Next-generation sequencing of root fungal communities in continuous cropping soybean

Li Bai1, 2, Haibing Sun1, Xueqi Zhang1, and Baiyan Cai1, 2*

1Heilongjiang University, School of Life Science, Harbin 150080, Heilongjiang, P.R. China.
2East University of Heilongjiang, Food and Environmental Engineering Department, Harbin 150086, P.R. China.
*Corresponding author (caibaiyan@126.com).

ABSTRACT

Soybean root rot is a major disease of soybean (Glycine max [L.] Merr.) under continuous cropping, which leads to dramatic variations in the rhizosphere microflora. Soybean was sown in a different field after year zero, and continuous cropping was applied for 1 or 2 yr. The objectives were to investigate the variation in fungal populations present or inhabiting soybean roots during 3 yr of monocropping using next-generation sequencing to compare the three sets of root samples and provide a theoretical basis for the following inoculation study of the different pathogens involved in root rot disease in soybean, variation in fungal populations, and incidence of root rot. Results showed that operational taxonomic units (OTUs) of the three samples were divided into 19 phyla, 169 families, and 235 genera. Ascomycota and Basidiomycota were the dominant phyla in the continuous cropping root samples. Continuous cropping could increase the relative abundance of some fungi, namely Fusarium, Rhizoctonia, and Thelebolus, which are associated with soybean root rot for 2 and 3 yr of cropping. Continuous cropping could also increase the abundance of Gymnoascus, Chrysosporium, Ctenomyces, Aphanoascus, and Aspergillus, which are soil pathogenic fungi that can cause other plant diseases.

Key words: Fungal populations, Glycine max, root rot fungi, soybean continuous cropping, taxonomic diversity.

INTRODUCTION

Soybean root rot is a major disease of soybean (Glycine max [L.] Merr.) under continuous cropping in Northeast China (Zhang et al., 2011; Liu et al., 2012). This disease is extremely harmful and difficult to prevent and control (Hartman et al., 2011; El-Abdean et al., 2013). In the early stages of plant development, it affects seedling growth and can even cause their death. In the maturity stage, the disease affects the growth and quantity of soybean root nodules, causes suboptimal plant growth, and reduces both pod number and grain weight, which lead to decreased commercial yield. The main root rot driver in cultivated soybean is continuous cropping, growing a crop species year after year with no fallow period.

Continuous cropping leads to dramatic variations in the rhizosphere microflora; there is a gradual transition between “bacterial-type”, high-fertilizer soil, and “fungal-type”, low-fertilizer soil, which feature microflora dominated by fungi (Vargas-Gil et al., 2011; Joa et al., 2014). Some of these fungi are potential pathogens that are more likely to infect various plants and cause disease. The pathogenic fungi of soybean root rot disease include Fusarium oxysporum, F. semitectum, Gliocladium roseum, Rhizoctonia solani, and Pythium spp. Soil fungal pathogens may also cause other plant diseases (Hashem et al., 2010); for example, corn ear rot caused by F. moniliforme in China and soilborne wilt disease in cucumber caused by F. oxysporum (Butler, 2013). Soybeans, corn, potatoes, and cotton have continuous cropping problems worldwide (Shi et al., 2010; Ma et al., 2015). Continuous cropping can reduce crop production.

The community structure and diversity of soil microbes have been the focus of soil ecology research in recent years. Traditional methods to study microbial diversity are plate counts and morphometric analyses, community level
physiological fingerprints, and various biomarkers (Hill et al., 2000). Modern research methods include real-time quantitative polymerase chain reaction (PCR), clone libraries, fluorescence in situ hybridization (FISH), and terminal-restriction fragment length polymorphism (T-RFLP) (Liebner et al., 2008; Eickhorst and Tippkötter, 2008; Chau et al., 2011). However, the abovementioned methods require pure cultures; otherwise, they cannot overcome the limitations of the species to provide enough data for a robust analysis. By contrast, next-generation high-throughput DNA sequencing offers a suitable way to obtain a large enough dataset by fast detection, especially for uncultured microorganisms occurring in soil, wastewater, human gut, sea, sludge, and other environments (Schlüter et al., 2008; Qin et al., 2010). This method does not rely on clone library construction, fluorescently labeled primers, or probes; other advantages are its large sequencing flux, good parallelism, and highly automated procedures.

In the present study, soybean was sown in a different field after year zero, and continuous cropping was applied for 1 or 2 yr. We used next-generation sequencing to analyze fungal diversity during the branching period of sown soybean that had been infected by fungi in the rhizosphere soil. By comparing the three sets of root samples, we investigated the variation in fungal populations present or inhabiting soybean roots during 3 yr of monocropping.

