

# Intergrated metagenomics and metabolomics analysis discovers nematicidal microbes, enzymes and metabolites from the plant rhizosphere microbiota

**Xiuyun Zhao**

Huazhong Agricultural University: Huazhong Agriculture University

**Changchun Lin**

Huazhong Agricultural University: Huazhong Agriculture University

**Wenfang Yuan**

Huazhong Agricultural University: Huazhong Agriculture University

**Song Ruan**

Huazhong Agricultural University: Huazhong Agriculture University

**Gaofu Qi** (✉ [qigaofu@mail.hzau.edu.cn](mailto:qigaofu@mail.hzau.edu.cn))

Huazhong Agricultural University <https://orcid.org/0000-0002-6396-2819>

---

## Research

**Keywords:** Meloidogyne incognita, rhizophere microbiota, metagenomics analysis, metabolomics analysis, acetophenone, Bacillus amyloliquefaciens

**Posted Date:** May 12th, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-187899/v2>

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

[Read Full License](#)

---

---

**Intergrated metagenomics and metabolomics analysis discovers nematicidal microbes,  
enzymes and metabolites from the plant rhizosphere microbiota**

Xiuyun Zhao, Changchun Lin, Wenfang Yuan, Song Ruan, Gaofu Qi†

Xiuyun Zhao: xiuyunzh@mail.hzau.edu.cn; Changchun Lin: 1966703629@qq.com; Wenfang  
Yuan: 1663503327@qq.com; Song Ruan: 476478609@qq.com

†Corresponding author. Email: qigaofu@mail.hzau.edu.cn

College of Life Science and Technology, Huazhong Agricultural University, Wuhan 430070,  
PR China

---

## Abstract

**Background:** Root-knot nematode *Meloidogyne incognita* infects root systems of many crops resulting in huge decrease of crop production. Nematicidal microorganisms provides a safe and effective strategy to control *M. incognita* infection. In order to find microorganisms with high activity and new nematicidal metabolites, we collected the *M. incognita* infected tobacco rhizosphere soils (RNI) and non-infected tobacco rhizosphere soils (NS), and investigated their microbial community and network via metagenomics and metabolomics analysis.

**Results:** Microbial networks of RNI soils were very different from the NS soils. Many nematicidal microorganisms were enriched in the NS soils, including isolates of *Aspergillus*, *Achromobacter*, *Acinetobacter*, *Bacillus*, *Burkholderia*, *Comamonas*, *Enterobacter*, *Lysobacter*, *Microbacterium*, *Paenibacillus*, *Pantoea*, *Pseudomonas*, *Streptomyces* and *Variovorax*. Enzymes analysis showed these nematicidal microorganisms can produce proteases, chitinase and lipases. The functions genes belonging to pathways of secondary metabolites biosynthesis and carbohydrate transport and metabolism were overrepresented in the rhizosphere microbiota of NS soils comparing with the RNI soils. 102 metabolites contents were significantly different between the RNI and NS rhizosphere microbiota. 35 metabolites were overrepresented in the NS soils comparing the RNI samples, including acetophenone. Acetophenone showed high nematicidal ( $LC_{50} = 0.66 \mu\text{g/ml}$ ) and avoidance activity against *M. incognita*. *Bacillus amyloliquefaciens* W1 could produce acetophenone. Liquid culture of W1 could kill 98.8% of *M. incognita* J2 juveniles after treatment for 24 h.

**Conclusions:** In general, the rhizosphere microbiota of NS soils could produce volatile materials, multiple enzymes and secondary metabolites against nematode. Collectively, the microbiota of NS and RNI rhizosphere differed significantly in microbial network structure, community composition, function genes and metabolites. Collectively, combination of multi-omics analysis and culture-dependent technology is powerful for finding nematicidal microorganisms and metabolites from soil.

Keywords: *Meloidogyne incognita*; rhizosphere microbiota; metagenomics analysis; metabolomics analysis; acetophenone; *Bacillus amyloliquefaciens*

## Background

The root-knot nematode *Meloidogyne incognita* can parasitize plant roots and causes dramatic

---

yield losses in many different crops worldwide [1], but the management options available for effectively controlling root-knot nematode are extremely limited. Although traditional control methods such as crop rotation, disease resistance breeding and use of chemical pesticides have reduced nematode diseases indeed, the crop yield losses caused by *M. incognita* are still enormous [2, 3]. Fortunately, some microbial agents can effectively control root-knot nematodes to reduce *M. incognita* infection. These microorganisms include rhizosphere bacteria, nematophagous fungi, specialized parasitic bacteria, actinomycetes, etc, [4, 5]. Among them, rhizosphere microbiota are especially effective. Rhizosphere microbiota seemed to regulate infection of root-knot nematode in plants [6-8].

Many microorganisms have been isolated and selected as biological agents for controlling root-knot nematodes [9-11]. However, the control effect of microbial agents are often unstable in field [12]. Many farmland soils in the Southern China are seriously degraded, resulting in unbalanced microbial communities and networks [13-15]. Thus, application of beneficial microorganisms to control plant disease in the degraded soils is important. Researchers often pay more attentions to screening of microbes for biocontrol of root-knot nematodes. The interaction relationships among root-knot nematodes, rhizosphere microbiota and biocontrol microbes are rarely understood in the soil ecosystem. Nowadays, managing microbial community for a better colonization and a higher biocontrol efficiency of beneficial microorganisms in soils becomes one of the research hotspots. Investigation of microbial interactions may help to resolve this bottle-neck problem. Microbial interactions encompass a spectrum ranging from antagonistic to cooperative, syntrophic and mutualistic interactions. Network analysis is favorable for exploring the organisation and dynamics of microbial interactions and niches [16, 17], and help us to find potential beneficial microorganisms and improve their colonization efficiency in soils.

Metagenomics and metabolomics analysis have enabled the study of microbial ecosystem structure to a greater depth and accuracy. Metagenomics sequencing technology is an interesting tool for selection of the desirable biocontrol microbes [18], and metabolomics is favorable for finding active metabolites produced by microorganisms [19]. The main object of this project is to study the important functional genes and key microorganisms related to the occurrence of root-knot nematode disease, and screen novel nematicidal microorganisms

---

and metabolites against root-knot nematodes based on the metagenomic and metabolomic analysis. This study will find new nematicidal metabolites and potential biocontrol strains for effective control of root-knot nematodes.

## **Methods**

### **Soil sampling and properties analysis**

Based on our previous disease investigation, fifteen root-knot nematode *M. incognita* infected tobacco fields (RNI) and fifteen non-infected tobacco (*Nicotiana tabacum* cultivar Yunyan 87) fields (NS) were located in Enshi City, Hubei Province, China (29°57'-29°58'N, 109°20'-109°22'E), with elevations ranging from 990 to 1,127 m and subtropical humid climate (table S1). The soil is yellow-brown soil (classified as Alfisols). Rhizosphere soil samples were collected from ten tobacco plants that showed galling in RNI group or healthy plants without gall in NS group, which are randomly distributed over each field (20 m × 33 m area) in August 2018 at the maturation stage of tobacco (from July 20<sup>th</sup> to September 10<sup>th</sup>). The tobaccos were removed gently from the field, then the loosely adherent soils were removed by vigorous shaking, and the soils adhering to root were collected with a brush. The collected rhizosphere soil of ten plants in each field was mixed together as a composite soil sample. Soil samples were sieved (2-mm mesh) to remove plant roots, debris and stones. Each soil sample (1 kg soil) was divided into two parts: one (50 g soil) was stored at - 80 °C for DNA and metabolites extraction within one week, and the other one (950 g soil) was stored at 4 °C for soil properties and root-knot nematode analysis. Quantification of *M. incognita* in soil samples was determined via real-time quantitative PCR [20].

Soil organic matter (SOM) content was assayed by the acidified potassium dichromate (K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>-H<sub>2</sub>SO<sub>4</sub>) heating method [21]. Alkali-hydrolyzable nitrogen (AN) content was determined by alkaline hydrolysis diffusion method [21]. Available phosphorus (AP) and potassium (AK) contents were determined photometrically by a flame spectrophotometer [21]. After suspending with deionized water (soil : water = 1 : 2.5, w/v), soil pH was determined using a pH meter (Mettler-Toledo FE20; Switzerland). Urease activity was detected by colorimetric determination of ammonium. Catalase activity was determined by colorimetric assay using K<sub>2</sub>Cr<sub>2</sub>O<sub>7</sub>/acetic acid reagent. Invertase activity was measured by 3,5-dinitrosalicylic acid colorimetry with sucrose as substrate. Phosphatase activity was

---

measured by *p*-nitrophenol colorimetry [22]. Soil water content (SWC), capillary moisture capacity (SCMC), bulk density (BD), total porosity (STP), capillary porosity (SCP), aeration porosity (SAP) and mean weight diameter (MWD) were routinely measured according to previous methods [21, 23].

### **Metagenomic sequencing data pretreatment and metagenome assembly**

For each soil sample, DNA was extracted from 1.0 g soils using the FastDNA Spin Kit (MP Biomedicals, USA). DNA concentration was measured using Qubit® dsDNA Assay Kit in Qubit® 2.0 Fluorometer (Life Technologies, CA, USA). 1 µg DNA per sample was used to construct library. Sequencing libraries were generated using NEBNext® Ultra™ DNA Library Prep Kit for Illumina (NEB, USA) and index codes were added to attribute sequences to each sample. Briefly, the DNA sample was fragmented by sonication (30s On/ 30s Off, 7 times) with a Bioruptor (Diagenode, Belgium) to a size of 350 bp, then DNA fragments were end-polished, A-tailed, and ligated with the full-length adaptor for PCR amplification. PCR products were purified by AMPure XP system (Beckman Coulter, USA) and libraries were analyzed for size distribution by Agilent2100 Bioanalyzer (Agilent Technologies Inc., USA) and quantified using real-time PCR. Clustering of the index-coded samples was performed on a cBot Cluster Generation System (Illumina®, San Diego, USA) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina HiSeq 2000 platform and paired-end reads were generated.

Preprocessing the raw data using Readfq (V8, <https://github.com/cjfields/readfq>) was conducted to acquire the clean data for subsequent analysis. The reads that are of host origin were filtered using Bowtie2.2.4 software [24]. The clean data was assembled and analyzed by MEGAHIT software (v1.0.4-beta) [25], then interrupted the assembled Scaffigs from N connection and leave the Scaffigs without N [26]. The unused PE reads was acquired by comparing clean data to Scaffolds using Bowtie2.2.4 software, further mixed assembly using the software SOAPdenovo V2.04/MEGAHIT v1.0.4-beta [27]. Scaffigs were obtained by breaking Scaffolds from N connection. Scaffigs shorter than 500 bp were filtered.

### **Gene prediction and abundance analysis**

The ORF in Scaffigs ( $\geq 500$  bp) assembled from both single and mixed assembly were all

---

predicted by MetaGeneMark V2.10 software [28]. For ORF prediction, the CD-HIT V4.5.8 software [29] was adopted to redundancy and obtained the unique initial gene catalogue [30]. Clean data of each sample was mapped to initial gene catalogue using Bowtie2.2.4 and reads number for genes mapped in each sample was got. We filtered the gene with the number of reads  $\leq 2$  in each sample and obtained the gene catalogue Unigenes eventually used for subsequently analysis [31]. Based on the number of mapped reads and the length of gene, the abundance information of each gene in each sample was counted.

### **Taxonomy prediction**

DIAMOND V0.9.9 software [32] was used to assign the Unigenes to the sequences of bacteria, fungi, archaea and viruses which are all extracted from the NR database of NCBI [32]. Number of genes and the abundance information of each sample in each taxonomy hierarchy (kingdom, phylum, class, order, family, genus, specie) were obtained. The abundance of a species in one sample equals the sum of the gene abundance annotated for the species. The gene number of a species in a sample equals the number of genes whose abundance are nonzero.

Principal component analysis (PCA) and non-metric multi-dimensional scaling (NMDS) analysis were performed based on the abundance data in R package. The difference between two groups of rhizosphere microbiota was tested by Anosim analysis. Metastats [33] and LEfSe analysis [34] were used to find the different species between groups using permutation test between groups [33]. Finally, random forest (RandoForest) was used to screen out important species.

### **Functional database annotations**

We adopted DIAMOND V0.9.9 software to blast Unigenes to functional database KEGG [35], eggNOG [36], and CAZy database [37] (e value  $\leq 1e-5$ ). For each sequence's blast result, the Blast Hit with the highest scoring annotated hit(s) was used for subsequent analysis. The relative abundance of different functional hierarchy was counted. The gene number of each sample in each taxonomy hierarchy was obtained. Based on the abundance of each taxonomy hierarchy, the following analyses were conducted. The annotated gene numbers were counted, and the general relative abundance and the abundance cluster heat map were exhibited

---

respectively. The decrease-dimension analysis of PCA and NMDS were conducted. Anosim analysis of the difference between groups was performed based on the functional abundance. The comparative analysis of metabolic pathways was done. Metatata and LEfSe analysis of the functional difference between groups were performed. We searched for carbohydrate-active enzymes (CAZymes) and secondary metabolite bio-synthetic gene clusters within the metagenome sequences using dbCAN [38] and antiSMASH [39], respectively.

### **Network construction and analysis**

The relative abundances of species or functional genes were used to construct microbial or functional networks using the Molecular Ecological Network Analysis Pipeline (MENAP), respectively [40, 41]. Briefly, species or gene abundances were normalized to the standardized relative abundances (SRA) [16]. A matrix of species SRA or genes SRA, a matrix of soil variables, and a species or genes annotation file were prepared. The SRA matrix was submitted to MENAP to construct the NS and RNI microbial or functional networks, respectively [40, 41]. Modules were detected by the greedy modularity optimization (a module separation algorithm). Three files were generated for network graph visualization by Cytoscape 3.7.2 software [42]. The network graph was represented using different species or genes (nodes) with positive or negative interactions (edges) [43, 44].

The topological role of each node (specie) was defined by two parameters: within-module connectivity ( $Z_i$ ) and among-module connectivity ( $P_i$ ).  $Z_i$  value described how well a node was connected to other nodes within its own module.  $P_i$  value described how well a node was connected to different modules [45]. Threshold values of  $Z_i$  and  $P_i$  for categorizing nodes were 2.5 and 0.62, respectively [46]. According to the threshold values of  $Z_i$  and  $P_i$ , the nodes were divided into four different categories. (i) Peripheral nodes (named specialists) with low  $Z_i$  ( $< 2.5$ ) and low  $P_i$  ( $< 0.62$ ) values. These nodes have only a few links and almost always are connected to the nodes within their own modules. (ii) Connectors (named generalists) have low  $Z_i$  values ( $< 2.5$ ) but high  $P_i$  values ( $> 0.62$ ). They are highly connected with other modules. (iii) Module hubs (also named generalists) with high  $Z_i$  values ( $> 2.5$ ) but low  $P_i$  values ( $< 0.62$ ), are highly connected with many nodes in their own modules. (iv) Network hubs (named super-generalists) have both high  $Z_i$  ( $> 2.5$ ) and  $P_i$  ( $> 0.62$ ) values. Generalists

---

(connectors, module hubs) and network hubs are the keystone microorganisms, which are critical in maintaining the network stability and microbial community structure and function [17].

### **Isolating nematicidal microorganisms from rhizospheric soils**

Ten grams of the NS rhizospheric soil was mixed with 90 ml of sterile water, serially diluted with sterile distilled water, then spread TSB agar plates (tryptone 1.5%, soy peptone 0.5%, NaCl 0.5%, and agar 1.5%, pH 7.2) and M9 agar plates (Na<sub>2</sub>HP<sub>4</sub> 6.74 g/L, KH<sub>2</sub>PO<sub>4</sub> 3g/L, NaCl 0.5 g/L, MgSO<sub>4</sub> 0.24 g/L, CaCl<sub>2</sub> 0.01 g/L, glucose 4 g/L, and agar 1.5%, pH 7.2) (dilution ratio 1, 1/2, 1/5, 1/10, 1/50, 1/100) and incubated at 37 °C for 48 h to isolate bacteria. Dilution was spread PDA plates (potato 20%, glucose 2%, agar 1.5%) and incubated at 28 °C for 5 days to isolate fungi. The bacteria and fungi with different morphology were isolated from plates for screening for nematicidal activity. We got 144 bacterial isolates and 20 fungal isolates.

For nematicidal activity assay, the isolated bacteria were incubated in LB medium at 37 °C and 200 rpm for 48 h, and the isolated fungi were incubated in PD medium at 28 °C and 200 rpm for 5 days. After incubation, the broth was centrifuged at 8,000 × g for 5 min for collecting supernatant. 200 µl of supernatant and ~ 50 numbers of *M. incognita* J2 juveniles were mixed together and then added to each well in 96-well plates. After incubation at 20 °C for 24 h, the mortality of nematode was counted. The nematodes without detectable movement were judged as dead. Tests were done in triplicates. LB or PD medium was added in the assay as control. Nematicidal activity was calculated as following formula [47]. Nematicidal activity =  $(A-B) / (1-B) \times 100\%$

Where A is mortality of nematode in the treatment group. B is mortality of nematode in the control group.

Nematicidal bacterial and fungal isolates were identified based on 16S rRNA or 18S rRNA genes sequence. Briefly, the bacterial or fungal genomic DNA was extracted using the DNA isolation kit (Tiangen Biotech CO., LTD, Beijing, China), then the bacterial 16S rRNA gene or fungal 18S rRNA gene was amplified, respectively [48, 49]. Sequence analysis and alignment were performed with Megalign and Clustal W, and the phylogenetic tree was

---

constructed using the software MrBayes 3.27 [50], then visualized with iTOL [51].

Protease activity was detected using the clearing zones method [52]. Single bacterial or fungal colony was incubated on the surface of selection medium (nonfat dried skimmed milk 5% (w/v), agar 2%) plates at 28 °C for 48 h, then the diameters of clearing zone (D2) and the bacterial or fungal colony (D1) were measured, respectively. The ratio of D2 to D1 was calculated as an indicator of protease activity. Chitinase activity was also detected using the clearing zones method on the surface of chitin medium (NaCl 1.5 g/L, KH<sub>2</sub>PO<sub>4</sub> 0.3 g/L, K<sub>2</sub>HPO<sub>4</sub> 0.7 g/L, FeSO<sub>4</sub>·H<sub>2</sub>O 0.02 g/L, MgSO<sub>4</sub> 0.5 g/L, agar 20 g/L, 250 ml of 2% self-made colloidal chitin, pH 7.0). Lipase activity was detected on the surface of selection medium (peptone 1%, yeast powder 0.5%, NaCl 1%, glyceryl triacetate 3%, bromcresol purple 0.004%, agar 2%) at 28 °C for 48 h. Diameters of the yellow clearing zone (D2) and the bacterial/fungal colony (D1) were measured, and the ratio of D2 to D1 was calculated as an indicator of lipase activity.

#### **Untargeted metabolomics analysis**

Metabolites of five NS samples with the least number of *M. incognita* and five RNI samples with the most number of *M. incognita* were measured by untargeted metabolomics analysis. For each soil sample, 100 mg soils were grounded with liquid nitrogen and the homogenate was suspended with chilled methanol and 0.1% formic acid. The samples were incubated on ice for 5 min and then were centrifuged at 15,000 g and 4 °C for 5 min. The supernatant was diluted to final concentration containing 60% methanol by LC-MS grade water. The samples were filtrated with 0.22 µm filter and then centrifuged at 15,000 g and 4 °C for 10 min. Finally, the filtrate was injected into the LC-MS/MS system. LC-MS/MS analyses were performed using a Vanquish UHPLC system (Thermo Fisher, Waltham City, Massachusetts) coupled with an Orbitrap Q Exactive HF-X mass spectrometer (Thermo Fisher, Waltham City, Massachusetts). Samples were injected onto an Hyperil Gold column (100 × 2.1 mm, 1.9 µm, Thermo Fisher Scientific Inc.) at a flow rate of 0.2 ml/min. The eluents for the positive polarity mode were eluent A1 (0.1% formic acid in water) and eluent B (methanol). The eluents for the negative polarity mode were eluent A2 (5 mM ammonium acetate, pH 9.0) and eluent B (methanol). The solvent gradient was set as follows: 2% B, 1.5 min; 2-100% B, 12.0

---

min; 100% B, 14.0 min; 100-2% B, 14.1 min; 2% B, 16 min. Q Exactive HF-X mass spectrometer was operated in positive/negative polarity mode with spray voltage of 3.2 kV, capillary temperature of 320 °C, sheath gas flow rate of 35 arb and aux gas flow rate of 10 arb. The raw data generated by LC-MS/MS were processed using the Compound Discoverer 3.0 (CD 3.0, Thermo Fisher, Waltham City, Massachusetts) to perform peak alignment, peak picking, and quantitation for each metabolite. Peak intensities were normalized to the total spectral intensity. The normalized data were used to predict the molecular formula based on additive ions, molecular ion peaks and fragment ions, and then peaks were matched with the mzCloud and ChemSpider databases to obtain the accurate qualitative and relative quantitative results. These compounds were annotated using the KEGG, HMDB, and Lipidmaps databases [53, 54].

Based on the relative quantification of different compounds, principal components analysis (PCA) and partial least squares discriminant analysis (PLS-DA) were performed at metaX [55]. Univariate analysis (*t*-test) was applied to calculate the statistical significance (*P*-value). The metabolites with  $VIP > 1$  and  $P\text{-value} < 0.05$  and fold change  $\geq 2$  or  $FC \leq 0.5$  were considered to be differential metabolites. Volcano plots were used to filter metabolites of interest which based on  $\text{Log}_2(FC)$  and  $-\text{log}_{10}(P\text{-value})$  of metabolites. For clustering heat maps, the data were normalized using z-scores of the intensity areas of differential metabolites and were plotted by Pheatmap package in R language. The functions of these metabolites and metabolic pathways were studied using the KEGG database.

*Bacillus amyloliquefaciens* W1 culture was isolated from NS soil samples and analyzed by untargeted metabolomics. W1 was incubated in 100 ml of M9 medium ( $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$  0.492 g/L,  $\text{CaCl}_2$  0.011 g/L,  $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$  17 g/L,  $\text{KH}_2\text{PO}_4$  3 g/L, NaCl 0.5 g/L,  $\text{NH}_4\text{Cl}$  1 g/L, glucose 4 g/L) at 37 °C and 180 rpm for 48 h. The culture was centrifuged at  $8,000 \times g$  for 15 min, and the supernatant was collected and mixed with equal volume of ethyl acetate, then centrifuged at  $8,000 \times g$  for 15 min. The extracted liquid was dried by rotary evaporation at 65 °C, then dissolved in 2 ml ethanol following filtration using 0.22  $\mu\text{m}$  membrane filters (Agela Technologies Inc., Delaware, USA). 100  $\mu\text{l}$  of sample was mixed with 400  $\mu\text{l}$  of 80% methanol and 0.1% formic acid by well vortexing, incubated on ice for 5 min, then centrifuged at  $15,000 \times g$  and 4 °C for 10 min. The supernatant was diluted to final

---

concentration containing 53% methanol by LC-MS grade water. After centrifugation at 15,000 × g and 4 °C for 10 min, the supernatant was collected for detection by the LC-MS/MS system and the data analysis was conducted as the above mentioned methods. The supernatant was also detected by gas chromatography (Agilent 7890A, California, USA) using DB-FFAP column (30m × 320µm × 0.25µm, Agilent, California, USA) at flow rate of 1 ml/min.

Acetophenone was used as standard.

#### **Analysis of nematicidal activity of metabolites**

*M. incognita* was maintained on tomato roots. Galls were peeled off from root and placed in distilled water at 20 °C until the second-stage juveniles (J2) hatched after two days. J2 juveniles were used in the following nematicidal bioassays.

