Peer-reviewed version available at *Int. J. Mol. Sci.* **2018**, *19*, 3421; doi:10.3390/ijms19113421

Composition and Genetic Diversity of the *Nicotiana tabacum* Microbiome in Different Topographic Areas and Growth Periods

Xiaolong Yuan [1], Xinmin Liu [1], Yongmei Du[1], Guoming Shen[1], Zhongfeng Zhang [1], Jinhai Li [2], Peng Zhang [1*]

1.Tobacco Research Institute of Chinese Academy of Agricultural Sciences, Qingdao, China, 266101.
2. China National Tobacco Corporation in Hubei province, Wuhan, China,430000.

*Correspondence: Peng Zhang, zhangpeng@caas.cn.

Abstract: Fungal endophytes are the most ubiquitous plant symbionts on earth and are phylogenetically diverse. Studies on the fungal endophytes in tobacco have shown that they are widely distributed in the leaves, stems, and roots, and play important roles in the composition of the microbial ecosystem of tobacco. Herein, we analyzed and quantified the endophytic fungi of healthy tobacco leaves at the seedling (SS), resettling growth (RGS), fast-growing (FGS), and maturing (MS) stages at three altitudes [600 (L), 1000 (M), and 1300 m (H)]. We sequenced the ITS region of fungal samples to delimit operational taxonomic units (OTUs) and phylogenetically characterize the communities. The result showed the number of clustering OTUs at SS, RGS, FGS and MS were greater than 170, 245, 140, and 164, respectively. At the phylum level, species in Ascomycota and Basidiomycota had absolute predominance, representing 97.8% and 2.0 % of the total number of species. We also found the number of unique fungi at the RGS and FGS stages were higher than those in the other two stages. Additionally, OTU richness was determined by calculating the Observed

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

Species, Shannon, Simpson, Chao1, ACE, Goods coverage and PD_whole_tree indices based on the total number of species. Our results showed RGS samples had the highest diversity indices. Furthermore, we found that the diversity of fungal communities tended to decrease with increasing tobacco growth altitude. The results from this study indicated that the tobacco harbors an abundant and diverse endophytic fungal population, which provides new opportunities for exploration their potential utilization.

Key words： *Nicotiana tabacum*; Topographic; Growth period; Fungal community composition; Genetic Diversity; ITS1 region

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

1. Introduction

Fungal endophytes are those microorganisms that live intercellularly or intracellularly in plants without causing host tissue damage [1]. There is a symbiotic relationship between endophytes and plant tissues, whereby plants can provide nutrition and protection for endophytes, and the endophytes in turn have significant influences on plant, protecting their host from biotic and abiotic stresses [2]. Previously, endophytic fungi have been isolated from the healthy strains of tobacco (*Nicotiana tabacum*) [3,4]. Studies of the fungal endophytes in tobacco have shown that they are widely distributed in the leaves, stems, and roots, and play important roles in the composition of the microbial ecosystem of tobacco [5]. Furthermore, these endophytes have various biological functions. Endophytic fungi in tobacco not only promote the growth of tobacco and enhance the capacity of resistance to diseases and pests and abiotic stresses (e.g., resistance to heavy metal and drought stress) but also play positive roles in decreasing the nitrosamine contents in tobacco and improving the quality of tobacco leaves. Therefore, it is desirable to determine the species distribution and characteristics of endophytic fungi in tobacco through a comprehensive investigation at different growth stages and altitudes.

Traditionally, the taxonomic identification and classification of fungi has been based mainly on their morphologies, growth conditions, and physiological and biochemical characteristics [6]. However, fungal taxa are complex, often lack discriminatory morphological characters, and are susceptible to environmental change. Moreover, most species are not amenable to in vitro culturing [7]. Consequently, it is

Peer-reviewed version available at *Int. J. Mol. Sci.* **2018**, *19*, 3421; doi:10.3390/ijms19113421

often difficult to identify species according their morphologies. With the development of molecular biology techniques, DNA sequence analysis has become widely applied in the taxonomic identification of fungi. In this regard, the nuclear ribosomal internal transcribed spacer (ITS) region and 18SrDNA sequences are widely used markers [8-11]. The fungal ITS region includes the ITS1, ITS2, and 5.8S genes. Typically, the 5.8S gene is highly conserved compared with the ITS1 and ITS2 sequences. Thus, the ITS1 and ITS2 genes are generally used for resolution at the genus and species levels [12-14]. High-throughput sequencing technologies can facilitate the identification of low abundance species and can reduce costs and have thus become the first choice for fungal community diversity analysis. Therefore, based on the characteristics of ITS1 region, we can construct pair-end libraries and determine sequences using technologies such as the Illumina HiSeq sequencing platform. Furthermore, we can analyze the abundance of species by clustering operational taxonomic units (OTUs) and determine the differences among samples by analyzing their alpha and beta diversities.

In this study, we performed a metagenomic analysis of fungal communities at four growth stages (seedling, resettling growth, fast-growing, and maturing stages) of healthy tobacco plants grown at three altitudes (600 m, 1000 m, and 1300 m) via high-throughput sequencing. Our main goal was to investigate the diversity and heritability of the *Nicotiana tabacum* microbiome. Specifically, we sought to (i) compare the fungal communities at different growth stages, and (ii) examine the effects of altitude on fungal community characteristics. The findings will not only provide a theoretical

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

basis for the control of tobacco diseases, but also can lay foundations for the development and study of endophytic fungi in tobacco.

2. Materials and Methods

2.1 Material collection

Samples were collected from the Modern Tobacco Agricultural Science and Technology Demonstration Garden on Wangcheng Slope, Enshi, Hubei (108°23′12″–110°38′08″E, 29°07′10″–31°24′13″N). The tobacco variety used was Yunyan87. We sampled healthy tobacco plants at four growth stages, the seedling stage (SS), resettling growth stage (RGS), fast-growing stage (FGS), and maturing stage (MS), grown at three altitudes (600, 1000, and 1300 m). We collected upper, middle, and lower leaves as samples and then mixed them. The samples from the seedling stage at 600 m (L), 1000 m (M), and 1300 m (H) were designated SSL1, SSL2, SSL3, SSM1, SSM2, SSM3, SSH1, SSH2, SSH3. Similarly, samples from the other three growth stages were designated RGSL1, RGSL2, RGSL3, RGSM1, RGSM2, RGSM3, RGSH1, RGSH2, and RGSH3; FGSL1, FGSL2, FGSL3, FGSM1, FGSM2, FGSM3, FGSH1, FGSH2, and FGSH3; and MSL1, MSL2, MSL3, MSM1, MSM2, MSM3, MSH1, MSH2, and MSH3. Every samples were collected in three replicates. All samples were immediately chilled on ice and then stored at -80°C as soon as possible until microbial DNA extraction.