**MATERIALS AND METHODS**

**Test materials**
The ‘Heinong48’ soybean is a high protein cultivar. Its fat content is 19.50% and protein content is 45.27%. The Soxhlet extraction method was used to detect fat content and the Kjeldahl method was used to detect protein content (Jung et al., 2003; Wan et al., 2011). This study was conducted at the Sugar Industry Research Institute Experimental Station of the Harbin Institute of Technology in Heilongjiang province, China.

**Sample collection and processing**
‘Heinong48’ seeds were sown in three fields at year zero (single planting year and there was no vegetation in the field before it was cleared for experimentation) and 1 and 2 yr of continuous cropping. Soybean roots were harvested at the branching stage from the 0-20 cm soil depth. Specifically, each year, five plants were randomly selected and their root samples (i.e., n = 5 per year) –under the same duration of continuous cropping– were cut and pooled together to produce a single mixed sample for analysis. These three root samples, from year zero and after 1 and 2 yr of continuous cropping, were designated as Fg1y, Fg2y, and Fg3y, respectively.

**Extraction of total root sample DNA**
Total DNA was extracted from the three root samples with the Plant DNA Isolation Reagent kit (Takara, Dalian, China) following the manufacturer’s instructions.

**PCR and 454 pyrosequencing**
For the internal transcribed spacer (ITS) region located between 18S rDNA and 28S rDNA, two fused primers were designed consisting of 5’-(454 A, B adaptor)-(specific primer)-3’: sequences were 5’-TCCGTAGGTGAACCTGCGG-3’ and 5’-TCCTCCGCTTATGATATGC-3’ for ITS1 and ITS4, respectively (White et al., 1990). The A adaptor-containing end was the sequencing end, and a 10 bp barcode was added between the adaptor and the forward primer sequence to allow for sample discrimination during sequencing (Liebner et al., 2008). The PCR was performed in a 20-μL reaction volume, which consisted of 4 μL 5 × buffer, 2 μL 2.5 mM dNTPs, 0.8 μL Upstream primer (5 μM), 0.8 μL downstream primer (5 μM), 0.4 μL polymerase, and 10 ng template DNA (Bai et al., 2015). The reaction system (20 μL) included the following: 94 ℃, 2 min (95 ℃, 30 s; 55 ℃, 30 s; 72 ℃, 30 s) × 30 and 72 ℃, 5 min. Three replicates of each sample were mixed and the target sequence was detected by 2%-AGE (agarose gel electrophoresis) (Bai et al., 2015).

The PCR products were purified with the AxyPrep DNA gel extraction kit (Axygen, Tewksbury, Massachusetts, USA) and their quantitative detection was carried out in the QuantiFluor fluorescence quantitative-ST system (PicoGreen dsDNA Quantitation Reagent, Cohasset, Massachusetts, USA). Pyrosequencing was performed as described by McKenna et al. (2008) on the Roche Genome Sequencer FLX platform (Roche, Basel, Switzerland) using the Roche GS FLX+ sequencing method manual_XLR70 kit.
Pyrosequencing data analysis
After sequencing, initial data were processed and an optimized sequence was obtained. The Seqcln software removed low quality sequences, adaptor and primer sequences. The Seqcln software was used again to detect the adaptor and trim end (http://sourceforge.net/projects/seqclean/), and the mothur software (http://www.mothur.org/wiki/Main_Page) filtered the sequence reads.

The optimized data sequences were evaluated by operational taxonomic unit (OTU) clusters and taxonomic analyses. Based on the OTU approach, alpha-diversity of the three samples was estimated in terms of community richness (Chao, Ace) and diversity (Shannon, Simpson) and sequencing coverage was calculated. Sequencing depth was detected. Based on the taxonomic approach, a statistical analysis was carried out of the fungal community composition in each category level (such as species or genus). The high quality sequences were compared with the SSUrRNA database to find species information and with credibility exceeding 80%. We analyzed the consistency of all sequences in each OTU at 97% similarity level to find species information using mothur software through the naive Bayesian classifier algorithm developed by the remote desktop protocol (RDP) (Wang et al., 2007).

The optimized sequences were OTUs clustered with mothur software, and the sequence was divided at 97% similarity (Schloss et al., 2009). Modified sequences were randomly drawn from each sample, and sequence numbers were plotted against the number of OTUs to draw the rarefaction curves using R language. For each OTU, its sequences were first counted and then OTUs were ranked from large to small in terms of abundance. The results were drawn in rank-abundance curves (using R language to make curves).