The nematicidal activity of metabolites were detected as previously described [47]. The nematicidal activities of different concentrations of acetophenone (0.412, 0.459, 0.515, 0.589, 0.687, 0.824, 1.03, and 2.06 µg/ml) were tested against *M. incognita* J2 juveniles for 24 h, respectively. The experiment was repeated in triplicates. LC<sub>50</sub> and LC<sub>90</sub> values of acetophenone were calculated. Assay for nematode avoidance of acetophenone and *Bacillus* isolates was conducted according to [56]. Water-agar medium (1.5% agar) was prepared in 10-cm square plates. Six sections (0.75 cm × 10 cm) were divided around the middle of the plates. A, B, C section are on one side, D, E, F section are on the other side. Two or four µL acetophenone were dropped on one side of the bottom of plate (F section), and 2 or 4 µL ddH<sub>2</sub>O were dropped on the other side of the bottom of plate (A section) as a control. For detecting *Bacillus* avoidance activity, each *Bacillus* isolate was incubated in LB liquid medium at 37 °C and 180 rpm for 48 h. The culture was centrifuged at 8,000 × g for 15 min, and the supernatant was collected. 100 µL supernatant were drooped on F section, and 100 µL ddH<sub>2</sub>O were dropped on A section. Approximately 50 larvae of *M. incognita* were transferred to the middle of plate. After 2 h, the nematodes in each section were counted and avoidance index was calculated. Influenced by repellent, *M. incognita* generally gathered at the A section. The avoidance index was calculated as following formula [56]. The range of avoidance index is between 0 and 1.

$$\text{Avoidance index} = [(A + B) - (E + F)] / N$$

where A, B, E, and F were the number of nematodes in the corresponding section. N was the

---

total number of nematodes in the plate.

### **Statistics analysis**

Differences in soil properties, compounds, microbial or gene abundances and network indexes between the NS and RNI soils were compared by a least-significant-difference (LSD) test in SPSS 20 program. Correlations between metabolite and nematode number was analyzed by Pearson correlation analysis with SPSS 20 software [22].

### **Results**

#### **Soil properties were different between the NS and RNI soils**

It was found that the number of *M. incognita* was very different between the non-infected tobacco rhizosphere soil (NS) and the root-knot nematode infected tobacco rhizosphere soil (RNI). The average number of *M. incognita* in NS was 0.98 nematodes per 100 mg soil; whereas the average number of *M. incognita* in RNI was 1091.76 nematodes per 100 mg soil (table S2). The number of *M. incognita* in RNI was significantly ( $P < 0.05$ ) higher than that in NS.

Soil properties were different between NS and RNI. The activities of invertase and catalase, pH, CEC, Fe, Ca, SWC, STP, SCMC and MWD were all significantly ( $P < 0.05$ ) higher in the NS soils compared to the RNI soils. Conversely, the phosphatase activity, AK and BD were all significantly ( $P < 0.05$ ) higher in the RNI soils than the NS soils. Catalase activity, CEC, Fe, Ca, STP and MWD were significantly ( $P < 0.05$ ) negatively correlated with the number of *M. incognita*, while AP and BD were significantly positively correlated with the number of *M. incognita* (table S3).

#### **Taxonomic diversity of rhizosphere microbiota and network structure**

The taxonomic diversity and functional potential of the rhizosphere microbiota were investigated in the tobacco maturation period. After metagenome sequencing and bioinformatics analysis (tables S4 and S5), taxonomic assignment of the microbial cell fraction from the tobacco rhizosphere showed that 85.7%, 0.2%, 2.1%, and 0.3% of the sequence reads corresponded to the domains Bacteria, Eukaryota, Archaea and virus, respectively (fig. S1A). Principal Co-ordinates Analysis showed significant difference (Anosim,  $P < 0.01$ ) between the microbial community structure and composition of the NS

---

and the RNI rhizosphere microbiota (fig. S2A).

*Proteobacteria* (35.8% - 41.4%), *Acidobacteria* (11.9% - 15%), *Actinobacteria* (5.9% - 8.0%) and *Gemmatimonadetes* (5.9% - 7.1%) dominated the bacterial community (fig. S2B). At the class level, top 10 were *Alphaproteobacteria*, *Gammaproteobacteria*, *Betaproteobacteria*, *Actinobacteria*, *Acidobacteriia*, *Deltaproteobacteria*, *Ktedonobacteria*, *Gemmatimonadetes*, *Chitinophagia* and *Nitrospira* (fig. S2C). At the family level, top 10 were *Sphingomonadaceae*, *Rhodanobacteraceae*, *Acidobacteriaceae*, *Gemmatimonadaceae*, *Bradyrhizobiaceae*, *Burkholderiaceae*, *Comamonadaceae*, *Chitinophagaceae*, *Ktedonobacteraceae* and *Streptomycetaceae* (fig. S2D). 88 species all became enriched in the NS soils comparing with the RNI soils (fig. 1A). *Aspergillus* belonging to *Ascomycota* was also enriched in the NS soils. Relative abundance of these enriched bacteria were positively correlated with the CEC, MWD, Ca and Fe contents, while were negatively correlated with the AK content and the number of *M. incognita* ( $P < 0.05$ ) (fig. 1B). We totally got 144 bacterial isolates and 20 fungal isolates from the NS rhizosphere soils, then these isolates were detected for their nematocidal activity respectively. Among them, 10 bacterial isolates showed low nematocidal activity (mortality < 8%) (table S6). 62 bacterial isolates and 2 fungal isolates showed high nematocidal activity (fig. 2). Among these bacteria, 47 isolates showed a mortality more than 80% towards *M. incognita*.

Co-occurrence network analysis revealed that network complexity in RNI soils was decreased compared to NS soils (fig. S3). A higher average degree (avgK) means a more complex network, and a smaller average path distance (GD) means all nodes in the network are closer (table 1, table 2). NS fungal and bacterial networks were more complex and their nodes were closer than the RNI fungal and bacterial networks. More positive edges were found in the NS networks comparing with the RNI networks. Overall, the microbial network of NS was composed of more positively interacted microbes to form a more clustered topology structure than the RNI networks.

Most of nodes in the two bacterial networks belonged to 29 bacterial phyla. Among them, *Actinobacteria*, *Bacteroidetes*, *Firmicutes* and *Proteobacteria* were more dominant than other phyla. Compared with the NS bacterial network, the relative proportions of *Actinobacteria*, *Cyanobacteria*, *Thaumarchaeota* and *Verrucomicrobia* were all decreased, whereas the

---

relative abundances of *Firmicutes*, *Planctomycetes* and *Proteobacteria* were all increased substantially in the RNI bacterial network. Only 24.4% (432/1771) of nodes were shared by these two bacterial networks, whereas most of nodes (689 for the NS bacterial network and 650 for the RNI bacterial network) were unshared and specific in their own networks (fig. 3A). We further analyzed the 403 nodes belonging to *Proteobacteria* (fig. S4A). 124 nodes belonging to *Proteobacteria* were enriched in the non-infected soils compared with the nematode infected soils.

Most of nodes in the two fungal networks belonged to 8 fungal phyla. Among them, *Ascomycota*, *Basidiomycota* and *Mucoromycota* were more dominant than other phyla. Compared with the NS fungal network, the relative proportions of *Basidiomycota*, *Chytridiomycota* and *Zoopagomycota* were all decreased, whereas the relative abundances of *Ascomycota* and *Microsporidia* were both increased in the RNI fungal network. 59.5% of nodes were shared, and 83 nodes (38 nodes in the NS network and 45 nodes in the RNI network) were unshared by these two fungal networks (fig. 3B). We further analyzed the nodes belonging to *Ascomycota* and *Basidiomycota*. Compared to the NS network, 27 nodes were enriched in the RNI fungal network (fig. S4B). Among them, 8 nodes belonged to *Ascomycota*, and 9 nodes belonged to *Basidiomycota*.

We further analyzed the keystone microorganisms in bacterial networks (fig. 4A). Seven module hubs and three connectors (generalists) were found in the NS bacterial network (table 3). Among these keystone microorganisms, we have isolated four strains belonged to *Acinetobacter* and one strain belonged to *Streptomyces* which all showed high nematicidal activity (mortality  $\geq 80\%$ ) against *M. incognita* (fig. 2A). Twelve module hubs and one connector (generalists) were found in the RNI bacterial network (table 3). One module hub and two connectors were found in the NS fungal network (fig. 4B). Three module hubs and two connectors were found in the RNI fungal network. Keystone microorganisms in the RNI networks were different from those in the NS networks. These keystone microorganisms played important roles in stabilizing microbial network, and different keystone microorganisms indicated different microbial networks in these two different soils.

### **Functional diversity of the rhizosphere microbiota**

---

Of the genes retrieved from the metagenome data, 58 to 61% were assigned to a known function (fig. S1B, C). Among 979,916 genes with associated functions, 3,641 genes and 1611 functions were significantly ( $P < 0.05$ ) enriched in the rhizosphere microbiota of the non-infected soils. These genes belonged mainly to the pathways classified as “Energy production and conversion”, “Amino acid transport and metabolism”, “Signal transduction mechanisms”, “Secondary metabolites biosynthesis, transport and catabolism” and “Carbohydrate transport and metabolism”.

Co-occurrence function network analysis revealed the average path distance (GD) of the NS function network was smaller than the RNI function network, which means that all the nodes in the NS network are closer than those in the RNI network. Average clustering coefficient (avgCC) and modularity of the NS network were higher than the RNI network, indicating the NS function network has more stability and resilience than the RNI network. AvgK of the RNI function network was higher than the NS network (table 4). For the NS functional network, 291 gene nodes (46.1%) were overrepresented in the NS condition, 65 gene nodes (10.3%) were overrepresented in the RNI condition (fig. S5A). Among those nodes enriched in the NS condition, 9 nodes belonged to Category Q (secondary metabolites biosynthesis, transport, and catabolism). One of them, COG2761 (dithiol-disulfide isomerase), is involved in polyketide biosynthesis. 6 nodes belonged to Category O (posttranslational modification, protein turnover, chaperones) (table S7). They were increased in the NS condition. In the RNI functional network, 256 nodes (36.9%) were overrepresented in the NS condition, 60 nodes (8.6%) were overrepresented in the RNI condition (fig. S5B). Among those nodes enriched in the RNI condition, COG3843 (virD2) and COG2948 (virB10) are involved in Type IV secretory pathway and belong to Category U (intracellular trafficking, secretion, and vesicular transport). Some gram-negative bacterial pathogens use Type IV secretion system that can translocate DNA or proteins into their eukaryotic host cells to promote infection.

For more detailed resolution of the specific functions associated with cluster of orthologous groups (COG) Category Q and Category G (carbohydrate transport and metabolism), we searched for carbohydrate-active enzymes (CAZymes) and secondary metabolite bio-synthetic gene clusters within the metagenome sequences. Among Category Q,

---

41 genes were significantly overrepresented ( $P < 0.05$ ) in the NS microbiota, which are all involved in polyketide biosynthesis (fig. S6A). Polyketides have been shown with antimicrobial and nematocidal activity. Among Category G, 6 genes were significantly overrepresented in the NS microbiota, including chitinase and collagenase (fig. S6B). We also found that proteases and lipases genes are enriched in the NS microbiota (fig. S6C, D). We detected the enzyme activity of the nematocidal microbes isolated from the NS rhizosphere soils. 20 nematocidal bacteria and one fungus showed protease activity. 3 nematocidal bacteria showed chitinase activity, and 16 nematocidal bacteria showed lipase activity (table 5).

Analysis of secondary metabolites revealed a total of 6866 biosynthetic gene clusters (BGCs) in the NS microbiota. Among them, BGCs associated with the biosynthesis of terpene, nonribosomal peptides (NRPS), arylpolyene, bacteriocin, betalactone and polyketides were more dominant than other BGCs. Genes abundances of 19 BGCs in the NS microbiota were higher than the RNI microbiota (table S8).

Co-occurrence network analysis of BGCs genes revealed rhizomide, anabaenopeptin, icosalide, TP1161, APE Vf, lysobactin, nosiheptide, malleobactin, flexirubin and bicornutin had high abundances in two groups (fig. S5C, D). Altogether, abundance of 49 BGCs in the NS microbiota were higher than in RNI microbiota. APE Vf, anabaenopeptin, bicornutin, enduracidin, glidobactin, icosalide, lysobactin, myxochromide and rhizomide were enriched in the NS microbiota as compared to the RNI microbiota (fig. S7). Abundances of APE Vf and bicornutin in the NS microbiota were significantly ( $P < 0.05$ ) higher than in the RNI microbiota.

### **Untargeted Metabolomics analysis of the rhizosphere microbiota**

PCA analysis showed that metabolites composition of the NS rhizosphere microbiota were different from the RNI rhizosphere microbiota (fig. S8A, B). Partial least squares discrimination analysis (PLS-DA) showed the RNI samples were well separated from the NS samples (fig. 5A, B). Different metabolites analysis showed 102 metabolites were significantly different between the RNI and NS rhizosphere microbiota (fig. 5C, D; fig. S9A, B). 67 metabolites were overrepresented in the RNI samples compared to the NS samples, such as alternariol. 35 metabolites were overrepresented in the NS samples comparing the

---

RNI samples, such as acetophenone (log2FoldChange = 1.083,  $P < 0.05$ ) and indole-3-acetic acid (log2FoldChange = 1.583,  $P < 0.05$ ) (fig. 5 E, F, G). We detected nematicidal activity of some potential compounds and found acetophenone had the highest nematicidal activity, while other compounds (e.g. aminolevulinic acid, indole-3-acetic acid, benzene, thiophanate methyl, carbendazim) show low nematicidal activity.

We further analyzed the correlation between the number of *M. incognita* and the metabolites of rhizosphere microbiota. A total of 20 metabolites were significantly ( $P < 0.05$ ) negatively correlated with the number of *M. incognita* (table S9), such as acetophenone. Acetophenone showed high nematicidal activity. After treatment of *M. incognita* J2 juveniles for 24 h, LC<sub>50</sub> and LC<sub>90</sub> value of acetophenone was 0.66 µg/ml and 1.01 µg/ml, respectively (fig. S10A). Indole-3-acetic acid showed a low nematicidal activity, 2.95% mortality of *M. incognita* was observed when indole-3-acetic acid was used at 1 mg/ml for treatment of nematode for 24 h. Other metabolites showed no nematicidal activity against *M. incognita* (data not shown). Besides, acetophenone can also repel *M. incognita*. The avoidance index of acetophenone was 0.11 and 0.18 for the dose of 2 and 4 µl, respectively (fig. S10B). This result indicated that acetophenone could repel nematode, its avoidance activity is moderate. We constructed the network including acetophenone, the microbes enriched in the NS microbiota and the keystone microorganisms in the NS network (fig. 6). *Bacillus* sp. positively interacted with acetophenone. *Rhizobium* negatively interacted with acetophenone. *Pantoea* and *Pseudomonas* positively interacted with *Bacillus*. We isolated 19 *Bacillus* stains and found that these *Bacillus* isolates showed different avoidance activity towards *M. incognita* (fig. S10C). Among them, *B. amyloliquefaciens* W1 showed the highest avoidance activity against *M. incognita*. We further detected metabolites in W1 liquid culture, and found W1 could produce acetophenone (table S10). Several other metabolites involved in the phenylalanine metabolism pathway were also detected in the W1 culture, such as 2-phenylacetamide, 2-hydroxyphenylacetic acid, phenylpyruvic acid and phenylacetaldehyde (table S10, fig. S10D, E). Ten genes involved in phenylalanine metabolism pathway could be found in our metagenomic dataset. By gas chromatography, we could detect the acetophenone produced by *B. amyloliquefaciens* W1 (fig. S10F, G).

---

## Discussion

Soil properties of the RNI soils were very different from NS soils. Catalase activity, pH, STP, MWD, Ca and Fe were all negatively correlated to the number of *M. incognita* significantly. Changes of these soil properties possibly are related to the aggravated root-knot nematode disease in the RNI soils. The lower pH value and less calcium and iron contents in the RNI soils may have aggravated the root-knot nematode disease by improving nematode eggs hatch, increasing density of the second instar larvae, and reducing plant's tolerance to pathogen infection (e.g. *M. incognita*) [57, 59-60]. Total porosity and mean weight diameter of the RNI soils were lower than NS soils. The compacted RNI soils may have reduced microbial abundance and activity, persistently altered microbiota structure and stunted crop growth [61, 62]. Actually, invertase and catalase activities in the RNI soils were lower than NS soils, suggesting that microbial activity in the RNI soils is lower than NS soils.

Microorganisms can attack and kill root-knot nematodes by diverse processes such as capturing, parasitizing, and producing toxins and enzymes. Previous reports have shown that nematicidal antagonists such as *Comamonas*, *Paenibacillus*, *Pseudomonas*, *Variovorax*, *Virgibacillus* and *Achromobacter xylosoxidans*, could produce nematicidal volatile organic compounds against *Meloidogyne* spp. [9, 63-65]. *Aspergillus*, *Lactobacillus brevis* WiKim0069 and *Lysobacter capsici* YS1215 showed strong nematicidal activity and could inhibit the egg hatch of *Meloidogyne* spp., and their nematicidal activities were derived mainly from organic acids [11, 66, 67]. Many soil microorganisms such as *Acinetobacter* spp., *Microbacterium* sp., *Enterobacter* spp., *Pantoea agglomerans*, *Pseudomonas putida* and *Sinorhizobium fredii* Sneb183, could induce systemic resistance in plants resulting in decreased the root-knot nematodes infestation [68-71]. *Alcaligenes faecalis* ZD02 showed toxicity against *M. incognita* by producing an extracellular serine protease, which can damage the intestines of nematode [52]. *Enterobacter asburiae* HK169 displayed promising nematicidal activity against *M. incognita*, and its nematicidal virulence factor was also determined as the proteolytic enzymes [72]. Rhizobacteria such as *Burkholderia gladioli*, *Delftia acidovorans* HK171, *Rhizobium etli* G12, *Pseudomonas* spp. and *Bacillus* spp. could significantly reduce the hatching of eggs, the numbers of second stage juvenile of *Meloidogyne* and root galls [10, 73-75]. *Escherichia coli* E2348/69 can secrete small

---

molecules including indole, indole-3-carboxaldehyde and indole-3-acetic acid, as virulence factors to kill nematode [76]. *Pseudoxanthomonas japonensis* ZKB-2 produced 3 - methoxycyclobutane - 1, 2 - dione to kill *M. incognita* [77]. Many nematocidal antibiotics and metabolites (e.g. jietacin, polyketides, lactones) were produced by *Streptomyces* spp. to reduce root galls and inhibit egg hatching [78-80]. Metagenomics analysis showed abundances of these nematocidal microorganisms in the NS soils were higher than in RNI soils, and from the NS soils we got 62 bacterial isolates and two fungal isolates belonging to these genera with high nematocidal activity. These nematocidal isolates may play key roles in controlling root-knot nematode. We found that abundances of 124 nodes belonging to *Proteobacteria* in the NS soils were higher than those in the RNI soils. Among them, *Pseudomonas fluorescens* can inhibit egg hatching and kill *M. incognita* juveniles [81]. The bacterial community richness and diversity in root-knot nematode diseased soils was significantly different from that in healthy soils [82]. Some studies have characterized the endophyte and rhizosphere microbial community and core microbiota in root-knot nematode infected plants [6-8, 83]. Our results are consistent to some findings in these studies: (1) the rhizosphere and endophyte microbial communities of plant root (especially Solanaceae) are affected by nemato-pathogenesis [7, 8, 83]; (2) potential biological control microorganisms such as *Arthrobacter*, *Bacillus*, *Lysobacter*, and *Pseudomonas* showed the large proportion in non-infested soil and the farm with low population of nematode [6, 8]; (3) In rhizosphere and endophyte community, the most abundant phylum is *Proteobacteria* [83]. In our study, *Proteobacteria* was also the most abundant phylum. We found that nematocidal microorganisms such as *Bacillus* and *Lysobacter* were enriched in the NS soils and isolates of *Bacillus*, *Lysobacter* and *Pseudomonas* showed nematocidal activity [6].

Higher average connectivity (*i.e.* average links per node in the network) means a more complex network. A smaller average geodesic distance means all the nodes in the network are closer [16, 41]. We found the NS fungal and bacterial networks were more complex and their nodes were closer than the RNI networks. The higher complexity of the NS networks indicated the NS rhizosphere microbiota has a greater potential for interactions and niche-sharing. Rhizosphere microorganisms are very important for protecting plants against pathogens infection. Highly connected networks, like those in the NS samples, are more

---

effective for the rhizosphere microbiota to defense pathogens invasion [15, 44]. On the contrary, the interactions or niche sharing are less pronounced in the RNI networks relative to the NS networks, indicating the ecological networks have shifted in the RNI soils compared to the NS soils. Changes in environmental properties (e.g. pH, hydrological characteristics, etc) can alter ecological networks [84, 85]. Resource and food availability are important drivers of network structures [86], so the increased complexity of microbial networks was likely due to the increased amount of carbon input in soil [16]. Positive interactions indicate the abundances of these species or genes changed following the same trend across different soil samples. Negative interactions indicate the abundances of those species or genes changed following the opposite trend in different soil samples [43, 44]. Interestingly, more positive edges were found in the NS networks than the RNI networks, indicating the potential for extensive mutualistic interactions among microorganisms assembles in the NS rhizosphere [87]. Negative co-occurrence patterns (co-exclusion) predominated in the RNI networks, suggesting that root-knot nematode infection triggered microbial competition and antagonistic interactions in the RNI soils [88]. Compared to the NS network, 27 nodes belonging to *Ascomycota* and *Basidiomycota* were enriched in the RNI fungal network. Among them, *Rhizophagus irregularis*, an arbuscular mycorrhizal fungi, can improve phosphorus uptake and water use efficiency in plants [89]. *Phanerochaete chrysosporium* is a decomposer in carbon cycle (such as lignin degradation) [90]. These results suggested that the composition of microbial networks had varied considerably in the RNI soils when compared with the NS soils. It seems that the microbial network topology structure and composition were substantially shaped in the root-knot nematode infection soils in this mountain ecosystem.

Ten keystone microorganisms were found in the NS bacterial network. These keystone species are critical in maintaining the NS network structure and function, and ecosystem stability [44, 46]. As previously reported by other studies, *Terriglobus saanensis* and *Granulicella mallensis* both belonging to *Acidobacteria*, can hydrolyze diverse carbohydrates including complex polysaccharides [91]. Xylanosome produced by *Oerskovia turbata* has been shown with powerful degradative ability towards xylan and hemicellulose [92]. *Streptomyces thermoautotrophicus* is able to fix N<sub>2</sub> [93]. We speculated that these keystone microorganisms could participate in carbon and nitrogen cycles, likely supplying carbon and

---

nitrogen resources for other microorganisms, in the NS networks. Occurrence of these keystone species may be indicative of the influence of soil nutrients and root exudates on microbial co-occurrence relationships in soil. The actinomycetes such as *Rubrobacter radiotolerans* produces two dimeric indole derivatives with acetylcholinesterase inhibitory activity [94]. The acetylcholinesterase inhibitors, *e.g.* oxamyl and avermectin, cause initial hyperactivity of nematode juveniles followed by a progressive decline in movement and then a final loss of activity [95]. Thereby, as keystone species, *R. radiotolerans* may play important roles in control of root-knot nematode disease. The keystone taxa shifted as conditions (land uses, pathogen infection) changed [44, 96]. Here, the keystone species of the RNI networks were different from those of the NS networks, indicating the keystone species of microbial network were altered in the RNI soil.

Microorganisms can kill nematodes by producing extracellular enzymes, including serine protease, metalloproteinase, collagenase, lipase and chitinase. These enzymes can degrade the nematode cuticle and eggshell [97, 98]. We found the abundances of chitinase, collagenase, lipase and protease genes in the NS microbiota were significantly higher than in RNI microbiota. These extracellular enzymes might act as important nematocidal virulence factors in the NS rhizosphere microbiota.

Polyketides, an important class of biologically active compounds, possess potent nematocidal activity [80]. Among nodes enriched in the NS function network, COG2761 (dithiol-disulfide isomerase) is involved in polyketide biosynthesis. Abundances of 41 Category Q genes involving in polyketide biosynthesis in the NS microbiota were higher than in RNI microbiota. The toxic secondary metabolites such as anabaenopeptins could reduce reproduction of the model nematode *Caenorhabditis elegans* and shorten its lifespan [99]. The secondary metabolites (*e.g.* bicornutin, enduracidin, glidobactins, icosalide, lysobactin, rhizomide) display antimicrobial activity, indicating these antibiotics are possible as nematocidal antibiotics to protect plants from nematode infection.

We found the metabolites of the NS microbiota were different from those of the RNI microbiota based on the untargeted metabolomics analysis. Alternariol, a kind of mycotoxin produced by the phytopathogenic fungi *Alternaria*, acts as a virulence and colonization factor during their infection of plants [100]. Interestingly, alternariol was enriched in the RNI

---

microbiota, indicating the tobacco plants were also infected by *Alternaria* in the RNI soils that was consistent with our investigations (data unshown). The co-infection of *Alternaria* and root-knot nematode may aggravate the death of plants.