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

2.2 DNA extraction and PCR amplification

Total genomic DNA of each sample was obtained from 100 mg material using the CTAB method [15]. The integrity of DNA was determined by 1% agarose gel electrophoresis. The purity and concentration of DNA were determined using a NanoDrop 2000c spectrophotometer. Each DNA sample was diluted to a final concentration of 1 ng/μL with sterile distilled water and this was used as a DNA template. The ITS1 sequences were amplified using the primer pair CTTGGTCATTTAGAGGAAGTAA and GCTGCGTTCTTCATCGATGC, which included barcodes and adapters. The PCR reactions were carried out in 20-μL reaction mixtures, containing 5.8 μL sterile distilled water, 10.0 μL 5× Phusion® High-Fidelity PCR Master Mix with GC Buffer (New England Biolabs), 1.0 μL forward primer (10 μM), 1.0 μL reverse primer (10 μM), 0.2 μL Taq DNA Polymerase, and 2.0 μL template DNA. The PCR amplification conditions were as follows: initial denaturation at 95°C for 2 min, 30 cycles of denaturation at 95°C for 30s, annealing at 50 °C for 20 s, elongation at 72°C for 60 s, and elongation at 72°C for 10 min, and then held at 4°C.

2.3 Library construction and sequencing

The PCR amplicons were detected using 2% agarose gel electrophoresis. All the amplicons were then pooled in equimolar ratios into a single tube. The mixed amplicons were further detected using 2% agarose gel electrophoresis and the target sequences were extracted using a QIAquick Gel Extraction Kit (Qiagen, 28704). The

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

libraries were constructed using a TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, 20015963). The libraries were quantified using a Qubit® 2.0 Fluorometer (Life Technologies) and qPCR, and then sequenced using the Illumina HiSeq 2500 platform. All sequence data have been submitted to the Sequence Read Archive (SRA: SRP145216) and are freely available at the NCBI (BioProject: PRJNA464258).

2.4 Data processing

For the sequencing reads, primers and barcodes were removed. The reads of each sample were then assembled using FLASH (V1.2.7) (http://ccb.jhu.edu/software/FLASH/) [16] to generate raw tags. Qiime（V1.7.0）(http://qiime.org/scripts/split_libraries_fastq.html) [17] was used with the following parameters for quality filtering: (a) tag truncating: raw tags from the continuous low-quality value (the default quality value < = 19) with the setting length (the default value is 3) were truncated at the first low-quality base site; (b) tag length filtering: removing the length <75% tags. Chimeras were checked against the Unite database (https://unite.ut.ee/) [18] using the UCHIME Algorithm (http://www.drive5.com/usearch/manual/uchime_algo.html) [19] and then removed. Finally, the effective tags were obtained.

2.5 OTU clustering and annotation

Uparse (Uparse v7.0.1001, http://drive5.com/uparse/) [20] was used to cluster all the effective tags. The effective tags with 97% identity can be clustered into the same OTUs. The OTUs with the highest frequency were selected as the representative of

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

OTU sequences. For a further determination of their function, these OTUs were annotated using the blast method with Qiime (Version 1.7.0) [21] based on the Unit database (https://unite.ut.ee/) [22]. Thereafter, the fungal communities at different taxonomic levels (kingdom, phylum, class, order, family, genus, and species) were counted. The phylogenetic relationship of all the OTUs were initially blasted using MUSCLE (Version 3.8.31, http://www.drive5.com/muscle/) [23], and then we analyzed their phylogenetic relationships using MrBayes 3 [24]. Finally, we performed homogenization processing of the data of the various samples for subsequent alpha and beta diversity analyses.

2.6 Alpha diversity

The Observed-species, Chao-1, Shannon, Simpson, ACE, Goods-coverage, and PD_whole_tree diversity indices were determined using Qiime (Version 1.7.0) [21]. Chao and ACE indices were used to determine community richness, and Shannon and Simpson indices were used to determine community diversity. These metrics are useful for estimating microbial diversity and richness. Coverage represents the sequencing depth. The PD_whole_tree index was used to compute Faith's phylogenetic diversity metric. R software (Version 2.15.3) was used to draw the dilution, rank abundance, and species accumulation curves, and to analyze the difference in alpha diversity indices among all groups.

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

2.7 Beta diversity

To calculate beta diversity, Unifrac distance metrics were calculated and used to construct UPGMA using Qiime software (Version 1.7.0) [21]. Subsequently, the ade4 and ggplot2 packages, WGCNA, stats and ggplot2 packages, and vegan packages of R software (Version 2.15.3) were used to draw principal coordinate analysis (PCoA), and NMDS diagrams, respectively. In order to obtain a better insight into the clustering of fungal communities, both weighted (takes into account changes in relative taxon abundance) and unweighted UniFrac metrics were used for beta diversity [25]. Metastats analysis were performed on different taxonomic levels (phylum, class, order, family, genus, and species), using the permutation test between groups.

3. Results

3.1 Sequencing results

Samples were successfully sequenced from 36 samples representing four growth periods (SS, RGS, FGS, and MS stages) and three altitudes (600, 1000, and 1300 m). The raw sequencing reads of samples ranged from 70,606 to 89,979 bp with an average length of approximately 310 bp (Table S1). Barcodes and primers were removed from the raw sequences prior to assembly. We assembled the raw reads into raw tags. Effective tags were finally obtained after qualification and removal of chimeras. In order to study the species diversity of a sample, the effective tags of samples with more than with 97% identity were clustered into OTUs. The

representativeness of OTUs were annotated for functional analysis. The number of clustering OTU for stages SS, RGS, FGS, and MS were greater than 170, 245, 140, and 164, respectively. The advantage species of each sample were annotated for OTUs and the result are shown in Figure 1.

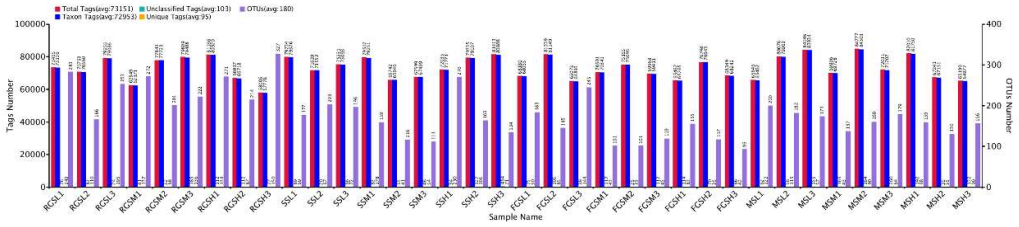

Figure 1. Statistics of operational taxonomic unit (OTU) clustering. The first ordinate is the number of tags. The red total tags (red) refer to the total number tags of each sample (equal to the effective tags used for OTU clustering and subsequent analysis). The unique tags (orange) refer to the unable OTU sequences. The taxon tags (blue) refer to the tags that were used to construct the OTUs and were annotated. Unclassified tags (green) did not have the annotation information of the number of tags. The second ordinate is the number of OTUs of each number.

