High-throughput sequencing-based analysis of the composition and diversity of endophytic bacterial community in seeds of upland rice

Zhishan Wang
University of Science and Technology Beijing
Yongqiang Zhu
Chinese National Human Genome Center at Shanghai
Ruixue Jing
University of Science and Technology Beijing
Xianyu Wu
University of Science and Technology Beijing
Ni Li
Hunan Hybrid Rice Research Center
Hai Liu
Hunan Hybrid Rice Research Center
Xiaoxia Zhang
Chinese Academy of Agricultural Sciences
Weiping Wang
Hunan Hybrid Rice Research Center
Yang Liu (ly81150@163.com)
University of Science and Technology Beijing https://orcid.org/0000-0003-3261-7243

Original article

Keywords: Upland rice, Endophytic bacteria, Diversity, Soil, Seed

Posted Date: April 27th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-23885/v1

Version of Record: A version of this preprint was published at Archives of Microbiology on September 29th, 2020. See the published version at https://doi.org/10.1007/s00203-020-02058-9.

Abstract

Upland rice is an ecotype crop formed by long-term domestication and evolution of rice in the dry land without water layer. Generally, its stem and leaf are thick and luxuriant, its leaf is wide and light, its root system is developed, its root hair is abundant, its osmotic pressure of root and cell juice concentration of leaf are high, and it is drought resistant, heat-resistant and water absorbing. The purpose of this study is to reveal the “core flora” of endophytes in upland rice seeds by studying the diversity and community structure of endophytes in upland rice seeds, and to reveal the impact of soil environment on the formation of endophyte community structure in upland rice seeds by comparing with soil environment microorganisms in upland rice habitats. In this study, the high-throughput sequencing technology based on the Illumina Hiseq 2500 platform was used to study the structure and diversity of endophytic bacterial communities using upland rice varieties collected in different places and soil samples from their unified planting sites as materials. There are 42 endophytic OTUs coexisted in the 14 samples. At the phylum level, the first dominant phyla was Proteobacteria (93.81–99.99%) in all 14 samples. At the genus level, *Pantoea* (8.77% -87.77%), *Pseudomonas* (1.15–61.58%), *Methylobacterium* (0.40–4.64%), *Sphingomonas* (0.26–3.85%), *Microbacterium* (0.01–4.67%) and *Aurantimonas* (0.04–4.34%), which are probably the core microflora in upland rice seeds, served as the dominant genera that coexisted in all upland rice seeds tested. Compared with the soil microbial community structure in the upland rice uniform planting site, it was found that it had little effect on the endophytic community structure in upland rice seeds. This study is of great significance for the isolation, screening, functional evaluation and re-action of some functional microorganisms in upland rice in order to improve its agronomic traits. It also provides a certain reference for the interaction between microorganisms and plants.

Introduction

Endophyte is a kind of microorganism that lives in the tissues and organs of healthy plants at certain or all stages without causing substantial damage to the host plants and establishing a harmonious joint relationship with plants. Endophytes usually enter the plant through the root area, but they can also enter through the aboveground part of the plant. Once they enter the root, endophytes can infect the adjacent plant tissues. Therefore, endophytes are ubiquitous in the plant host, including the aboveground and underground parts of the plant, even seeds, thus having a positive impact on the plant development (Zinniel et al. 2015; Chebotar et al. 2015). Endophytes can provide a variety of benefits to host plants, especially through multiple functions to improve the growth and health of plants, so that they can be protected from pathogens (Agler et al. 2016; Liu et al. 2017a,b; Xu et al. 2019; Afzal et al. 2019; Sánchez-Cruz et al. 2019). Under different environmental conditions, the communication and interaction between endophyte and plant is stronger than that between rhizosphere bacteria (Coutinho et al. 2015). A number of studies have shown that plant varieties, genotypes and geographical location have a great impact on the establishment of microbial diversity and community structure in plants (Liu et al. 2017a, 2019; Edwards et al. 2015). Although endophyte was discovered more than 100 years ago, its existence has been ignored until 1930s. As a new microbial resource, endophyte has been widely concerned. Especially

in recent years, the interaction between endophyte and plant has gradually become a research hotspot in the field of plant science, agronomy, thremmatology and ecology (Vandenkoornhuyse et al. 2015; Wei et al. 2017).

As a special reproduction of gymnosperms and angiosperms, seeds can survive for decades in suspension. When the external environment is suitable, they will germinate and grow rapidly into new plants (Nelson 2004; Steinbrecher and Leubner-Metzger 2017). It has been proved that there are a lot of special endophytic or epiphytic microflora in seeds besides the basic nutrients needed for development, and the composition and vertical transmission of these microflora may directly or indirectly affect seed germination, plant growth and health (Links et al. 2014; Haimin et al. 2014; Nelson 2017; Shade et al. 2017; Truyens et al. 2015). Truyens et al. (2015) found that the main endophytic bacteria in seeds belong to γ - Proteobacteria phylum, followed by Actinobacteria, Firmicutes and Bacteroidetes phyla. Microorganisms such as *Bacillus*, *Paenibacillus*, *Pseudomonas*, *Micrococcus*, *Staphylococus*, *Acinetobacter* and *Pantoea* are the most easily detected and isolated in various plant seeds (Zhang et al. 2018). Compared to the researches on rhizosphere and foliar microbial communities, the understanding of endophytes in seeds is still limited.

With the development and promotion of biotechnology, the research methods of endophytes in plant seeds have also changed. In the past reports, the researchers used the methods of conventional culture-dependent method (Liu et al. 2009; Jiang et al. 2013), 16S rDNA clone library technique (Liu et al. 2012; Zou et al. 2012; Chen et al. 2014; Haimin et al. 2014) and deformable gradient gel electrophoresis (DGGE) to study endophytes (Hardoim et al. 2012). In recent years, with the rise of high-throughput sequencing technology, it has been widely used in plant seed endophyte community structure and diversity analysis. Liu et al. (2017a; 2020) analyzed the endophytic community structure and diversity of different corn seeds by high-throughput sequencing technology, and then speculated the impact of endophytic plants on the environmental adaptability and vertical propagation of different corn varieties, providing basis for the cultivation of Corn Varieties in the future. Rybakova et al. (2017) analyzed the microbial composition of *Brassica napus* seeds by high-throughput sequencing technology, and found that the microbial structure mainly depends on the cultivated varieties, and will affect the interaction between symbionts and pathogens. Wang et al. (2016) used the same method to preliminarily reveal the core Actinobacteria in rice roots, stems and grains.

Upland rice is an ecological crop formed by long-term domestication and evolution of rice in the dry land without water layer. Compared with traditional rice crops, upland rice has stronger ability of drought resistance and drought tolerance, can adapt to the arid climate, and can greatly save labor and production costs (Kumar and Ladha, 2011; Xia et al. 2019). However, up to now, the researches on upland rice mainly focus on the functional genes, genetic diversity and growth promoting microorganisms (Lyu et al. 2014; Tuhina-Khatun et al. 2015; Ferrari et al. 2018; Braga et al. 2018; de Sousa et al. 2018), and there are few studies on endophytes of upland rice at home and abroad, especially the high-throughput sequencing technology.

In order to gain a deeper understanding of the community structure and diversity of endophytic bacteria in upland rice seeds, further summarize the “core flora” of endophytic bacteria in upland rice and infer whether the soil environment is related to the formation of endophytic bacteria in upland rice seeds and plant growth and development, high-throughput sequencing and related bioinformatics analysis based on the Illumina Hiseq 2500 platform in this study were used to study 14 different varieties of upland rice seed samples and soil samples collected from the Sanya Division Farm in Hainan Province, where they were grown intensively. It is worth mentioning that 11 of the 14 different varieties of upland rice seed resources are collected in different regions of the country, and the other 3 are imported varieties. The samples are not only precious and rare, but also very representative.

Materials And Methods

Upland rice seeds and soil sampling

The 14 upland rice seeds collected in different places used in this research were provided by Hunan Hybrid Rice Research Center. Detailed information about the rice seeds is shown in Table 1. All samples were uniformly planted on the Sanya Division Farm in Hainan Province, and seed samples were collected for processing after maturity. Soil samples were taken using a five-point sampling method from the periphery. First, remove the surface topsoil, use an alcohol burning shovel to dig out the soil layer of 5 ~ 20 cm underground, then remove the visible impurities and make the soil pass the 2mm sieve, mix them evenly, store them in five sterile centrifuge tubes, and transport them back to the laboratory under 0℃ for DNA extraction.

Sample surface sterilization and treatment

One replicate of each sample were collected in this study. Firstly, the husks of each upland rice seed sample were removed by a small sheller. Then, under aseptic conditions the following operations were performed in the order listed: husked seeds were washed 3 times with prepared sterile water; 2.5 g of seeds were placed in a clean and sterile 50 mL tube containing 25 mL of phosphate buffer (per litre: 7.15 g of $NaH_2PO_4 \cdot 2H_2O$, 22.04 g of $Na_2HPO_4 \cdot 12H_2O$, 200 µL of Silwet L-77) (Zhang et al. 2018; Liu et al. 2019), and the seeds were sonicated twice by an Ultrasonic Processor Scientz-IID sonicator (NingBo Scientz Biotechnology Co., Ltd., China) at low power (237.5 W; 950 W×25%) in an ice bath for 5 min (alternating thirty 2-s bursts and thirty 2-s rests) (Lundberg et al. 2012; Zhang et al. 2018; Liu et al. 2019). To validate that the surface was sterilized, sterile tweezers were used to press surface-sterilized seeds into LB medium (LUQIAO), and the samples were incubated at 30 °C for 72 h.

DNA extraction

Five gram of surface-sterilized upland rice seeds from each sample was frozen with liquid nitrogen and was quickly ground into a fine powder with a pre-cooled sterile mortar, and then the DNA was extracted using the FastDNA® SPIN Kit for Soil (MP Biomedicals, Solon, OH, USA) following the manufacturer’s instructions of the Kit. Soil total genomic DNA was extracted from fully mixed 20 g of soil using the

FastDNA® SPIN Kit for Soil (MP Biomedicals, Solon, OH, USA) following the manufacturer's instructions. The DNA was extracted immediately after sampling in order to maintain uniformity for comparing the bacterial communities.

Amplicon library preparation and sequencing

All PCR amplifications were performed using TransStartFastPfu DNA Polymerase (TransGen,Beijing, China). For rice seeds, 799F (5'-AACAGGATTAGATACCCTG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') was used for the first-round amplification. Then the 750bp fragment amplified from endophytic bacteria was cut and used as the template for the second-round amplification for V6-V8 region (968F: 5'-AACGCGAAGAACCTTAC-3' and 1378R: 5'-CGGTGTGTACAAGGCCCGGGAACG-3'). For soil, the V6-V8 region was amplified using 968F (5'-AACGCGAAGAACCTTAC-3') and 1378R (5'-CGGTGTGTACAAGGCCCGGGAACG-3') which contain five to eight bases barcode. All of the thermocycling steps are as follows: 5 min at 95 ℃; 20 cycles of 45 sec at 95 ℃, 30 sec at57 ℃, and 30 sec at 72 ℃. The amplified products were purified and mixed in equivalent amounts. After using the TruePrep® DNA Library Prep Kit V2 for Illumina (Vazyme, China) to perform the library, all samples were sequenced on MiSeq instrument using the MiSeq® Reagent Kit v3 (600 cycle) (Illumina).

Sequence data processing

The assembly of paired FASTQ files was performed by Mothur (version 1.39.0) (Schloss et al. 2011). Briefly, paired sequence reads were assembled after removing raw reads with ambiguous bases or low quality, such as read length < 50 bp, average Qscore < 25, or reads not matching the primer (pdiffs = 0) and barcode (bdiffs = 0). The high-quality DNA sequences were aligned to SILVA reference database (V119) (Quast et al. 2013), and using chimera.uchime module to remove chimerasequences. Then the reads were classified and grouped into OTUs (Operational taxonomic units) under the threshold of 97% identity.