Acetophenone was overrepresented in the NS samples. Two derivatives of acetophenone, 4-nitroacetophenone and 4-iodoacetophenone, were previously reported with nematicidal activity against *M. incognita*, with EC<sub>50</sub> values of 12 and 15 mg/ L respectively [101]. In this study, we found that acetophenone was also very effective for killing nematodes. Not only that, its nematicidal activity (LC<sub>50</sub> = 0.66 mg/L) was much higher than its derivatives, how acetophenone kill nematode is needed to be further investigated. The nematicidal activity of acetophenone was even higher than the popularly used commercial nematicide avermectin (LC<sub>50</sub> = 2.0 mg/L) [102]. Moreover, as nematodes resistance to avermectins is on the rise, alternatives or new products with enhanced potency and high activity are urgently needed [103, 104]. Considering that acetophenone has the advantages of simple structure, high efficiency and safety [105], it is a very potential nematicide candidate in the future. Besides for directly killing nematode, we found that acetophenone could strongly repel *M. incognita* as a repellent of root-knot nematode. Interestingly, acetophenone was also reported with strong repellency activity to some pests such as the kissing bug *Rhodnius prolixus* at a low dose (1%) [106], and as an anti-attractant for the pest western pine beetle, *Dendroctonus brevicomis* [107]. Addition to acetophenone, indole-3-acetic acid (IAA) was also overrepresented in the NS samples. It is known that IAA produced by microbes can promote plant growth. Thereby, overrepresent of IAA in the NS rhizosphere soils may be favorable for improving tobacco growth.

Co-occurrence network analysis found *Bacillus* positively interacted with acetophenone. *B. amyloliquefaciens* W1 culture could repel and kill 98.8% of *M. incognita* J2 juveniles and could produce acetophenone. This study firstly found the nematicidal activity of acetophenone produced by *B. amyloliquefaciens*. This new discovery has broken new ground in exploration of the nematicidal metabolites from microbes through combining multi-omics analysis as well as culture-dependent technology. Interestingly, several intermediary metabolites in phenylalanine metabolism pathway were also found in the W1 culture. We supposed that acetophenone may be one of the metabolites in phenylalanine metabolism, and

---

the pathway of acetophenone production was needed to be investigated. We found that *Pantoea* and *Pseudomonas* positively interacted with *Bacillus*, thereby these nematicidal bacteria possibly cooperate together to kill nematodes by multiple nematicidal mechanisms. For example, a previous study demonstrated that a native consortium of five bacterial isolates could better protect their host plant from a fungal sudden-wilt disease than a single isolate [108]. Therefore, a consortium of several different nematicidal isolates was possibly more effective for controlling *M. incognita* than a single isolate, but this point needed to be proven in our future research.

### **Conclusion**

We used multi-omic methods to screen new nematicidal compounds and microorganisms. Rhizosphere microbial composition, function and network were altered accompanying with root-knot nematode infection. Abundances of nematicidal metabolites, antibiotics and extracellular enzymes' genes in the NS microbiota were higher than the RNI microbiota. Nematicidal microbes were also enriched in the NS rhizosphere microbiota and could produce extracellular nematicidal enzymes. Based on metabolomics analysis, acetophenone was identified as a new nematicidal compound with high efficiency to kill nematodes. *Bacillus amyloliquefacens* W1 could produce acetophenone and showed high nematicidal and avoidance activity against *M. incognita*.

### **Declarations**

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

Not applicable.

#### **Consent for publication**

Not applicable.

#### **Availability of data and material**

All sequencing data has been submitted to the NCBI repository and can be accessed via the following accession numbers: whole-genome shotgun sequencing PRJNA657468, PRJNA657384, PRJNA657444.

---

## Competing interests

The authors declare that they have no competing interests.

## Funding

This work was supported by the National Natural Science Foundation of China (31870030).

## Authors' contributions

Conceptualization: X.Z., and G.Q. Investigation: C.L., W.Y., and S.R. Data curation: X.Z., and C.L. Writing-review and editing: X.Z., and G.Q. Funding acquisition: G.Q. The author(s) read and approved the final manuscript.

## Acknowledgements

Thank you to Dr. Wang R. and Tan J for the assistance with soil samples collecting.

## References

1. Davies LJ, Elling AA. Resistance genes against plant-parasitic nematodes: a durable control strategy? *Nematology*. 2015, 17: 249-263.
2. Parsons J, Matthews W, Iorizzo M, Roberts P, Simon P. *Meloidogyne incognita* nematode resistance QTL in carrot. *Mol Breeding*. 2015, 35: 114.
3. German E, Mauricio T, Eduardo S, Mario A. Chemical control of *Meloidogyne* spp. in grapevines (*Vitisvinifera*). *J Appl Biosciences*. 2019, 136: 13896-13908.
4. Huang Y, Ma L, Fang DH, Xi JQ, Zhu ML, Mo MH, et al. Isolation and characterisation of rhizosphere bacteria active against *Meloidogyne incognita*, *Phytophthora nicotianae* and the root knot-black shank complex in tobacco. *Pest Manag Sci*. 2015, 71: 415-22.
5. Du B, Xu Y, Dong H, Li Y, Wang J. *Phanerochaete chrysosporium* strain B-22, a nematophagous fungus parasitizing *Meloidogyne incognita*. *PLoS One*. 2020, 15: e0216688.
6. Castillo JD, Vivanco JM, Manter DK. Bacterial microbiome and nematode occurrence in different potato agricultural soils. *Microb Ecol*. 2017, 74: 888-900.
7. Tian B, Cao Y, Zhang K. Metagenomic insights into communities, functions of endophytes, and their associates with infection by root-knot nematode, *Meloidogyne incognita*, in tomato roots. *Sci Rep*. 2015, 5: 17087.

- 
8. Zhou D, Feng H, Schuelke T, De Santiago A, Zhang Q, Zhang J, et al. Rhizosphere microbiomes from root knot nematode non-infested plants suppress nematode infection. *Microb Ecol.* 2019, 78: 470-481.
  9. Cheng W, Yang J, Nie Q, Huang D, Yu C, Zheng L, et al. Volatile organic compounds from *Paenibacillus polymyxa* KM2501-1 control *Meloidogyne incognita* by multiple strategies. *Sci Rep.* 2017, 7: 16213.
  10. Turatto MF, Dourado FDS, Zilli JE, Botelho GR. Control potential of *Meloidogyne javanica* and *Ditylenchus* spp. using fluorescent *Pseudomonas* and *Bacillus* spp. *Braz J Microbiol.* 2018, 49: 54-58.
  11. Seo HJ, Park AR, Kim S, Yeon J, Yu NH, Ha S, et al. Biological control of root-knot nematodes by organic acid-producing *Lactobacillus brevis* WiKim0069 isolated from Kimchi. *Plant Pathol J.* 2019, 35: 662-673.
  12. Goswami BK, Pandey RK, Singh S, Rathour KS. Biopesticides: an ecofriendly, sustainable and cost effective approach for integrated disease and insect pest management of agricultural crops. In: Ahamad S., editor. *Plant Disease Management for Sustainable Agriculture*. Daya Publishing House; Delhi: 2009. pp. 292-298.
  13. Deketelaere S, Tyvaert L, França SC, Höfte M. Desirable traits of a good biocontrol agent against *Verticillium* wilt. *Front Microbiol.* 2017, 8: 1186.
  14. Chen S, Qi G, Luo T, Zhang H, Jiang Q, Wang R, et al. Continuous-cropping tobacco caused variance of chemical properties and structure of bacterial network in soils. *Land Degrad Dev.* 2018, 29: 4106-20.
  15. Qi G, Ma G, Chen S, Lin C, Zhao X. Microbial network and soil properties are changed in bacterial wilt-susceptible soil. *Appl Environ Microbiol.* 2019, 85: e00162-19.
  16. Zhou J, Deng Y, Luo F, He Z, Yang Y. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO<sub>2</sub>. *mBio.* 2011, 2: e00122-11.
  17. Faust K, Raes J. Microbial interactions: from networks to models. *Nat Rev Microbiol.* 2012, 10: 538-550.
  18. Zhou X, Wang JT, Wang WH, Tsui CK, Cai L. Changes in bacterial and fungal microbiomes associated with tomatoes of healthy and infected by *Fusarium oxysporum* f. sp. *lycopersici*. *Microb Ecol.* 2020.

- 
19. Grim CM, Luu GT, Sanchez LM. Staring into the void: demystifying microbial metabolomics. *FEMS Microbiol Lett.* 2019, 366: fnz135.
  20. Zhao YL, Ruan WB, Yu L, Zhang JY, Fu JM, Shain EB, et al. Combining maxRatio analysis with real-time PCR and its potential application for the prediction of *Meloidogyne incognita* in field samples. *J Nematol.* 2010, 42: 166-72.
  21. Bao SD. Soil and agricultural chemistry analysis (eds Jiang, R. F. et al.) (China Agriculture Press, ISBN 978-7-109-06644-1, 2013, pp 23-106.
  22. Wang R, Zhang H, Sun L, Qi G, Chen S, Zhao X. Microbial community composition is related to soil biological and chemical properties and bacterial wilt outbreak. *Sci Rep.* 2017, 7: 343.
  23. Bielanders CL, De Backer LW, Delvaux B. Particle density of volcanic soils as measured with a gas pycnometer. *Soil Sci Soc Am J.* 1990, 54: 822-826.
  24. Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, et al. Gut metagenome in European women with normal, impaired and diabetic glucose control. *Nature.* 2013, 498: 99-103.
  25. Li D, Liu C-M, Luo R, Sadakane K, Lam T-W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics.* 2015, 31: 1674-1676.
  26. Qin N, Yang F, Li A, Prifti E, Chen Y, Shao L, et al. Alterations of the human gut microbiome in liver cirrhosis. *Nature.* 2014, 513: 59-64.
  27. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience.* 2012, 1: 18.
  28. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li J, Sunagawa S, et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat Biotechnol.* 2014, 32: 822-828.
  29. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics.* 2012, 28: 3150-3152.
  30. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006, 22: 1658-9.
  31. Li J, Jia H, Cai X, Zhong H, Feng Q, Sunagawa S, et al. An integrated catalog of

- 
- reference genes in the human gut microbiome. *Nat Biotechnol.* 2014, 32: 834-841.
32. Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. *Nat Methods.* 2015, 12: 59-60.
33. White JR, Nagarajan N, Pop M. Statistical methods for detecting differentially abundant features in clinical metagenomic samples. *PLoS Comput Biol.* 2009, 5: e1000352.
34. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011, 12: 1-18.
35. Kanehisa M, Goto S, Sato Y, Kawashima M, Furumichi M, Tanabe M. Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res.* 2014, 42: D199-D205.
36. Powell S, Forslund K, Szklarczyk D, Trachana K, Roth A, Huerta-Cepas J, et al. eggNOG v4.0: nested orthology inference across 3686 organisms. *Nucleic Acids Res.* 2014, 42: D231-9.
37. Cantarel BL, Coutinho PM, Rancurel C, Bernard T, Lombard V, Henrissat B. The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. *Nucleic Acids Res.* 2009, 37: D233-238.
38. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 2012, 40: W445-51.
39. Weber T, Blin K, Duddela S, Krug D, Kim HU, Brucoleri RE, et al. antiSMASH 3.0-A comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 2015, 43: W237-W243.
40. Luo F, Zhong JX, Yang YF, Scheuermann RH, Zhou JZ. Application of random matrix theory to biological networks. *Phys Lett A.* 2006, 357: 420-423.
41. Deng Y, Jiang YH, Yang YF, He ZL, Luo F, Zhou JZ. Molecular ecological network analyses. *BMC Bioinformatics.* 2012, 13: 113.
42. Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with Cytoscape 3. *Current Protocols in Bioinformatics.* 2014, 47: 8.13.1-8.13.24.
43. Bascompte J, Stouffer DB. The assembly and disassembly of ecological networks. *Philos Trans R Soc Lond B Biol Sci.* 2009, 364: 1781-1787.
44. Lu L, Yin S, Liu X, Zhang W, Gu T, Shen Q, et al. Fungal networks in yield-invigorating

- 
- and -debilitating soils induced by prolonged potato monoculture. *Soil Biol Biochem.* 2013, 65: 186-194.
45. Guimera R, Nunes Amaral LA. Functional cartography of complex metabolic networks. *Nature.* 2005, 433: 895-900.
46. Olesen J, Bascompte J, Dupont Y, Jordano P. The modularity of pollination networks. *Proc Natl Acad Sci USA.* 2007, 104: 19891-19896.
47. Gao H, Qi G, Yin R, Zhang H, Li C, Zhao X. *Bacillus cereus* strain S2 shows high nematicidal activity against *Meloidogyne incognita* by producing sphingosine. *Sci Rep.* 2016, 6: 28756.
48. Borneman J, Hartin RJ. PCR primers that amplify fungal rRNA genes from environmental samples. *Appl Envir Microbiol.* 2000, 66: 4356-4360.
49. Larousse M, Rancurel C, Syska C, Palero F, Etienne C, Industri B, et al. Tomato root microbiota and *Phytophthora parasitica*-associated disease. *Microbiome.* 2017, 5: 56.
50. Ronquist F, Teslenko M, van der Mark P, Ayres DL, Darling A, Höhna S, et al. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol.* 2012, 61: 539-42.
51. Letunic I, Bork P. Interactive Tree Of Life v2: Online annotation and display of phylogenetic trees made easy. *Nucleic Acids Res.* 2011, 39: W475-W478.
52. Ju S, Lin J, Zheng J, Wang S, Zhou H, Sun M. *Alcaligenes faecalis* ZD02, a novel nematicidal bacterium with an extracellular serine protease virulence factor. *Appl Environ Microbiol.* 2016, 82: 2112-2120.
53. Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS, Shockcor J, et al. Global metabolic profiling procedures for urine using UPLC-MS. *Nature protocols.* 2010, 5: 1005-1018.
54. Kieffer DA, Piccolo BD, Vaziri ND, Liu S, Lau WL, Khazaeli M, et al. Resistant starch alters gut microbiome and metabolomics profiles concurrent with amelioration of chronic kidney disease in rats. *Am J Physiol Renal Physiol.* 2016, 310: F857-71.
55. Wen B, Mei Z, Zeng C, Liu S. metaX: a flexible and comprehensive software for processing metabolomics data. *BMC Bioinformatics.* 2017, 18: 183.
56. Luo T, Hou S, Yang L, Qi G, Zhao X. Nematodes avoid and are killed by *Bacillus*

- 
- mycoides*-produced styrene. J Invertebr Pathol. 2018, 159: 129-136.
57. Tian X, Zhao X, Mao Z, Xie B. Variation and dynamics of soil nematode communities in greenhouses with different continuous cropping periods. Horticult Plant J. 2020, 6: 301-312.
58. Jiang T, Zhan X, Xu Y, Zhou L, Zong L. Roles of calcium in stress-tolerance of plants and its ecological significance. Ying Yong Sheng Tai Xue Bao. 2005, 16: 971-6.
59. Springer YP, Hardcastle BA, Gilbert GS. Soilcalcium and plant disease in serpentine ecosystems: a test of the pathogen refuge hypothesis. Oecologia. 2007, 151: 10-21.
60. Mahender A, Swamy BPM, Anandan A, Ali J. Tolerance of iron-deficient and -toxic soil conditions in rice. Plants. 2019, 8: 31.
61. Shah AN, Tanveer M, Shahzad B, Yang G, Fahad S, Ali S, et al. Soil compaction effects on soil health and cropproductivity: an overview. Environ Sci Pollut Res Int. 2017, 24: 10056-10067.
62. Hartmann M, Niklaus PA, Zimmermann S, Schmutz S, Kremer J, Abarenkov K, et al. Resistance and resilience of the forest soil microbiome to logging-associated compaction. ISME J. 2014, 8: 226-44.
63. Xu YY, Lu H, Wang X, Zhang KQ, Li GH. Effect of volatile organic compounds from bacteria on nematodes. Chem Biodivers. 2015, 12: 1415-21.
64. Wolfgang A, Taffner J, Guimarães RA, Coyne D, Berg G. Novel strategies for soil-borne diseases: exploiting the microbiome and volatile-based mechanisms toward controlling *Meloidogyne*-based disease complexes. Front Microbiol. 2019, 10: 1296.
65. Huang D, Yu C, Shao Z, Cai M, Li G, Zheng L, et al. Identification and characterization of nematicidal volatile organic compounds from deep-sea *Virgibacillus dokdonensis* MCCC 1A00493. Molecules. 2020, 25. pii: E744.
66. Lee YS, Naning KW, Nguyen XH, Kim SB, Moon JH, Kim KY. Ovicidal activity of lactic acid produced by *Lysobacter capsici* YS1215 on eggs of root-knot nematode, *Meloidogyne incognita*. J Microbiol Biotechnol. 2014, 24: 1510-5.
67. Lee S, Lee K, Chun H, Ha S, Gwak H, Kim H, et al. Process development of oxalic acid production in submerged culture of *Aspergillus niger* F22 and its biocontrol efficacy against the root-knot nematode *Meloidogyne incognita*. Bioprocess Biosyst Eng. 2018, 41: 345-352.

- 
68. Munif A, Hallmann J, Sikora RA. Induced systemic resistance of selected endophytic bacteria against *Meloidogyne incognita* on tomato. Meded Rijksuniv Gent Fak Landbouwkde Toegep Biol Wet. 2001, 66: 663-9.
69. Tian F, Wang Y, Zhu X, Chen L, Duan Y. Effect of *Sinorhizobium fredii* strain Sneb183 on the biological control of soybean cyst nematode in soybean. J Basic Microbiol. 2014, 54: 1258-63.
70. Topalović O, Elhady A, Hallmann J, Richert-Pöggeler KR, Heuer H. Bacteria isolated from the cuticle of plant-parasitic nematodes attached to and antagonized the root-knot nematode *Meloidogyne hapla*. Sci Rep. 2019, 9: 11477.
71. Topalović O, Bredenbruch S, Schleker ASS, Heuer H. Microbes attaching to endoparasitic phytonematodes in soil trigger plant defense upon root penetration by the nematode. Front Plant Sci. 2020, 11: 138.
72. Oh M, Han JW, Lee C, Choi GJ, Kim H. Nematicidal and plant growth-promoting activity of *Enterobacter asburiae* HK169: genome analysis provides insight into its biological activities. J Microbiol Biotechnol. 2018, 28: 968-975.
73. Mahdy M, Hallmann J, Sikora RA. Influence of plant species on the biological control activity of the antagonistic rhizobacterium *Rhizobium etli* strain G12 toward the root-knot nematode, *Meloidogyne incognita*. Meded Rijksuniv Gent Fak Landbouwkde Toegep Biol Wet. 2001, 66: 655-62.
74. Han JW, Oh M, Choi GJ, Kim H. Genome sequence of *Delftia acidovorans* HK171, a nematicidal bacterium isolated from tomato roots. Genome Announc. 2017, 5. pii: e01746-16.
75. Khanna K, Sharma A, Ohri P, Bhardwaj R, Abd Allah EF, Hashem A, et al. Impact of plant growth promoting rhizobacteria in the orchestration of *Lycopersicon esculentum* mill. Resistance to plant parasitic nematodes: a metabolomic approach to evaluate defense responses under field conditions. Biomolecules. 2019, 9. pii: E676.
76. Bommarius B, Anyanful A, Izrayelit Y, Bhatt S, Cartwright E, Wang W, et al. A family of indoles regulate virulence and Shiga toxin production in pathogenic *E. coli*. PLoS One. 2013, 8: e54456.

- 
77. Hu Y, Li J, Li J, Zhang F, Wang J, Mo M, et al. Biocontrol efficacy of *Pseudoxanthomonas japonensis* against *Meloidogyne incognita* and its nematostatic metabolites. FEMS Microbiol Lett. 2019, 366.
78. Tsuzuki K, Yan FS, Otoguro K, Omura S. Synthesis and nematocidal activities of jietacin A and its analogs. J Antibiot (Tokyo). 1991, 44: 774-84.
79. Li JS, Zhang H, Zhang SY, Wang HY, Zhang J, Chen AL, et al. New macrocyclic lactones with acaricidal and nematocidal activities from a genetically engineered strain *Streptomyces bingchenggensis* BCJ60. J Asian Nat Prod Res. 2017, 19: 339-346.
80. Feng Y, Yu Z, Zhang S, Xue Z, Huang J, Zhang H, et al. Isolation and characterization of new 16-membered macrolides from the aveA3 gene replacement mutant strain *Streptomyces avermitilis* TM24 with acaricidal and nematocidal activities. J Agric Food Chem. 2019, 67: 4782-4792.
81. Siddiqui IA, Haas D, Heeb S. Extracellular protease of *Pseudomonas fluorescens* CHA0, a biocontrol factor with activity against the root-knot nematode *Meloidogyne incognita*. Appl Environ Microbiol. 2005, 71: 5646-9.
82. Huang K, Jiang Q, Liu L, Zhang S, Liu C, Chen H, et al. Exploring the key microbial changes in the rhizosphere that affect the occurrence of tobacco root-knot nematodes. AMB Express. 2020, 10: 72.
83. Yergaliyev TM, Alexander-Shani R, Dimerets H, Pivonia S, Bird DM, Rachmilevitch S, et al. Bacterial community structure dynamics in *Meloidogyne incognita*-infected roots and its role in worm-microbiome interactions. mSphere. 2020, 5: e00306-20.
84. Tylianakis JM, Tscharntke T, Lewis OT. Habitat modification alters the structure of tropical host-parasitoid food webs. Nature. 2007, 445: 202-205.
85. Barberan A, Bates ST, Casamayor EO, Fierer N. Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J. 2012, 6: 343-351.
86. Foster EA, Franks DW, Morrell LJ, Balcomb KC, Parsons KM, van Ginneken A, et al. Social network correlates of food availability in an endangered population of killer whales. *Orcinus orca* Anim Behav. 2012, 83: 731-736.
87. Ren D, Madsen JS, Sorensen SJ, Burmolle M. High prevalence of biofilm synergy among bacterial soil isolates in cocultures indicates bacterial interspecific cooperation. ISME J.

- 
- 2015, 9: 81-89.
88. Berry D, Widder S. Deciphering microbial interactions and detecting keystone species with co-occurrence networks. *Front Microbiol.* 2014, 5: 1-14.
89. Le Pioufle O, Ganoudi M, Calonne-Salmon M, Ben Dhaou F, Declerck S. *Rhizophagus irregularis* MUCL 41833 improves phosphorus uptake and water use efficiency in maize plants during recovery from drought stress. *Front Plant Sci.* 2019, 10: 897.
90. Hong CY, Ryu SH, Jeong H, Lee SS, Kim M, Choi IG. *Phanerochaete chrysosporium* multienzyme catabolic system for in vivo modification of synthetic lignin to succinic acid. *ACS Chem Biol.* 2017, 12: 1749-1759.
91. Rawat SR, Männistö MK, Bromberg Y, Häggblom MM. Comparative genomic and physiological analysis provides insights into the role of *Acidobacteria* in organic carbon utilization in Arctic tundra soils. *FEMS Microbiol Ecol.* 2012, 82: 341-55.
92. Dou TY, Liu WJ, Chen J. Isolation, substrate specificity, and subunit characterization of the xylanosomes produced by *Oerskovia turbata* JCM 3160. *Curr Microbiol.* 2020, 77: 924-930.
93. Ribbe M, Gadkari D, Meyer O. N<sub>2</sub> fixation by *Streptomyces thermoautotrophicus* involves a molybdenum-dinitrogenase and a manganese-superoxide oxidoreductase that couple N<sub>2</sub> reduction to the oxidation of superoxide produced from O<sub>2</sub> by a molybdenum-CO dehydrogenase. *J Biol Chem.* 1997, 272: 26627-33.
94. Li JL, Huang L, Liu J, Song Y, Gao J, Jung JH, et al. Acetylcholinesterase inhibitory dimeric indole derivatives from the marine actinomycetes *Rubrobacter radiotolerans*. *Fitoterapia.* 2015, 102: 203-7.
95. Wright DJ, Birtle AJ, Roberts IT. Triphasic locomotor response of a plant-parasitic nematode to avermectin: inhibition by the GABA antagonists bicuculline and picrotoxin. *Parasitology.* 1984, 88: 375-82.
96. Lupatini M, Suleiman A, Jacques R, Antonioli Z, Ferreira A, Kuramae EE, et al. Network topology reveal high connectance levels and few key microbial genera within soils. *Front Environ Sci.* 2014, 2: 1-11.
97. Yang J, Liang L, Li J, Zhang K. Nematicidal enzymes from microorganisms and their applications. *Appl Microbiol Biotechnol.* 2013, 97: 7081-95.

