.5             |      |      |
|                | LJTC-2       |                  | 10.5           | 10.1             |      |      |
|                | LJTC-3       |                  | 13.0           | 12.5             |      |      |
|                | QBTC-1       |                  | 12.2           | 11.7             |      |      |
|                | QBTC-2       |                  | 12.8           | 12.3             |      |      |
|                | QBTC-3       |                  | 12.9           | 12.3             |      |      |

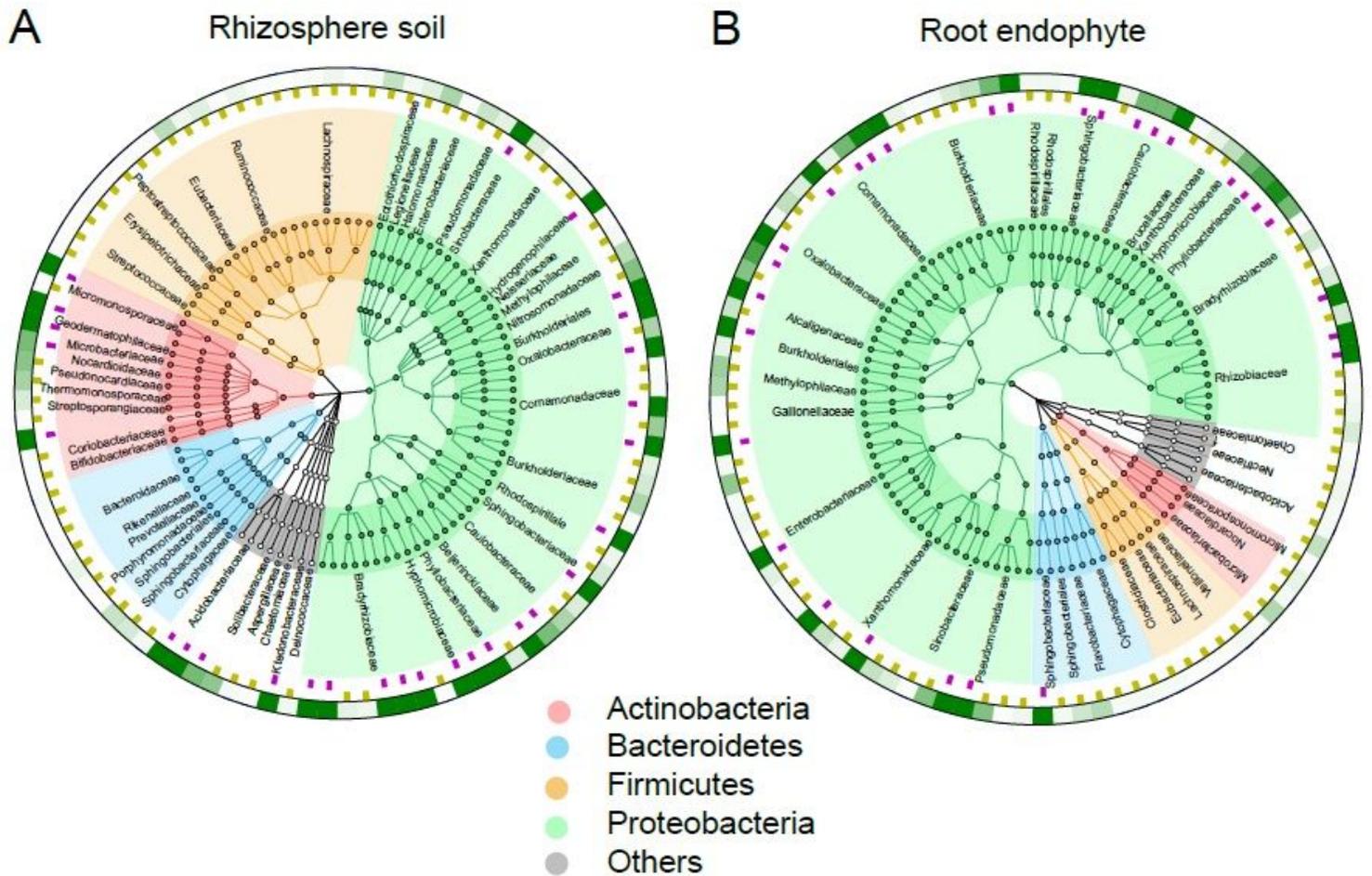
Note: Raw reads are the data volume of paired-end sequencing, and Clean reads are sequences that are still paired after quality control and de-hosting.