Data statistics

Community richness, evenness and diversity analysis (Shannon, Simpsonenven, ACE, Chao and Good's coverage) were performed using Mothur. Both PCoA and NMDS were analyzed based on thetayc matrix by mothur. The t-test (with 95% confidence intervals) was used to determine whether the means of evaluation indices were statistical difference, and p-value < 0.05 were considered as significant standard. Taxonomy was assigned using the online software RDP classifier (Wang et al. 2007) at default parameter (80% threshold) based on the Ribosomal Database Project (Cole et al. 2009). Genera and family abundance differences between samples were analyzed by Metastats (White et al. 2009). Spearmancorrelation coefficient between two variables was calculated usingthe R command "cor.test".

Results

Quality control of sequencing data

Based on the barcode and front-end primer information, the quality control sequences were separated into 14 sets of sequence files, and a total of 629,504 high-quality sequences were obtained, with an average of 44,964 sequences per sample (Table 2), and a minimum of 28,388 reads was applied as the criteria for data normalization.

The diversity of endophytic bacteria in upland rice seeds

At 97% sequence similarity level, the 16S rRNA gene obtained was divided into 11485 OTUs according to the different distances between the sequences, and the diversity indexes, including ACE, Chao, Shannon andSimpson values, are shown in Table 3. ACE and Chao values are used for sample abundance assessment, while Shannon and Simpson values are used for sample diversity assessment. In general, 19H001, 19H002, 19H005, 19H009, 19H010, 19H016 and 19H020 have significant differences in ACE and Chao values from other samples (Table 3), while 19H001, 19H005, 19H009, 19H010, 19H016 and 19H020 have significant differences in Shannon and Simpson values from other samples (Table 3), and the endophytic bacterial diversity and richness of sample 19H005 were greatest, while the richness is lowest in sample 19H010. And rarefaction curve and Rank abundance curve also showed that the diversity and richness of sample 19H010 was significantly lower than others (Fig. 1 and Fig. 2).

In order to show the shared endophytic OTUs between duplicate samples, the statistical results found that 42 endophytic OTUs coexisted in the 14 sample seeds. And each sample contained its unique OTUs. The proportions of unique OTUs were 51.28%, 33.88%, 40.49%, 53.69%, 46.98%, 46.83%, 50.38%, 55.38%, 45.57%, 40.96%, 48.15%, 57.22%, 43.89%, 51.90% in sample 19H001, 19H002, 19H004, 19H005, 19H009, 19H010, 19H016, 19H020, 19H023, 19H025, 19H026, 19H027, 19H028 and 19H031, respectively, which indicated that differences in upland rice genotypes to some extent have an impact on endophyte composition in seeds. Furthermore, PCA results can reflect similarity in the endophytic community structures among the different upland rice seed genotypes. As shown in Figure 3, the different samples were able to separate from each other on both the PC1 (10.09%) and PC2 (9.20%) axes in PCA, but the separation distance is not very large, which illustrated that the endophyte community structures in the fourteen upland rice seed samples were different, but the differences were not significant.

Endophytic bacterial community structures in upland rice seeds

As shown in Figure 4, at the phylum level, Proteobacteria was the dominant endophytes that coexisted, in different proportions, in each of the different samples of upland rice seeds. The results indicated that the abundances distribution of bacterial endophytes had some discrepancies among the different upland rice seed samples, which was in line with the results obtained in the heat map in Figure 5. In the different upland rice seed samples, the abundance of the dominant phyla Proteobacteria ranged from 93.81 to 99.99%. And the abundance of the dominant phyla Actinobacteria ranged from 0.06 to 5.47% in the

different upland rice seed samples except sample 19H002. At the genus level, every upland rice seed sample also had dominant endophytes, including mainly *Pantoea* (8.77%-87.77%), *Pseudomonas* (1.15-61.58%), *Methylobacterium* (0.40-4.64%), *Sphingomonas* (0.26-3.85%), *Microbacterium* (0.01-4.67%) and *Aurantimonas* (0.04-4.34%) (Fig. 6). The dominant genera of every upland rice seed sample is shown in Table 4, and the results indicated that the abundances distribution of bacterial endophytes had distinct discrepancies among the different upland rice seed samples. According to the above results, the commonly dominant genera in upland rice seeds were *Pantoea*, *Pseudomonas*, *Methylobacterium*, *Sphingomonas*, *Microbacterium*, *Aurantimonas*, *Agrobacterium*, *Curtobacterium*, *Erwinia*, *Buttiauxella* and *Magnetospirillum.*

Community structure and diversity of bacteria in soil samples

Based on the barcode and front-end primer information, the quality control sequences were separated into 5 sets of sequence files, and a total of 266,660 high-quality sequences were obtained, with an average of 53,332 sequences per sample (Table 5), and a minimum of 43,869 reads was applied as the criteria for data normalization.

At 97% sequence similarity level, the 16S rRNA gene obtained was divided into 63521 OTUs according to the different distances between the sequences, and the diversity indexes, including ACE, Chao, Shannon andSimpson values, are shown in Table 6. In general, with the exception of HNSYSB_3 and HNSYSB_5, ACE, Chao, Shannon and Simpson values have no significant differences among the soil samples (Table 6), and the endophytic bacterial diversity and richness of sample HNSYSB_5 were greatest, while the richness is lowest in sample HNSYSB_3. And rarefaction curve (Fig. 7) also showed that the diversity and richness of sample HNSYSB_3 was significantly lower than others, while sample HNSYSB_5 was significantly higher than others. Despite this, the results of Rank abundance curve of five soil samples (Fig. 8) showed that the richness and uniformity of the species among these soil samples have no significant differences.

All five soil samples were used to create the Venn diagram shown in Fig. 9, which was used to investigate whether shared endophytic OTUs exist. At a 97% similarity level, the numbers of OTUs for sample HNSYSB_1, HNSYSB_2, HNSYSB_3, HNSYSB_4 and HNSYSB_5 was 16120, 16517, 12877, 14975 and 18637 respectively. Among them 1323 endophytic OTUs coexisted in the five soil samples. And each sample contained its unique OTUs. The proportions of unique OTUs were 69.58%, 70.48%, 66.91%, 67.79% and 74.24% in sample HNSYSB_1, HNSYSB_2, HNSYSB_3, HNSYSB_4 and HNSYSB_5, respectively. At the phylum level, Proteobacteria, Chloroflexi, Acidobacteria and Actinobacteria were the dominant endophytes that coexisted, and Proteobacteria (24.93-31.24%) was the first dominant phyla shared in all five samples, and other dominant shared phyla belonged to Chloroflexi (14.36-20.35%), Acidobacteria (10.50-20.68%), and Actinobacteria (3.85-24.50%) (Fig. 10). At the genus level, there were a large number of endophytic bacteria without clear classification status in the soil samples, which was most likely due to the complexity of the soil itself (Fig. 11).

The similarity of the multi-sample community structure of each sample at the OTU =0.03 level was calculated based on the thetayc algorithm, and then the 5 samples were hierarchically clustered to reflect whether the bacterial communities were different among the soilsamples. The results showed that the similarity of community structure between sample HNSYSB_1 and HNSYSB_5 was relatively large, while the similarity of community structure between sample HNSYSB_3 and others was smaller, and the results obtained by PCA analysis were also consistent with it (Fig. 12a). In addition, the heat map showed that there were some differences in the abundance of the five groups of soil samples (Fig. 12b).

Comparison of upland rice seed and soil samples

Although Proteobacteria was the main dominant phyla in both upland rice seed samples and soil samples, PCA and NMDS analysis showed that there were obvious differences in endophytic bacterial community structure among them (Fig. 13a,b). And the heat map results of the top ten dominant bacterial genera of 19 samples also showed that there were significant differences between upland rice seed samples and soil samples (Fig. 14). From this we can infer that the soil samples from the breeding ground of upland rice did not have a substantial effect on the formation of endophytic bacteria in upland rice seeds.

Discussion

Rice, as the most important cereal crop in the world, has four ecosystems: irrigation, rainfed lowland, deep water and rainfed highland. Although most of rice cultivation is based on deep water ecosystem and irrigation ecosystem, drought has always been one of the most disastrous pressure sources in rice cultivation, causing serious loss of rice annual output (Luo 2010; Guimarães et al. 2016). Compared with traditional rice planting, upland rice ecosystem has stronger water-saving ability and climate adaptability, and can greatly reduce greenhouse gas emissions. However, due to the problems of water supply, weed infection, diseases and pests, lack of suitable cultivated varieties, its productivity is limited by nearly 50% of the production potential, so upland rice ecosystem accounts for the proportion of global rice production area is very small, about 10%, mainly concentrated in some mountainous areas of Asia, Africa and Latin America, growing under aerobic and rain fed conditions (Galinato et al. 1999; Bernier et al. 2008; Kumar and Ladha, 2011; Bridhikitti and Overcamp, 2012). The related research of upland rice is of great significance for rice breeding and drought resistance. In addition to the genetic level and external environment, the endophyte of upland rice is also one of the very important factors, especially in recent years, more and more reports show that the endophyte of upland rice has a very important impact on plant health and growth, so this study uses high-throughput sequencing technology to explore endophytic bacteria of upland rice seeds in many regions and reveal their community structure and diversity, which is not only a unique perspective innovation, but also of great significance for the follow-up research of upland rice.

Endophytic bacteria are common in many plants, but up to now, most of the researches on endophytic Bacteria in rice seeds have been focused on rice, and there are few reports on endophytic Bacteria in upland rice seeds. In this research, 629,504 high-quality sequences and 11485 OTUs of endophytic bacteria were obtained via high-throughput sequencing, which could basically reflect the composition of the upland rice seed endophytes. The major microbial groups are similar among the 14 upland rice seed samples, and the results showed that 42 endophytic OTUs were shared by all the fourteen upland rice seed samples. Previous reports have shown that Proteobacteria, Firmicutes, Bacteroidetes and Actinobacteria were commonly found in plant seeds (Truyens et al. 2015; Hardoim et al. 2012), and Proteobacteria was also the main dominant phyla of endogenous bacteria in upland rice seeds in this study.

The major microbial groups and community structure in this study are similar among the fourteen upland rice samples. The overlapping areas of the Venn diagram are often used to represent shared microbiota between multiple samples from a particular ecological scale (Vandenkoornhuyse et al. 2015), but in this study due to the large number of samples, the Venn diagram could not be drawn. The Venn diagram data (Supplementary Table S1) of the fourteen genotypes showed that 0.37% of the endophyte OTUs was shared by all the samples, of which abundance accounted for > 2.04% of the total OTUs in each genotype. The major shared genera were *Pantoea*, *Pseudomonas*, *Methylobacterium*, *Sphingomonas*, *Microbacterium*, *Aurantimonas*, *Agrobacterium*, *Curtobacterium*, *Erwinia*, *Buttiauxella* and *Magnetospirillum. Pantoea* (8.77%-87.77%) was the first dominant genera shared in all fourteen samples, and other dominant shared genera belonged to *Pseudomonas* (1.15–61.58%), *Methylobacterium* (0.40–4.64%), *Sphingomonas* (0.26–3.85%),*Microbacterium* (0.01–4.67%) and *Aurantimonas* (0.04–4.34%). According to related research reports, *Pantoea* is the dominant genus of many plants, including rice, *Salvia miltiorrhiza*, maize, turf grass, mulberry, hemp and so on (Zhang et al. 2018; Chen et al. 2018; Liu et al. 2017a; Chen et al. 2020; Xu et al. 2019; Scott et al. 2018), and some species in *Pantoea* are often reported to promote plant growth. For example, *Pantoea agglomerans* C1 exhibits high biotechnological potential as plant growth-promoting bacterium in heavy metal polluted soils (Luziatelli et al. 2020), and *Pantoea agglomerans* Pa can not only promote considerable growth of wheat seedlings, high chlorophyll content, lower accumulation of proline, and favored K^+ accumulation in the inoculated plants, but also has significant salt stress relief and PGP activities. *Pantoea* sp. Strain 1.19 isolated from rice rhizosphere has effective capacity to promote growth of Legumes and Nonlegumes (Megías et al. 2017). In addition, Nascimento et al. (2020) also found that *Pantoea phytobeneficialis* MSR2 possess an astonishing number of plant growth promotion genes as a rare group of Pantoea strains, including those involved in nitrogen fixation, phosphate solubilization, 1-aminocyclopropane-1-carboxylic acid deaminase activity, indoleacetic acid and cytokinin biosynthesis, and jasmonic acid metabolism.