- 
98. Aballay E, Prodan S, Zamorano A, Castaneda-Alvarez C. Nematicidal effect of rhizobacteria on plant-parasitic nematodes associated with vineyards. *World J Microbiol Biotechnol.* 2017, 33: 131.
99. Lenz KA, Miller TR, Ma H. Anabaenopeptins and cyanopeptolins induce systemic toxicity effects in a model organism the nematode *Caenorhabditis elegans*. *Chemosphere.* 2019, 214: 60-69.
100. Wenderoth M, Garganese F, Schmidt-Heydt M, Soukup ST, Ippolito A, Sanzani SM, et al. Alternariol as virulence and colonization factor of *Alternaria alternata* during plant infection. *Mol Microbiol.* 2019, 112: 131-146.
101. Caboni P, Aissani N, Demurtas M, Ntalli N, Onnis V. Nematicidal activity of acetophenones and chalcones against *Meloidogyne incognita* and structure-activity considerations. *Pest Manag Sci.* 2016, 72: 125-30.
102. Yi C, Chen J, Wei C, Wu S, Wang S, Hu D, et al.  $\alpha$ -Haloacetophenone and analogues as potential antibacterial agents and nematicides. *Bioorg Med Chem Lett.* 2020, 30: 126814.
103. Liu KL, He L, Wang JJ, Zhao ZM. Advances in the research on the pest resistance to avermectins. *Chin Bull Entomol.* 2007, 44: 194-200.
104. Ghosh R, Andersen EC, Shapiro JA, Gerke JP, Kruglyak L. Natural variation in a chloride channel subunit confers avermectin resistance in *C. elegans*. *Science.* 2012, 335: 574-8.
105. Api AM, Belsito D, Botelho D, Bruze M, Burton Jr GA, Buschmann J, et al. RIFM fragrance ingredient safety assessment, acetophenone, CAS Registry Number 98-86-2. *Food Chem Toxicol.* 2018, 118 Suppl 1: S162-S169.
106. Franco TA, Xu P, Brito NF, Oliveira DS, Wen X, Moreira MF, et al. Reverse chemical ecology-based approach leading to the accidental discovery of repellents for *Rhodnius prolixus*, a vector of Chagas diseases refractory to DEET. *Insect Biochem Mol Biol.* 2018, 103: 46-52.
107. Erbilgin N, Gillette NE, Mori SR, Stein JD, Owen DR, Wood DL. Acetophenone as an anti-attractant for the western pine beetle, *Dendroctonus brevicomis* LeConte (Coleoptera: Scolytidae). *J Chem Ecol.* 2007, 33: 817-23.
108. Santhanam R, Menezes RC, Grabe V, Li D, Baldwin IT, Groten K. A suite of

complementary biocontrol traits allows a native consortium of root-associated bacteria to protect their host plant from a fungal sudden-wilt disease. *Mol Ecol.* 2019, 28: 1154-1169.

**Table 1 Topological properties of the co-occurrence fungal network analysis**

Network properties	NS	RNI
Number of nodes	160	167
Number of edges	1251	581
Positive edges	905	522
Negative edges	345	58
Modularity	0.242	0.443
Number of module	6	13
Average degree (avgK)	15.637	6.958
Average clustering coefficient (avgCC)	0.374	0.239
Average path distance (GD)	2.879	4.376

**Table 2 Topological properties of the co-occurrence bacterial network analysis**

Network properties	NS	RNI
Number of nodes	1121	1082
Number of edges	2947	2465
Positive edges	2737	2212
Negative edges	209	252
Modularity	0.744	0.786
Number of module	199	175
Average degree (avgK)	5.258	4.556
Average clustering coefficient (avgCC)	0.232	0.189
Average path distance (GD)	6.908	8.004

**Table 3 Keystone microorganisms in bacterial and fungal networks**

Condition	Species	Categories	Phylum	Module	Relative abundance
NS	<i>Acinetobacter gernerii</i>	module hub	<i>Proteobacteria</i>	NS6	0.0001
	<i>Granulicella mallensis</i>	module hub	<i>Acidobacteria</i>	NS1	0.0587
	<i>Luteimonas mephitis</i>	module hub	<i>Proteobacteria</i>	NS1	0.0484
	<i>Niveispirillum irakense</i>	module hub	<i>Proteobacteria</i>	NS2	0.0038
	<i>Streptomyces thermoautotrophicus</i>	module hub	<i>Actinobacteria</i>	NS0	0.0168
	<i>Terriglobus saanensis</i>	module hub	<i>Acidobacteria</i>	NS1	0.0248
	<i>Thiohalobacter thiocyanaticus</i>	module hub	<i>Proteobacteria</i>	NS5	0.0338
	<i>Gloeobacter violaceus</i>	connector	<i>Cyanobacteria</i>	NS6	0.009
	<i>Oerskovia turbata</i>	connector	<i>Actinobacteria</i>	NS4	0.0001
	<i>Rubrobacter radiotolerans</i>	connector	<i>Actinobacteria</i>	NS2	0.0048
	<i>Ceraceosorus bombacis</i>	module hub	<i>Basidiomycota</i>	NS2	0.00002
	<i>Bifiguratus adelaidae</i>	connector	<i>Mucoromycota</i>	NS2	0.0005

	<i>Ogataea parapolyomorpha</i>	connector	<i>Ascomycota</i>	NS2	0.00002
RNI	<i>Actibacterium atlanticum</i>	module hub	<i>Proteobacteria</i>	RNI1	0.00004
	<i>Actinomadura macra</i>	module hub	<i>Actinobacteria</i>	RNI3	0.0031
	<i>Actinopolymorpha alba</i>	module hub	<i>Actinobacteria</i>	RNI3	0.0099
	<i>Chitinophaga terrae</i>	module hub	<i>Bacteroidetes</i>	RNI1	0.0146
	<i>Enterovibrio calviensis</i>	module hub	<i>Proteobacteria</i>	RNI13	0.00003
	<i>Gracilibacillus orientalis</i>	module hub	<i>Firmicutes</i>	RNI31	0.00003
	<i>Halorubrum distributum</i>	module hub	<i>Euryarchaeota</i>	RNI9	0.00004
	<i>Marvinbryantia formatexigens</i>	module hub	<i>Firmicutes</i>	RNI12	0.00005
	<i>Megamonas hypermegale</i>	module hub	<i>Firmicutes</i>	RNI8	0.00003
	<i>Rubrobacter xylanophilus</i>	module hub	<i>Actinobacteria</i>	RNI6	0.0258
	<i>Sphingomonas sanxanigenens</i>	module hub	<i>Proteobacteria</i>	RNI4	0.0048
	<i>Streptomyces celluloflavus</i>	module hub	<i>Actinobacteria</i>	RNI1	0.00004
	<i>Streptacidiphilus rugosus</i>	connector	<i>Actinobacteria</i>	RNI6	0.0058
	<i>Amanita muscaria</i>	module hub	<i>Basidiomycota</i>	RNI0	0.00018
	<i>Paxillus involutus</i>	module hub	<i>Basidiomycota</i>	RNI0	0.00028
	<i>Trichoderma harzianum</i>	module hub	<i>Ascomycota</i>	RNI2	0.00002
	<i>Hebeloma cylindrosporum</i>	connector	<i>Basidiomycota</i>	RNI3	0.00004
	<i>Smittium mucronatum</i>	connector	<i>Zoopagomycota</i>	RNI4	0.00001

**Table 4 Topological properties of the co-occurrence functional network analysis**

Network properties	NS	RNI
Number of nodes	631	693
Number of edges	5720	8450
Positive edges	3488	5166
Negative edges	2231	3283
Modularity	0.520	0.307
Number of module	40	35
Average degree (avgK)	18.130	24.387
Average clustering coefficient (avgCC)	0.333	0.296
Average path distance (GD)	5.806	6.035

**Table 5 The enzyme activities of nematocidal microorganisms**

Strains	Protease	Chitinase	Lipase
<i>Aspergillus sojae</i> R3	0.15 ± 0.09	0	0
<i>Pseudomonas trivialis</i> E25	0.48 ± 0.03	0	0
<i>Pseudomonas simiae</i> E15	1.2 ± 0.08	0	0
<i>Pseudomonas mosselii</i> E11	1.52 ± 0.19	0	0
<i>Bacillus amyloliquefaciens</i> W1	1.18 ± 0.19	0	2.74 ± 0.17
<i>Paenibacillus polymyxa</i> R4	0.43 ± 0.03	0	0.83 ± 0.06
<i>Microbacterium arabinogalactanolyticum</i> L10	0.58 ± 0.02	0	0.56 ± 0.1
<i>Bacillus muralis</i> N17	1.35 ± 0.2	0	0.93 ± 0.12
<i>Bacillus stratosphericus</i> N15	0.67 ± 0.01	0	0
<i>Bacillus cohnii</i> N1	0.12 ± 0.03	0	0.11 ± 0.02

<i>Microbacterium jejuense</i> G15	1.39 ± 0.1	0	0.57 ± 0.22
<i>Variovorax paradoxus</i> G7	0	0	1.06 ± 0.1
<i>Microbacterium trichothecenolyticum</i> G2	0	0.12 ± 0.01	1.18 ± 0.09
<i>Bacillus methylotrophicus</i> M12	1.08 ± 0.33	0	2.03 ± 0.25
<i>Bacillus safensis</i> B16	0.87 ± 0.23	0	0
<i>Bacillus cereus</i> B13	1.09 ± 0.09	0.11 ± 0.01	1.13 ± 0.18
<i>Bacillus toyonensis</i> B8	1.16 ± 0.09	0	0
<i>Bacillus licheniformis</i> W15	0	0.19 ± 0.06	0
<i>Bacillus subtilis</i> J6	1.0 ± 0.07	0	2.11 ± 0.19
<i>Bacillus siamensis</i> J4	1.53 ± 0.3	0	2.66 ± 0.08
<i>Acinetobacter bayly</i> B12	0	0	0.08 ± 0.01
<i>Burkholderia cepacia</i> B8	1.04 ± 0.2	0	0
<i>Bacillus aryabhatai</i> O13	0.58 ± 0.14	0	1.14 ± 0.29
<i>Bacillus aryabhatai</i> K2	0.85 ± 0.06	0	2.32 ± 0.03
<i>Pantoea dispersa</i> P3	0.79 ± 0.07	0	1.39 ± 0.1

### Figure legends

#### Figure 1 Microbes with potential nematicidal activity are enriched in the non-infected soils.

(A) Comparing abundance of microbes between the NS and RNI soils. (B) Correlation of enriched bacteria with soil properties.

**Figure 2 Nematicidal bacteria isolated from the NS rhizosphere soils.** (A) Nematicidal activity of bacteria and fungal isolates. (B) Phylogenetic tree of the bacterial isolates. Colors depict the taxonomic classification. Green bar charts represent the nematicidal activity of each isolate.

**Figure 3 Relative abundances of different nodes in NS and RNI networks.** (A) Proportion of nodes belonging to different phylogenetic groups in the NS and RNI bacterial networks; Venn diagrams indicate the number of nodes shared and not shared by the NS and RNI bacterial networks. (B) Proportion of nodes belonging to different phylogenetic groups in the NS and RNI fungal networks; Venn diagrams indicate the number of nodes shared and not shared by the NS and RNI fungal networks.

**Figure 4 Zi-Pi plot indicates the distribution of nodes based on their topological roles in bacterial network (A) and fungal network (B).** Each symbol represents a node in the non-infected soil (NS) network (red triangle) or root-knot infected soil (RNI) network (blue square). The threshold values of Zi and Pi for categorizing nodes are 2.5 and 0.62, respectively.

---

**Figure 5 Metabolomics analysis of microbiota in the NS and RNI rhizosphere soils.** (A) and (B): PLS-DA analysis in negative polarity mode and positive polarity mode, respectively. (C) and (D): Different metabolites in negative and positive polarity mode, respectively. (E) and (F): Z-score of different metabolites in negative and positive polarity mode, respectively. C<sub>21</sub>H<sub>28</sub>O<sub>6</sub> represents Bisphenol A Bis(2,3-dihydroxypropyl) Ether. C<sub>20</sub>H<sub>32</sub>O<sub>4</sub> represents 5-(4-Carboxy-3-methylbutyl)-1,4a-dimethyl-6-methylenedecahydro-1-naphthalenecarboxylic acid. C<sub>20</sub>H<sub>34</sub>O<sub>3</sub> represents 5-[(1S,2R,4aR)-5-(Hydroxymethyl)-1,2,4a-trimethyl-1,2,3,4,4a,7,8,8a-octahydro-1-naphthalenyl]-3-methylpentanoic acid. C<sub>22</sub>H<sub>34</sub>O<sub>4</sub> represents 7(S),17(S)-Dihydroxy-8(E),10(Z),13(Z),15(E),19(Z)-docosapentaenoic acid. C<sub>13</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> represents Acetyl-N-formyl-5-methoxykynurenamine. C<sub>14</sub>H<sub>16</sub>N<sub>2</sub>O<sub>4</sub> represents Ethyl 1-(2-furylcarbonyl)-3-propyl-1H-pyrazole-4-Carboxylate. C<sub>27</sub>H<sub>26</sub>N<sub>2</sub>O<sub>5</sub> represents N1-[4-(6-Methyl-1,3-benzothiazol-2-yl) phenyl]-2-cyclopentyl-2-phenylacetamide. (G)

**Nematicidal activity of different metabolites against *M. incognita*.**

**Figure 6 Interactions between acetophenone and different microbes.** Nodes of microbes are indicated with green color, and node of acetophenone is indicated with yellow color. Red lines and blue lines indicate negative and positive interactions between two nodes, respectively. The two nodes interacting with acetophenone are labeled.

**Figure S1 Overview of annotated genes in the metagenome after quality filtering.** (A) Percentage of reads taxonomically classified at Domain level. (B) Percentage of annotated reads using the KEGG database. (C) Percentage of annotated reads using the eggNOG database.

**Figure S2 Overview and distribution of microbial reads extracted from the metagenome.** (A) Beta diversity of microbial community in root-knot nematode infected soil (RNI) and non-infected soil (NS). Unigenes were used to calculate Bray-Curtis distances and conduct Principal Co-ordinates Analysis (87.1% of the overall variance). Significance of the PCoA analysis was assessed through Anosim ( $P < 0.01$ ). (B-D) Histograms showing the taxonomic distribution at Phylum, Class and Family level, respectively.

**Figure S3 Taxonomic network inference of the rhizosphere microbiota.** (A) and (B): Co-occurrence network based on correlation of 16S rRNA reads (species level) extracted from metagenomes of rhizosphere soil in the two conditions: non-infected soil (NS), and root-knot

---

infected soil (RNI), respectively. (C) and (D): Co-occurrence network based on correlation of 18S rRNA reads (species level) extracted from metagenomes of the NS and RNI rhizosphere microbiota, respectively. The size of each node is proportional to the number of connections (node degree). Nodes belonging to different phylum are differently colored. Red lines and blue lines indicate negative and positive interactions between different nodes, respectively. Modules with more than five nodes are included. Generalists were indicated.

**Figure S4 Sub-network of nodes belonging to *Proteobacteria* (A), or *Ascomycota* and *Basidiomycota* (B).** The circle sizes represent the mean read relative abundance of the differentially abundant taxa. Bacterial or fungal species that are significantly enriched in the comparison between NS and RNI are indicated in green for NS and in blue for RNI; nonsignificant taxa are indicated in yellow. Red color lines represent positive interaction. Blue color lines represent negative interaction.

**Figure S5 Function network inference of the rhizosphere microbiota.** (A) and (B): Co-occurrence network based on correlation of COG genes in the two conditions: non-infected soils (NS), and root-knot infected soils (RNI), respectively. The size of each node is proportional to the abundance of nodes. Nodes belonging to different category are differently colored. Red lines and blue lines indicate negative and positive interactions between different nodes, respectively. Node shapes are based on change of node abundance: circle, non-change nodes in any treatment; triangle, overrepresented in the NS condition; diamond, overrepresented in the RNI condition. (C) and (D): Co-occurrence network based on correlation of BGCs genes in the NS and RNI microbiota. Nodes are shaped by BGC types and the size is correlated with the abundance of each node. Node color is based on change of node abundance. Yellow, non-change nodes in any treatment; red: enriched in the RNI microbiota, green: enriched in the NS microbiota.

**Figure S6 Heatmaps depict relative abundance of the significantly enriched genes involved in polyketide biosynthesis (A), CAZymes (B) or enzymes (C, D) in the NS microbiota as compared to the RNI microbiota.**

**Figure S7 Comparing gene abundances between the NS and RNI BGCs network.**

**Figure S8 Metabolomics analysis of rhizosphere microbiota in the NS and RNI soils.** (A), (B): PCA analysis in negative polarity mode and positive polarity mode, respectively.

---

**Figure S9 The hierarchical clustering of different metabolites in negative polarity mode (A and positive polarity mode (B, respectively.**

**Figure S10 Nematicidal and avoidance activities of metabolites and microbes. (A)** Nematicidal activity of different concentrations of acetophenone against *M. incognita*. **(B)** Avoidance of acetophenone against *M. incognita*. **(C)** Avoidance activity of *Bacillus* isolates against *M. incognita*. **(D)** Phenylalanine metabolism pathway from KEGG. The genes which could be found in metagenomic dataset were labeled by red letter. **(E)** MS spectra for metabolites in phenylalanine metabolism pathway detected in *B. amyloliquefaciens* W1 liquid culture by untargeted metabolomics analysis. **(F)** Detecting standard acetophenone by gas chromatography. **(G)** Detecting acetophenone produced by *B. amyloliquefaciens* W1 detected by gas chromatography.

Table S1 Soil sample sites and properties

Table S2 Properties of non-infected (NS) and root-knot nematode infected (RNI) soils

Table S3 Correlation coefficients and significances between soil variables and number of *M. incognita*

Table S4 Metagenome sequence statistics after filtering

Table S5 CDS statistics of each sample

Table S6 Bacterial isolates showed low nematicidal activity

Table S7 Gene nodes enriched in the NS function network

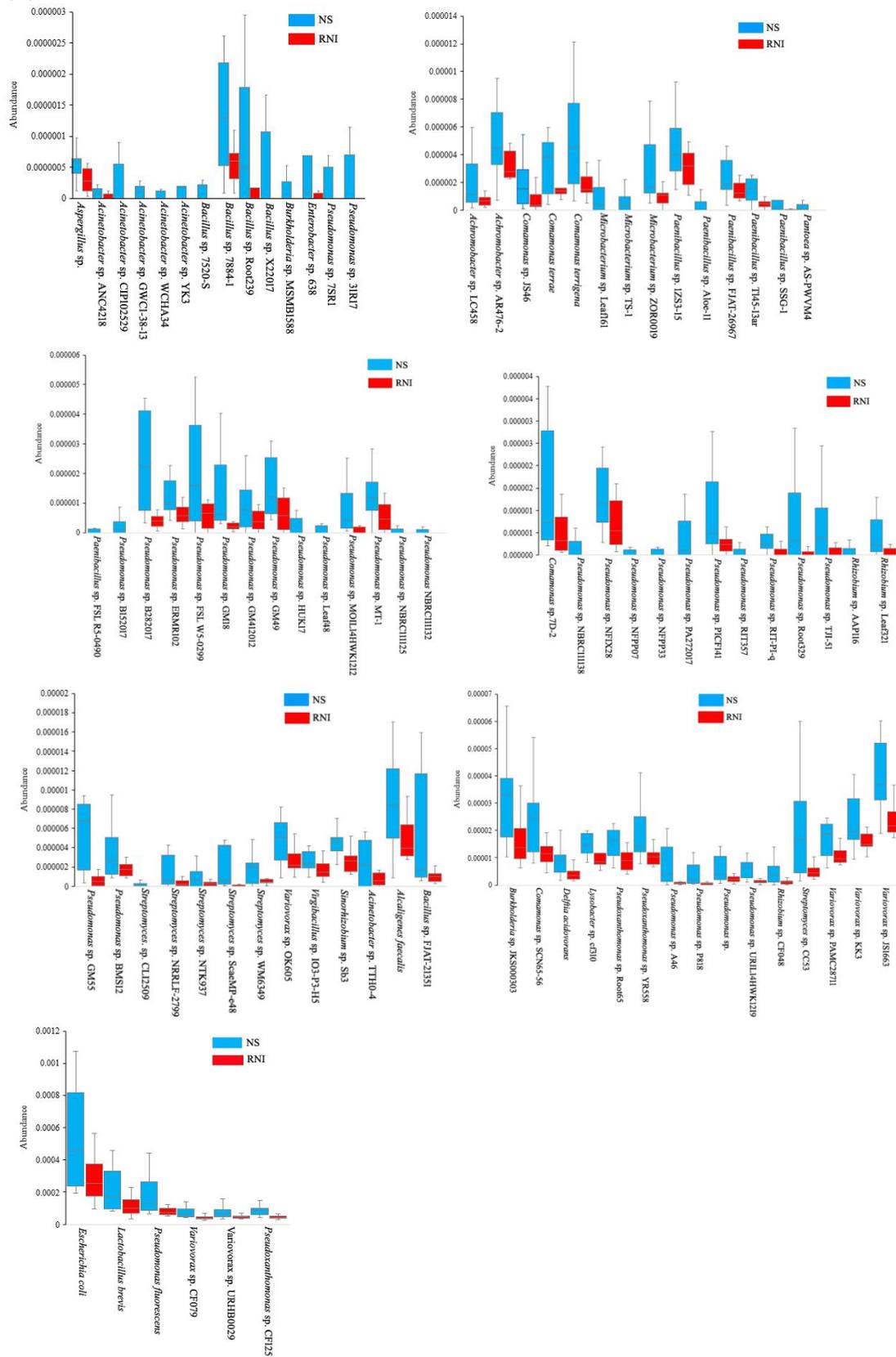
Table S8 The abundances of biosynthetic gene clusters (BGCs) in two groups

Table S9 Correlation of metabolites and number of *M. incognita*

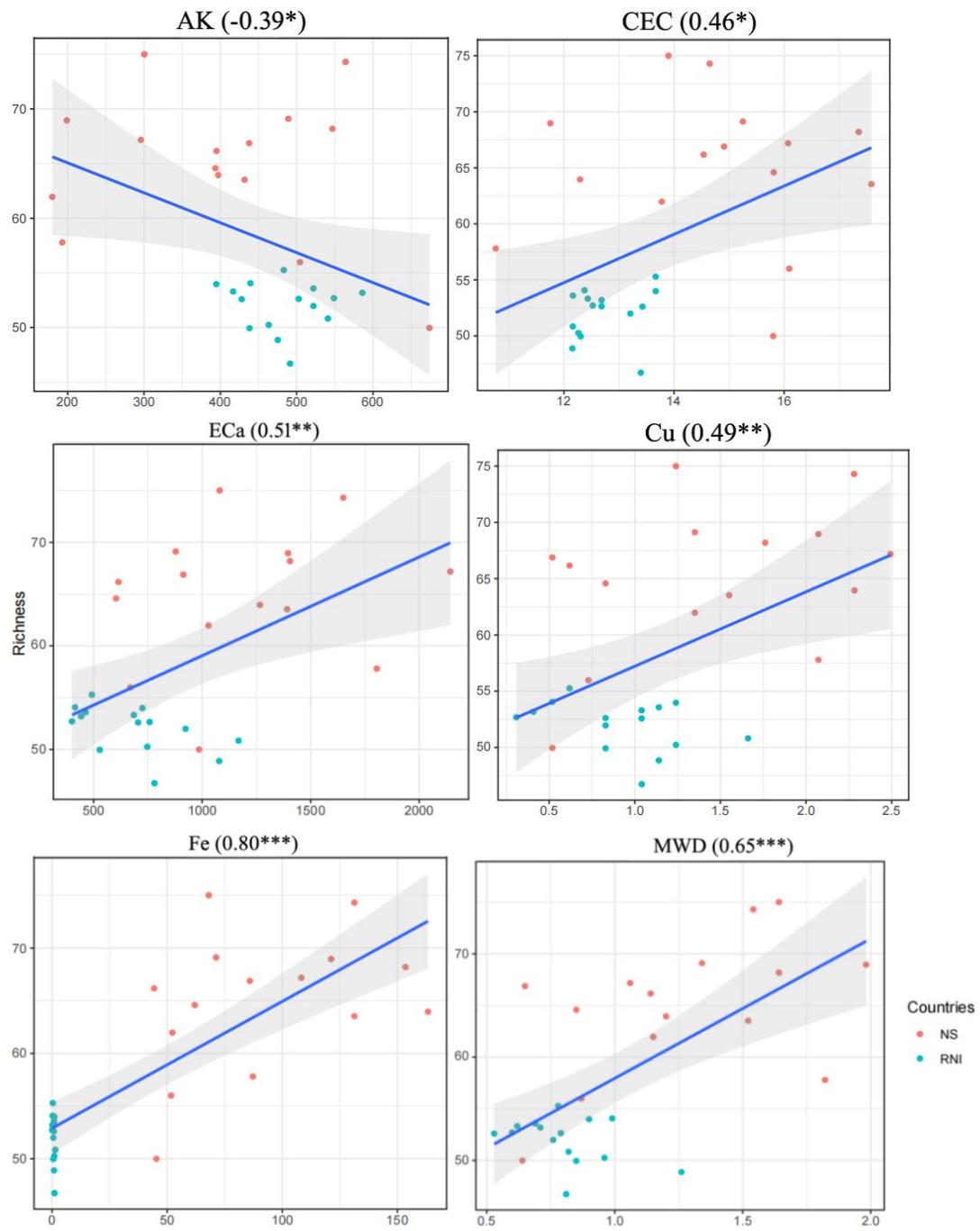
Table S10 Metabolites in *Bacillus amyloliquefaciens* W1 culture

**Figure 1** Microbes with potential nematicidal activity are enriched in the non-infected soils. **(A)** Comparing abundance of microbes between the NS and RNI soils. **(B)** Correlation of enriched bacteria with soil properties.