3.2 Fungal taxonomic composition

For the annotation of all OTUs, we screened the identified species according to the classification results of each sample. At the phylum level, species in Ascomycota and Basidiomycota showed virtually absolute predominance, representing 97.8% and 2.0% of the total number of species, respectively. The top species at the class level were dominated by Leotiomycetes (90.74%), followed by Dothideomycetes (6.01%) and Agaricomycetes (0.85%). Erysiphales, Pleosporales, Capnodiales, Polyporales, and Agaricales, comprising approximately 90.71%, 4.97%, 0.58%, 0.27%, and 0.30%, respectively, of the samples, were the dominant groups at the order level. At the family level, Erysiphaceae accounted for 90.71% of the entire community, whereas

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

Pleosporaceae, Davidiellaceae, Didymosphaeriaceae, Meruliaceaea, and Leptosphaeriaceae together contributed to 5.19% of the total community in all samples. In terms of genera, the abundances of *Microidium*, *Cladosporium*, *Pseudopithomyces*, *Irpex*, and *Lysurus* were higher than those of other genera (Figure 2). Network analysis provides us with a new perspective for studying complex microbial community structures and functions. Given that relationships may differ under different environmental conditions, the advantage of dominant species and the species with close interactions can be directly determined [26]. Through analyzing the correlation index of all the sample calculations, with a cutoff = 0.6 to filter the absolute value of correlation coefficient, coupled with species abundance, we constructed a figure showing the network of samples at the genus level. In our study, *Undeniomyces* and *Rhodotoruta* in the Basidiomycota were the genera with the highest abundance and were found to be located in the "hub" position of the network. *Ascomycota, Microidium, Trimmatostroma, Ophiognomonia, Trichomerium,* and *Paraphaeosphaeria* were the top five genera in terms of richness. Moreover, we found that *Undeniomyces* had a negative relationship with *Microidium* (Figure 3).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 August 2018 doi:10.20944/preprints201808.0268.v1

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

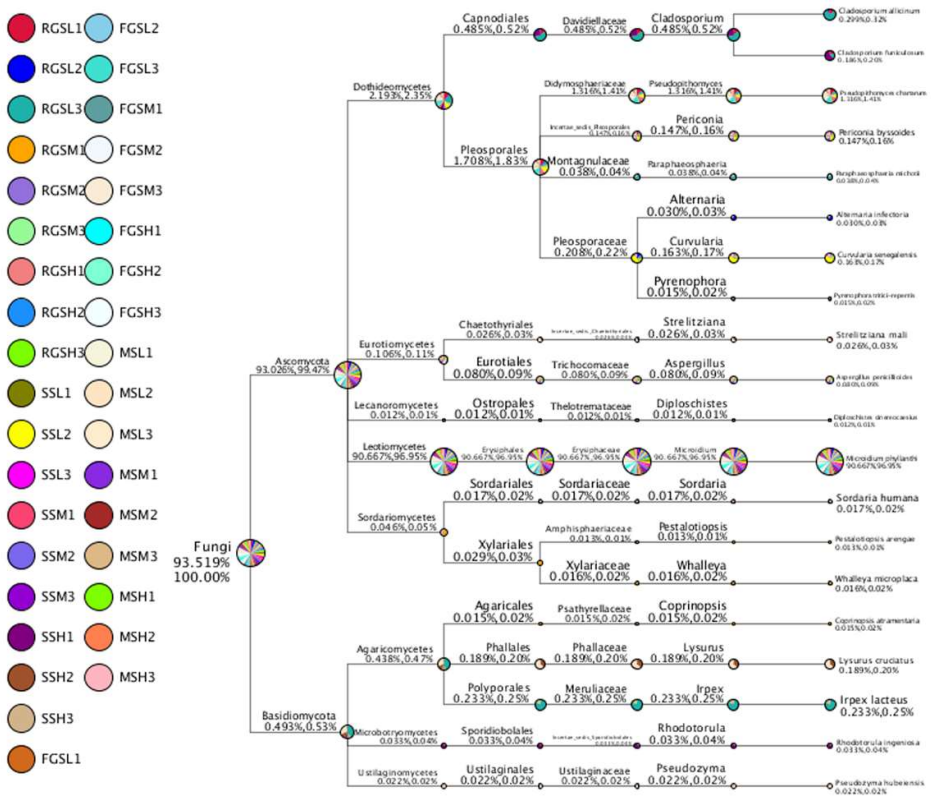

Figure 2 Taxonomic tree of all samples

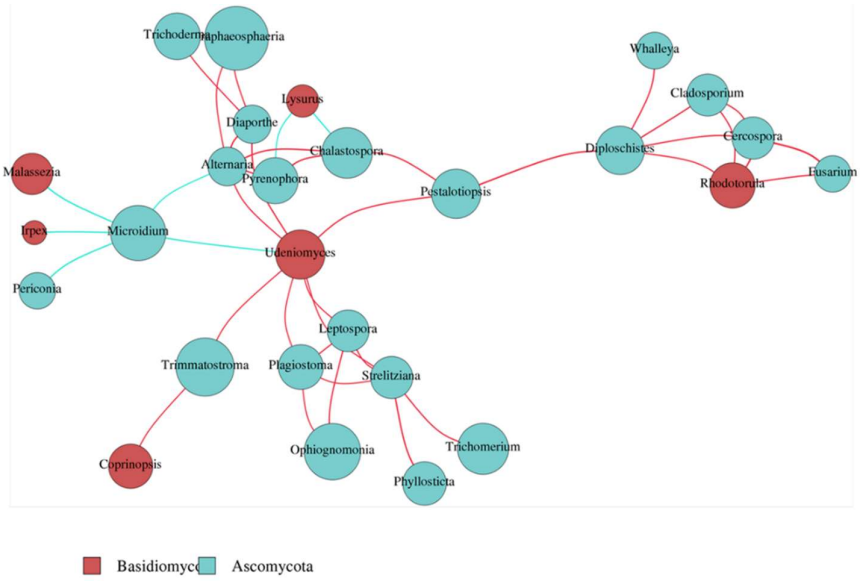

Figure 3. Network analysis at the genus level. Different nodes represent different genera, and node size represents the average relative abundance. Nodes with the same color represent the same phylum. The thickness of the attachment is the correlation coefficient with the absolute value. Red lines indicate a positive correlation and blue lines indicate negative correlation.

3.3 Dissimilarity of fungal communities among different growth stages

A comparison of the OTUs of the different samples indicated that SS samples shared 98 common OTUs, and that samples SSL, SSM, and SSH had 120, 52, and 140 unique OTUs, respectively. The 98 common OTUs were annotated to 58 species. For the RGS samples, 151 core OTUs were found at different altitudes. RGSL, RGSM, and RGSH samples contained 134, 92, and 174 unique OTUs, respectively. Further, 76 and 112 common OTUs were found in the FGS and MS samples (Figure 4). Thus, among the four growth stages examined, we identified a higher abundance of unique OTUs at the RGS and FGS stages.