Other dominant genus *Pseudomonas, Methylobacterium*, *Sphingomonas*, *Microbacterium*, *Aurantimonas*, *Agrobacterium, Curtobacterium, Erwinia, Buttiauxella and Magnetospirillum* are also common endophytic bacterial groups in plant seeds. And *Pseudomonas, Sphingomonas* and *Microbacterium* as endophytic dominant genera were also isolated or detected from seeds of rice, peanuts, browntop millet, Marama

bean and so on (Liu et al. 2019; Zhang et al. 2018; Jiang et al. 2013; Sobolev et al. 2013; Verma and White 2018; Chimwamurombe et al. 2016). It has been reported that many species of *Pseudomonas, Sphingomonas* and *Microbacterium*can promote plant growth by improving the nitrogen-fixing ability of plants or secreting secondary metabolites with antibacterial and plant growth benefits. For example, nitrogen-fixing *Pseudomonas stutzeri* A15 isolated from the rhizosphere and inner layer of rice has a significant growth promotion effect on rice seedlings under greenhouse conditions, and the nitrogen-fixing effect of this strain is better than that of ordinary chemical nitrogen fertilizers (Pham et al. 2017). *Pseudomonas chlororaphis* and *Pseudomonas aurantiaca* isolated from cacti, cotton and gramineous plants can produce secondary metabolites and promote plant growth (Shahid et al. 2017). Study by Chu et al. (2019) found that the strain *Pseudomonas* PS01 isolated from corn rhizosphere can increase the germination rate of *Arabidopsis* seeds under high salt stress. It was found that the strains *Sphingomonas trueperi* NNA-14, *Sphingomonas trueperi* NNA-19, *Sphingomonas trueperi* NNA-17, *Sphingomonas trueperi* NNA-20 isolated from the rhizosphere soil and stems and roots of giant reed and switchgrass can promote plant growth, of which NNA-14 Significantly increased the root length, specific surface area and fine root number of corn, increased the N, Ca, S, B, Cu and Zn content of corn, while NNA-19 significantly increased the root dry weight and root tip number of wheat, and increased wheat Calcium content (Xu et al. 2018). The functions of *Sphingomonas* sp. cra20, which can improve the growth rate of *Arabidopsis* plants under drought stress, promote the development plasticity of *Arabidopsis* root system and change the microbial community structure in rhizosphere, have also been proved by the researchers (Luo et al. 2019). In addition, Liu et al. (2015) Isolated two yellow bacterial strains NBD5^T and NBD8 from the nori branch, and identified them as a new species of *Sphingomonas*, which provided precious biological resources for further study on the relationship between *Sphingomonas* and plant growth. Study have found that the volatiles from root-associated bacteria of the genus *Microbacterium* can promote the growth of different plant species, and can promote plant growth without direct and long-term contact between bacteria and plants, and this process may be through regulation Sulfur and nitrogen metabolism is achieved (Cordovez et al. 2018). At the same time, research shows that secondary metabolites produced by some species of *Microbacterium* have a significant inhibitory effect on some pathogenic bacteria of the host plant (Lopes et al. 2015; Mannaa et al. 2017; Savi et al. 2018). Researchers often isolate *Methylobacterium* from plants such as rice, barley, and legumes (Liu et al.2019; Chen et al. 2019; Tani et al. 2015; Andrews and Andrews 2017). And several species of *Methylobacterium* have been reported as plant growth-promoting bacteria (PGPB) and Hg-resistant bacteria (Durand et al. 2017; Antunes et al. 2017; Grossi et al. 2020). In addition, Tani et al. (2015) found that *Methylobacterium* can not only use methanol emitted by plants as a carbon source and energy source, promote plant maturity, increase particle size, but also show super-strong species level on a given plant species Selection pressure is critical to achieving growth promotion. And some species of *Buttiauxella* can improve plant growth and cadmium accumulation, and have strong radiation tolerance (Wu et al. 2018; Beblo-Vranesevic et al. 2018). It is worth mentioning that many species of *Curtobacterium* can improve the drought resistance and metal tolerance of plants (Silambarasan et al. 2019; Bourles et al. 2019). It can be seen that some species in the three genera *Sphingomonas*, *Curtobacterium* and *Buttiauxella* have certain significance for the study of drought resistance of upland rice. The two genera, *Aurantimonas* and

Erwinia, have also been identified in plants (Liu et al. 2016; Li et al. 2018; Borruso et al. 2017). Some species of *Aurantimonas* have certain antibacterial activity, and researchers have previously isolated an *Aurantimonas aggregata* sp. nov. in extreme environment (Pereira et al. 2017; Li et al. 2017). *Erwinia amylovora* often lead to the fire blight disease, which causes huge economic losses to apples and pears (Borruso et al. 2017).

Although researchers have done little research on the endophytes of upland rice seeds, the research on the endophytes of rice seeds is very mature. Liu et al. (2019) used high-throughput sequencing technology to investigate the diversity and community structure of endophytic bacteria in super hybrid rice “Shenliangyou 5814” and its parents, and found that the dominant genus shared by the three are *Pantoea*, *Methylobacterium*, *Sphingomonas*, *Rhizobium* and *Microbacterium* and *Pseudomonas*. Zhang et al. (2018) used the same method to explore five different genotypes of rice seeds, and the results showed that the common dominant genera were *Pantoea*, *Acinetobacter*, *Xanthomonas*, *Bacillus*, *Flavobacterium*, *Stenotrophomonas*, *Neorhizobium* and *Pseudomonas*. Walitang et al. (2019) also introduced some bacterial species of *Herbaspirillum*, *Microbacterium*, *Curtobacterium*, *Stenotrophlomonas*, *Xanthonas* and *Enterobacter* in their research, and indicated that they may be part of the core endophytic microflora that is prevalent and dominant in rice seeds. After comparing the above results with the results of this study, it is found that there are both common dominant genera and some unique dominant genera between endophytes of upland rice seeds and the first dominant genus is *Pantoea*, which further reveals the evolutionary homology of upland rice and rice from the plant endophyte level. It can be seen that there is an adaptive process in the evolution of the interaction between microorganisms and plants. Based on this, we can also infer that upland rice and rice may choose to retain the basic microorganisms needed for survival and development in the evolution process, and retain some microorganisms in the environment that can make themselves better adapt to the environment.

Members of the plant microbiota can be transmitted either horizontally (acquired from the surrounding environment) or vertically (acquired directly from the parent) (Gundel et al. 2011; Shade et al. 2017). Then, as the seed of one of the important organs of the plant and an important source of endophytes, where does its endophytes come from? A large number of studies have shown that the community assembly of seed endophytic bacteria may be affected by its parent, and that plants can autonomously select endophytic bacteria and spread them to the offspring through seeds (Liu et al. 2019; Liu et al. 2020; Walitang et al. 2018; Truyens et al. 2015). In addition, there are related studies that show that the the bacterial community in the soil environment can affect the composition of endogenous bacterial communities in other organs of the plant by affecting the microbial composition of the plant rhizosphere (Yang et al. 2017; Haruna et al. 2018). However, the relative importance of vertical and horizontal transmission in the assembly process of plant microbial communities is unclear (Shade et al. 2017). From our results, the endophytic bacteria, especially the dominant bacteria, of the 14 upland rice seed samples were not directly related to the dominant bacteria of the five soil samples in their habitat. And the results of PCA and NMDS also show significant differences between them. We suspect that this result may be due to the different origin of these upland rice seed samples. As we all know, it often takes a long

time for bacteria to colonize and exist in a certain plant, and this process is also affected by various environmental factors including nutrients, water, light, inorganic salts and so on. Therefore, compared with the parents of these upland rice seed samples and their native soils that have grown longer, the soil in which we uniformly plant these seed samples has less effect on the formation of endophytic bacteria.

Conclusion

Pantoea, *Pseudomonas*, *Methylobacterium*, *Sphingomonas*, Microbacterium and *Aurantimonas*served as major core endophytic bacteria in fourteenupland rice seed samples in this study. Endophytic diversity and community structure of 14 upland rice seed samples were not significantly different, but there were differences in abundance. It is found that there was no significant relationship between the bacterial diversity and community composition of soil samples from the uniform planting site of seed samples and the endophytic diversity and community structure of upland rice seed samples. It is worth mentioning that this is the first time that high-throughput sequencing technology has been used to explore the diversity and community structure of endophytic bacteria in upland rice seeds, and the study also found that there is a large similarity between the dominant bacteria of upland rice seeds and the dominant rice bacteria. In addition, the two genus *Curtobacterium* and *Buttiauxella* that have been reported to be resistant to radiation, drought and metal are first identified in upland rice seeds. It can be seen that in order to better adapt to the environment during the evolution of upland rice, the endophytic diversity and community structure of upland rice may be more inclined to retain some microorganisms that are compatible with the environment on the basis of the microorganisms needed for basic survival, which will also provide a new idea for the understanding of the interaction between the whole plant and microorganisms. The information collected in this study provides some basic information for the next step of microbial isolation, at the same time, finding out the microbial information in upland rice is of great significance for the separation, screening, functional evaluation of some functional microorganisms in the future as well as re reaction on plant resources to improve its agronomic properties.

Abbreviations

OTU: Optical Transform Unit; DGGE: Denaturing Gradient Gel Electrophoresis; PCA: Principal Coordinates Analysis; NMDS: Non-metric Multidimensional Scaling

Declarations

Acknowledgments

I would like to thank the research team of Prof. Zheng Huajun of Shanghai-MOST Key Laboratory of Health and Disease Genomics of Chinese National Human Genome Center at Shanghai for its platform and technical support in high-throughput sequencing and bioinformatics analysis.

Authors' Contributions

Wang ZS and Zhu YQ designed and participated in all experimental procedures, performed data analysis, and drafted the manuscript. Jing RX and Wu XY participated in the samples collection and preparation. Li N and Liu H participated in the plant samples cultivation. Zhang XX, Wang WP and Liu Y supervised the study and critically revised the manuscript. All authors read and approved the final manuscript.

Funding

The research was supported by the Fundamental Research Funds for the Central Universities (No. FRF-TP-18-012A1) and the Open Research Fund of State Key Laboratory of Hybrid Rice (Hunan Hybrid Rice Research Center) (No. 2017KF04).

Availability of Data and Materials

All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Compliance with ethical standards

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest

The authors declare that they have no competing interests.

References

Afzal I, Shinwari ZK, Sikandar S, Shahzad S (2019) Plant Beneficial Endophytic Bacteria: Mechanisms, Diversity, Host Range and Genetic Determinants. Microbiological Research 221:36-49

Agler MT, Ruhe J, Kroll S, Morhenn C, Kemen EM (2016) Microbial hub taxa link host and abiotic factors to plant microbiome variation. PLoS Biology 14(1):e1002352.

Andrews M, Andrews ME (2017) Specificity in Legume-Rhizobia Symbioses. International journal of molecular sciences 18(4)

Antunes JEL, Lyra MCCP, Ollero FJ, Freitas ADS, Figueiredo MVB (2017) Diversity of plant growth-promoting bacteria associated with sugarcane. Genetics and molecular research: GMR 16(2):gmr16029662

Beblo-Vranesevic K, Bohmeier M, Perras AK, Schwendner P, Rabbow E, Moissl-Eichinger C, Cockell CS, Vannier P, Marteinsson VT, Monaghan EP, Ehrenfreund P, Garcia-Descalzo L, Gomez F, Malki M, Amils R, Gaboyer F, Westall F, Cabezas P, Walter N, Rettberg P (2018) Lack of correlation of desiccation and radiation tolerance in microorganisms from diverse extreme environments tested under anoxic conditions. FEMS microbiology letters 365(6)

Bernier J, Atlin GN, Serraj R, Kumar A, Spaner D (2008) Breeding upland rice for drought resistance. J Sci Food Agric 88(6):927– 939

Braga LF, Oliveira FA, Couto EAPD, Santos KFDN, Ferreira EPB, Martin-Didonet CCG (2018) Polyphasic characterization of bacteria obtained from upland rice cultivated in cerrado soil. Braz. J. Microbiol. 49(1):20-28

Bridhikitti A, Overcamp TJ (2012) Estimation of Southeast Asian rice paddy areas with different ecosystems from moderate-resolution satellite imagery. Agriculture Ecosystems & Environment 146:113–120

Borruso L, Salomone-Stagni M, Polsinelli I, Schmitt AO, Benini S (2017) Conservation of *Erwinia amylovora* pathogenicity-relevant genes among erwinia genomes. Archives of Microbiology 199(1):1335-1344.