## Figures



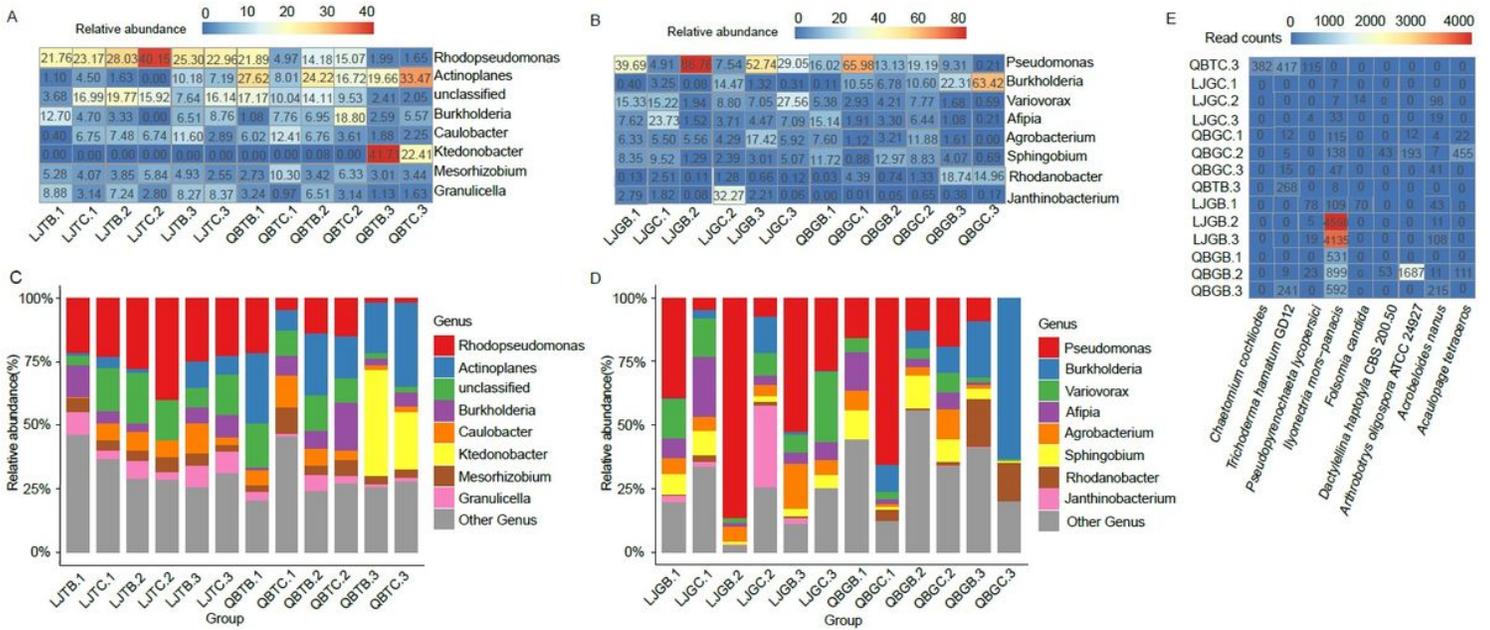
**Figure 1**

Characteristics of healthy and root rot diseased *Panax notoginseng* plants. (A) Diseased and healthy *P. notoginseng* plant parts. (B) Richness index box plot based on the rhizosphere soil and root endophyte bacteria species abundance. (C) Box plots of the top five species in each of the rhizosphere soil and root endophyte groups: healthy rhizosphere soil (TC), diseased rhizosphere soil (TB), healthy root endophytes (GC), and diseased root endophytes (GB)