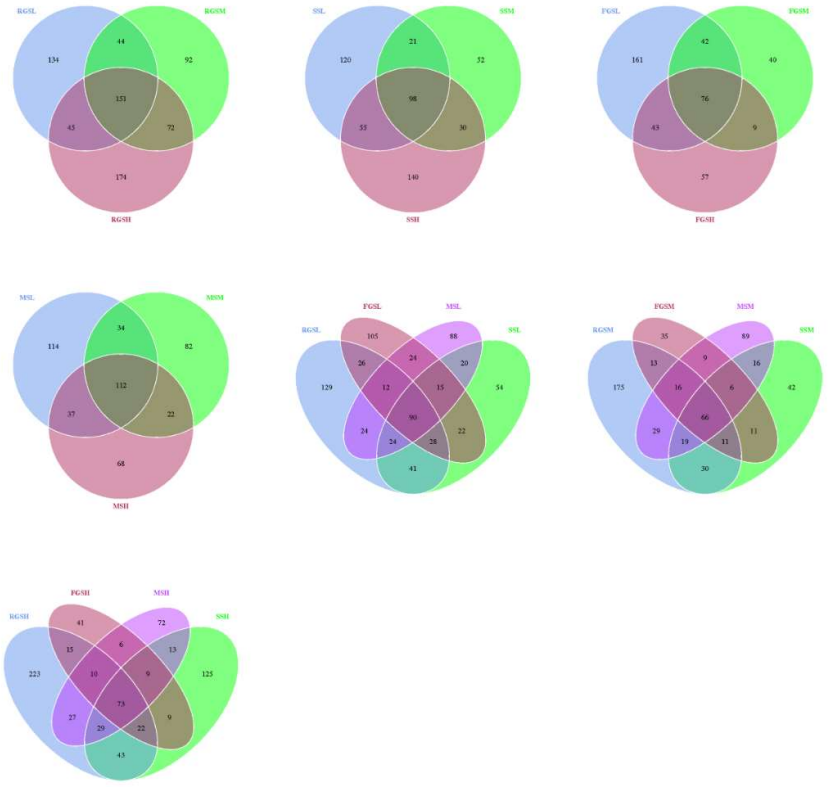

Figure 4. Venn diagram of all samples. Each circle in the figure represents a sample. The number in the circle represents the operational taxonomic unit (OTU) number. The overlap number represents the common OTUs among samples, whereas the no overlap number represents the unique OTUs in the sample.

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

3.4 Dissimilarity of fungal communities among different growth altitudes

When we compared the OTUs of samples from different growth stages at 600 m, we found that 90 core OTUs were common to all four stages. We selected the top five species for comparison. At the seedling stage, *Microidium phyllanthi, Pseudopithomyces chartarum, Curvularia senegalensis, Periconia byssoides,* and *Pseudozyma aphidis* were the most abundance species at 600 m, with percentages of 90.43%, 1.92%, 1.22%, 0.38%, and 0.24%, respectively. Similarly, *Microidium phyllanthi, Pseudopithomyces chartarum, Pseudozyma aphidis, Candida albicans* and *Penicillium oxalicum* were the five most abundant species at 1000 m, representing 94.80%, 0.36%, 0.13%, 0.07%, and 0.07% of the total species, respectively. At 1300 m, *Microidium phyllanthi* (90.97%) was the most abundant species, followed by *Cladosporium funiculosum* (1.39%), *Lysurus cruciatus* (0.77%), *Cladosporium allicinum* (0.03%), and *Rhodotorula ingeniosa* (0.03%). In the resettling growth stage, *Microidium phyllanthi* (69.19%), *Cladosporium allicinum* (0.03%), *Pseudopithomyces chartarum* (0.29%), *Irpex lacteus* (0.23%), and *Cladosporium funiculosum* (0.08%) were the dominant species at 600 m. At 1000 m, *Microidium phyllanthi*, *Pseudopithomyces chartarum*, *Periconia byssoides*, *Curvularia senegalensis* and *Whalleya microplaca* comprised approximately 84.16%, 0.02%, 0.02%, 0.02%, and 0.02%, respectively, of the total species. At the fast-growing and maturing stages, *Microidium phyllanthi* and *Pseudopithomyces chartarum* were dominant species at all three altitudes, comprising more than 94% of the total species. (Figure 5). Totoally, we

found that the diversity of fungal communities tended to decrease with increasing tobacco growth altitude.

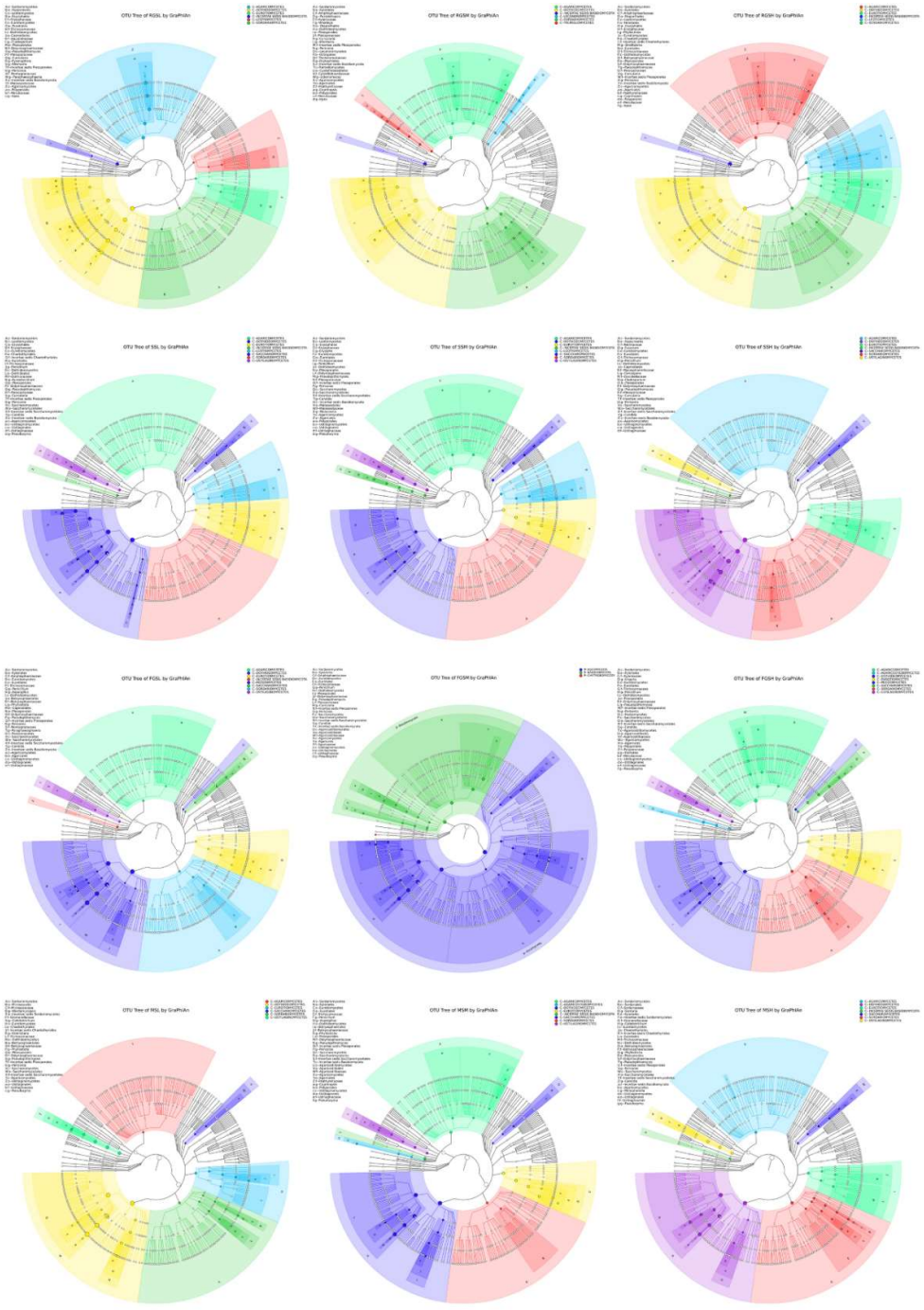

Figure 5. Hierarchy tree based on the operational taxonomic unit (OTU) annotation results. The circles from the inside to outside represent the classification of different levels, and species abundance is proportional to the size of the circle. The different

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

colors represent different phyla and the abundances of the top 40 species are shown in solid circles.