Bourles A, Guentas L, Chalkiadakis E, Majorel C, Juillot F, Cavaloc Y, Burtet-Sarramegna V, Medevielle V, Jourand P, Amir H (2019) New Caledonian ultramafic conditions structure the features of *Curtobacterium citreum* strains that play a role in plant adaptation. Canadian Journal of Microbiology 65(12):880-894

Chebotar V, Malfanova N, Shcherbakov A, Ahtemova G, Borisov AY, Lugtenberg B, Tikhonovich I (2015) Endophytic bacteria in microbial preparations that improve plant development. Appl. Biochem. Microbiol.

51:271–277.

Chen H, Wu H, Yan B, Zhao H, Liu F, Zhang H, Sheng Q, Miao F, Liang Z (2018) Core Microbiome of Medicinal Plant *Salvia miltiorrhiza* Seed: A Rich Reservoir of Beneficial Microbes for Secondary Metabolism? International journal of molecular sciences 19(3):672-

Chen Q, Meyer WA, Zhang Q, White JF (2020) 16S rRNA metagenomic analysis of the bacterial community associated with turf grass seeds from low moisture and high moisture climates. PeerJ 8:e8417

Chen T, Chen Z, Ma GH, Du BH, Xu K (2014) Diversity and potential application of endophytic bacteria in ginger. Genetics and molecular research: GMR 13(3):4918-4931

Chen WM, Cai CY, Li ZH, Young CC, Sheu SY (2019) *Methylobacterium oryzihabitans* sp. nov., isolated from water sampled from a rice paddy field. International journal of systematic and evolutionary microbiology 69:3843-3850

Chimwamurombe PM, Grönemeyer JL, Reinhold-Hurek B (2016) Isolation and characterization of culturable seed-associated bacterial endophytes from gnotobiotically grown marama bean seedlings. FEMS Microbiology Ecology 92(6):fiw083

Chu TN, Tran BTH, Van Bui L, Hoang MTT (2019) Plant growth-promoting rhizobacterium *Pseudomonas* PS01 induces salt tolerance in *Arabidopsis thaliana*. BMC research notes 12:11

Cole JR, Wang Q, Cardenas E, Fish J, Chai B, Farris RJ, Kulam-Syed-Mohideen AS, McGarrell DM, Marsh T, Garrity GM, Tiedje JM (2009) The Ribosomal Database Project: improved alignments and new tools for rRNA analysis. Nucleic Acids Res 37(Database issue):D141-5

Cordovez V, Schop S, Hordijk K, Dupre de Boulois H, Coppens F, Hanssen I, Raaijmakers JM, Carrion VJ (2018) Priming of Plant Growth Promotion by Volatiles of Root-Associated *Microbacterium* spp. Applied and environmental microbiology 84(22)

Coutinho BG, Licastro D, Lucia Mendonça-Previato, Miguel Cámara, Venturi V (2015) Plant-influenced gene expression in the rice endophyte burkholderia kururiensis m130. Molecular Plant-Microbe Interactions 28(1):10-21.

de Sousa TP, de Souza ACA, de Filippi MCC, Lanna AC, Cortes MV, Pinheiro HA, da Silva GB (2018) Bioagents and silicon promoting fast early upland rice growth. Environ Sci Pollut Res Int 25(4):3657-3668.

Durand A, Maillard F, Alvarez-Lopez V, Guinchard S, Bertheau C, Valot B, Blaudez D, Chalot M (2017) Bacterial diversity associated with poplar trees grown on a Hg-contaminated site: community characterization and isolation of Hg-resistant plant growthpromoting bacteria. Sci Total Environ 622–623:1165–1177

Edwards J, Johnson C, Santosmedellín C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl Acad Sci 112:911–920

Ferrari S, Pagliari P, Trettel J (2018) Optimum Sowing Date and Genotype Testing for Upland Rice Production in Brazil. Scientific reports 8:8227

Galinato MI, Moody K, Piggin CM (1999). Upland rice weeds of south and southeast asia.

Grossi CEM, Fantino E, Serral F, Zawoznik MS, Fernandez Do Porto DA, Ulloa RM (2020) *Methylobacterium* sp. 2A Is a Plant Growth-Promoting Rhizobacteria That Has the Potential to Improve Potato Crop Yield Under Adverse Conditions. Front Plant Sci 11:71

Guimarães CM, Castro AP, Stone LF, Oliveira JP (2016) Drought tolerance in upland rice: identification of genotypes and agronomic characteristics. Acta Scientiarum. Agron Maringá 38(2):201–206

Gundel PE, Rudgers JA, Ghersa CM (2011) Incorporating the processof vertical transmission into understanding of host–symbiont dynamics. Oikos 120:1121-1128

Haimin C, Hongxia W, Bin Y, Hongguang Z, Fenghua L, Haihua Z, Qing S, Fang M, Zongsuo L (2018) Core microbiome of medicinal plant *salvia miltiorrhiza* seed: a rich reservoir of beneficial microbes for secondary metabolism? International Journal of Molecular Sciences 19(3):672-

Hardoim PR, Hardoim CC, van Overbeek LS, van Elsas JD (2012) Dynamics of seed-borne rice endophytes on early plant growth stages. PLoS ONE 7(2):e30438

Haruna E, Zin NM, Kerfahi D, Adams JM (2018) Extensive overlap of tropical rainforest bacterial endophytes between soil, plant parts, and plant species. Microbial ecology 75(1):88-103

Jiang XY, Gao JS, Xu FH, Cao YH, Tang X, Zhang XX (2013) Diversity of endophytic bacteria in rice seeds and their secretion of indole acetic acid (in Chinese). Acta Microbiol Sin 53:269–275

Kumar V, Ladha JK (2011) Direct seeding of rice. recent developments and future research needs. Advances in Agronomy 111:297-413.

Li AZ, Lin LZ, Zhang MX, Zhu HH (2017) *Aurantimonas aggregata* sp. nov., isolated from deep-sea sediment. International journal of systematic and evolutionary microbiology 67:5056-5061

Links MG, Demeke T, Grafenhan T, Hill JE, Hemmingsen SM, Dumonceaux TJ (2014) Simultaneous profifiling of seed-associated bacteria and fungi reveals antagonistic interactions between microorganisms within a shared epiphytic microbiome on *Triticum* and *Brassica* seeds. New Phytol 202:542-553

Liu BB, Wang HF, Li QL, Zhou XK, Zhang YG, Xiao M, Li QQ, Zhang W, Li WJ (2016) *Aurantimonas endophytica* sp. nov., a novel endophytic bacterium isolated from roots of Anabasis elatior (C. A. Mey.) Schischk. International journal of systematic and evolutionary microbiology 66:4112-4117

Liu L, Liu Y, Song W (2009) Indigenous bacterial community diversity in hybrid rice (*Oryza sativa* L.) seed (in Chinese). Biotechnol Bull 1:95–111

Liu Y, Zuo S, Xu L, Zou Y, Song W (2012) Study on diversity of endophytic bacterial communities in seeds of hybrid maize and their parental lines. Arch Microbiol 194:1001–1012

Liu Y, Yao S, Lee YJ, Cao Y, Zhai L, Zhang X, Su J, Ge Y, Kim SG, Cheng C (2015) *Sphingomonas morindae* sp. nov., isolated from Noni (Morinda citrifolia L.) branch. International journal of systematic and evolutionary microbiology 65:2817-2823

Liu Y, Wang RH, Li YH, Cao YH, Chen CY, Qiu CZ, Bai FR, Xu TJ, Zhang X, Dai WK, Zhao JR, Cheng C (2017a) High-throughput sequencing-based analysis of the composition and diversity of endophytic bacterial community in seeds of “beijing ”hybrid maize planted in China. Plant Growth Regul 81:317–324

Liu Y, Bai FR, Li N, Wang WP, Cheng C (2017b) Identification of endophytic bacterial strain RSE1 from seeds of super hybrid rice Shenliangyou 5814 (*Oryza sativa* L.) and evaluation of its antagonistic activity. Plant Growth Regul 82:403–408

Liu Y, Xu PP, Yang FZ, Li M, Yan H, Li N, Zhang XX, Wang WP (2019) Composition and diversity of endophytic bacterial community in seeds of super hybrid rice ‘Shenliangyou 5814’ (*Oryza sativa* L.) and its parental lines. Plant Growth Regulation 87:257–266

Liu Y, Yan H, Zhang X, Zhang R, Li M, Xu T, Yang F, Zheng H, Zhao J (2020) Investigating the endophytic bacterial diversity and community structures in seeds of genetically related maize (*Zea mays* L.) genotypes. 3 Biotech 10(1):27

Li Y, Xu G, Lin C, Wang X, Piao CG (2018) *Aureimonas populi* sp. nov., isolated from poplar tree bark. International journal of systematic and evolutionary microbiology 68:487-491

Lopes RB, Costa LE, Vanetti MC, de Araújo EF, de Queiroz MV(2015) Endophytic bacteria isolated from common bean (*phaseolus vulgaris*) exhibiting high variability showed antimicrobial activity and quorum sensing inhibition. Current Microbiology 71(4):509-16

Luo LJ (2010) Breeding for water-saving and drought-resistance rice (WDR) in China. J. Exp. Bot. 61:3509–3517

Luo Y, Wang F, Huang Y, Zhou M, Gao J, Yan T, Sheng H, An L (2019) *Sphingomonas* sp. Cra20 Increases Plant Growth Rate and Alters Rhizosphere Microbial Community Structure of *Arabidopsis thaliana* Under Drought Stress. Frontiers in microbiology 10:1221

Luziatelli F, Ficca AG, Cardarelli M, Melini F, Cavalieri A, Ruzzi M (2020) Genome Sequencing of *Pantoeaagglomerans* C1 Provides Insights into Molecular and Genetic Mechanisms of Plant Growth-Promotion and Tolerance to Heavy Metals. Microorganisms 8(2)

Lyu J, Li B, He W, Zhang S, Gou Z, Zhang J, Meng L, Li X, Tao D, Huang W, Hu F, Wang W (2014) A genomic perspective on the important genetic mechanisms of upland adaptation of rice. BMC plant biology 14:160

Mannaa M, Oh JY, Kim KD (2017) Biocontrol Activity of Volatile-Producing *Bacillus megaterium* and *Pseudomonas protegens* against *Aspergillus flavus* and Aflatoxin Production on Stored Rice Grains. Mycobiology 45:213-219

Megías E, Reis Junior FB, Ribeiro RA, Megías M, Ollero FJ, Hungria M (2017) Genome sequence of *Pantoea* sp. strain 1.19, isolated from rice rhizosphere, with the capacity to promote growth of leg-umes and nonlegumes. Genome Announc 5(30):e00707–e00717

Nascimento FX, Hernandez AG, Glick BR, Rossi MJ (2020) The extreme plant-growth-promoting properties of *Pantoeaphytobeneficialis* MSR2 revealed by functional and genomic analysis. Environmental microbiology 22(4):1341-1355

Nelson EB (2004) Microbial dynamics and interactions in the spermosphere. Ann Rev Phytopathol 42:271–309

Nelson EB (2017) The seed microbiome: Origins, interactions, and impacts. Plant Soil 422:7–34

Pereira LB, Palermo BRZ, Carlos C, Ottoboni LMM (2017) Diversity and antimicrobial activity of bacteria isolated from different Brazilian coral species. FEMS microbiology letters 364(16)

Pham VTK, Rediers H, Ghequire MGK, Nguyen HH, De Mot, René, Vanderleyden J, Spaepen S (2017) The plant growth-promoting effect of the nitrogen-fixing endophyte*Pseudomonas stutzeria*15. Archives of Microbiology 199(3):513-517

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res 41:D590-596.