In the taxonomy analysis, all the sequences in each OTU were analyzed at a 97% similarity level to determine their species identity. The number of common and unique OTUs in the multiple samples was counted, and their distributions were described by Venn diagram (using R language to make curves). Multiple samples were compared for each category, and a histogram conveyed their community structure at the genus level (using Excel to make curves).

Phylogenetic analysis
The MEGAN 4 software was used to analyze fungal species and species richness of samples contained in the data (MEtaGenome ANalyzer, http://ab.inf.uni-tuebingen.de/software/megan/). By querying the National Center for Biotechnology Information (NCBI), the dendrogram of each sample’s composition was drawn, which was an in-depth analysis of the fungal community structure and phylogeny (Bai et al., 2015).

RESULTS
Polymerase chain reaction (PCR) root sample results
The DNA of each root sample was amplified with the ITS primers. Figure 1 shows PCR amplification results. The target sequence was approximately 750 bp by the 2%-AGE test. The size of the PCR products was correct and the concentration appropriate; hence, they were suitable for sequencing.

Figure 1. Specific polymerase chain reaction (PCR) amplification results for three soybean root samples.

DL2000: Marker; CK: negative control.
Three root samples, from year zero and after 1 and 2 yr of continuous cropping, were designated as Fg1y, Fg2y, and Fg3y, respectively.
Fungal community
Multiple parallel sequencing methods were used to study the fungal composition of the soybean roots sampled from the same field at year zero, and continuously cropped for 1 and 2 yr. After the three root sample were sequenced, 20 630 valid ITS sequences were obtained with a mean length of 527 bp; after filtering, 18 384 optimized sequences (c. 90%) with a mean length of 538 bp were retained. Sequence data for each sample are summarized in Table 1.

Sequencing results were deemed reasonable and sufficient, and they could be used with confidence to analyze the diversity of soybean root fungi. After cluster analysis, optimized sequences were divided into 757 OTUs. The alpha-diversity values of the three samples are listed in Table 2. The coverage of these three libraries was > 98.7%, which would indicate that the sequencing results were valid, and thus representative of real fungal diversity. The larger the Shannon index or the smaller the Simpson index, the higher community diversity was. As a result, continuous cropping could increase root sample diversity, but microbial diversity peaked after 1 yr (Fg2y) and not after 2 yr (Fg3y).

Modified sequences were randomly drawn from each sample, and sequence numbers were plotted against the number of OTUs to plot the rarefaction curves (Figure 2). With the increased number of sequences, curves tended to flatten. This indicated that data was fitted to the analysis because more data would provide only a few more OTUs. The OTU upper limits of Fg2y and Fg3y were 136 and 123, respectively, which exceeded the OTU upper limit (100) of Fg1y. This result was consistent with the diversity indexes (Table 2), and indicated that an increase in community diversity was caused by continuous cropping, although it was more pronounced for 1 yr.

The rank-abundance curve is another way of examining species diversity because it conveys two key aspects of diversity: the abundance and evenness of the species. In the horizontal direction, increased species abundance corresponds to a greater width of the curve. Seen vertically, the shape of the curve—or its smoothness—reflects the evenness of the species in the samples; the smoother the curve, the more uniform the distribution of the species is. The rank-abundance curves of the three root samples eventually became flat at the end, which indicated an even distribution of species (Figure 3). However, the Fg2y and Fg3y curves were wider than the Fg1y curve along the x-axis, especially so for Fg2y in which species abundance peaked.

Taxonomic composition
There were different OTUs in different roots, and the abundance of each OTU varied in each sample (no rank and unclassified were included). These OTUs were classified as 19 phyla, 169 families, and 235 genera. The dominant phyla were Ascomycota and Basidiomycota; except for no rank or unclassified, their proportion of taxa in Fg1y, Fg2y, and Fg3y

| Table 1. Optimized sequential data statistics. |
|-----------------|-----------------|-----------------|-----------------|
| Sample          | Valid           | Trimmed         | Percent         |
| Fg1y            | 7082            | 6426            | 90.74%          |
| Fg2y            | 7331            | 6461            | 88.13%          |
| Fg3y            | 6217            | 5497            | 88.42%          |

Three root samples, at year zero and after 1 and 2 yr of continuous cropping, were designated as Fg1y, Fg2y, and Fg3y, respectively.