3.5 Alpha diversity

Qiime was used to compare the alpha diversity of all the samples [21]. Rank abundance curves were drawn based the relative abundance of each sample and ranked from more to less of the OTU numbers, which can reflect the richness and evenness of species in the samples (Figure 6). Additionally, the OTU richness was determined by calculating the Observed Species, Shannon, Simpson, Chao1, ACE, Goods coverage, and PD_whole_tree indices based on the total number of species. The alpha diversity indices are shown in Table 1. The number of Observed Species was highest in the RGSH samples at 238.33 ± 61.61, was lowest in the FGSM samples at 88.67 ± 12.66. The Shannon diversity index gives more weight to the numerically dominant OTUs (Garcia-Mazcorro J F et al., 2017). Our result showed that RGSL samples had the highest Shannon indices with an average value of 2.21 ± 1.14. In contrast, the lowest Shannon index values (0.90 ± 0.27) were obtained for the FGSH samples. The Chao1 index is also an accurate estimator of richness. We found that RGSH (261.76 ± 59.95) and FGSM (99.51 ± 17.63) samples had the highest and lowest Chao1 values, respectively. Similarly, values for the ACE and PD_whole_tree indices were highest for RGSH samples and lowest for FGSM samples, with the values of 268.72 ± 60.71 and 105.90 ± 23.83 and 65.16 ± 8.26 and 31.27 ± 5.26, respectively. The Simpson indices of all the samples ranged from 0.25 to 0.52, whereas we found little difference in the values of Goods coverage of all samples, all of which were nearly equal to one.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 August 2018 doi:10.20944/preprints201808.0268.v1

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

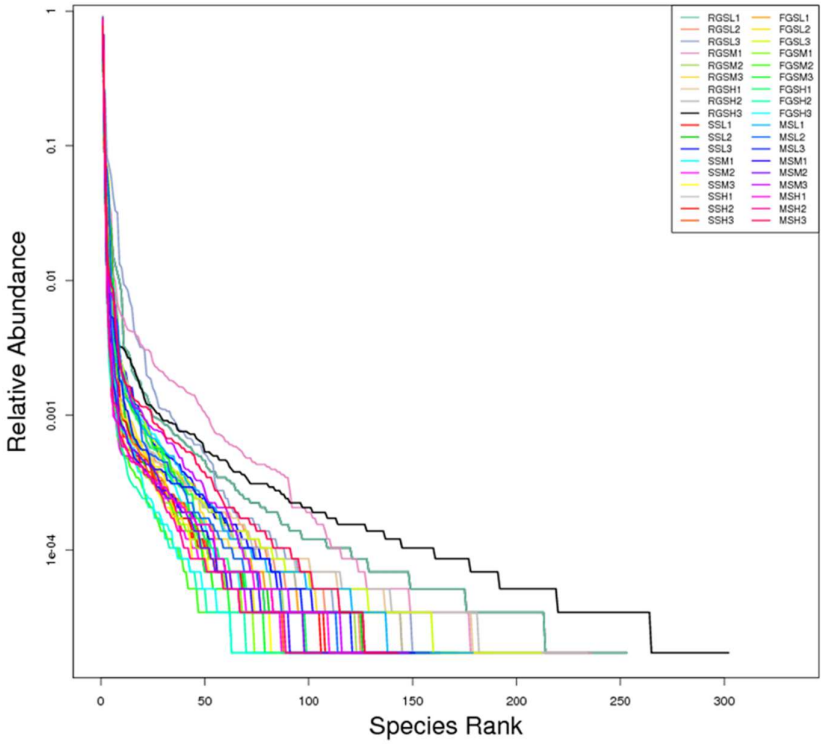

Figure 6 Rank abundance plots of fungal community composition in *Nicotiana tabacum*.

3.6 Beta diversity

Beta diversity is a comparative analysis of microbial community constitutions for different samples. To determine beta diversity, we initially obtained information on species abundance based on the annotation results of all the species and the abundance information of OTUs. Thereafter, we calculated the Unifrac distance based on the phylogenetic relationship among OTUs [17]. Unifrac distance refers the sample distance calculated from the information on samples sequences. Distance matrixes can be obtained for constructing weighted Unifrac distance [25]. Finally, PCoA and NMDS were used to show the differences between different samples. PCoA based on unweighted Unifrac distances demonstrated that MSL, MSM, and MSH samples

tended to cluster together. A similar pattern was observed in other three stages (Figure 7). The highest variations in the microbiota of different samples were 12.34% (PC1) and 8.98% (PC2), representing a strong separation based on the different stages and attitudes (Figure 7a). NMDS is a ranking method that is applicable to ecology. Nonlinear NMDS models can overcome the shortcomings of linear models and reflect the nonlinear structure of ecology data [27]. NMDS analysis results based on OTU levels were shown in Figure 7b. These results indicated that different growth stages play an important role in shaping fungal communities in *N. tabacum*. The stress of all samples in our study was 0.194, which can accurately reflect the differences among the samples. The result showed that there were high similarities among the samples collected at different growth stages. On the basis of the similarities of fungal communities, samples were divided into three groups, namely, those containing samples RGSL, samples RGSM and RGSH; and the remaining samples formed a community These clusters confirmed that growth stage is the most important factor influencing the degree of difference in beta diversity among the fungal communities of the multiple samples. Both the results of PCOA and NMDS showed that the higher the elevations are, the smaller the differences between samples are.In order to clearly compare the beta diversity among samples, we generated boxplots to show the variation among all samples. The weighted UniFrac distance indicated that the RGSL sample had the highest variation (Figure 8). Furthermore, we identified a trend indicating that the diversity of fungal communities tended to become lower with increasing tobacco growth altitude. Interestingly, the fungal communities in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 August 2018 doi:10.20944/preprints201808.0268.v1

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

samples are highly similar, apart from samples FGSH and RGSL (P = 0.0053), FGSM and RGSL (P = 0.0151), MSH and RGSL (P = 0.0285), MSM and RGSL (P = 0.0218), and RGSL and SSM (P=0.0239), which suggests that the fungal communities in these samples are extremely differentiated from each other.