Rybakova D, Mancinelli R, Wikström M, Birch-Jensen AS, Postma J, Ehlers RU, Goertz S, Berg G (2017) The structure of the brassica napus seed microbiome is cultivar-dependent and affects the interactions of symbionts and pathogens. Microbiome 5(1):104

Sánchez-Cruz R, Tpia Vázquez I, Batista-García RA, Méndez-Santiago EW, Sánchez-Carbente MDR, Leija A, Lira-Ruan V, Hernández G, Wong-Villarreal A, Folch-Mallol JL (2019) Isolation and characterization of endophytes from nodules of Mimosa pudica with biotechnological potential. Microbiol Res 218:76-86

Savi DC, Shaaban KA, Gos FMW, Thorson JS, Glienke C, Rohr J (2019) Secondary metabolites produced by *Microbacterium* sp. LGMB471 with antifungal activity against the phytopathogen Phyllosticta citricarpa. Folia microbiologica 64:453-460

Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS One 6(12):e27310.

Scott M, Rani M, Samsatly J, Charron JB, Jabaji S (2018) Endophytes of industrial hemp (*Cannabis sativa* L.) cultivars: identification of culturable bacteria and fungi in leaves, petioles, and seeds. Canadian journal of microbiology 64:664-680

Shade A, Jacques MA, Barret M (2017) Ecological patterns of seed microbiome diversity,transmission, and assembly. Curr. Opin. Microbiol. 37:15–22

Shahid I, Rizwan M, Baig DN, Saleem RS, Malik KA, Mehnaz S (2017) Secondary Metabolites Production and Plant Growth Promotion by *Pseudomonas chlororaphis* and *P. aurantiaca* Strains Isolated from Cactus, Cotton, and Para Grass. Journal of microbiology and biotechnology 27:480-491

Silambarasan S, Logeswari P, Valentine A, Cornejo P (2019) Role of *Curtobacterium herbarum* strain CAH5 on aluminum bioaccumulation and enhancement of Lactuca sativa growth under aluminum and drought stresses. Ecotoxicology and environmental safety 183:109573

Sobolev VS, Orner VA, Arias RS (2013) Distribution of bacterial endophytes in peanut seeds obtained from axenic and control plant material under field conditions. Plant Soil 371:367–376

Steinbrecher T, Leubner-Metzger G (2017) The biomechanicsof seed germination. Journal of Experimental Botany 68(4):765-783

Tani A, Sahin N, Fujitani Y, Kato A, Sato K, Kimbara K (2015) *Methylobacterium* Species Promoting Rice and Barley Growth and Interaction Specificity Revealed with Whole-Cell Matrix-Assisted Laser Desorption/Ionization-Time-of-Flight Mass Spectrometry (MALDI-TOF/MS) Analysis. PloS one 10:e0129509

Truyens S, Weyens N, Cuypers A, Vangronsveld J (2015) Bacterial seed endophytes: genera, vertical transmission and interaction with plants. Environ Microbiol Rep 7:40–50

Tuhina-Khatun M, Hanafi MM, Rafii Yusop M, Wong MY, Salleh FM, Ferdous J (2015) Genetic Variation, Heritability, and Diversity Analysis of Upland Rice (*Oryza sativa* L.) Genotypes Based on Quantitative Traits. BioMed research international 2015:290861

Vandenkoornhuyse P, Quaiser A, Duhamel M, Le Van A, Dufresne A (2015) The importance of the microbiome of the plant holobiont. New Phytol 206:1196–1206

Verma SK, White JF (2018) Indigenous endophytic seed bacteria promote seedling development and defend against fungal disease in browntop millet (*Urochloa ramosa* L.). J Appl Microbiol 124(3):764–778

Walitang DI, Chan-Gi K, Sunyoung J, Yeongyeong K, Tongmin S (2018) Conservation and transmission of seed bacterial endophytes across generations following crossbreeding and repeated inbreeding of rice at different geographic locations. MicrobiologyOpen 8(3):e00662

Wang Q, Garrity GM, Tiedje JM, Cole JR (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Appl Environ Microbiol 73(16):5261-5267

Wang W, Zhai Y, Cao L, Tan H, Zhang R (2016) Illumina-based analysis of core actinobacteriome in roots, stems, and grains of rice. Microbiol Res 190:12–18

Wei Z, Jousset A (2017) Plant breeding goes microbial. Trends in Plant Science 22(7):555-558

White JR, Nagarajan N, Pop M (2009) Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLoS Comput Biol 5(4):e1000352

Wu K, Luo J, Li J, An Q, Yang X, Liang Y, Li T (2018) Endophytic bacterium *Buttiauxella* sp. SaSR13 improves plant growth and cadmium accumulation of hyperaccumulator Sedum alfredii. Environmental science and pollution research international 25:21844-21854

Xia H, Luo Z, Xiong J, Ma X, Lou Q, Wei H, Qiu J, Yang H, Liu G, Fan L, Chen L, Luo L (2019) Bi-directional selection in upland rice leads to its adaptive differentiation from lowland rice in drought resistance and productivity. Molecular Plant 12(2):170-184

Xu J, Kloepper JW, Huang P, McInroy JA, Hu CH (2018) Isolation and characterization of N2 -fixing bacteria from giant reed and switchgrass for plant growth promotion and nutrient uptake. Journal of basic microbiology 58:459-471

Xu W, Wang F, Zhang M, Ou T, Wang R, Strobel G, Xiang Z, Zhou Z, Xie J (2019) Diversity of cultivable endophytic bacteria in mulberry and their potential for antimicrobial and plant growth-promoting activities. Microbiol Res 229:126328.

Yang L, Danzberger J, Scholer A, Schroder P, Schloter M, Radl V (2017) Dominant Groups of Potentially Active Bacteria Shared by Barley Seeds become Less Abundant in Root Associated Microbiome. Frontiers in plant science 8:1005

Zhang J, Zhang C, Yang J, Zhang R, Zhang X (2018) Insights into endophytic bacterial community structures of seeds among various *Oryza sativa*L. rice genotypes. Journal of Plant Growth Regulation 38(1):1-10.

Zinniel DK, Lambrecht P, Harris NB, Feng Z, Kuczmarski D, Higley P, Ishimaru CA, Arunakumari A, Barletta RG, Vidaver AK (2002) Isolation and characterization of endophytic colonizing bacteria from agronomic

crops and prairie plants. Appl. Environ. Microbiol. 68:2198–2208

Zou YY, Liu L, Liu Y, Zhao L, Deng QY, Wu J, Zhuang W, Song W (2012) Diversity of indigenous bacterial communities in *Oryza sativa* seeds of different varieties (in Chinese). Chin J Plant Ecol 36:880–890

Tables

Table 1 Information on fourteen varieties of rice seeds

Sample ID	Variety names	Breeding places
19H001	Heibiao	Korea
19H002	IAC150/76	Brazil
19H004	IRAT266	Ivory Coast
19H005	Funingzipi	Funing, Hebei Province
19H009	Hanmadao	Huaibin, Henan Province
19H010	Shanjiugu	Gao, Sichuan Province
19H016	Gongju 73	Menglian, Yunnan Province
19H020	Youzhan	Zhenning, Guizhou Province
19H023	Xianggu	Jinping, Yunnan Province
19H025	Menjiading 2	Ledong, Hainan Province
19H026	Feienuo 2	Anlong, Guihzou Province
19H027	Xishen 15	Baoting, Hainan Province
19H028	Honggenghangu	Donglan, Guangxi Province
19H031	Qingke	Yuanyang, Yunnan Province

Table 2 Statistics of the upland rice seed sample

Sample	Read number	Minlength	Maxlength	Average length
19H001	48392	393	579	520
19H002	65245	325	578	513
19H004	31618	325	581	517
19H005	37442	305	583	530
19H009	52548	368	586	529
19H010	42982	319	576	519
19H016	52617	313	583	529
19H020	54247	328	586	548
19H023	44539	306	581	524
19H025	53561	324	583	515
19H026	30328	326	582	525
19H027	28388	320	576	538
19H028	45249	358	580	526
19H031	42348	330	588	534

Table 3 Statistical results of endophytic bacterial alpha-diversity in each sample

Sample	OTUs	ACE	Chao	Shannon	Simpson
19H001	1946	12185.599	6457.800	3.961	0.087
19H002	791	3730.646	2140.011	2.527	0.250
19H004	1188	7263.790	3882.446	2.593	0.314
19H005	2058	12499.316	6619.603	3.773	0.105
19H009	1509	10215.829	5721.426	3.202	0.156
19H010	694	4252.855	2261.800	1.770	0.432
19H016	1979	11021.393	5966.005	3.691	0.130
19H020	1784	11161.881	5671.730	3.067	0.245
19H023	1400	7816.201	4328.940	2.800	0.257
19H025	1062	6185.812	3330.750	2.112	0.460
19H026	1186	6905.181	3956.122	2.356	0.309
19H027	1066	6992.703	3998.958	2.494	0.255
19H028	1219	6546.548	3819.954	2.848	0.173
19H031	1000	5935.459	3436.371	2.506	0.204

Note: The difference of bacterial alpha-diversity based on a 16S rDNA sequence assignment datasetwith a 97% sequence similarity threshold in upland rice seeds

Table 4 Statistical results of dominant genera in each upland rice seed sample

Taxonomy	19H001	19H002	19H004	19H005	19H009	19H010	19H016
Pantoea	24.62%	58.22%	69.90%	50.80%	42.51%	87.77%	54.68%
Pseudomonas	24.15%	15.23%	4.68%	9.88%	19.54%	2.27%	5.71%
Methylobacterium	4.07%	0.96%	3.61%	2.50%	2.34%	0.40%	4.64%
Sphingomonas	3.85%	2.09%	2.20%	2.03%	2.02%	0.41%	2.80%
Microbacterium	4.67%	——	0.16%	2.26%	0.62%	0.01%	2.08%
Aurantimonas	0.71%	0.43%	0.66%	0.26%	0.49%	0.04%	1.58%
Taxonomy	19H020	19H023	19H025	19H026	19H027	19H028	19H031
Pantoea	54.18%	73.66%	80.72%	77.38%	8.77%	70.86%	82.14%
Pseudomonas	5.56%	1.34%	1.15%	3.32%	61.58%	7.98%	2.04%
Methylobacterium	2.79%	2.27%	1.16%	0.70%	1.63%	0.74%	1.12%
Sphingomonas	2.95%	0.72%	0.71%	0.80%	0.94%	0.38%	0.26%
Microbacterium	1.19%	0.36%	0.90%	0.18%	0.01%	1.21%	0.72%
Aurantimonas	4.34%	1.52%	0.93%	1.53%	0.08%	0.40%	0.30%

Table 5 Statistics of the soil sample

Sample	Read number	Minlength	Maxlength	Average length
HNSYSB_1	58945	301	589	502
HNSYSB_2	56380	301	591	502
HNSYSB_3	62854	301	591	503
HNSYSB_4	43869	302	585	503
HNSYSB_5	44612	301	580	505

Table 6 Statistical results of endophytic bacterial alpha-diversity in each soil sample

Sample	ACE	Chao	Shannon	Simpson
HNSYSB_1	278232.169	108798.382	8.468	0.00123
HNSYSB_2	290087.026	111407.030	8.557	0.00106
HNSYSB_3	150919.097	65612.592	7.970	0.00232
HNSYSB_4	227790.518	95095.037	8.351	0.00129
HNSYSB_5	411216.023	151144.140	8.769	0.00103

Note: The difference of bacterial alpha-diversity based on a 16S rDNA sequence assignment datasetwith a 97% sequence similarity threshold in soil sample.