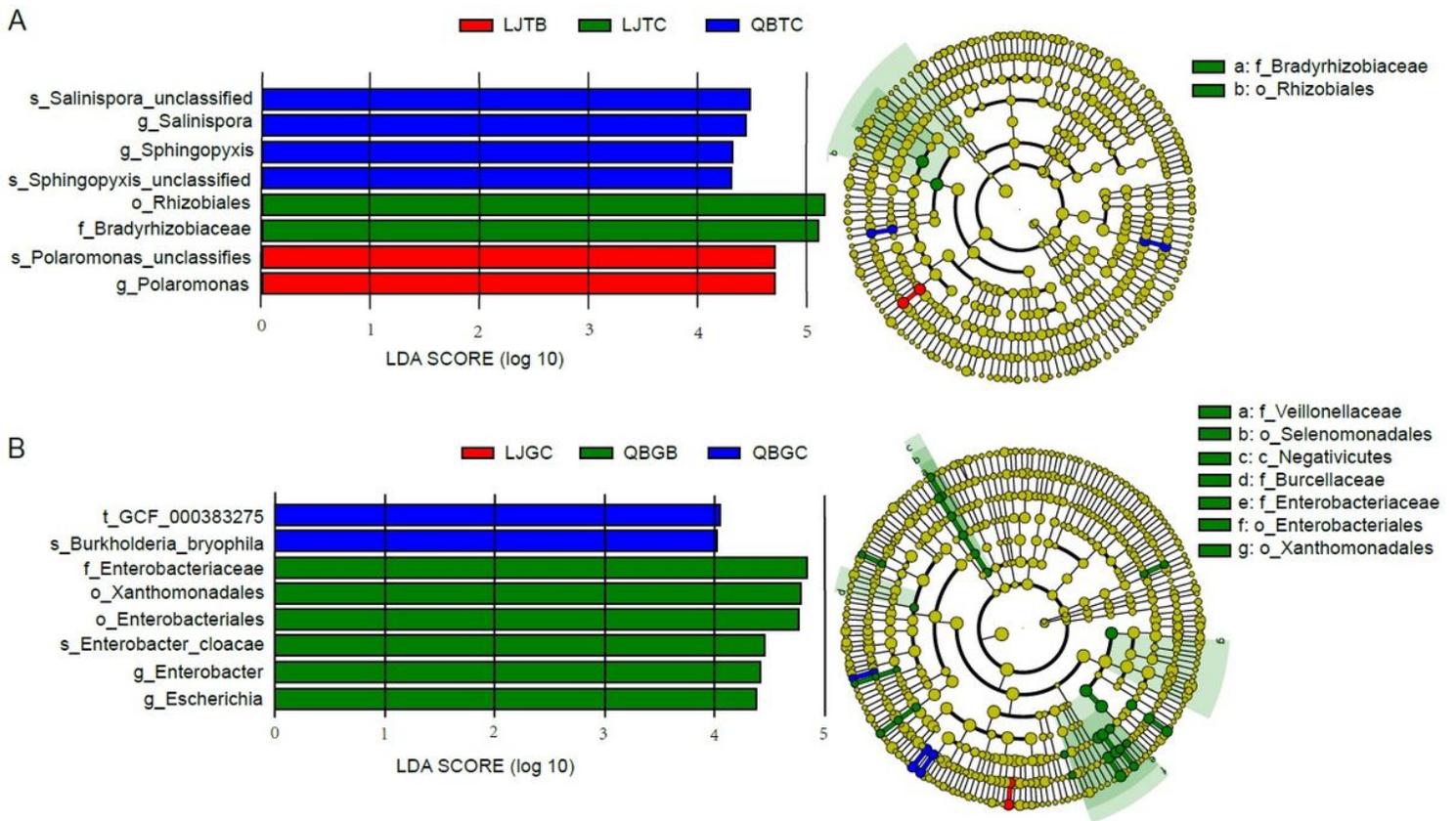
**Figure 2**

Species composition circle diagrams based on the top 100 most abundant species in 12 rhizosphere soil samples (A) and 12 root endophyte samples (B) of healthy and root rot infected *Panax notoginseng* plants. The colors represent the four phyla that contain these species. The purple squares in the first ring indicate the species with an abundance greater than 5 per thousand. The yellow squares in the second ring indicate the species with an abundance less than 5 per thousand. The third ring indicates the average value for all samples, where the different shades of green show the different species abundances