<table>
<thead>
<tr>
<th>Table 2. Diversity and richness indexes of the fungal communities in the three root samples with &lt; 97% sequence identity.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample</td>
</tr>
<tr>
<td>Fg1y</td>
</tr>
<tr>
<td>Fg2y</td>
</tr>
<tr>
<td>Fg3y</td>
</tr>
</tbody>
</table>

Ace and Chao: Community richness estimators; Coverage: sampling coverage; Shannon and Simpson: community diversity estimators.
Three root samples, from year zero and after 1 and 2 yr of continuous cropping, were designated as Fg1y, Fg2y, and Fg3y, respectively.
were 79.45%, 58.79%, and 78.86% for Ascomycota, respectively, and 16.44%, 32.14%, and 16.57% for Basidiomycota, respectively. Leotiomycetes was the dominant class in Ascomycota. The proportion of taxa in Fg1y, Fg2y, and Fg3y was 54.55%, 26.43%, and 38.66%, respectively.

To visually display the numbers, similarity, and overlap of the three samples, a Venn diagram was drawn (Figure 4). Forty-eight common OTUs were shared in the continuous cropping soybean root samples; 30 were also shared in Fg1y. The result indicated that continuous cropping could increase root fungal diversity, with 18 OTUs produced by continuous cropping (Chen et al., 2012).

**Figure 2.** Rarefaction curves of the operational taxonomic units (OTUs) clustered at < 97% sequence identity for three soybean root samples.

**Figure 3.** Operational taxonomic unit (OTU) rank-abundance curves for OTUs in three soybean root samples.
The Venn diagram in Figure 4 shows the quantitative variation of the samples because it cannot show the taxonomic differences of the OTUs. Hence, the community composition and the relative abundance at the genus level of each sample were analyzed (Figure 5). The number of fungal genera was higher in Fg2y and Fg3y than in Fg1y. Nevertheless, root rot pathogenic fungi occurred in all three samples. Infection was stable, although pathogenic fungi in the three samples were not identical. Ascomycota and Basidiomycota were the dominant fungi in soybean root under continuous cropping. The relative abundances of Mortierellales and Pezizomycotina clearly increased in the root samples from year zero, (0.05% and 0.19%, respectively) to 1 yr after continuous cropping (0.19% and 0.52%, respectively).

Phylogenies were inferred based on the OTU abundances and their corresponding taxonomic information. Using this classification system, fungal evolution and abundance relationships of the three samples are shown in Figures 6a to 6c. The fungal composition of each sample varied widely in terms of the dominant genera and content of each genus. Continuous cropping could increase the relative abundance of some fungi, namely Fusarium, Rhizoctonia, and Thelebolus, which

Figure 4. Venn diagram of operational taxonomic units (OTUs) in three soybean root samples.

Figure 5. Composition of the different fungal communities at the genus level for three soybean root samples.
were associated with soybean root rot in after 2 and 3 yr of cropping. Continuous cropping could also increase the abundance of *Gymnoascus*, *Chrysosporium*, *Ctenomyces*, *Aphanoascus*, and *Aspergillus*, which are soil pathogenic fungi that can cause other plant diseases.

By contrast, the abundance of *Verticillium*, –absent in the year zero sample– clearly increased in the root sample that was continuously cropped for 2 yr. These results suggest that as the duration of continuous cropping increased, the growth and infection of soybean root rot pathogenic fungi were evident.

**DISCUSSION**

**Root rot incidence**

Previous studies have shown that root rot incidence is closely related to the soil microbial community, and that soybean continuous cropping affects the diversity of rhizosphere fungi (Berendsen et al., 2012; Bai et al., 2015; Cui et al., 2016). We showed the variation in fungal populations present or inhabiting soybean roots under 3 yr of monocropping. We used high-throughput sequencing techniques to characterize the fungal community of three root samples: control and continuous cropping for 1 and 2 yr. Although high-throughput sequencing techniques were successfully used to study microbial diversity in rhizosphere soil (Acosta-Martínez et al., 2008; Buée et al., 2009; Jumpponen et al., 2010), few have been used as directly for root systems. Our results not only provide a theoretical basis for the following inoculation study
of the different pathogens involved in soybean root rot disease, the variation in fungal populations, and root rot incidence, but they also provide a sound knowledge base for disease prevention and control measures that could significantly increase soybean yield and quality.