Figures

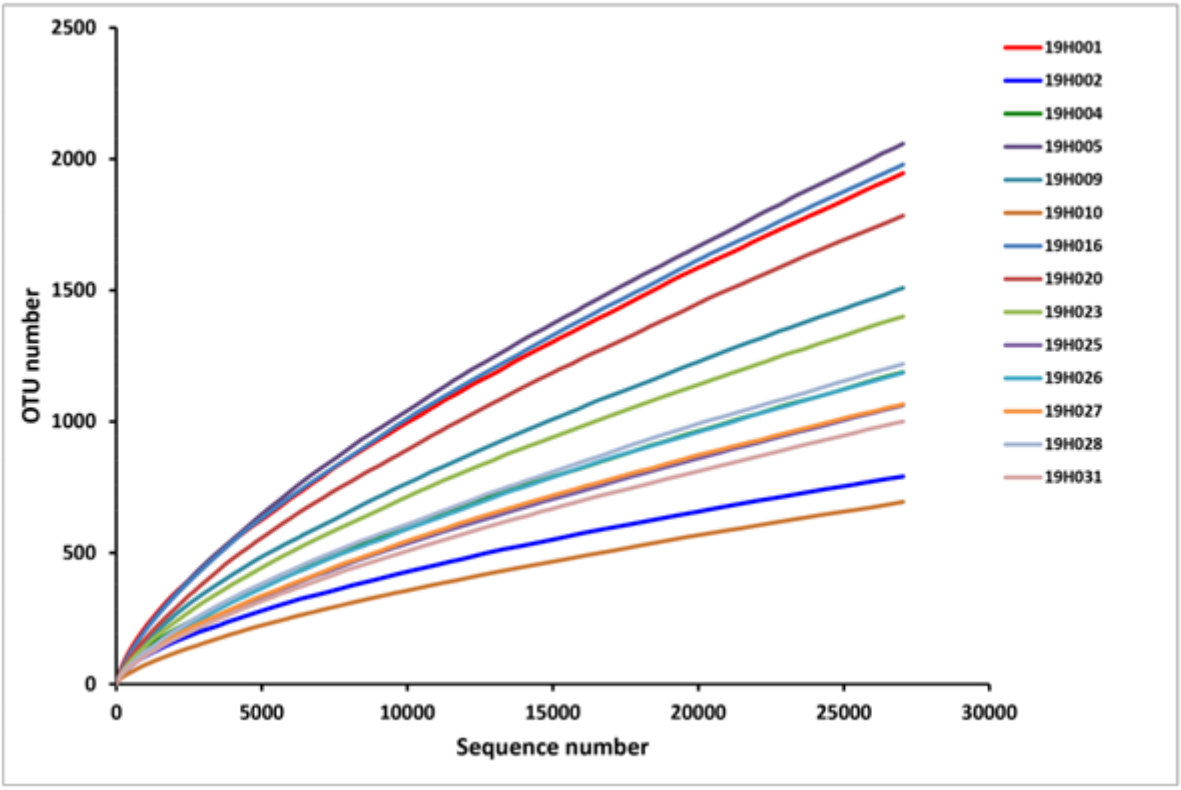

Figure 1

Rarefaction curve of OTUs in the fourteen rice samples. Rarefaction curves present therelationship between number of samples and bacterial species richness in OTUs.

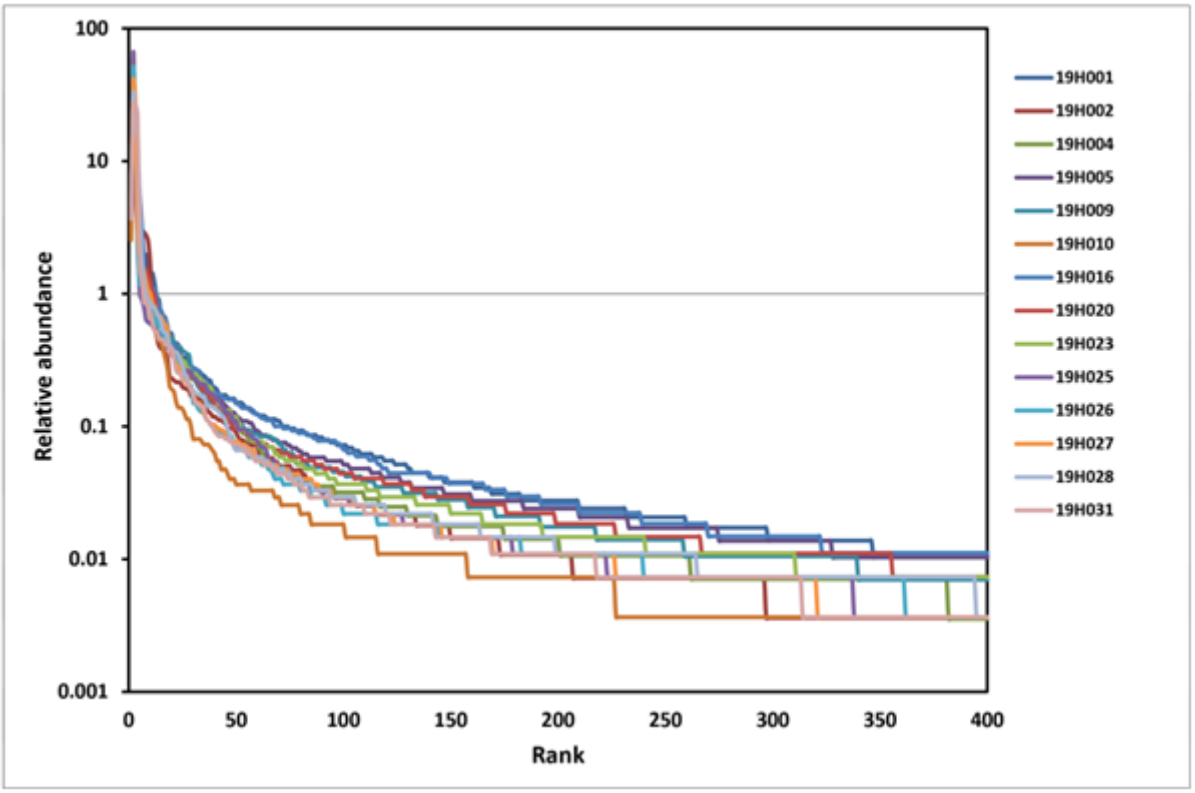

Figure 3

Rank abundance curve of 14 samples. Rank abundance curve is used to observe the richness and evenness of species.

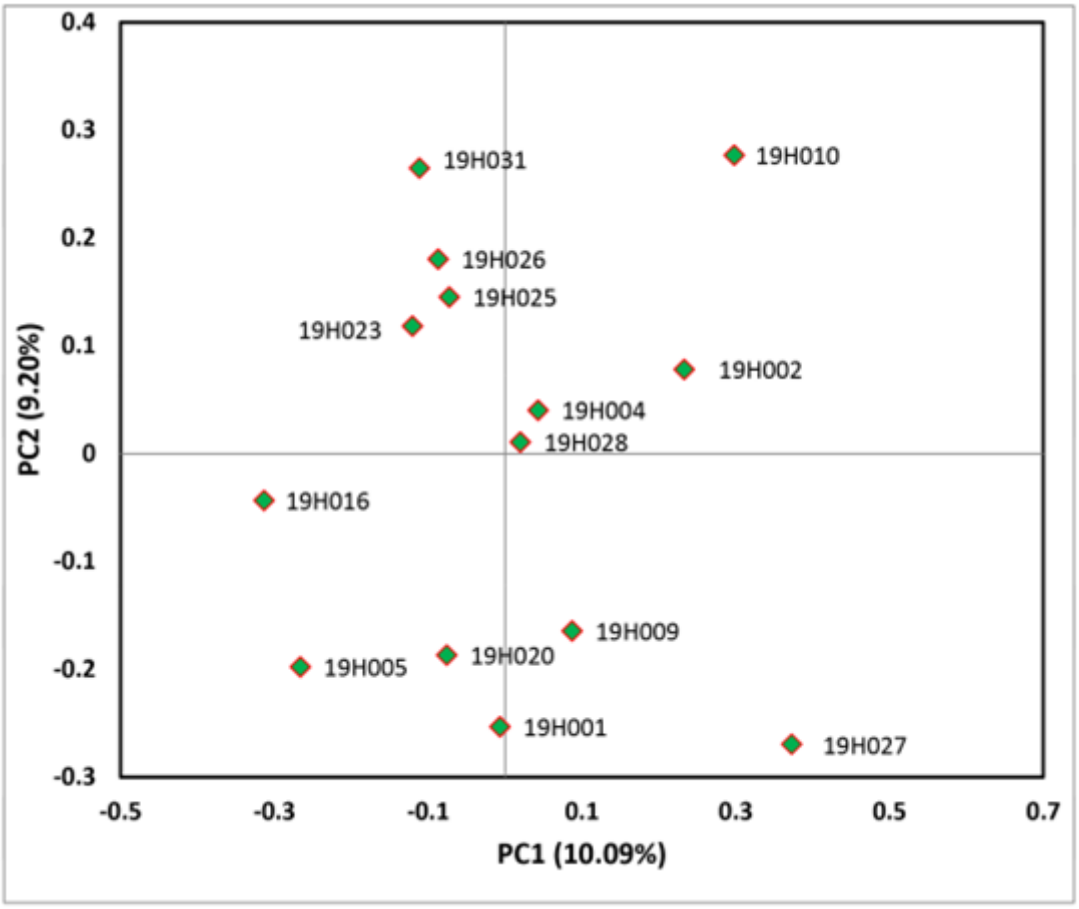

Figure 5

Principal Coordinates Analysis (PCA). Ecological differences (two-dimensional) between the different groups and samples in the case of mixed samples at OTU = 0.03. The abscissa and ordinate represent the contribution rate of the principal components 1 and 2 to the distribution of the samples. Each point in the figure represents a sample, and the points of the same color come from the same group.

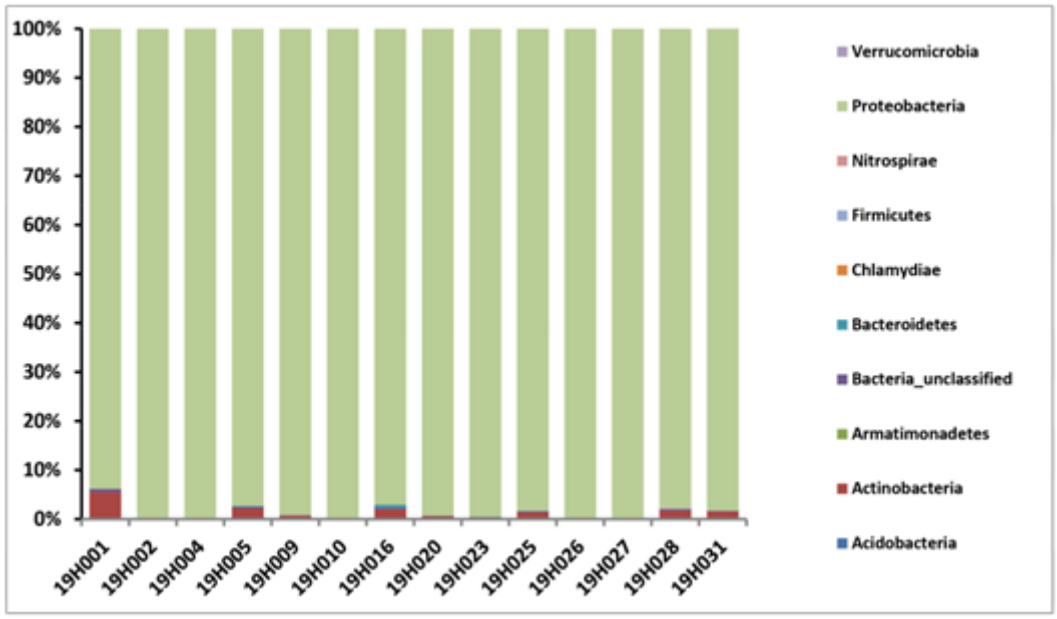

Figure 7

Relative abundance of shared/unshared phyla in each upland rice seed sample.