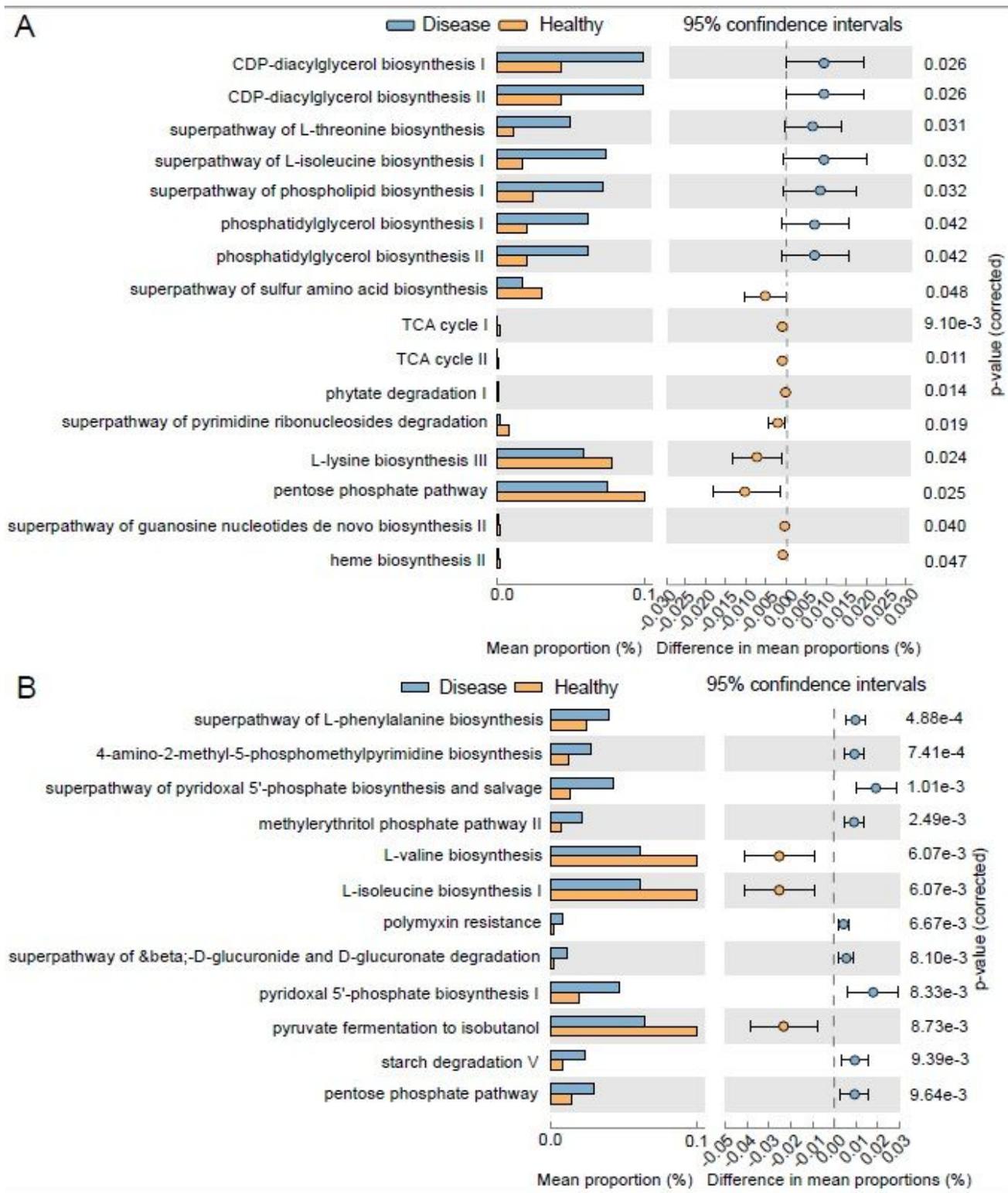
**Figure 3**

Genus composition heat maps and corresponding histograms for the microbial communities associated with healthy and root rot infected *Panax notoginseng* plants. A and C show the relative abundance of the top eight species in each sample of the rhizosphere soil groups. B and D show the relative abundance of the top eight species in each sample of the root endophyte groups. E shows the top nine species of the eukaryotes present in the samples. The sample names represent the sample locations: Lijiong (LJ) and Qiubei (QB), the sample groups: healthy rhizosphere soil (TC), diseased rhizosphere soil (TB), healthy root endophytes (GC), and diseased root endophytes (GB), and the replicates: 1, 2, and 3