**(A)**

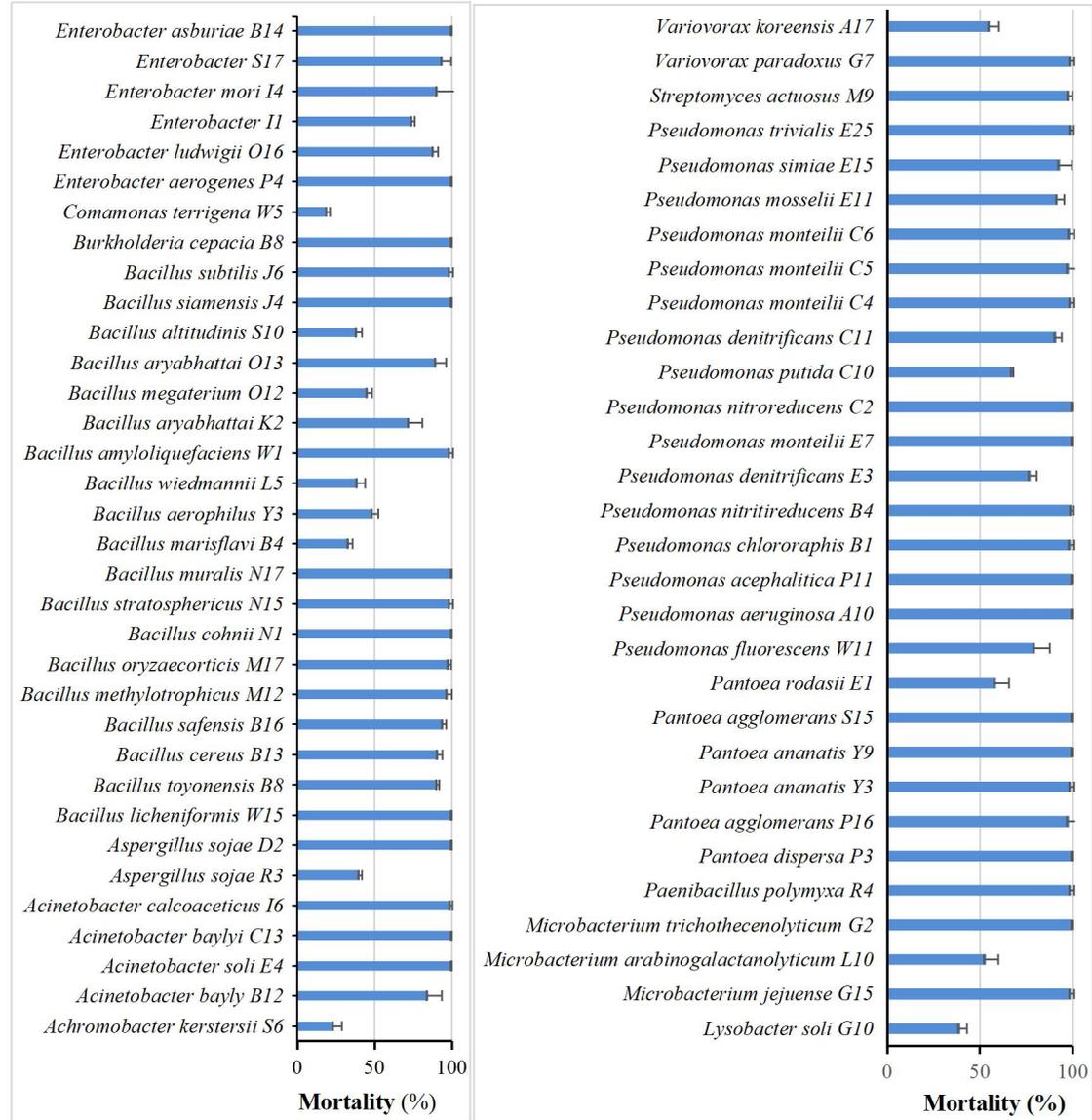


(B)



**Figure 2 Nematicidal bacteria isolated from the NS rhizosphere soils. (A)** Nematicidal activity of bacteria and fungal isolates. **(B)** Phylogenetic tree of the bacterial isolates. Colors depict the taxonomic classification. Green bar charts represent the nematicidal activity of each isolate.

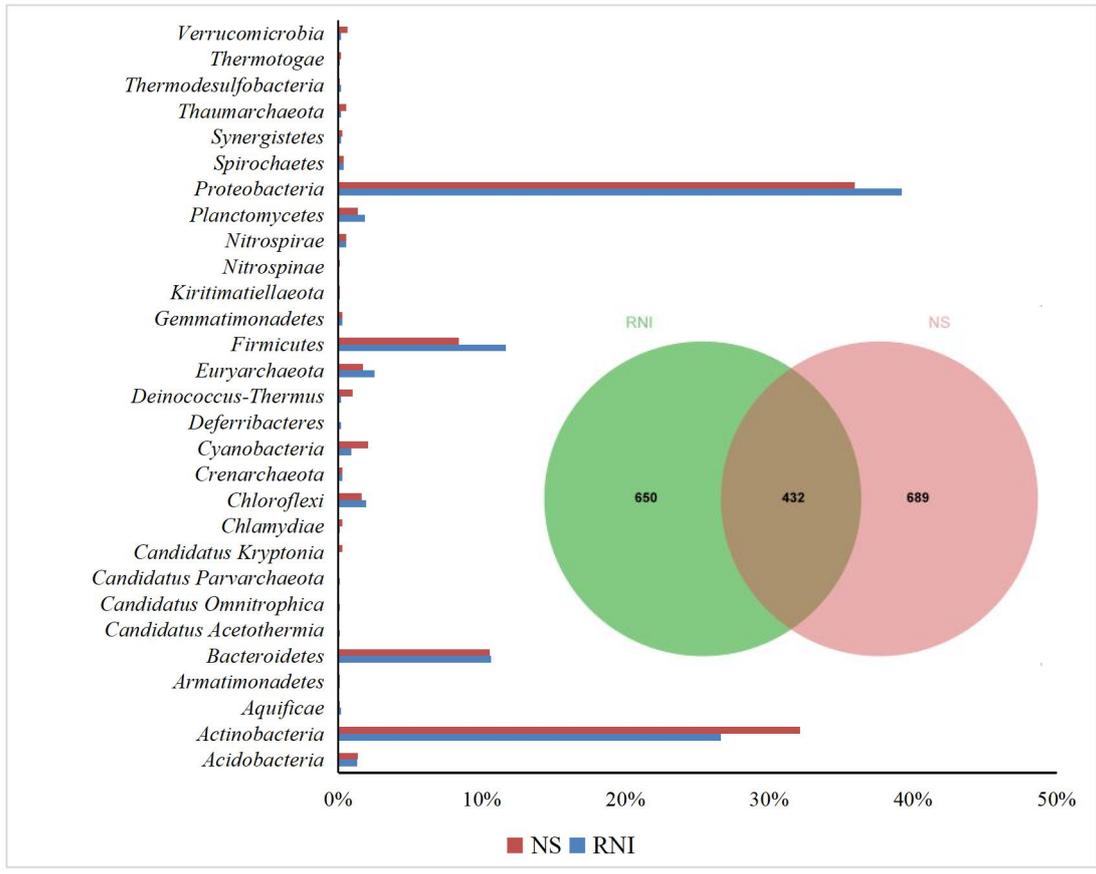
**(A)**



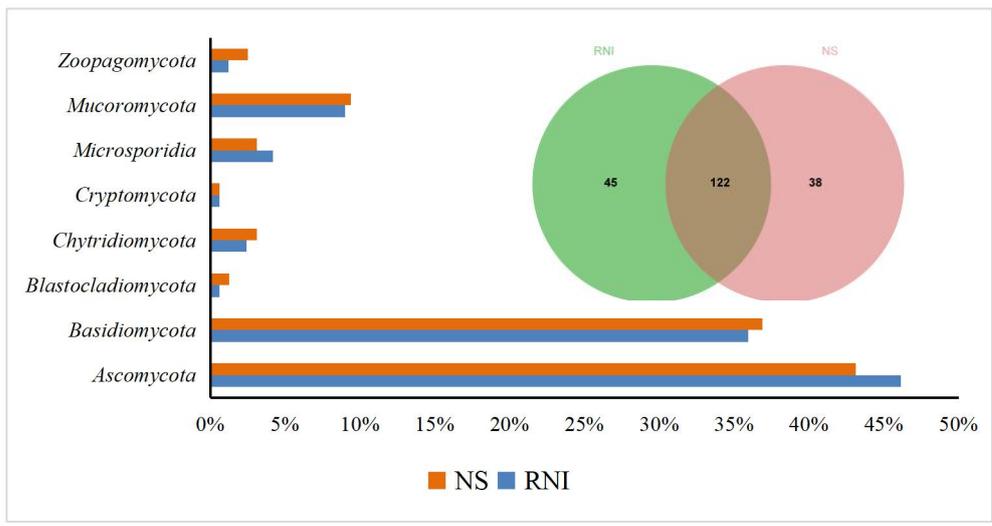


**Figure 3 Relative abundances of different nodes in NS and RNI networks.** (A) Proportion of nodes belonging to different phylogenetic groups in the NS and RNI bacterial networks; Venn diagrams indicate the number of nodes shared and not shared by the NS and RNI bacterial networks. (B) Proportion of nodes belonging to different phylogenetic groups in the NS and RNI fungal networks; Venn diagrams indicate the number of nodes shared and not shared by the NS and RNI fungal networks.

(A)

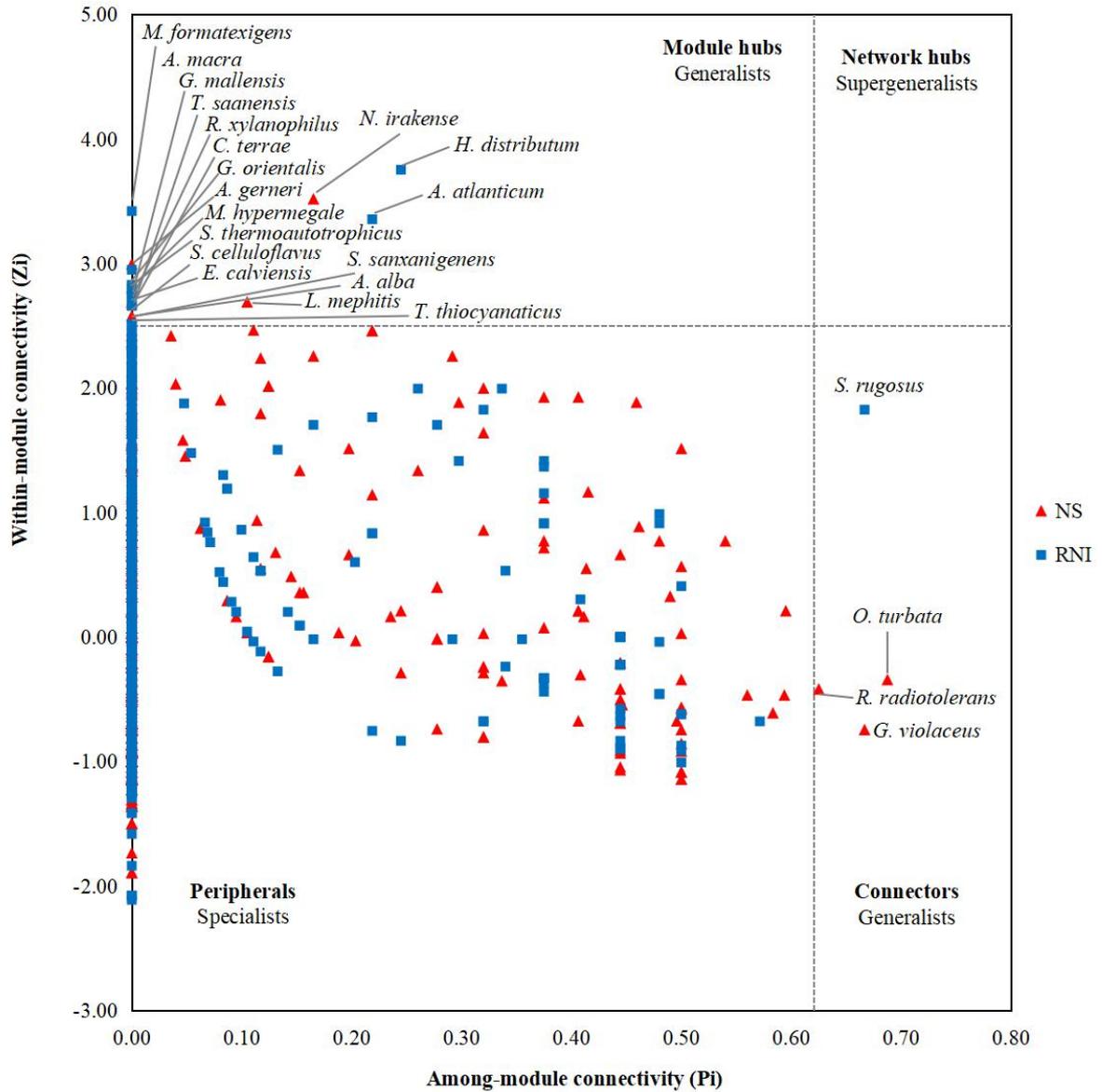


(B)

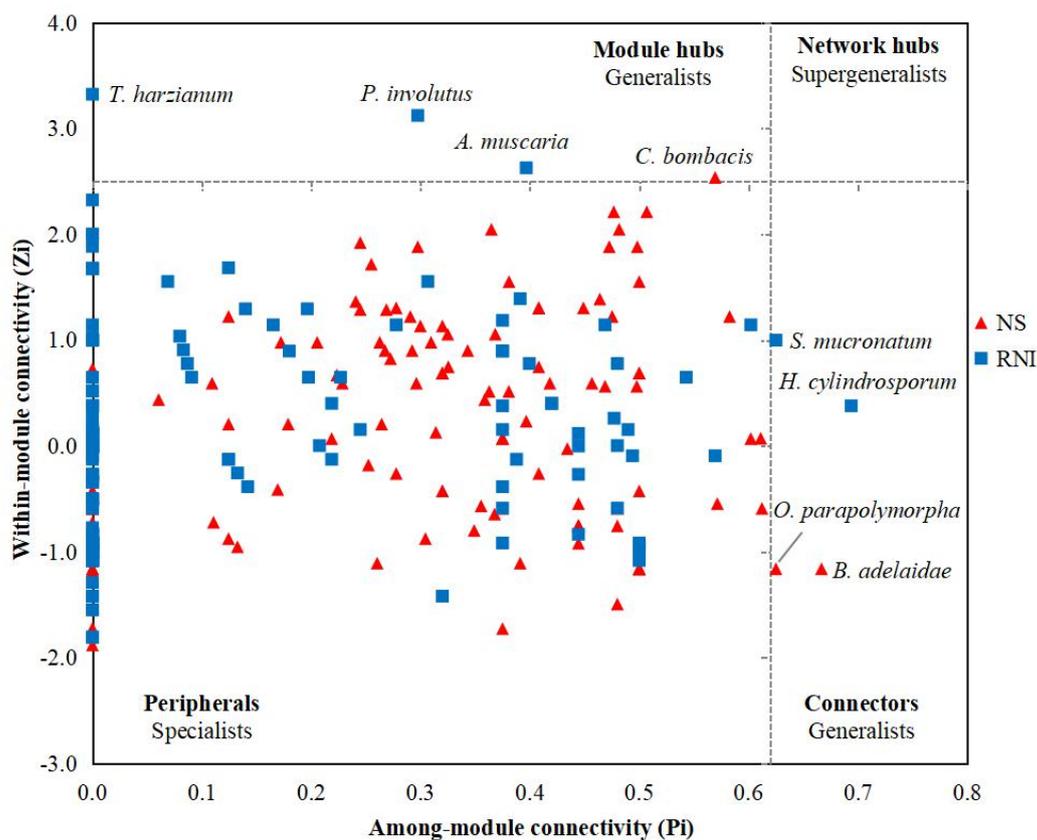


**Figure 4 Zi–Pi plot indicates the distribution of nodes based on their topological roles in bacterial network (A) and fungal network (B).** Each symbol represents a node in the non-infected soil (NS) network (red triangle) or root-knot infected soil (RNI) network (blue square). The threshold values of Zi and Pi for categorizing nodes are 2.5 and 0.62, respectively.

(A)

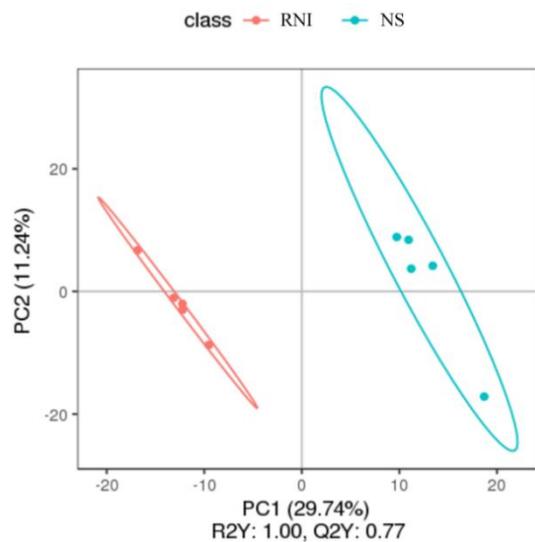


(B)

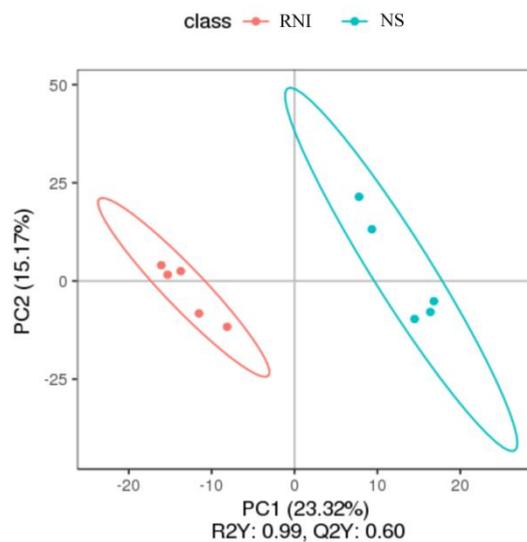


**Figure 5** Metabolomics analysis of microbiota in the NS and RNI rhizosphere soils. (A) and (B): PLS-DA analysis in negative polarity mode and positive polarity mode, respectively. (C) and (D): Different metabolites in negative and positive polarity mode, respectively. (E) and (F): Z-score of different metabolites in negative and positive polarity mode, respectively. (G) Nematicidal activity of different metabolites against *M. incognita*.

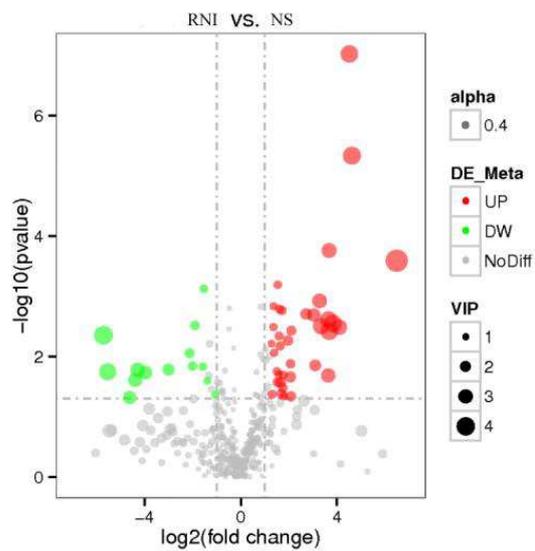
(A)



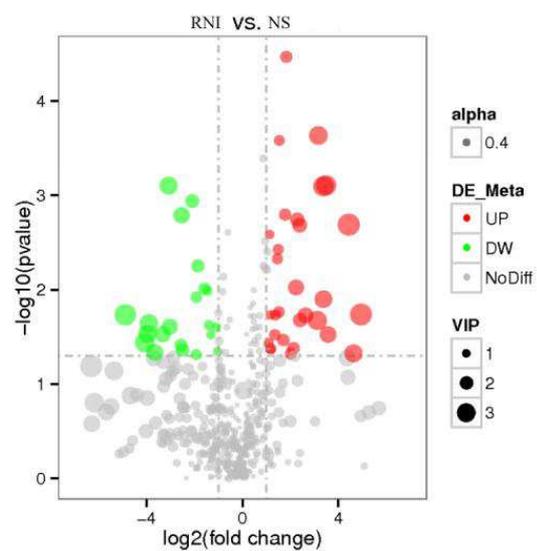
(B)



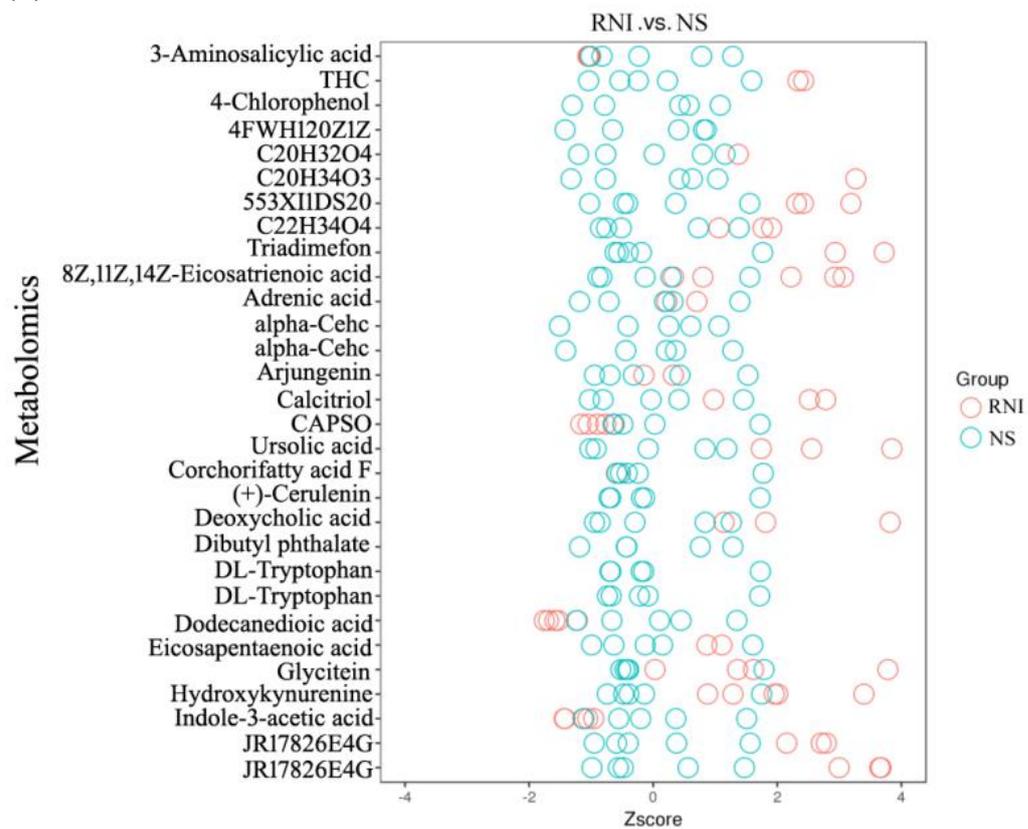
(C)