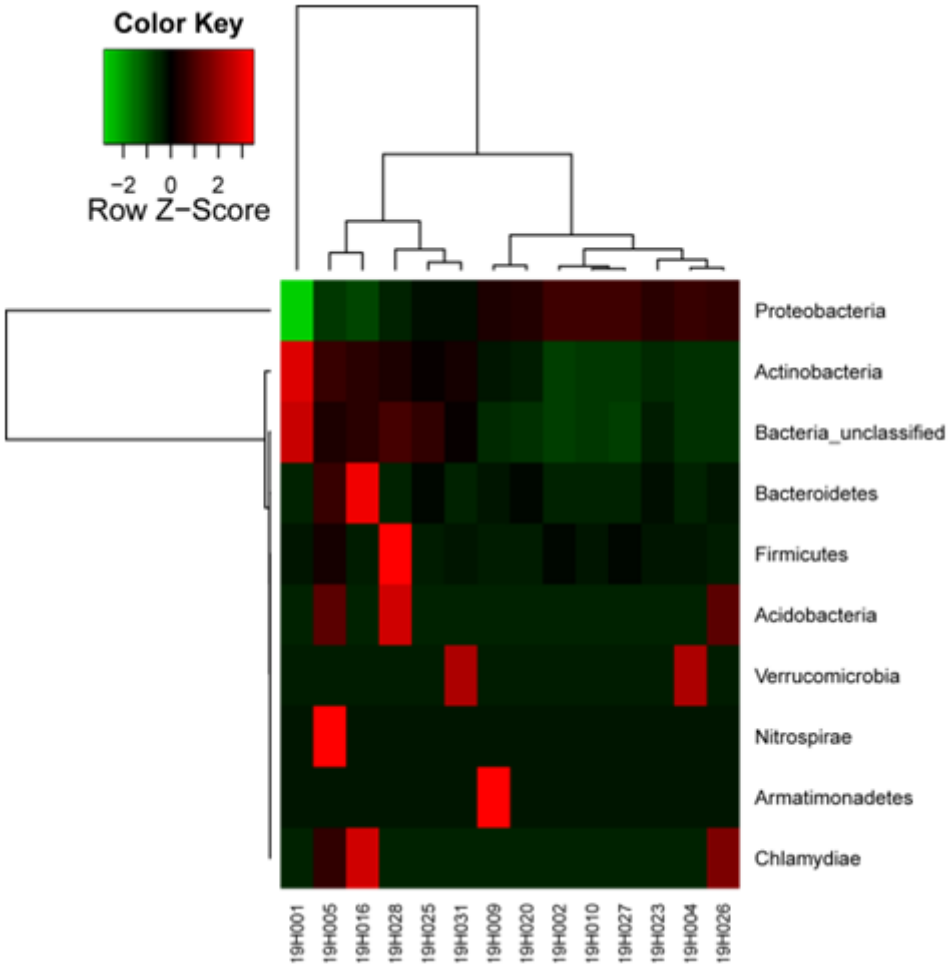

Figure 9

Classification of samples with OTU = 0.03 (phylum, top20).

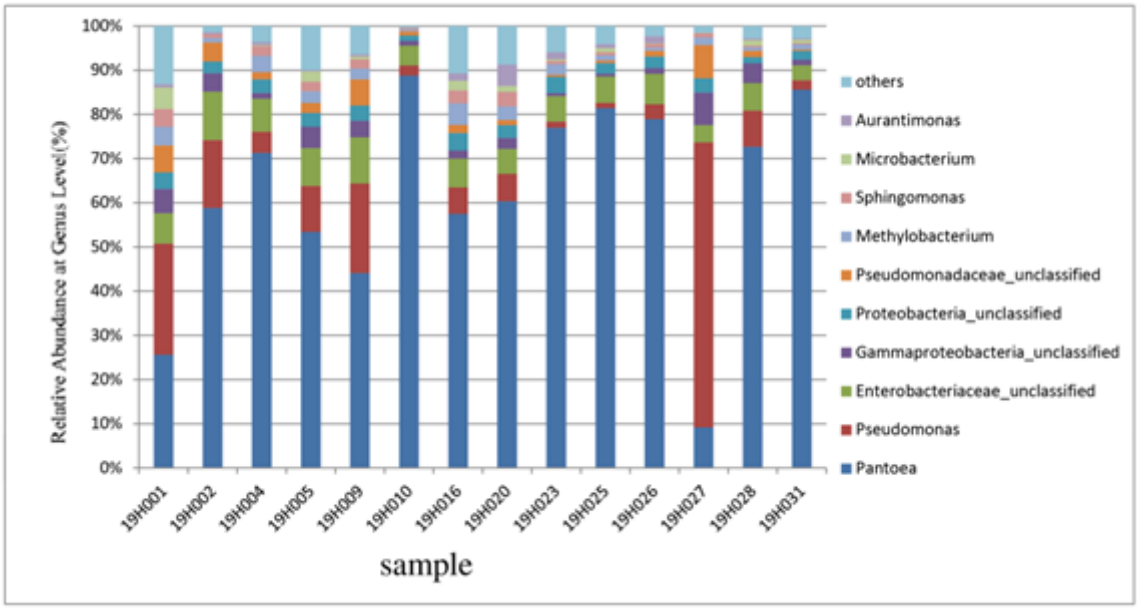

Figure 11

Relative abundance of shared/unshared genus in each upland rice seed sample.

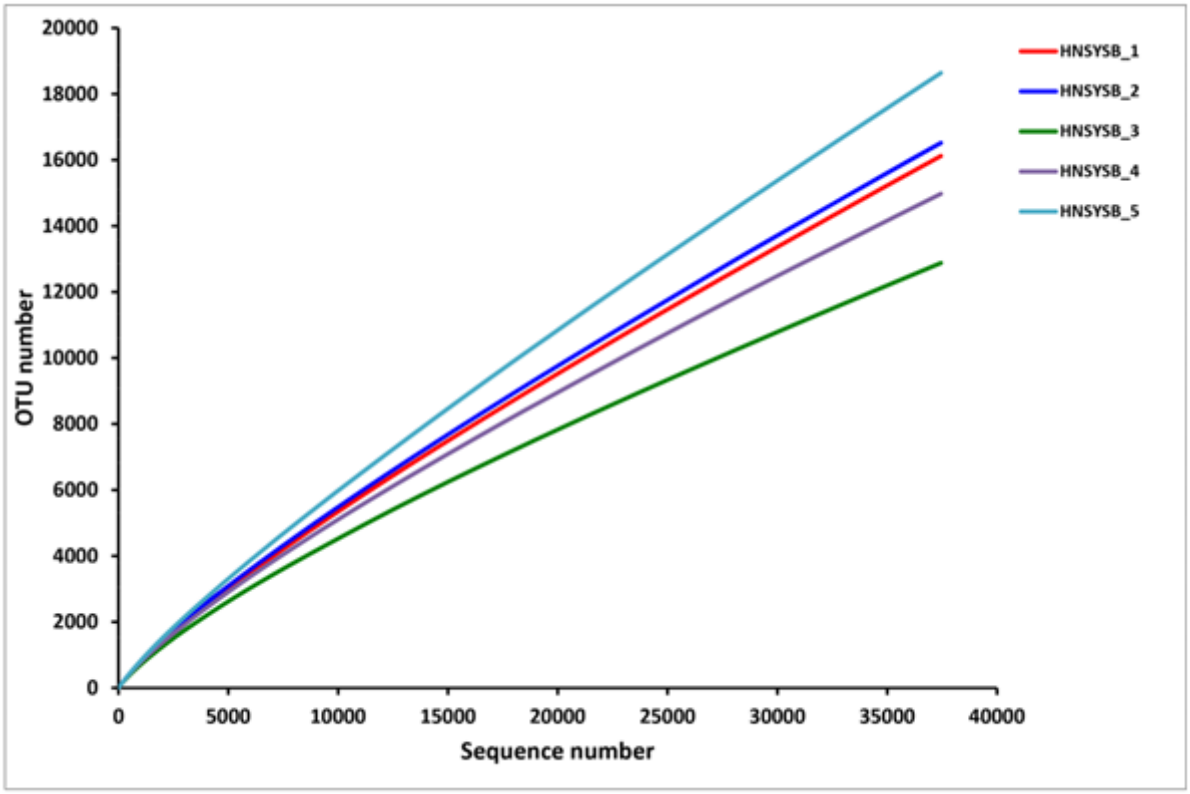

Figure 13

Rarefaction curve of OTUs in the fivesoil samples. Rarefaction curves present the relationship between number of samples and bacterial species richness in OTUs.

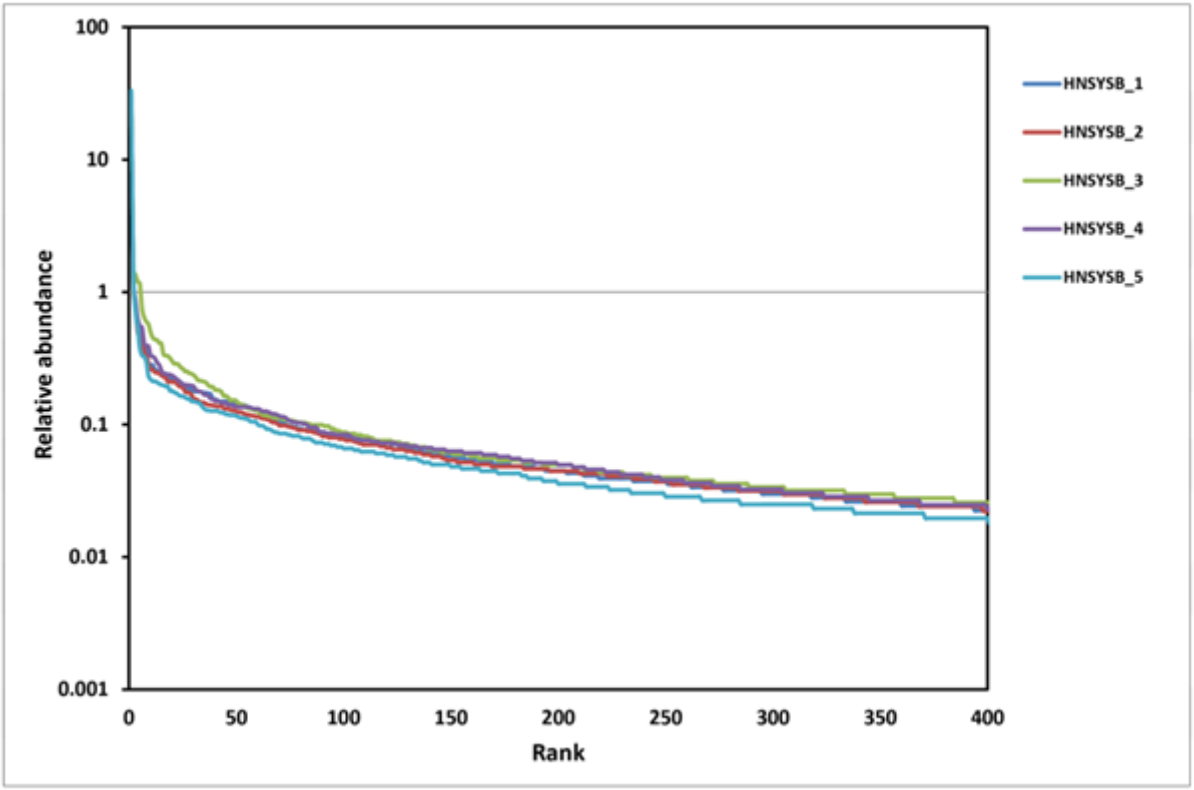

Figure 15

Rank abundance curve of five soil samples. Rank abundance curve is used to observe the richness and evenness of species.

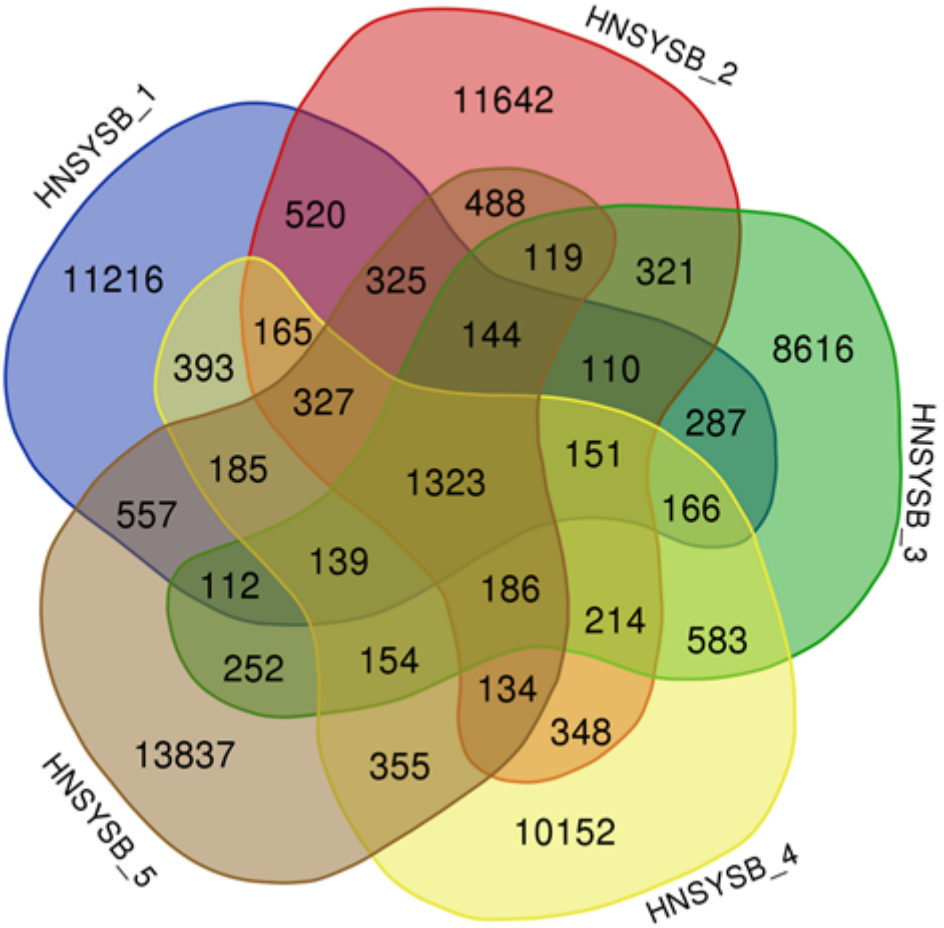

Figure 17

Venn Diagram of OTUs in the fivesoil samples. Venn diagram shows the difference in species-level OTU recovery between these five individual seed type. Values in the parentheses indicate numbers inside each region indicate the number of unique or shared OTUs.