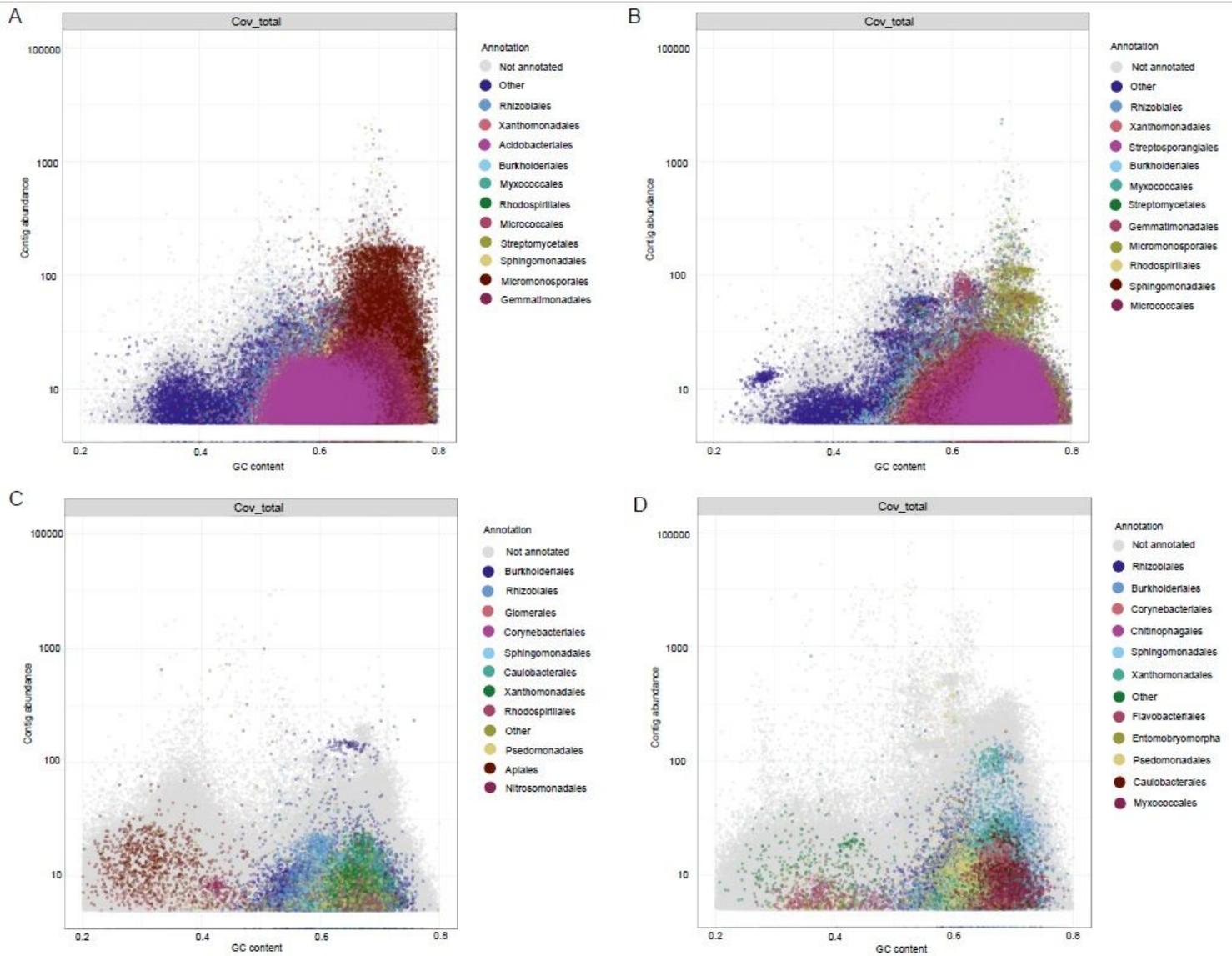
**Figure 4**

Species differences in the rhizosphere soil (A) and root endophyte (B) microbial communities associated with healthy and root rot infected *Panax notoginseng* plants. The circle charts, determined using LEfSe, display the species with significant differences (LDA > 4) between the groups. The colors in the histograms and circle charts represent the different groups according to the sample locations: Lijiang (LJ) and Qiubei (QB), and the sample groups: healthy rhizosphere soil (TC), diseased rhizosphere soil (TB), healthy root endophytes (GC), and diseased root endophytes (GB)



**Figure 5**

Differences in the functional composition of the rhizosphere soil and root endophyte microbial communities associated with healthy and root rot infected *Panax notoginseng* plants. A and B show the functional pathways with significant differences between healthy and diseased rhizosphere soils and healthy and diseased root endophytes, respectively