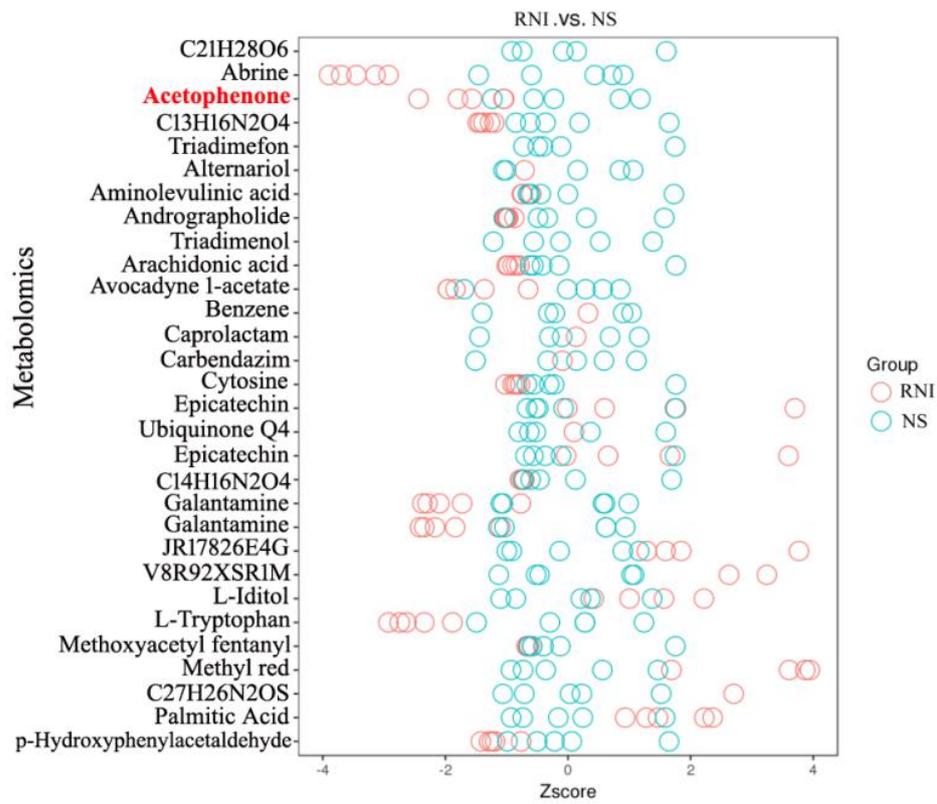
(D)



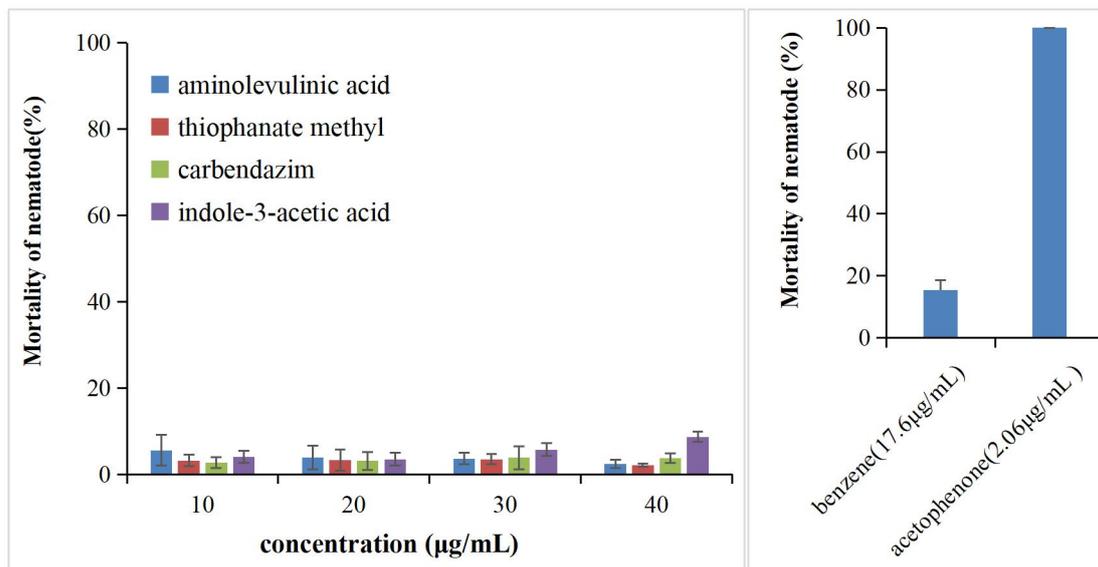
(E)



(F)



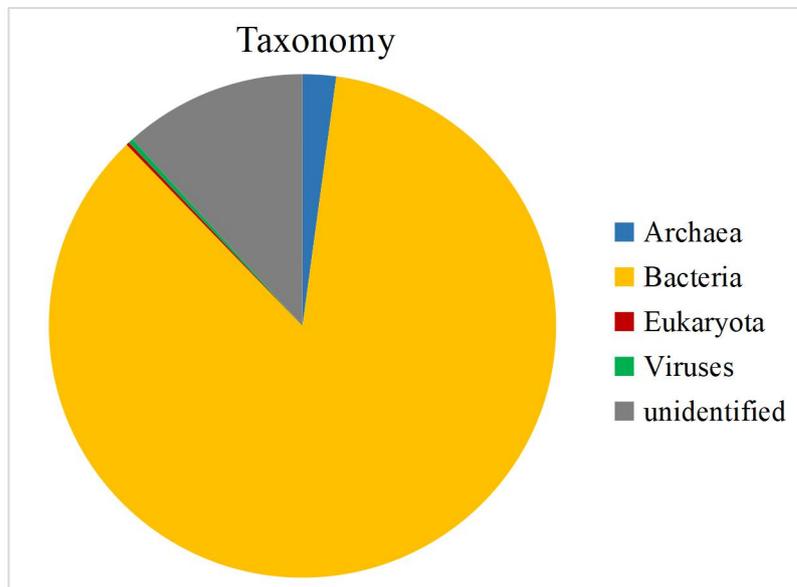
(G)



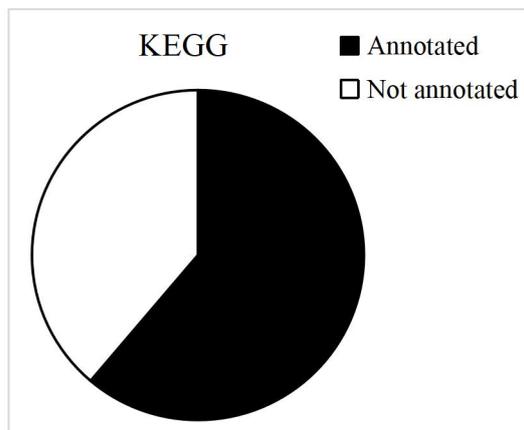


**Figure S1 Overview of annotated genes in the metagenome after quality filtering. (A)** Percentage of reads taxonomically classified at Domain level. **(B)** Percentage of annotated reads using the KEGG database. **(C)** Percentage of annotated reads using the eggNOG database.

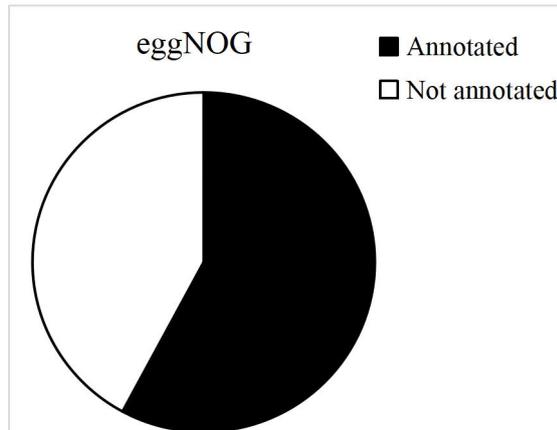
**(A)**



**(B)**



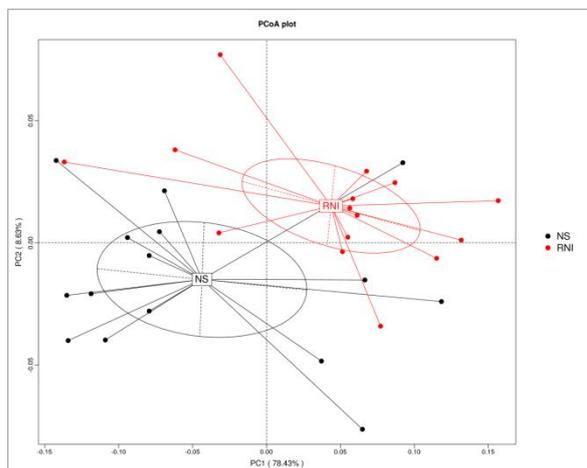
**(C)**



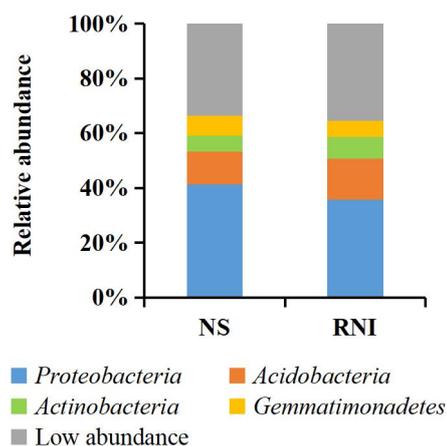
**Figure S2 Overview and distribution of microbial reads extracted from the metagenome.**

(A) Beta diversity of microbial community in root-knot nematode infected soil (RNI) and non-infected soil (NS). Unigenes were used to calculate Bray-Curtis distances and conduct Principal Co-ordinates Analysis (87.1% of the overall variance). Significance of the PCoA analysis was assessed through Anosim ( $P < 0.01$ ). (B-D) Histograms showing the taxonomic distribution at Phylum, Class and Family level, respectively.

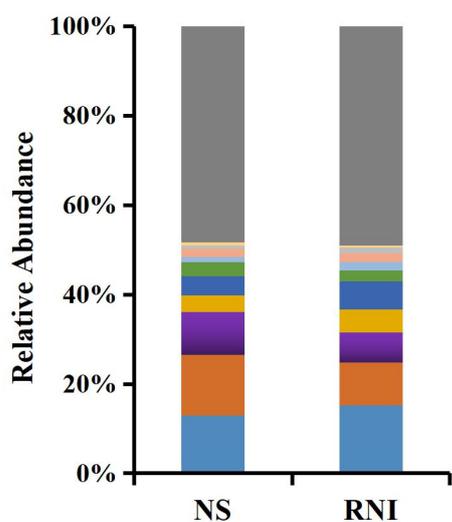
(A)



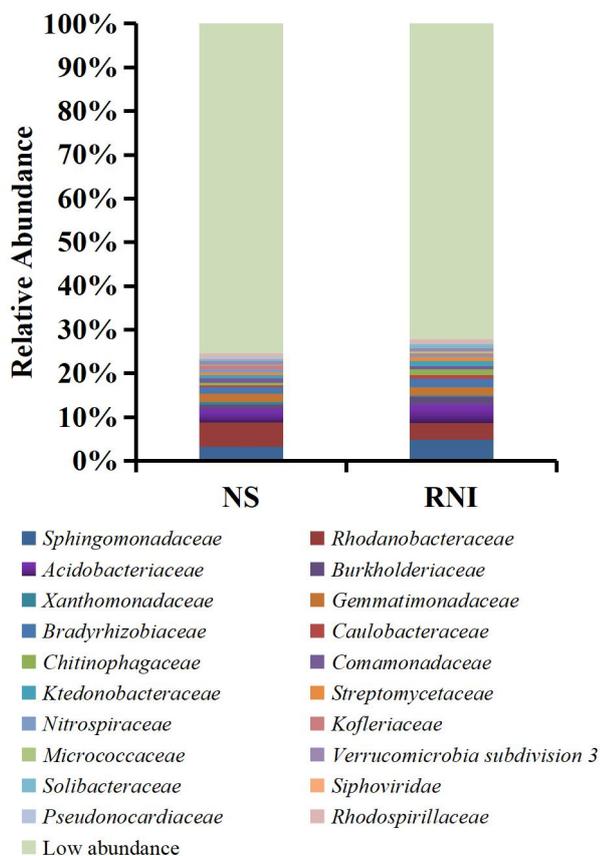
(B)



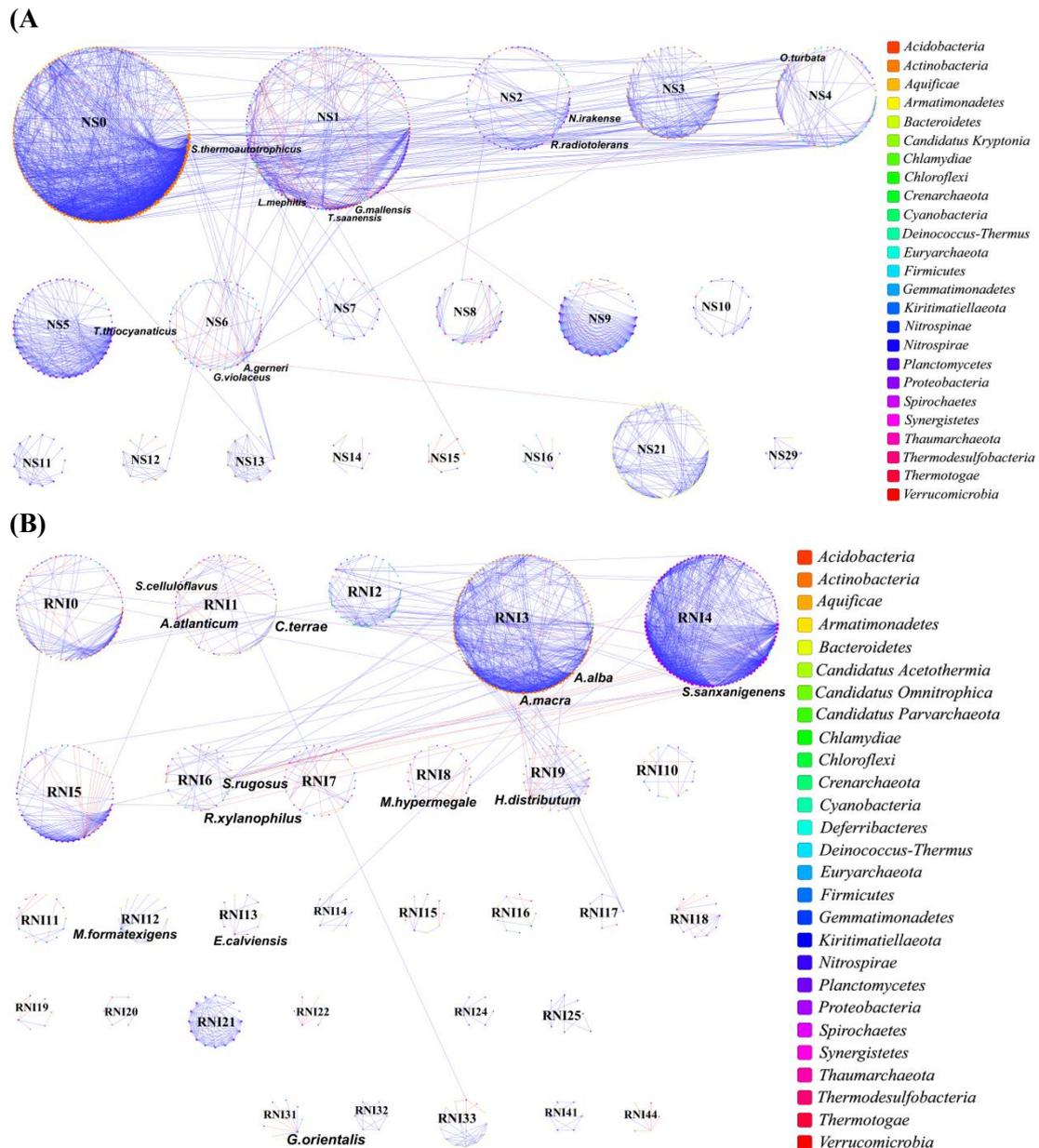
(C)



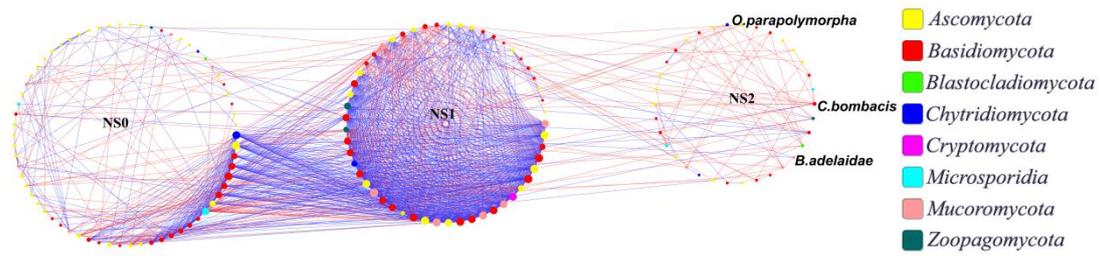
(D)



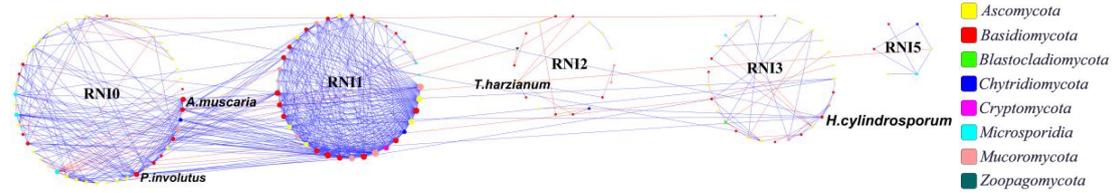
**Figure S3 Taxonomic network inference of the rhizosphere microbiota. (A) and (B):** Co-occurrence network based on correlation of 16S rRNA reads (species level) extracted from metagenomes of rhizosphere soil in the two conditions: non-infected soil (NS), and root-knot infected soil (RNI), respectively. **(C) and (D):** Co-occurrence network based on correlation of 18S rRNA reads (species level) extracted from metagenomes of the NS and RNI rhizosphere microbiota, respectively. The size of each node is proportional to the number of connections (node degree). Nodes belonging to different phylum are differently colored. Red lines and blue lines indicate negative and positive interactions between different nodes, respectively. Modules with more than five nodes are included. Generalists were indicated.



(C)

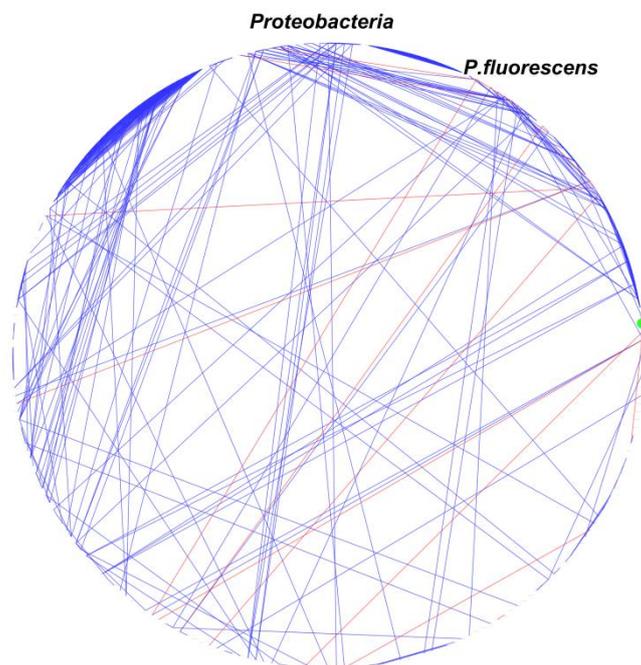


(D)

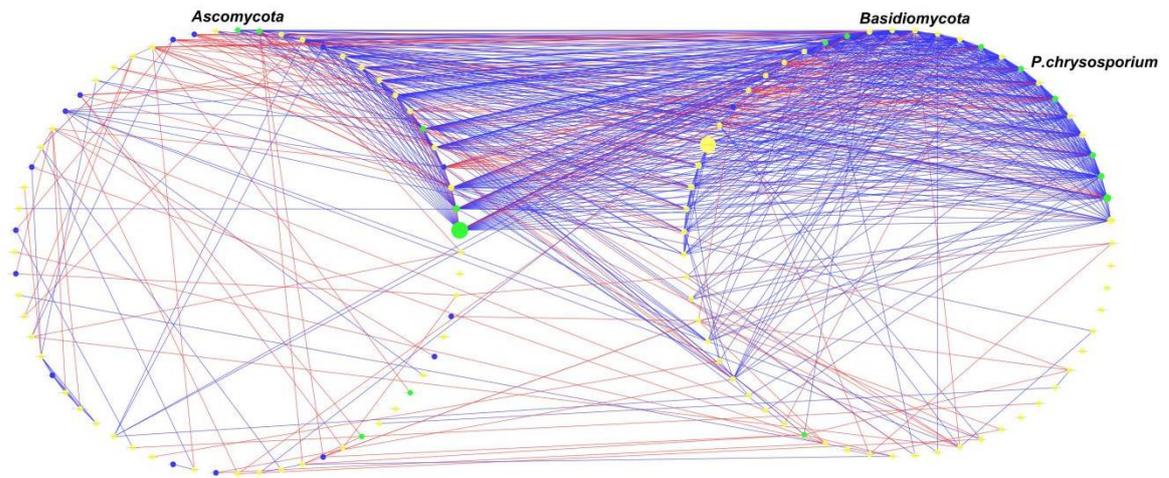


**Figure S4 Sub-network of nodes belonging to *Proteobacteria* (A), or *Ascomycota* and *Basidiomycota* (B).** The circle sizes represent the mean read relative abundance of the differentially abundant taxa. Bacterial or fungal species that are significantly enriched in the comparison between NS and RNI are indicated in green for NS and in blue for RNI; nonsignificant taxa are indicated in yellow. Red lines represent negative interaction. Blue lines represent positive interaction.