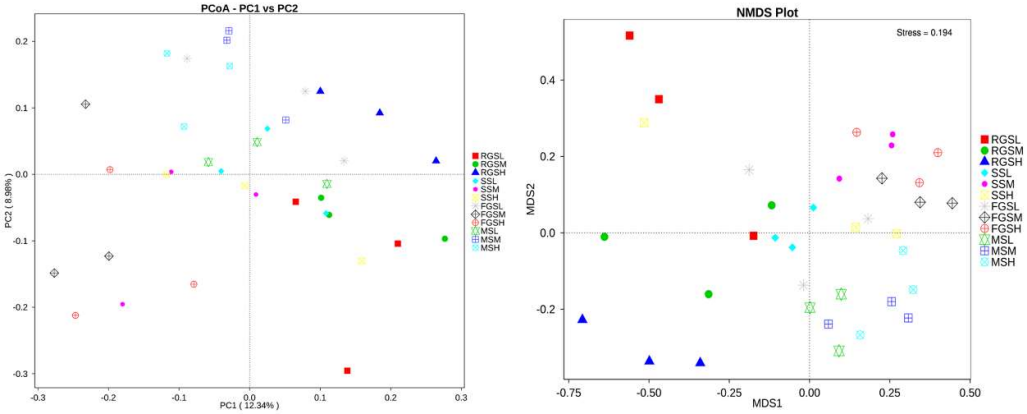

Figure 7. The beta diversity indices of different samples in *Nicotiana tabacum* (a) Principal coordinate analysis (PCoA) based on the weighted and unweighted Unifrac distances. (b) NMDS analysis results based on operational taxonomic unit (OTU) levels. Each point represents one sample and the distance between points indicates the difference degree.

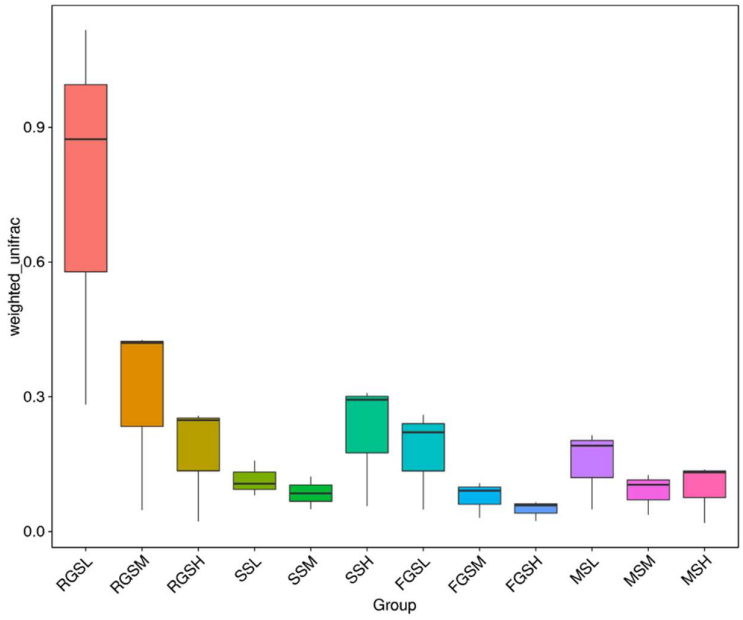

Figure 8. Box plots of beta diversity based on the weighted and unweighted Unifrac distances.

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

4. Discussion

Endophytic fungi live within plant tissues causing no harm to their host species and are considered plant mutualists. They receive nutrition and protection from the host plant, whereas the host plant may benefit from enhanced competitive abilities [2]. Furthermore, it has been demonstrated that endophytic fungi play important roles in providing nutrients to their hosts, adapting hosts to their environments, defending hosts from biotic and abiotic stresses, and promoting plant community biodiversity [28]. Recently, the diversity of endophytic fungi has been assessed by analyzing the ITS region using high-throughput sequencing techniques. For example, in 2012, Li et al. investigated the diversity of endophytic fungi in five flowering plant species based on morphological characteristics and ITS analysis [29]. In 2017, Liu et al. analyzed and compared the fungal species richness, diversity, and community composition of *Dendrobium officinale* in seven different geographic areas in China using a high-throughput sequencing technique [30]. The endophytic fungi of *N. tabacum* play important roles in the microbial ecosystem of fields in which this crop it is cultivated. It has been reported that they are widely distributed and have various biological functions. For example, they can contribute to the growth of their hosts by increasing the ability to adapt different ecological environments and by improving their quality and reducing harmful substances in tobacco [3](Spurr et al., 1975). Moreover, it has been confirmed that metabolites of tobacco endophytic fungi have considerable potential to be utilized in the field of medicine, also in the development of plant growth accelerators and biopesticides [31]. Previous studies on the endophytic fungi

of tobacco have mainly focused on the isolation of samples from different tissues; however, due to the lack of a comprehensive study on the endophytic fungi in tobacco, the fungal communities in this plant remains elusive. Therefore, to investigate the biodiversity, distribution, and community structures of the endophytic fungi in tobacco, we systematically analyzed these fungi at different growth stages and different altitudes.

In earlier studies of the fungi associated with tobacco, Welty et al. isolated fungi from tobacco leaves [32], whereas Norse and Spurr isolated endophytes from tobacco leaf tissues and found that they showed diverse morphologies. *Alternaria* spp., *Penicillium*, *Aspergillus*, and *Cladosporium* spp. have been found to be frequent endophytes in *Nicotiana* spp.[3,5], whereas English and Mitchell found that *Penicillium*, *Aspergillus*, *Trichoderma*, and *Circnella* are the dominant populations [33]. In 2012, Li et al. analyzed the community structure and diversity of fungi in tobacco at the mature stage of growth by amplifying the 18S and ITS genes (Li X et al., 2011). Although these previous studies undoubtedly laid the foundations for our study on the fungal communication in tobacco, these studies were neither systematic nor comprehensive, and it is therefore highly probable that information on certain uncultured fungi is lacking.

In the present study, we performed a comprehensive comparison on the fungal communities in tobacco among four growth stages (SS, RGS, FGS, and MS) at three altitudes (600, 1000, and 1300 m), which included uncultured fungi, in anticipation of determining fungal composition and variation at different growth stages and

different altitudes. We initially compared the fungal communities among the different growth stages at the same growth altitudes. Among the communities of fungi in our study, *Microidium phyllanthi* was dominant in all the growth stages. Network analysis also showed that *Microidium* occurs at most of the nodes. *Microidium phyllanthi* accounted for more than 90% of the total species in SS, FGS, and MS samples, whereas at the RGS stage, this species accounted for 69% of all species at the low altitude site. Similar fungal compositions were detected at the intermediate altitude. However, at the high-altitude site, *Microidium phyllanthi* was found to account for more than 95% of species in FGS and MS samples, which was higher than that in SS and RGS samples. Moreover, we found that *Pseudopithomyces chartarum* is another core fungus associated with all the examined growth stages. These finding imply that these two species may play critical roles in maintaining the structure and function of fungal communities in tobacco. However, the function of *Microidium phyllanthi* is still not known. Our results showed that the RGS stage is the growth stage associated with the highest abundance of fungal species. The high richness at the RGS stage may be due to the fact that fungal growth rate will be accelerated after tobacco is transplanted into fields, at which time plant vegetative growth predominates, and there is a high requirement for water and nutrients. Therefore, a diverse group of fungal species can readily form at the RGS stage compared with other stages. A further reason may be that after the seedlings have been transplanted and the fungicidal treatment discontinued, endophytic fungi can become established in the leaves [3]. In contrast, the lowest abundance of endophytes was detected at the MS stage. In

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

addition, we found higher numbers of unique fungi at the RGS and FGS stages compared with the other two stages. These two stages coincide with the period of rapid plant growth. Endophytic fungi may thus be associated with the promotion of growth in tobacco, which would be consistent with the findings of a previous study [31]. The general trend we observed in the variation of endophytes in tobacco is that the young tissues or fast growth stages tend to have higher diversity than the mature tissues, which is consistent with the findings of Espinosa Garcia et al., who demonstrated that endophyte diversity tended to increase from the 1- to 3-year-old leaves of *Sequoia sempervirens* and then decrease in fully mature needles that were 8 years old or older [34]. This is further supported by the findings of Houlden et al. (2008), who detected a decrease in the diversity of symbiotic fungal species with plant maturity, which has also been found in mycorrhizal associations as well as rhizosphere communities [35,36]. Our results also showed clear differences in the relative abundances of certain species among the four growth stages examined, suggesting that some endophyte species may preferentially proliferate at a certain stage and play ecological roles that are distinct from those of other endophytes.