**Figure 6**

The binning analysis results based on the microbial communities associated with healthy and root rot infected *Panax notoginseng* plants. The abscissa is the GC (healthy root endophytes) content of the contig, the ordinate is contig abundance, one point represents a contig, and contigs of the same color come from the same Order. Healthy rhizosphere soil (A), Diseased rhizosphere soil (B), Healthy root endophytes (C), Diseased root endophytes (D)

## Supplementary Files

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

- [SupplementaryFigure1.pdf](#)
- [SupplementaryFigure2.pdf](#)
- [SupplementaryFigure3.pdf](#)

- [SupplementaryFigure4.pdf](#)
- [SupplementaryFigure5.pdf](#)
- [SupplementaryFigure6.pdf](#)
- [SupplementaryFigure7.pdf](#)
- [SupplementaryFigure8.pdf](#)
- [SupplementaryFigure9.pdf](#)
- [SupplementaryFigure10.pdf](#)
- [SupplementaryFigure11.pdf](#)
- [SupplementaryTable1.xlsx](#)
- [SupplementaryTable2.xlsx](#)
- [SupplementaryTable3.xlsx](#)
- [SupplementaryTable4.xlsx](#)
- [SupplementaryTable5.xlsx](#)
- [SupplementaryTable6.xlsx](#)
- [SupplementaryTable7.xlsx](#)
- [SupplementaryTable8.xlsx](#)
- [SupplementaryTable9.xlsx](#)