(A)

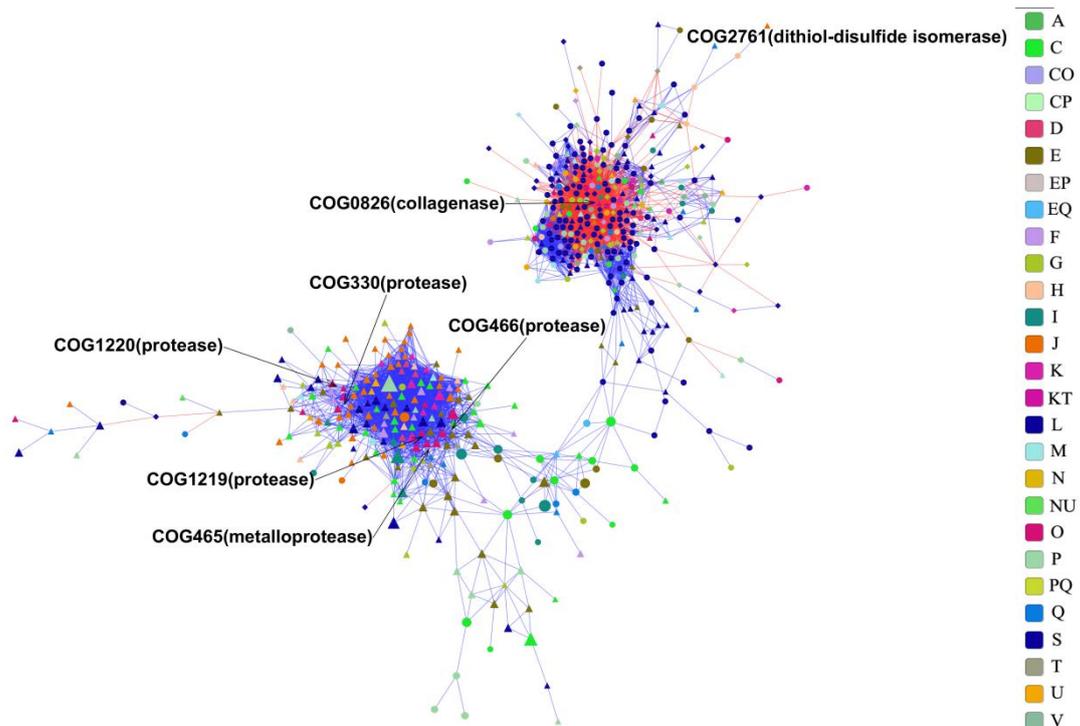


(B)

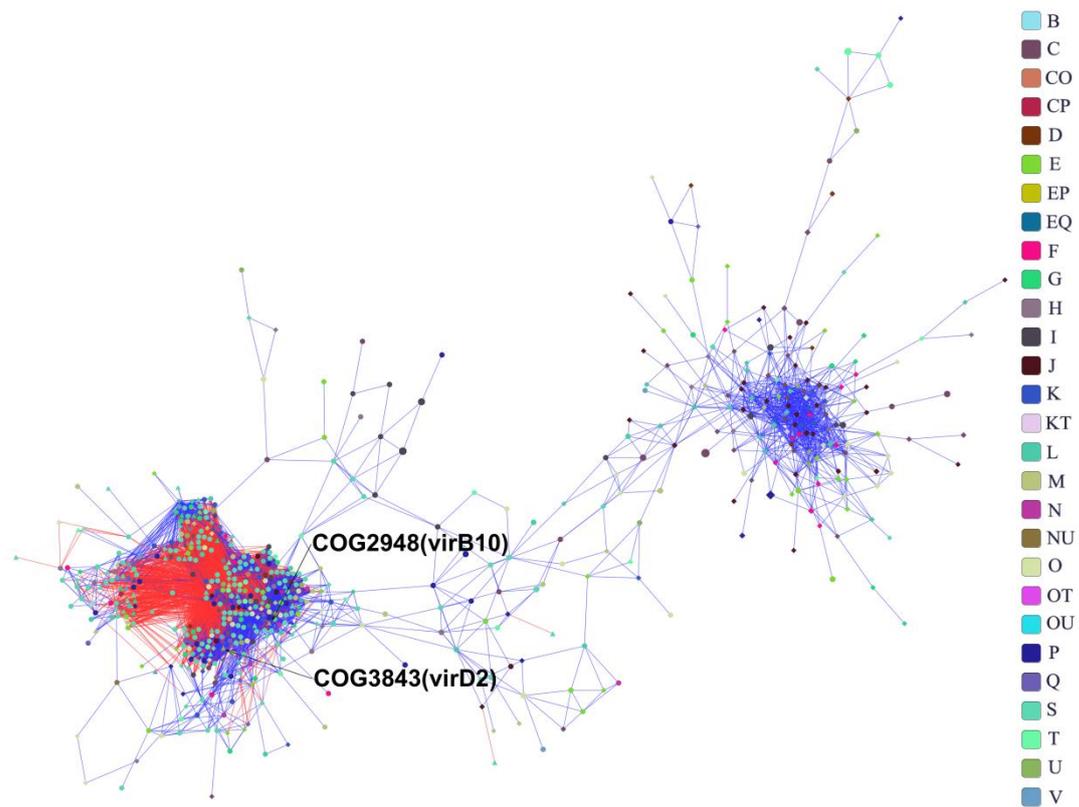


**Figure S5 Function network inference of the rhizosphere microbiota.** (A) and (B): Co-occurrence network based on correlation of COG genes in the two conditions: non-infected soils (NS), and root-knot infected soils (RNI), respectively. The size of each node is proportional to the abundance of nodes. Nodes belonging to different category are differently colored. Red lines and blue lines indicate negative and positive interactions between different nodes, respectively. Node shapes are based on change of node abundance: circle, non-change nodes in any treatment; triangle, overrepresented in the NS condition; diamond, overrepresented in the RNI condition. (C) and (D): Co-occurrence network based on correlation of BGCs genes in the NS and RNI microbiota. Nodes are shaped by BGC types and the size is correlated with the abundance of each node. Node color is based on change of node abundance. Yellow, non-change nodes in any treatment; red: enriched in the RNI microbiota, green: enriched in the NS microbiota.

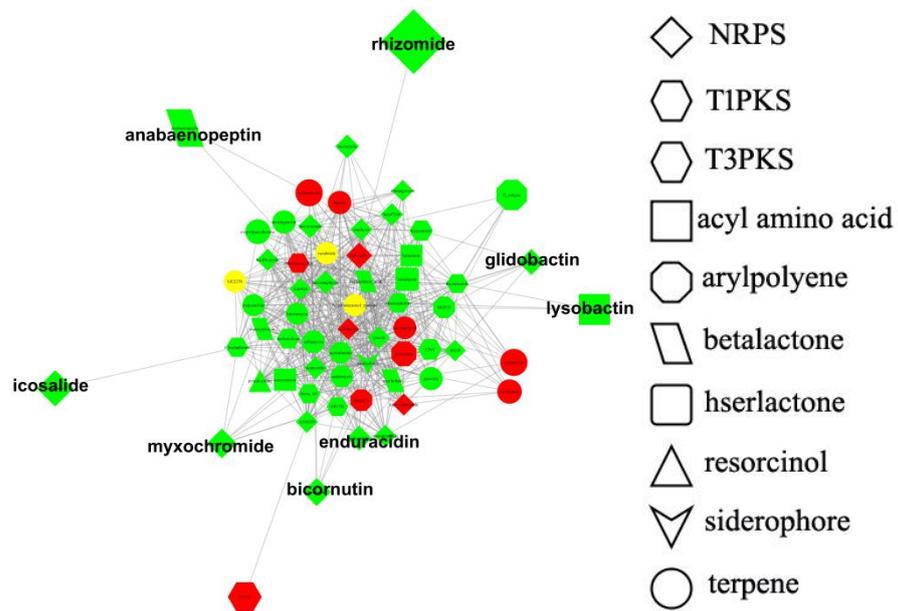
(A)



(B)



(C)



(D)

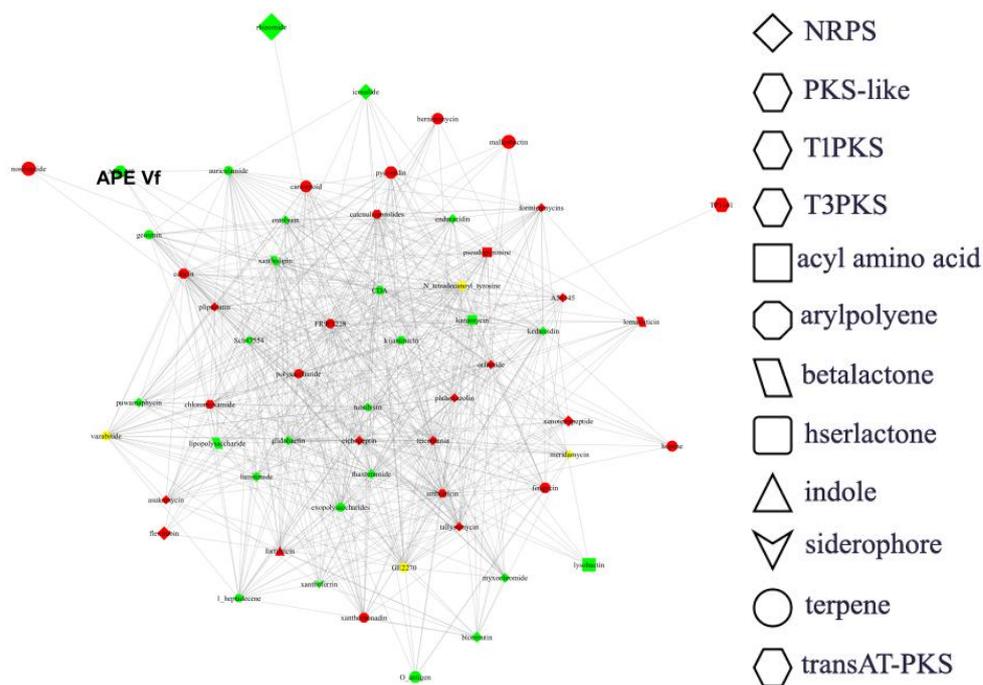
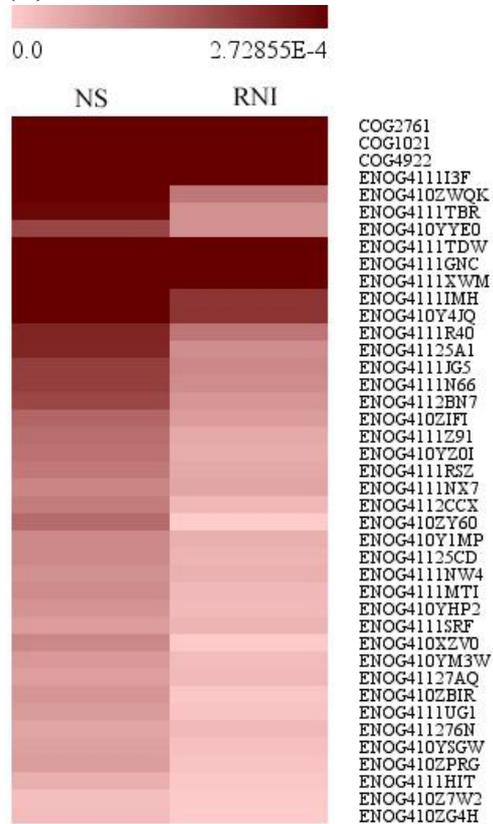
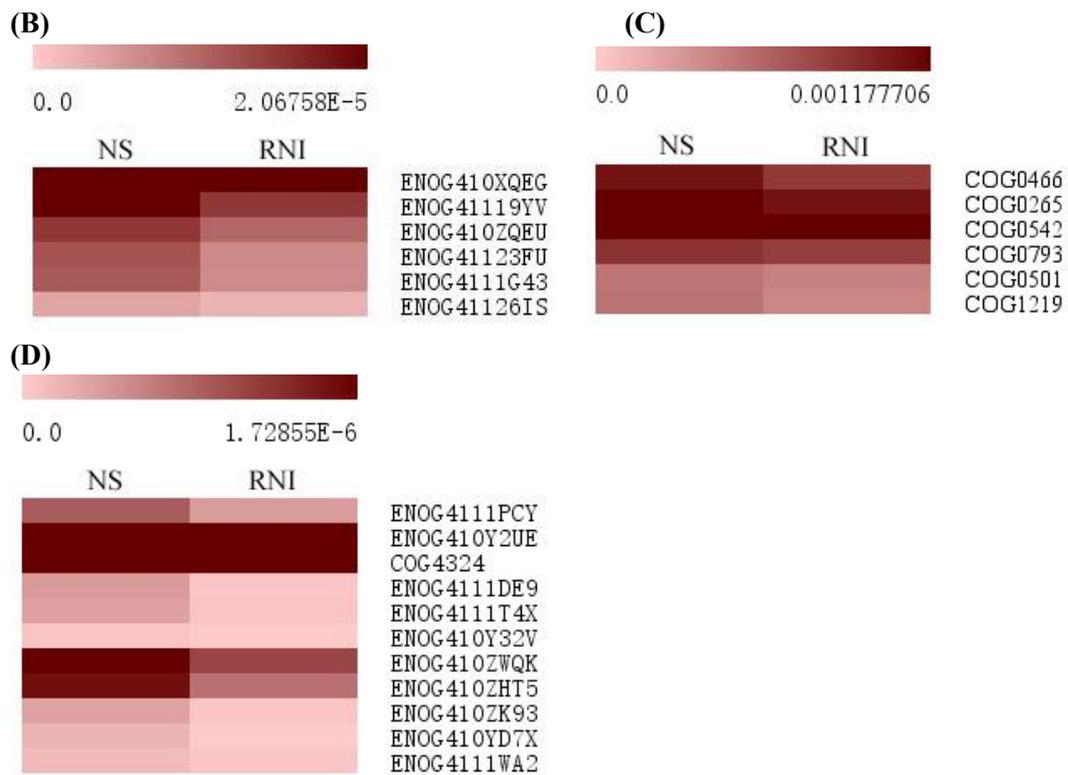


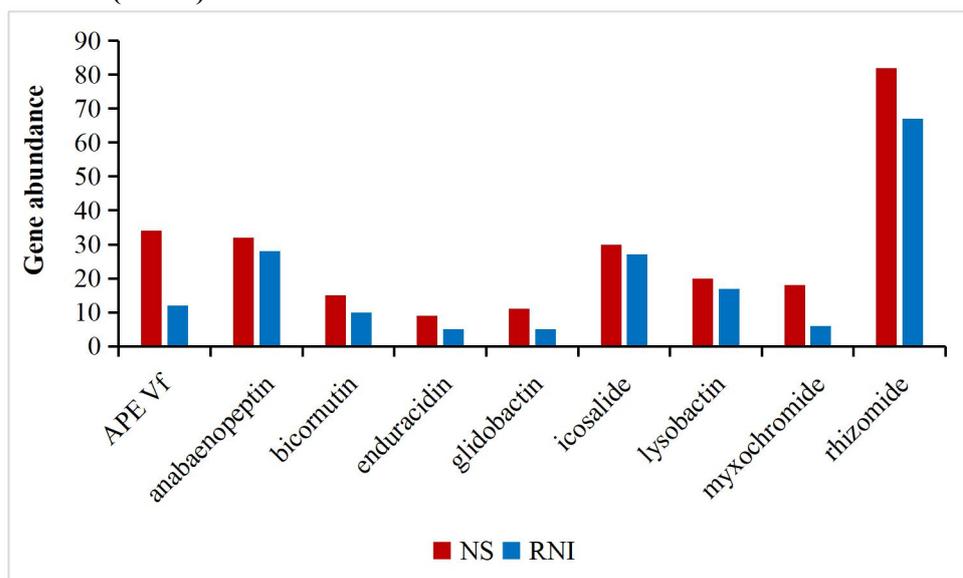
Figure S6 Heatmaps depict relative abundance of the significantly enriched genes involved in polyketide biosynthesis (A), CAZymes (B) or enzymes (C, D) in the NS microbiota as compared to the RNI microbiota.

(A)

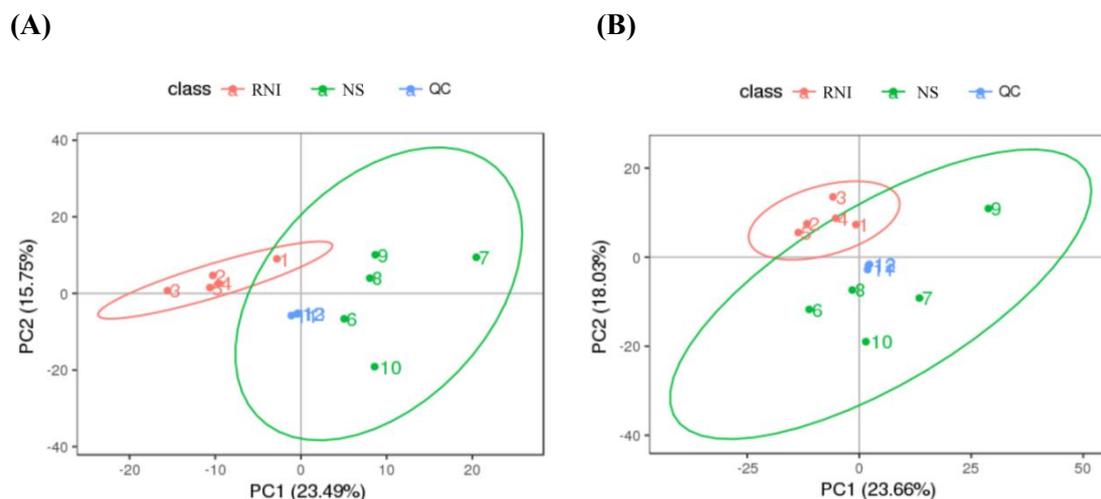




**Figure S7 Comparing gene abundances between the NS and RNI biosynthetic gene clusters (BGCs) network.**



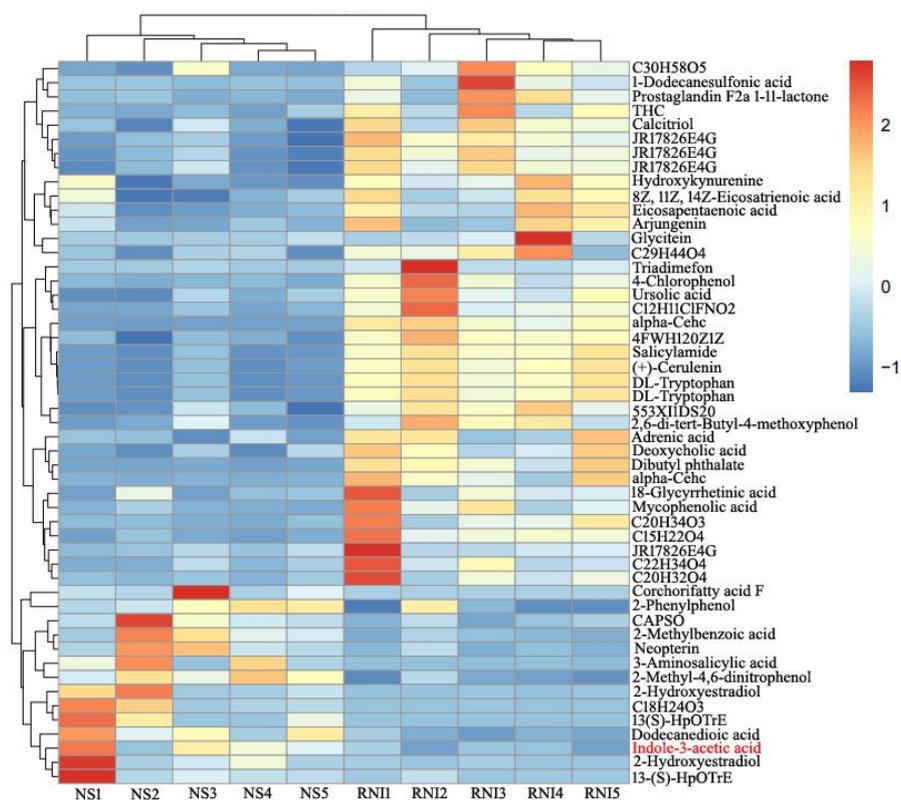
**Figure S8 Metabolomics analysis of rhizosphere microbiota in the NS and RNI soils. (A), (B):** PCA analysis in negative polarity mode and positive polarity mode, respectively.



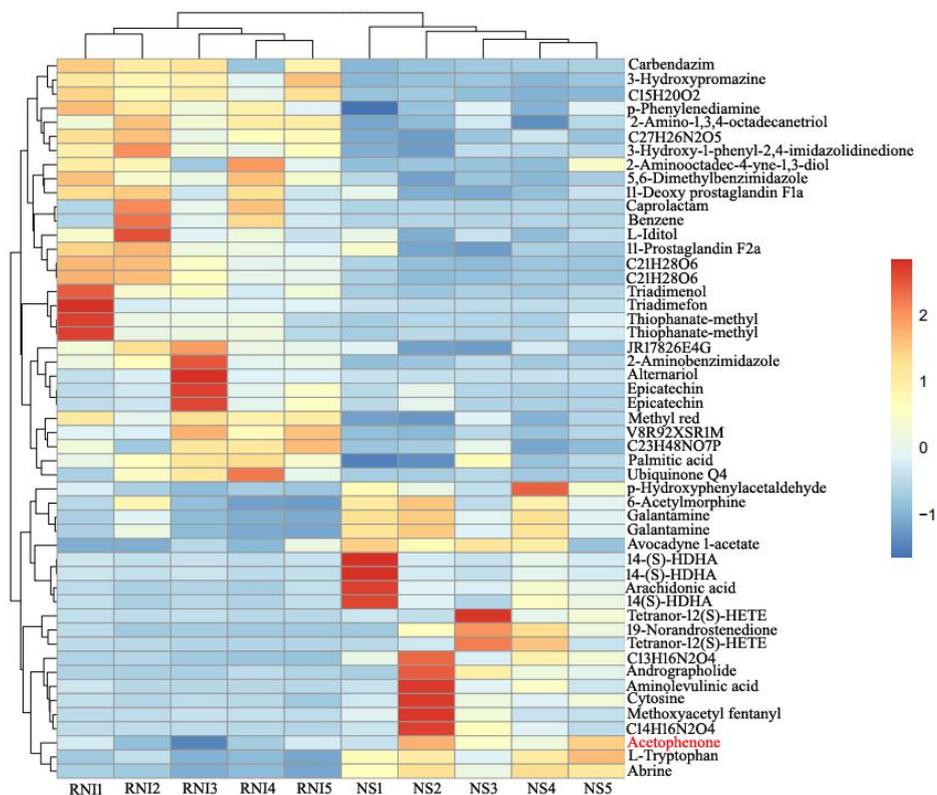
**Figure S9 The hierarchical clustering of different metabolites in negative polarity mode (A) and positive polarity mode (B), respectively.**

C30H58O5 represents 1-(dodecanoyloxy)-3-hydroxy-2-propanyl pentadecanoate. C29H44O4 represents 2,3-dimethoxy-5-methyl-6-[(2E,6E,10E)-3,7,11,15-tetramethyl-2,6,10,14-hexadecatetraen-1-yl]-1,4-benzenediol. C12H11ClFNO2 represents 3-[(2-chloro-6-fluoroanilino)methylidene]pentane-2,4-dione. C20H34O3 represents 5-[(1S,2R,4aR)-5-(Hydroxymethyl)-1,2,4a-trimethyl-1,2,3,4,4a,7,8,8a-octahydro-1-naphthalenyl]-3-methylpentanoic acid. C15H22O4 represents 1,9b-dihydroxy-6,6,9a-trimethyl-5,5a,6,7,8,9,9a,9b-octahydronaphtho[1,2-c]furan-3(1H)-one. C22H34O4 represents 7(S),17(S)-dihydroxy-8(E),10(Z),13(Z),15(E),19(Z)-docosapentaenoic acid. C20H32O4 represents 5-(4-carboxy-3-methylbutyl)-1,4a-dimethyl-6-methylenedecahydro-1-naphthalenecarboxylic acid. C18H24O3 represents 3-[(2E)-3,7-dimethyl-2,6-octadien-1-yl]-5-methoxy-2-methyl-1,4-benzoquinone. C15H20O2 represents 3-[(2E)-3,6-Dimethyl-2,5-heptadien-1-yl]-1,2-benzenediol. C27H26N2OS represents N1-[4-(6-Methyl-1,3-benzothiazol-2-yl)phenyl]-2-cyclopentyl-2-phenylacetamide. C21H28O6 represents 8-hydroxy-2,7,7,11,15-pentamethyl-5,12,14-trioxapentacyclo[9.8.0.0.2,8.0.4,6.0]13,18]nonadec-13(18)-ene-3,17-dione. C23H48NO7P represents 1-{[(2-Aminoethoxy)(hydroxy)phosphoryl]oxy}-3-hydroxy-2-propanyl stearate. C13H16N2O4 represents acetyl-N-formyl-5-methoxykynurenamine. C14H16N2O4 represents ethyl 1-(2-furylcarbonyl)-3-propyl-1H-pyrazole-4-carboxylate. Acetophenone and indole-3-acetic acid were highlighted by red letters because of their nematicidal activity and promoting-growth activity, respectively.