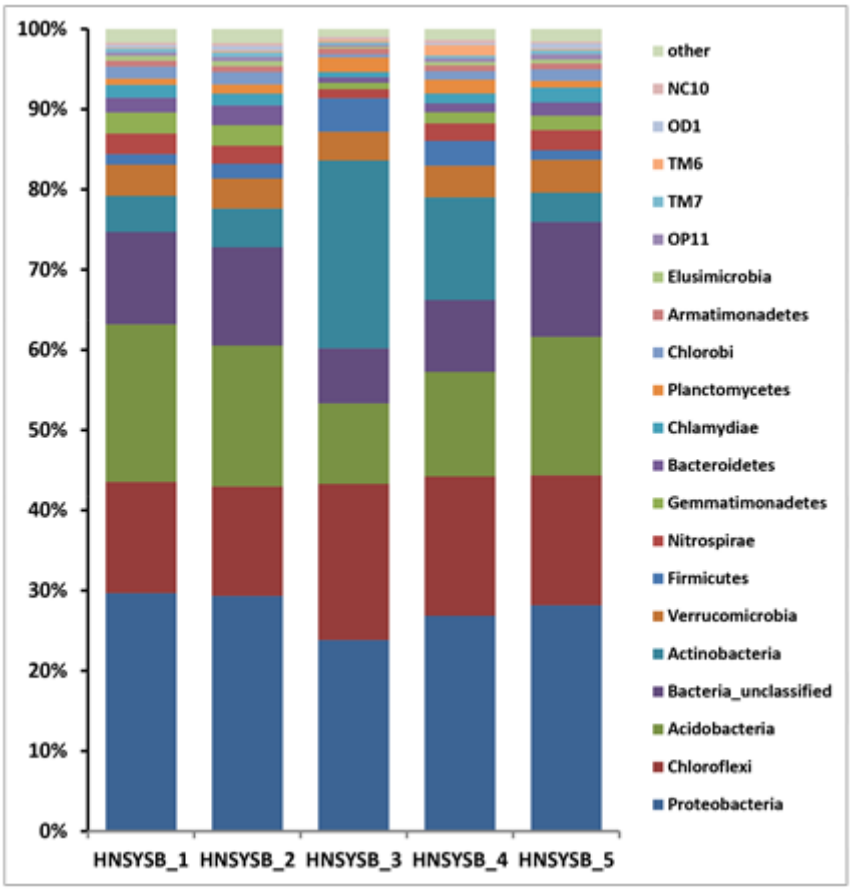

Figure 19

Relative abundance of shared/unshared phyla in each soil sample.

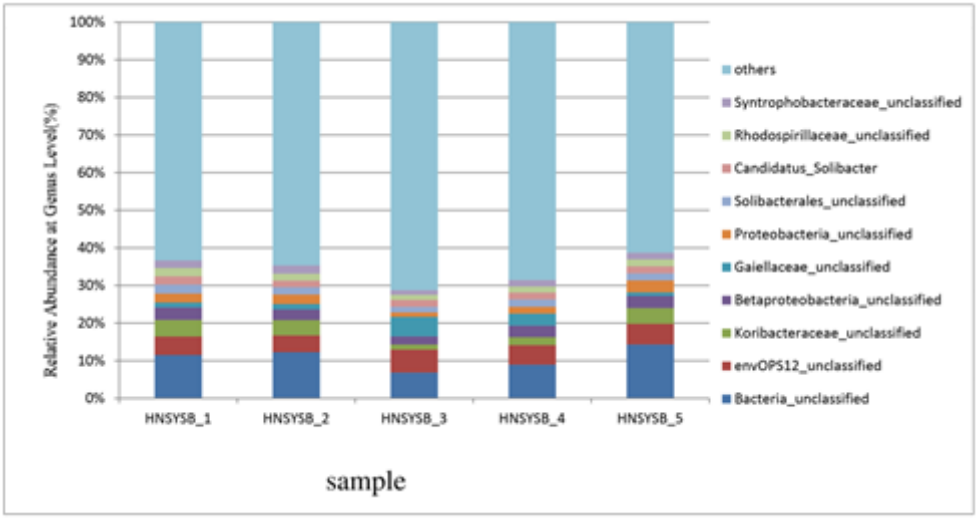

Figure 21

Relative abundance of shared/unshared genus in each soil sample.

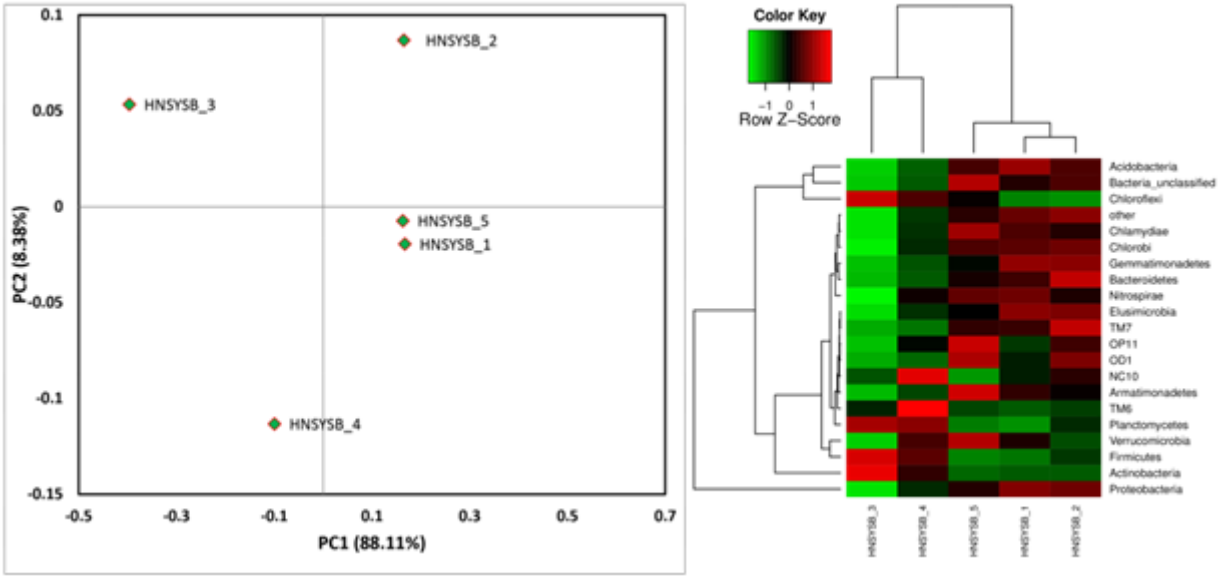

Figure 23

Inspection and quantification of endophytic bacterial beta diversity and community structure in each sample. a Principal Coordinates Analysis (PCA). b Heat map depicting the statistic significant phylums between five groups.

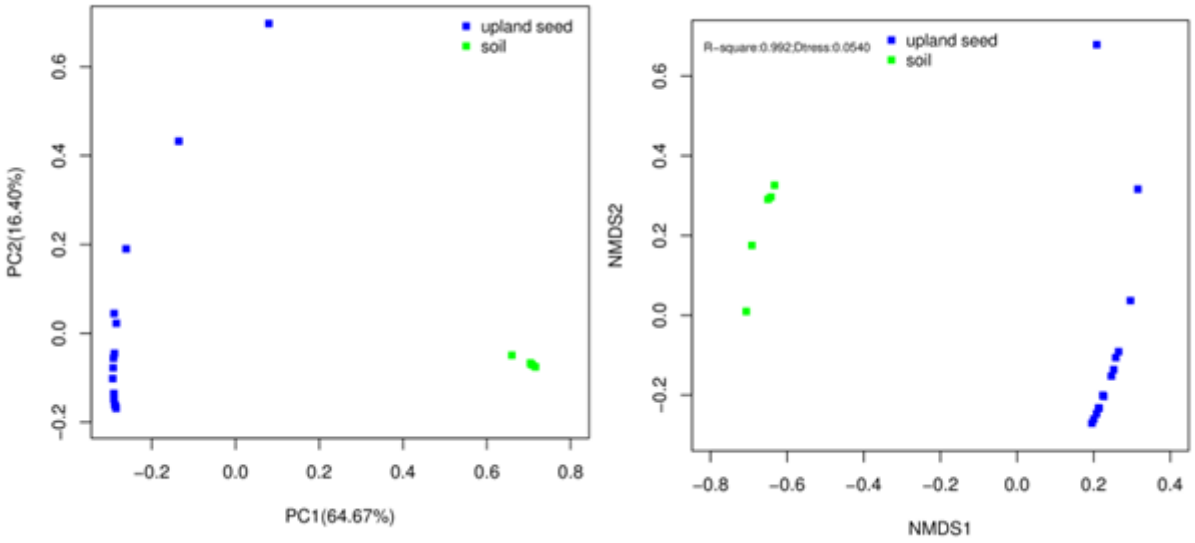

Figure 25

Analysis of upland rice seed samples and soil samples of by Principal Coordinates Analysis (PCA) (a) and Non-metric Multidimensional Scaling (NMDS) (b).

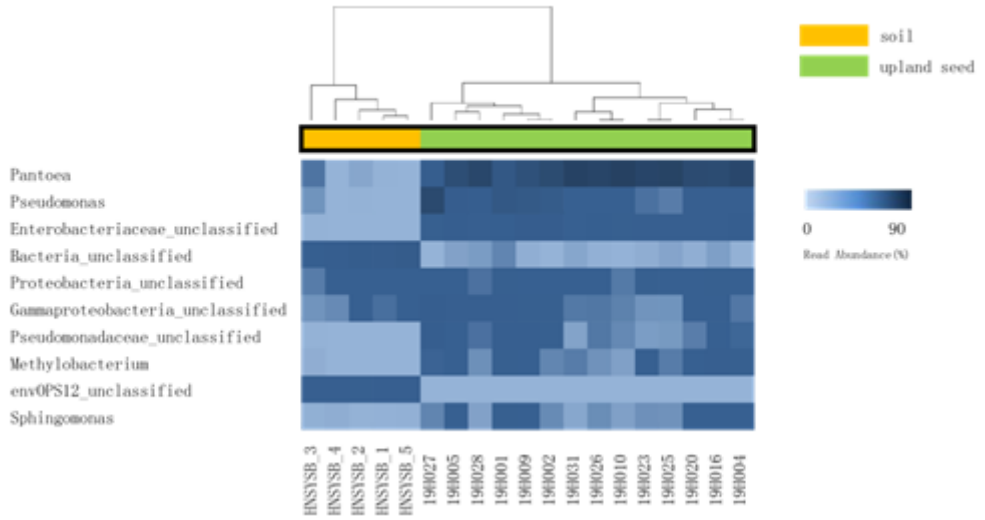

Figure 27

Heat map of the top 10 dominant endophytic bacteria in upland rice seed and soil samples.

Supplementary Files

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

- Supplementarydata.doc
- Supplementarydata.doc