At the SS and RGS stages, the diversity of endophytic fungi initially tended to decrease and then increased with an increase in altitude. In contrast, at the FGS and MS stages, the diversity of endophytic fungi in tobacco showed a decreasing trend with increasing altitude, which was paralleled by a decreasing trend in alpha diversity. Overall, we identified a trend whereby the diversity of fungal communities tended to decrease with an increase in the altitude at which tobacco is grown. This observation

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

is consistent with the findings of previous studies that have shown that altitude can influence the ecosystem by regulating the living environment through climatic factors and soil formation [37,38]. Similarly, previous studies have demonstrated that the physical and chemical properties of soil change with an increase of altitude, and that this is associated with a decrease in microbial community abundance and diversity[39,40]). Our results showed that there were significant differences in the functional diversity of fungal communities at different elevations.

5. Conclusion

In this study, we analyzed and quantified the endophytic fungi of healthy tobacco leaves from seedling, resettling growth, fast-growing, and maturing stages at three altitudes (600, 1000, and 1300 m) using a high-throughput sequencing technique. Our results indicated that tobacco is host to an abundant community of endophytic fungi, which provides new opportunities for exploring for their potential utilization. In future studies, more tissues and environmental factors should be integrated to better characterize the distributions and functions of the endophytic fungal community in tobacco.

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

Table 1 Median alpha diversity indices

Sample name	Observed species	Shannon	Simpson	Chao1	ACE	Goods coverage	PD_whole tree
SSL	161.33±11.59	1.24±0.23	0.31±0.06	188.74±6.69	202.87±18.15	0.99±0	48.24±2.25
SSM	108.67±15.95	0.98±0.36	0.26±0.10	120.22±22.99	126.28±27.77	0.99±0	39.10±6.89
SSH	161.67±24.08	1.15±0.54	0.29±0.14	183.57±23.35	193.26±18.65	0.99±0	49.96±9.17
RGSL	204±54.69.00	2.21±1.14	0.52±0.28	232.33±65.59	246.19±76.29	0.99±0	56.10±11.99
RGSM	196.33±36.56	1.75±0.88	0.40±0.17	231.17±54.97	247.21±51.28	0.99±0	52.26±4.58
RGSH	238.33±61.61	1.28±0.60	0.29±0.12	261.76±59.95	268.72±60.71	0.99±0	65.16±8.26
FGSL	164.67±44.77	1.20±0.42	0.30±0.12	188.69±60.70	196.37±57.17	0.99±0	54.48±7.73
FGSM	88.67±12.66	0.94±0.27	0.25±0.06	99.51±17.63	105.90±23.83	1.00±0	31.27±5.26
FGSH	102.33±26.08	0.90±0.27	0.25±0.07	122.94±29.65	127.68±34.26	0.99±0	35.47±6.91
MSL	162.67±14.57	1.29±0.30	0.33±0.09	196.60±28.75	198.75±18.52	0.99±0	50.46±5.90
MSM	134.13±18.52	1.08±0.34	0.29±0.09	159.54±29.82	169.47±30.37	0.99±0	51.84±3.10
MSH	127.67±19.30	1.05±0.40	0.28±0.09	136.94±21.49	141.81±20.66	1.00±0	48.14±6.40

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

Author contributions

Peng Zhang and Zhongfeng Zhang designed experiments; Xiaolong Yuan and Jinhai Li. carried out experiments. Xinmin Liu, Yongmei Du and Guoming Shen analyzed experimental results and Xiaolong Yuan wrote the manuscript.

Funding

This work was supported by the Natural Science Foundation of China (grant number 31700295) and from the Science Foundation for Young Scholars of Institute of Tobacco Research of Chinese Academy of Agricultural Sciences (grant number 2016A02).

Peer-reviewed version available at Int. J. Mol. Sci. 2018, 19, 3421; doi:10.3390/ijms19113421

Reference

1. Hyde, K.D.; Soytong, K. The fungal endophyte dilemma. *Fungal Diversity* **2008**, *33*, 163-173.
2. Saikkonen, K.; Faeth, S.H.; Helander, M.; Sullivan, T.J. Fungal endophytes: A continuum of interactions with host plants. *Annual Review of Ecology & Systematics* **1998**, *29*, 319-343.
3. Spurr, H.W. Characterization of endophytic fungi in healthy leaves of nicotiana spp. *Phytopathology* **1975**, *65*, 417.
4. Hussein. Diversity of the rhizosphere soil culture-dependent fungi of mature tobacco. *_titleAmerican Journal of Microbiology* **2011**, *2*, 9-14.
5. Norse, D. Fungi isoalted from surface-sterilized toabcco leaves. *Transactions of the British Mycological Society* **1972**, *58*, 515-518.
6. Guo, L.D.; Hyde, K.D.; Liew, E.C.Y. Identification of endophytic fungi from livistona chinensis based on morphology and rdna sequences. *New Phytologist* **2010**, *147*, 617-630.
7. Makimura, K.; Tamura, Y.; Mochizuki, T.; Hasegawa, A.; Tajiri, Y.; Hanazawa, R.; Uchida, K.; Saito, H.; Yamaguchi, H. Phylogenetic classification and species identification of dermatophyte strains based on DNA sequences of nuclear ribosomal internal transcribed spacer 1 regions. *Journal of Clinical Microbiology* **1999**, *37*, 920-924.
8. Schoch, C.L.; Seifert, K.A.; Huhndorf, S.; Robert, V.; Spouge, J.L.; Levesque, C.A.; Chen, W.; Consortium, F.B. Nuclear ribosomal internal transcribed spacer (its) region as a universal DNA barcode marker for fungi. *Proceedings of the National Academy of Sciences of the United States of America* **2012**, *109*, 6241-6246.
9. Nilsson, R.H.; Hyde, K.D.; Pawłowska, J.; Ryberg, M.; Tedersoo, L.; Aas, A.B.; Alias, S.A.; Alves, A.; Anderson, C.L.; Antonelli, A. Improving its sequence data for identification of plant pathogenic fungi. *Fungal Diversity* **2014**, *67*, 11-19.
10. Gargas, A.; Taylor, J.W. Polymerase chain reaction (pcr) primers for amplifying and sequencing nuclear 18s rdna from lichenized fungi. *Mycologia* **1992**, *84*, 589-592.
11. Bachy, C.; Dolan, J.R.; Deschamps, P.; Moreira, D. Accuracy of protist diversity assessments: Morphology compared with cloning and direct pyrosequencing of 18s rrna genes and its regions using the conspicuous tintinnid ciliates as a case study. *Isme Journal* **2013**, *7*, 244-255.
12. Pryce, T.M.; Palladino, S.; Kay, I.D.; Coombs, G.W. Rapid identification of fungi by sequencing the its1 and its2 regions using an automated capillary electrophoresis system. *Medical Mycology* **2003**, *41*, 369.
13. Brookman, J.L.; Mennim, G.; Trinci, A.P.J.; Theodorou, M.K.; Tuckwell, D.S. Identification and characterization of anaerobic gut fungi using molecular methodologies based on ribosomal its1 and 18s rrna. *Microbiology* **2000**, *146 (Pt 2)*, 393-403.
14. Blaalid, R.; Kumar, S.; Nilsson, R.H.; Abarenkov, K.; Kirk, P.M.; Kauserud, H. Its1 versus its2 as DNA metabarcodes for fungi. *Molecular Ecology Resources* **2013**, *13*, 218.
15. Porebski, S.; Bailey, L.G.; Baum, B.R. Modification of a ctab DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant Molecular Biology Reporter* **1997**, *15*, 8-15.
16. Magoč, T.; Salzberg, S.L. Flash: Fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **2011**, *27*, 2957-2963.

17. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Peã±A, A.G.; Goodrich, J.K.; Gordon, J.I. Qiime allows analysis of high-throughput community sequencing data.
18. Haas, B.J.; Gevers, D.; Earl, A.M.; Feldgarden, M.; Ward, D.V.; Giannoukos, G.; Ciulla, D.; Tabbaa, D.; Highlander, S.K.; Sodergren, E. Chimeric 16s rrna sequence formation and detection in sanger and 454-pyrosequenced pcr amplicons. *Genome Research* **2011**, *21*, 494.
19. Edgar, R.C.; Haas, B.J.; Clemente, J.C.; Quince, C.; Knight, R. Uchime improves sensitivity and speed of chimera detection. *Bioinformatics* **2011**, *27*, 2194.
20. Edgar, R.C. Uparse: Highly accurate otu sequences from microbial amplicon reads. *Nature Methods* **2013**, *10*, 996-998.
21. Altschul, S.; Gish, W.; Miller, W.; Myers, E.; Lipman, D. Basic local alignment search tool. J. Mol. Biol. **1990**.
22. Kõljalg, U.; Nilsson, R.H.; Abarenkov, K.; Tedersoo, L.; Taylor, A.F.; Bahram, M.; Bates, S.T.; Bruns, T.D.; Bengtssonpalme, J.; Callaghan, T.M. Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology* **2013**, *22*, 5271-5277.
23. Edgar, R.C. Muscle: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **2004**, *32*, 1792-1797.
24. Ronquist, F.; Huelsenbeck, J.P. Mrbayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **2003**, *19*, 1572-1574.
25. Lozupone, C.A.; Hamady, M.; Kelley, S.T.; Knight, R. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Applied & Environmental Microbiology* **2007**, *73*, 1576-1585.
26. Qin, J.; Li, Y.; Cai, Z.; Li, S.; Zhu, J.; Zhang, F.; Liang, S.; Zhang, W.; Guan, Y.; Shen, D. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **2012**, *490*, 55-60.
27. Rivas, M.N.; Burton, O.T.; Wise, P.; Zhang, Y.Q.; Hobson, S.A.; Lloret, M.G.; Chehoud, C.; Kuczynski, J.; Desantis, T.; Warrington, J. A microbiota signature associated with experimental food allergy promotes allergic sensitization and anaphylaxis. *Journal of Allergy & Clinical Immunology* **2013**, *131*, 201-212.
28. Pandey, R.; Mishra, A.K.; Tiwari, S.; Singh, H.N.; Kalra, A. Enhanced tolerance of mentha arvensis against meloidogyne incognita (kofoid and white) chitwood through mutualistic endophytes and pgprs. *Journal of Plant Interactions* **2011**, *6*, 247-253.
29. Li, H.Y.; Li, D.W.; He, C.M.; Zhou, Z.P.; Mei, T.; Xu, H.M. Diversity and heavy metal tolerance of endophytic fungi from six dominant plant species in a pb–zn mine wasteland in china. *Fungal Ecology* **2012**, *5*, 309-315.
30. Liu, W.; Zhou, Z.; Liu, Y.; Hu, X.; Guo, Y.; Li, J. Application of high-throughput internal transcribed spacer rrna metagenomics analysis in deciphering endophytic fungi diversity of dendrobium officinale. *Journal of Biobased Materials & Bioenergy* **2017**, *11*, 106-118.
31. Sherameti, I.; Shahollari, B.; Venus, Y.; Altschmied, L.; Varma, A.; Oelmüller, R. The endophytic fungus piriformospora indica stimulates the expression of nitrate reductase and the starch-degrading enzyme glucan-water dikinase in tobacco and arabidopsis roots through a homeodomain transcription factor that binds to a conserved motif in. *Journal of Biological Chemistry* **2005**, *280*, 26241-26247.

32. Welty, R.E.; Lucas, G.B.; Fletcher, J.T.; Yang, H. Fungi isolated from tobacco leaves and brown-spot lesions before and after flue-curing. *Applied Microbiology* **1968**, *16*, 1309.
33. English, J.T.; Mitchell, D.J. Development of microbial communities associated with tobacco root systems. *Soil Biology & Biochemistry* **1988**, *20*, 137-144.
34. Espinosa-Garcia, F.J.; Langenheim, J.H. The endophytic fungal community in leaves of a coastal redwood population diversity and spatial patterns. *New Phytologist* **1990**, *116*, 89-97.
35. Husband, R.; Herre, E.A.; Turner, S.L.; Gallery, R.; Young, J.P. Molecular diversity of arbuscular mycorrhizal fungi and patterns of host association over time and space in a tropical forest. *Molecular Ecology* **2002**, *11*, 2669-2678.
36. Houlden, A.; Timms-Wilson, T.M.; Day, M.J.; Bailey, M.J. Influence of plant developmental stage on microbial community structure and activity in the rhizosphere of three field crops. *Fems Microbiology Ecology* **2008**, *65*, 193-201.
37. Hanawalt; R., B.; Whittaker; R., H. Altitudinally coordinated patterns of soils and vegetation in-the san jacinto mountains, california. *Soil Science* **1976**, *121*.
38. Bangroo, S.A.; Najar, G.R.; Rasool, A. Effect of altitude and aspect on soil organic carbon and nitrogen stocks in the himalayan mawer forest range. *Catena* **2017**, *158*, 63-68.
39. Zhang, S.; Chen, D.; Sun, D.; Wang, X.; Smith, J.L.; Du, G. Impacts of altitude and position on the rates of soil nitrogen mineralization and nitrification in alpine meadows on the eastern qinghai–tibetan plateau, china. *Biology & Fertility of Soils* **2012**, *48*, 393-400.
40. Zhang, L.; Wang, A.; Yang, W.; Xu, Z.; Wu, F.; Tan, B.; Liu, Y.; Chen, L. Soil microbial abundance and community structure vary with altitude and season in the coniferous forests, china. *Journal of Soils & Sediments* **2016**, *17*, 1-11.