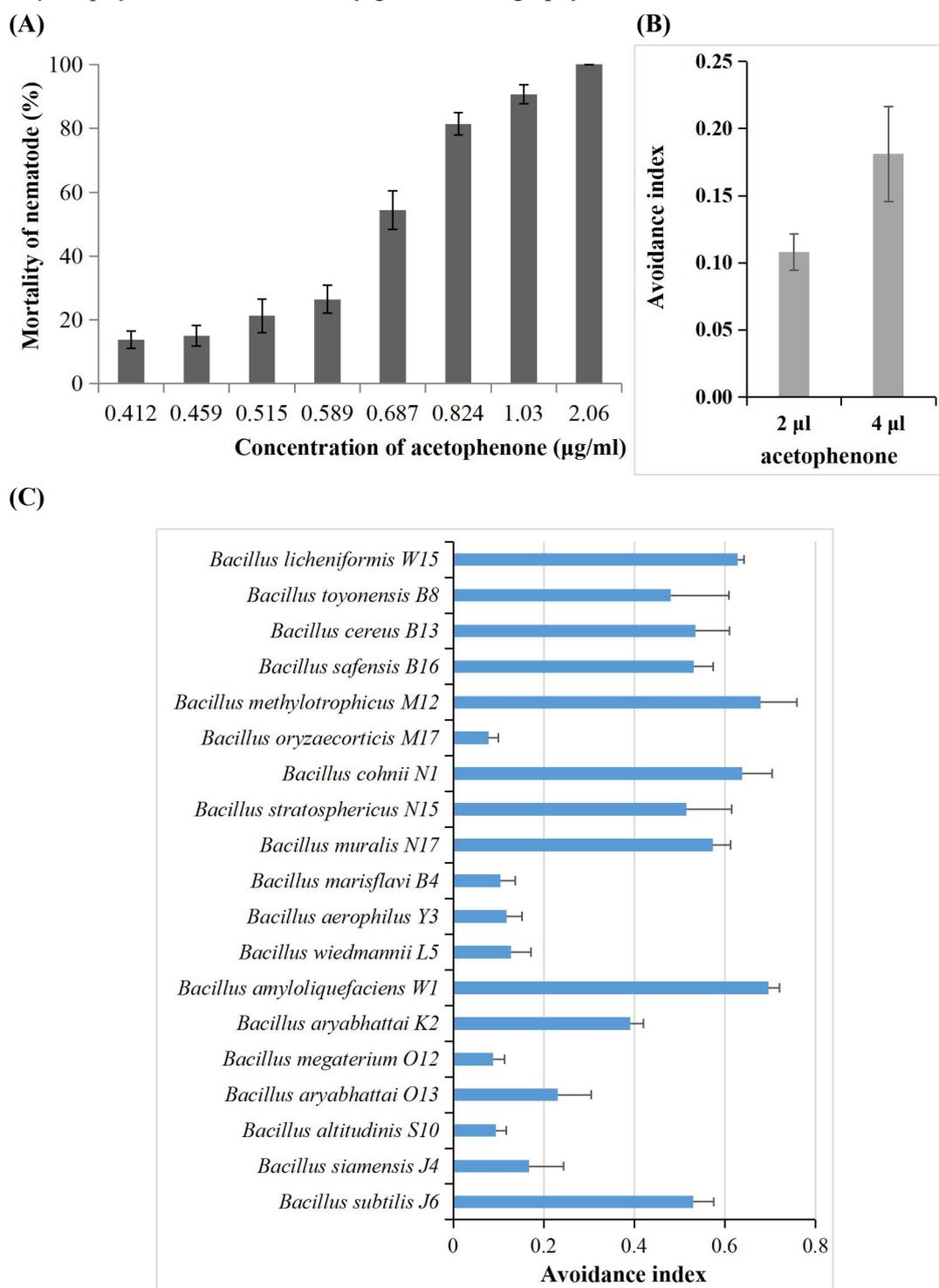
(A)



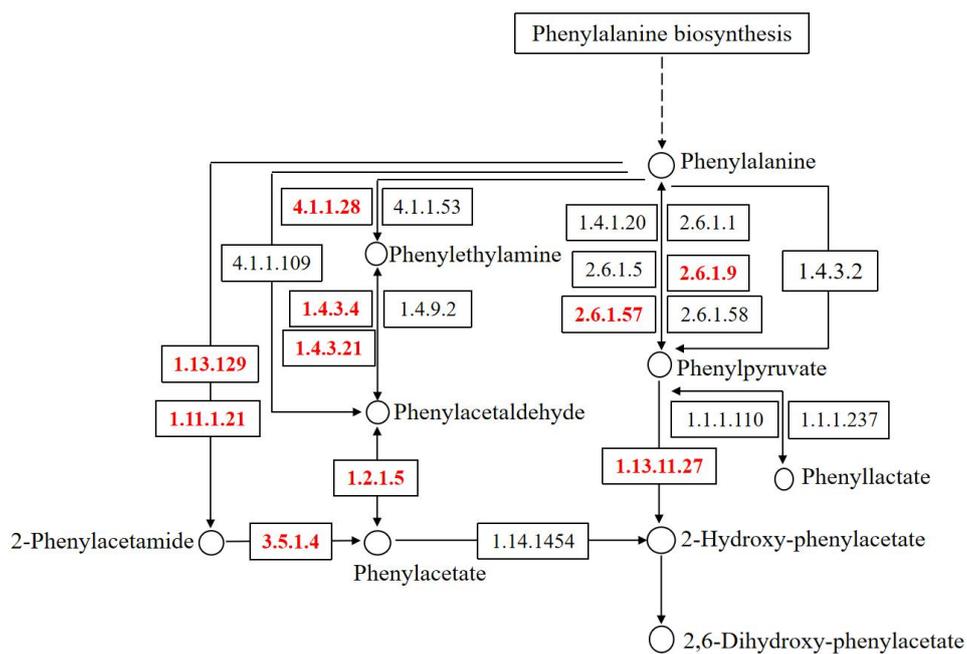
(B)



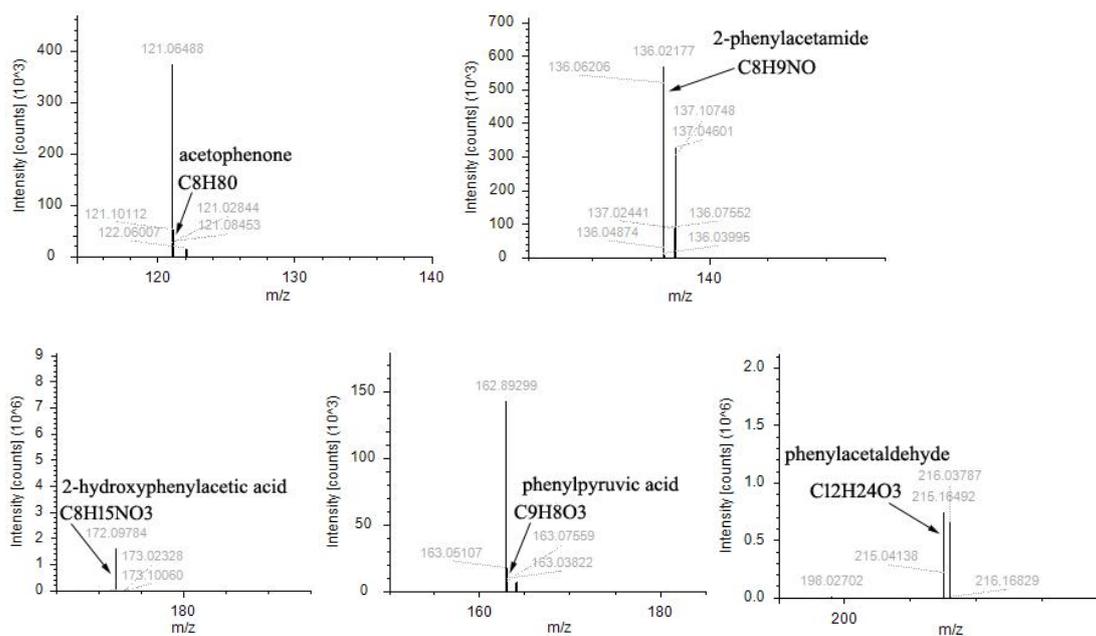
**Figure S10 Nematicidal and avoidance activities of metabolites and microbes.** (A) Nematicidal activity of different concentrations of acetophenone against *M. incognita*. (B) Avoidance of acetophenone against *M. incognita*. (C) Avoidance activity of *Bacillus* isolates against *M. incognita*. (D) Phenylalanine metabolism pathway from KEGG. The genes which could be found in metagenomic dataset were labeled by red letter. (E) MS spectra for metabolites in phenylalanine metabolism pathway detected in *B. amyloliquefaciens* W1 liquid culture by untargeted metabolomics analysis. (F) Detecting standard acetophenone by gas chromatography. (G) Detecting acetophenone produced by *B. amyloliquefaciens* W1 detected by gas chromatography.



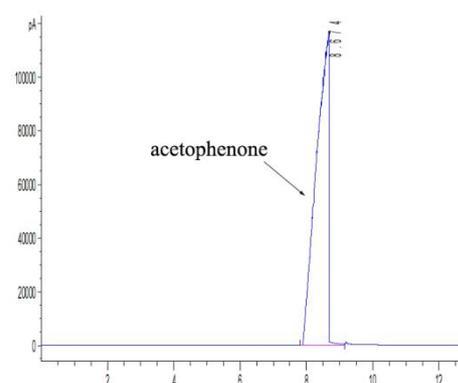
(D)



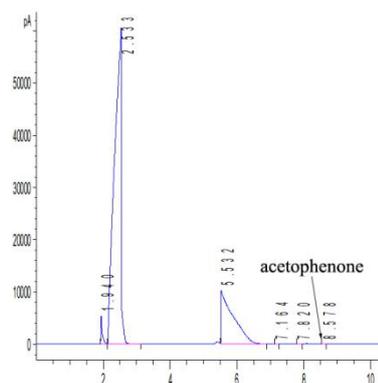
(E)



(F)



(G)



**Table S1 Soil sample sites and properties**

Soil samples	Elevations (m)	longitude	latitude	soil type	pH	SOM (g/kg)	Pesticide treatment	Planting history <sup>a</sup>
NS1	1118	109.21.2	29.57.9	yellow-brown soil	5.38	25.10	Cyhalothrin	10 years
NS2	1118	109.21.2	29.57.9	yellow-brown soil	5.35	25.21	Cyhalothrin	10 years
NS3	1119	109.21.2	29.57.10	yellow-brown soil	4.64	23.75	Cyhalothrin	10 years
NS4	1120	109.21.2	29.57.10	yellow-brown soil	5.00	24.16	Cyhalothrin	10 years
NS5	1116	109.21.1	29.57.11	yellow-brown soil	5.59	25.99	Cyhalothrin	10 years
NS6	1116	109.21.1	29.57.1	yellow-brown soil	4.96	22.71	Cyhalothrin	10 years
NS7	1127	109.20.55	29.57.18	yellow-brown soil	5.99	22.53	Cyhalothrin	10 years
NS8	1125	109.20.59	29.57.10	yellow-brown soil	5.22	22.61	Cyhalothrin	10 years
NS9	1122	109.21.18	29.57.32	yellow-brown soil	5.07	27.04	Cyhalothrin	10 years
NS10	1120	109.20.59	29.57.10	yellow-brown soil	5.92	27.28	Cyhalothrin	10 years
NS11	1118	109.20.58	29.57.10	yellow-brown soil	4.80	23.66	Cyhalothrin	10 years
NS12	1104	109.20.59	29.57.19	yellow-brown soil	5.00	30.66	Cyhalothrin	10 years
NS13	1104	109.21.0	29.57,18	yellow-brown soil	4.51	28.84	Cyhalothrin	10 years
NS14	1099	109.20.59	29.57.19	yellow-brown soil	4.89	22.39	Cyhalothrin	10 years
NS15	1098	109.20.58	29.57.59	yellow-brown soil	5.16	29.02	Cyhalothrin	10 years
RNI1	1012	109.22.38	29,58.26	yellow-brown soil	4.82	21.17	Cyhalothrin	10 years
RNI2	1012	109.22.38	29,58.26	yellow-brown soil	5.05	25.72	Cyhalothrin	10 years
RNI3	1011	109.22.38	29,58.26	yellow-brown soil	4.67	27.42	Cyhalothrin	10 years
RNI4	1007	109.22.39	29.58.28	yellow-brown soil	4.73	27.59	Cyhalothrin	10 years
RNI5	1009	109.22.39	29.58.28	yellow-brown soil	4.80	25.44	Cyhalothrin	10 years
RNI6	1006	109.22.39	29.58.28	yellow-brown soil	4.74	21.64	Cyhalothrin	10 years
RNI7	1013	109.22.43	29.58.40	yellow-brown soil	4.85	24.47	Cyhalothrin	10 years
RNI8	1013	109.22.43	29.58.40	yellow-brown soil	4.90	26.25	Cyhalothrin	10 years
RNI9	1011	109.22.43	29.58.39	yellow-brown soil	4.58	25.44	Cyhalothrin	10 years
RNI10	1012	109.22.43	29.58.39	yellow-brown soil	4.97	24.91	Cyhalothrin	10 years
RNI11	994	109.22.41	29.58.36	yellow-brown soil	4.75	26.58	Cyhalothrin	10 years
RNI12	992	109.22.41	29.58.36	yellow-brown soil	4.87	28.72	Cyhalothrin	10 years
RNI13	990	109.22.41	29.58.36	yellow-brown soil	4.83	24.39	Cyhalothrin	10 years
RNI14	990	109.22.42	29.58.37	yellow-brown soil	4.96	24.21	Cyhalothrin	10 years
RNI15	991	109.22.42	29.58.37	yellow-brown soil	5.38	21.48	Cyhalothrin	10 years

**a: years of continuously planting tobacco.**

**Table S2 Properties of non-infected (NS) and root-knot nematode infected (RNI) soils**

Soil properties	NS	RNI
<b>Number of <i>M. incognita</i></b>	<b>0.98 ± 1.01 b</b>	<b>1091.76 ± 833.87 a</b>
<b>Urease (mg / g)</b>	<b>0.18 ± 0.11 a</b>	<b>0.13 ± 0.04 a</b>
<b>Phosphatase (mg / g)</b>	<b>1.23 ± 0.51 b</b>	<b>1.95 ± 1.05 a</b>
<b>Invertase (mg / g)</b>	<b>5.23 ± 3.89 a</b>	<b>2.61 ± 1.07 b</b>
<b>Catalase (mg / g)</b>	<b>0.95 ± 0.20 a</b>	<b>0.77 ± 0.14 b</b>
<b>pH</b>	<b>5.17 ± 0.43 a</b>	<b>4.86 ± 0.19 b</b>

AK (mg / kg)	400.28 ± 145.44 b	483.48 ± 54.21 a
AP (mg / kg)	44.29 ± 8.42 a	54.89 ± 19.66 a
AN (mg / kg)	121.52 ± 33.77 a	117.39 ± 8.73 a
SOM (g / kg)	25.40 ± 2.66 a	25.03 ± 2.25 a
CEC (cmol / kg)	14.70 ± 1.95 a	12.75 ± 0.57 b
Fe (mg / kg)	91.86 ± 40.18 a	0.93 ± 0.36 b
Mn (mg / kg)	57.80 ± 25.29 a	42.98 ± 18.98 a
Zn (mg / kg)	6.61 ± 3.81 a	5.05 ± 1.82 a
Ca (mg / kg)	1188.04 ± 448.88 a	687.82 ± 236.91 b
Mg (mg / kg)	68.82 ± 35.09 a	51.83 ± 20.87 a
B (mg / kg)	0.27 ± 0.14 a	0.25 ± 0.06 a
SWC (%)	13.64 ± 1.28 a	12.68 ± 0.77 b
BD (g / cm <sup>3</sup> )	1.20 ± 0.07 b	1.25 ± 0.03 a
STP (%)	54.28 ± 2.15 a	52.75 ± 1.05 b
SCMC (%)	15.83 ± 1.66 a	14.54 ± 1.01 b
SCP (%)	18.97 ± 1.56 a	18.12 ± 1.21 a
SAP (%)	35.31 ± 2.38 a	34.63 ± 1.63 a
MWD (mm)	1.27 ± 0.41 a	0.80 ± 0.18 b

All data are presented as the mean ± SE. The properties which were significantly higher in NS soils than in RNI soils are represented by red letters. The properties which were significantly lower in NS soils than in RNI soils are represented by green letters. Different letters in the same line indicate significant ( $p < 0.05$ ) difference between NS and RNI soils.

**Table S3 Correlation coefficients and significances between soil variables and number of *M. incognita***

	Catalase	AP	CEC	Fe	Ca	BD	STP	MWD
Number of <i>M. incognita</i>	-0.38(0.04)	0.39(0.03)	-0.40(0.03)	-0.59(0.001)	-0.44(0.02)	0.37(0.047)	-0.36(0.049)	-0.43(0.02)

The data in brackets were the p values.

**Table S4 Metagenome sequence statistics after filtering**

<b>Data Clean</b>	
Total Raw Data	4,211,442 Mbp
Average Raw Data	135,852.97 Mbp
<b>Assembly and Mix-Assembly</b>	
Scaffolds (Average)	344610
Total length (nt)	9,998,862,533 bp
Average length (nt)	935.97 bp
Longest length (nt)	284,861 bp
N50 length (nt)	893.84 bp
N90 length (nt)	542.84 bp
Scaftigs (Average)	344610
Total length (nt)	9,998,862,533 bp

---

Average length (nt)	936 bp
N50 length (nt)	894 bp
N90 length (nt)	543 bp
<b>Gene Prediction</b>	
Total ORFs	16725240
Average ORFs	539524
Gene catalogue	7677442
Complete ORFs	1,812,302(23.61%)
Total length (Mbp)	4105.43
Average length (bp)	534.74
GC percent	63.42%
<b>Taxonomic Annotation</b>	
Gene catalogue	7677442
Annotated on NR	5,678,146(73.96%)
Annotated on Unclassified	34.95%
Annotated on Kingdom level	65.05%
Annotated on Phylum level	57.52%
Annotated on Class level	52.52%
Annotated on Order level	48.42%
Annotated on Family level	43.30%
Annotated on Genus level	39.22%
Annotated on Species level	29.79%
Assigned Phyla(top 5)	Proteobacteria
Sign_diff Phyla(top 5)	k__Bacteria;p__Proteobacteria
<b>Functional Annotation</b>	
Gene catalogue	7677442
Annotated on KEGG	4,698,328(61.20%)
Annotated on KO	2,616,196(34.08%)/8,019
Annotated on EC	1,660,255(21.63%)/2,760
Annotated on pathway	1,687,236(21.98%)/405
Annotated on eggNOG	4,442,108(57.86%)
Annotated on OG	4,442,108(57.86%)/45,236
Annotated on CAZy	255,379(3.33%)
<b>CARD Annotation</b>	
Gene catalogue	7677442
Annotated on CARD	1722
Annotated AROs	735

---

**Table S5 CDS statistics of each sample**

Treatment	Sample ID	ORFs number	Total length (Mbp)	Average length (bp)
Non-infected soils	NS1	437,836	226.38	517.04
	NS2	390,789	203.19	519.96
	NS3	759,756	407.73	536.66
	NS4	509,777	259.85	509.72
	NS5	312,173	152.04	487.04
	NS6	579,246	302.07	521.49
	NS7	392159	202.63	516.7
	NS8	777797	429.97	552.8
	NS9	483135	255.6	529.04
	NS10	236114	115.98	491.19
	NS11	578,225	313.07	541.43
	NS12	517,214	271.54	525.01
	NS13	784483	418.17	533.05
	NS14	404746	196.39	485.23
	NS15	151614	70.72	466.43
Root-knot nematode infected soils	RNI1	742392	402.52	542.2
	RNI2	564414	296.43	525.2
	RNI3	704704	395.78	561.63
	RNI4	904021	496.51	549.23
	RNI5	946039	528.48	558.62
	RNI6	582316	311.18	534.39
	RNI7	534738	276.24	516.59
	RNI8	288109	144.52	501.61
	RNI9	539190	301.05	558.34
	RNI10	570671	301.03	527.5
	RNI11	517614	272.63	526.7
	RNI12	511731	266.75	521.26
	RNI13	604914	319.65	528.42
	RNI14	547648	285.91	522.06
	RNI15	331647	163.74	493.73

**Table S6 Bacterial isolates showed low nematocidal activity**

Strains	Mortality of <i>M. incognita</i>
<i>Pseudomonas plecoglossicida</i> SX4	7.04 ± 1.55
<i>Pseudomonas entomophila</i> SX5	3.86 ± 1.69
<i>Pseudomonas putida</i> SX6	5.02 ± 1.86
<i>Pseudomonas lutea</i> SX17	1.71 ± 1.30
<i>Paenarthrobacter nicotinovorans</i> SX-31	0.36 ± 1.24
<i>Ensifer adhaerens</i> SX32	1.36 ± 1.01
<i>Arthrobacter nitroguajacolicus</i> SX33	1.40 ± 1.03
<i>Pseudomonas pseudoalcaligenes</i> SX-40	3.42 ± 1.80

<i>Curtobacterium citreum</i> SM-17	4.83 ± 1.87
<i>Paenarthrobacter ureafaciens</i> SM-18	1.78 ± 0.22

**Table S7 Gene nodes enriched in the NS function network**

Genes	Category	Function	Module	Relative Abundance
COG2175	Category Q	dithiol-disulfide isomerase	NS0	0.0289
COG3508	Category Q	homogentisate 1,2-dioxygenase	NS5	0.0227
COG4664	Category Q	trap dicarboxylate transporter	NS1	0.0216
COG4663	Category Q	extracellular solute-binding protein	NS1	0.0188
COG2368	Category Q	4-hydroxyphenylacetate	NS0	0.0117
COG2761	Category Q	DSBA oxidoreductase	NS6	0.0098
COG2854	Category Q	Toluene tolerance family protein	NS4	0.0081
COG4665	Category Q	Tripartite ATP-independent periplasmic transporter	NS0	0.0072
ENOG410XSW0	Category Q	amidohydrolase	NS0	0.0056
COG0466	Category O	protease	NS1	0.1075
COG0465	Category O	metalloprotease	NS1	0.0849
COG0330	Category O	protease	NS1	0.0739
COG1219	Category O	protease	NS1	0.0514
COG1220	Category O	protease	NS1	0.0306
COG0826	Category O	collagenase	NS0	0.0072

**Table S8 The abundances of biosynthetic gene clusters (BGCs) in two groups**

BGC	Abundance in NS	Abundance in RNI
	microbiota	microbiota
terpene	1977	1447
NRPS	1398	1221
arylpolyene	839	389
bacteriocin	703	453
betalactone	555	343
polyketides	665	477
acyl amino acids	105	55
hglEKS	98	67
hserlactone	82	69
resorcinol	71	46
siderophore	48	38
phosphonate	39	31
ectoine	31	11
ladderane	20	12
thiopeptide	18	4
indole	15	12
TfuA	12	3
butyrolactone	12	3
LAP	10	6

**Table S9 Correlation of metabolites and number of *M. incognita***

Name of metabolites	Number of <i>M. incognita</i>
2-Methyl-4,6-dinitrophenol	-0.725 (0.02)
Dodecanedioic acid	-0.688 (0.03)
2-hydroxymyristic acid	-0.649 (0.04)
11-Deoxycortisol	-0.840 (0.002)
Indole-3-acetic acid	-0.640 (0.046)
2-Methylbenzoic acid	-0.657 (0.04)
10-Hydroxydecanoic acid	-0.822 (0.003)
2-Naphthalenesulfonic acid	-0.655 (0.04)
8-Amino-7-oxononanoic acid	-0.824 (0.003)
Acetyl-N-formyl-5-methoxykynurenamine	-0.664 (0.04)
6-Acetylmorphine	-0.638 (0.047)
Abrine	-0.820 (0.004)
6-O-acetyl-beta-D-glucose	-0.736 (0.02)
Dimethyl octanediimidate	-0.637 (0.048)
Oxatomide	-0.754 (0.01)
L-Tryptophan	-0.806 (0.005)
Acetophenone	-0.634 (0.049)
Triphenylphosphine oxide	-0.665 (0.04)
Avocadyne 1-acetate	-0.736 (0.02)
Galantamine	-0.764 (0.01)

**Table S10 Metabolites in *Bacillus amyloliquefaciens* W1 culture**

Name	Formula	Molecular Weight	RT [min]	Relative quantification
acetophenone	C8H8O	120.15	10.86	372347971.8 ± 88929054.2
2-phenylacetamide	C8H9NO	135.16	12.69	12758259.5 ± 692367.1
2-hydroxyphenylacetic acid	C8H15NO3	173.11	6.71	174202911.7 ± 56866399.8
phenylpyruvic acid	C9H8O3	164.05	0.71	63467095.7 ± 5268587.4
phenylacetaldehyde	C12H24O3	216.17	11.16	70325975.3 ± 40011957.3