This result shows that continuous cropping clearly increased fungal diversity in soybean root, but that this effect differed between years. Continuous cropping significantly reduces soil pH, making the soil conducive to fungal growth while inhibiting the reproduction of bacteria and actinomycetes. As a result, fungi eventually become the dominant species in rhizosphere soil (Vargas-Gil et al., 2011). The number of fungal genera was higher in Fg2y and Fg3y than in Fg1y. It showed that continuous cropping could induce great changes in the microbial flora of soybean root, which was infected by rhizosphere fungi. Most fungi that were found belonged to genera in the phylum Ascomycota, class Leotiomycetes (Larson et al., 2006). Other studies have shown that most fungi in Leotiomycetes can cause severe plant diseases (Li et al., 2013).
PCR primer and program selection
To ensure the accuracy and reliability of subsequent data analyses, two conditions should be established: use low-cycle number amplification as much as possible and check that the number of cycles amplified in each sample is consistent. With this in mind, we selected a representative sample for preliminary experiments to ensure that the majority of samples were amplified at the lowest number of cycles to produce suitable product concentration.

In the results, the high relative abundance of *Papilionoideae* species, which are not fungi, was likely caused by a defect in the PCR primers; they directly amplified all eukaryote DNA and not only of fungi. The primer did not effectively exclude the plant host data, but it did not affect fungal diversity analysis (Wensing et al., 2010).

Comparison with other community studies
The information provided by high-throughput sequencing is more effective than traditional methods. The ratios of unclassified fungi and no rank fungi increased with sequencing depth. A given OTU is defined as no rank or unclassified
when there is no taxonomic name for it in NCBI, the taxonomic information is below the reliability threshold, or there is no taxonomic information at all. This suggests that high-throughput sequencing can help to discover new fungal species in crop soil and also detect rare and infrequent fungi (i.e., those with very low abundance). Several of our study results were consistent with previous work. For example, the presence of Mortierellales and Pezizomycotina increased in the root samples when soybean was under continuous cropping for 1 yr. Similarly, Mortierella was the dominant pathogenic fungus for the stubble of continuously cropped Chinese pepper (Zhao et al., 2008). This result suggests a close positive correlation between increased fungi and continuous cropping. Pezizomycotina species are ecologically diverse and function in key ecological processes and symbioses, including plant decay, although some function as plant pathogens. However, their other functions on the impediment of continuous cropping require in-depth research. With the rapid development of biological research methods, new techniques are being increasingly applied to study soil microbes (e.g., proteomics research). Proteomics facilitate the analysis of the expression level of a protein in different environments, such as planted crops. It will be an effective method to study the pathogenic mechanism of pathogenic fungi. When the application of new metaproteomics and high-throughput sequencing are combined, they will open a promising new chapter in the study of soybean rhizosphere microbes under continuous cropping.

**Data accessibility**
The datasets generated during the current study are available in the Sequence Read Archive (SRA)-National Center for Biotechnology Information (NCBI) repository (BioProject: PRJNA395107).

**CONCLUSIONS**
To further study root infection by fungi in rhizosphere soil, we used high-throughput sequencing techniques to characterize the fungal community of three root samples: control and continuous cropping for 1 and 2 yr. Results showed that continuous cropping clearly increased fungal diversity in soybean root, but that this effect differed between years. The number of fungal genera was higher in the second and third years of monocropping than 1 yr of soybean cropping. This shows that continuous cropping could induce great changes in the microbial flora of soybean root, which was infected by rhizosphere fungi. Ascomycota and Basidiomycota were the dominant phyla in the continuous cropping root samples. The relative abundance of some fungi, such as Fusarium, Rhizoctonia, and Thelebolus, associated with soybean root rot disease noticeably increased. Gymnoascus, Chrysosporium, Ctenomyces, Aphanoascus, and Aspergillus, which were soil pathogenic fungi that can cause other plant diseases, showed an observable increase that was induced by the continuous cropping treatment.

**ACKNOWLEDGEMENTS**
This study was supported by grants from the National Natural Science Foundation (31570487).

**REFERENCES**

Zhang, X., Li, X., and Zhang, H. 2011. Ecological risk of long-term chlorimuron-ethyl application to soil microbial community:

White, T.J., Bruns, T., Lee, S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for

Pantoea stewartii

Wensing, A., Zimmermann, S., and Geider, K. 2010. Identification of the corn pathogen


Shi, P., Gao, Q., Wang, S.P., and Zhang, Y. 2010. Effects of continuous cropping of corn and fertilization on soil microbial


Soxhlet extraction. Port Health Control 2011-02.


White, T.J., Bruns, T., Lee, S., and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for


Zhang, X., Li, X., and Zhang, H. 2011. Ecological risk of long-term chlorimuron-ethyl application to soil microbial community:
