Field fungal diversity in freshly harvested japonica rice

ABSTRACT

Rice is a major food crop in China and Japonica rice production in Heilongjiang Province ranks No.1 in total annual rice production in the country. Rice is prone to invasion by fungi and mycotoxins produced by the fungi are proven to be serious threats to human health. The objective of the present study was to investigate fungal diversity of freshly harvested rice in the four main cultivation regions of Heilongjiang Province in order to find the abundance (delete) difference of dominant fungi among the four regions. Through high throughput sequencing we detected *Ascomycota* accounts for absolute dominant phylum; *Dothideomycetes*, *Sordariomycetes*, *Tremellomycetes*, *Microbotryomycetes*, and *Eurotiomycetes* were dominant classes; *Capnodiales, Hypocreales, and Pleosporales* were the main orders; *Cladosporiaceae, Pleosporaceae, Nectriaceae, Clavicipitaceae, Tremellaceae, Phaeosphaeriaceae, Trimorphomycetaceae, Sporidiobolaceae, Bionectriaceae,* and *Trichocomaceae* were major family; *Cladosporium*, *Epicoccum*, *Fusarium*, and *Alternaria* were the most abundant phylotypes at genus level; *Epicoccum_nigrum*, *Gibberella_zeae*, and *Fusarium_proliferatum* were the dominant fungal species. Great fungal diversity was observed in the rice samples harvested in the four major Japonica rice-growing regions in Heilongjiang province. However, no significant difference in diversity was observed among the four regions, likely due to the relatively close geographical proximity leading to very similar climatic conditions. Since some of the fungal species produce mycotoxins, it is necessary to take precautions to ensure the rice is stored under safe conditions to prevent the production of mycotoxins. This is the first report on investigation of field fungal diversity in freshly harvested Japonica rice in Heilongjiang Province in China.

Keywords: Field fungi; diversity; japonica rice; high through-put sequencing

1. INTRODUCTION

Rice (*Oryza sativa L.*) is a major food crop in China and more than 65% of the populace consumes rice as staple food. China ranks No.1 in total annual rice production in the world and accounts for around 1/3 of the global paddy rice production [1]. Heilongjiang province ranks 1st in Japonica rice cultivation in China with total production of 30 million tons in 2018 [2]. In addition, rice cultivated in Heilongjiang province is very famous for its high quality and excellent flavor due to optimal environmental conditions suitable for rice growing. The rice produced in this province is well received throughout the country and is even exported to many regions around the world. However, rice is prone to invasion by fungi and contamination by their mycotoxins. Fungi play a key role in rice safety and understanding the fungi community structure is of great importance when taking

appropriate measures to ensure rice safety. Fungi infect rice crops early in the field and may produce mycotoxins during this period. Consumers are concerned about this issue and consequently it is necessary to investigate the status of fungi contamination of rice in Heilongjiang province. Rice is widely cultivated in China under different climatic conditions and is extensively contaminated by various fungi. However, little information is currently available on the fungal diversity of field fungi, especially aflatoxigenic fungal contamination of rice in the main cultivating regions of Heilongjiang province. The objective of this study was to investigate fungal diversity of freshly harvested rice in the four main cultivation regions of Heilongjiang province through high-throughput sequencing and FUNGuild in order to find the abundance difference of dominant fungi among the four regions.

2. MATERIALS AND METHODS

2.1 Materials

Twelve rice samples were harvested from four regions in Heilongjiang's major rice cultivation areas indicated in Fig. 1: Wuchang city (three samples, three repetitions, marked as WC-1, WC-2, WC-3, WC-4, WC-5, WC-6, WC-7, WC-8, and WC-9), Jiamusi city (three samples, three repetitions, marked as JMS-1, JMS-2, JMS-3, JMS-4, JMS-5, JMS-6, JMS-7, JMS-8, and JMS-9), Zhaoyuan county (three samples, three repetitions, marked as ZY-1, ZY-2, ZY-3, ZY-4, ZY-5, ZY-6, ZY-7, ZY-8, and ZY-9), and Tailai county (three samples, three repetitions, marked as TL-1, TL-2, TL-3, TL-4, TL-5, TL-6, TL-7, TL-8, and TL-9). During September 26-29 of 2017, around 2 kg of each sample was cut using a reaping hook from rice fields, put into plastic bags, and sealed tightly. After arriving at the lab, 30 spikes of rice were manual threshed from each sample and three 50 g paddy rice samples were weighed from each sample into 1000 mL Erlenmeyer flasks with a 500 mL PBS buffer (KH_2PO_4 0.27 g, NA_2HPO_4 1.42 g, NaCl 8 g, KCl 0.2 g, diluted with 800 mL distilled water, adjusted pH value of 7.4, constant volume of 1 L□ and sterilized). They were labeled as three replicates of one sample. These samples were shaken with a vibrator for 30 minutes, subjected to sucking filtration, and filtered through 0.45 um water membranes. The residues were collected from the membranes using medicinal ladles and transferred into 1 mL microcentrifuge tubes and preserved by cryopreservation using liquid nitrogen. All operations were conducted in a sterile room and masks and gloves were worn to guarantee the samples would not be contaminated by environmental fungi. The samples were then transported to Guangzhou Gene Denovo Bio-Tech Ltd. Co. (Guangzhou, China) under dry ice conditions to perform high throughput sequencing of the PCR products. The obtained data was assembled into sequence tags and subject to BLAST in GenBank for microbe classification, followed by OTU, and diversity and inter-sample comparative analyses.

Fig. 1 Distribution of samples collecting locations in Heilongjiang province of China. a. Illustration of the geographical location of Heilongjiang province in China. b. Distribution of samples collecting locations in Heilongjiang province

2.2 DNA extraction and PCR amplification

Microbial DNA was extracted from stool samples using the E.Z.N.A. stool DNA Kit (Omega Bio-Tek, Norcross, GA, U.S.) according to manufacturer's protocols. The ITS region of the Eukaryotic ribosomal RNA gene was amplified by PCR: 95 °C for 2 min, followed by 27 cycles at 98 °C for 10 s, 62 °C for 30 s, and 68 °C for 30 s and a final extension at 68 °C for 10 min using primers ITS3_KYO2F 5'-GATGAAGAACGYAGYRAA -3' and ITS4R 5'- TCCTCCGCTTATTGATATGC-3', where the barcode is an eight-base sequence unique to each sample. PCR reactions were performed in triplicate 50 μL mixture containing 5 μL of 10 × KOD Buffer, 5 μL of 2.5 mM dNTPs, 1.5 μL of each primer (5 μM), 1 μL of KOD Polymerase, and 100 ng of template DNA.

2.3 Illumina Hiseq2500 sequencing

Amplicons were extracted from 2% agarose gels and purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, U.S.) according to the manufacturer's instructions and quantified using QuantiFluor™-ST (Promega, U.S.). Purified amplicons were pooled in equimolar and paired-end sequenced (2 × 250) on an Illumina platform according to the standard protocols.

2.4 Bioinformatics analysis

Bioinformatics analysis was conducted according to Fig.2.

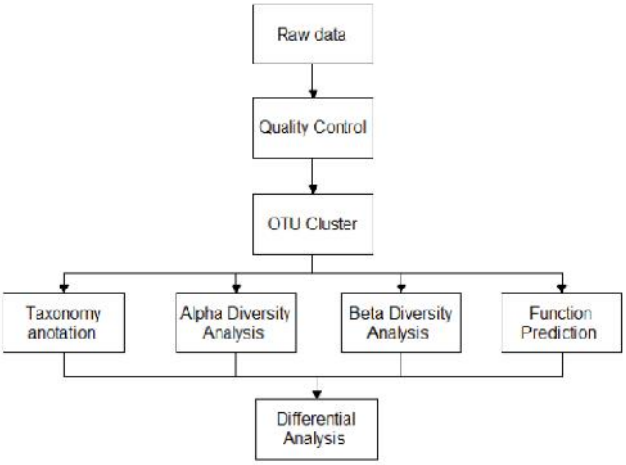


Fig.2 ITS analysis flow chart

2.5 Quality control and reads assembly

2.5.1 Reads filtering

Raw data containing adapters or low-quality reads would affect the following assembly and analysis. Hence, to get high-quality clean reads, raw reads were further filtered according to the following rules:

1) Removing reads containing more than 10% of unknown nucleotides (N);

2) Removing reads containing less than 80% of bases with quality (Q-value)≤ 20.

2.5.2 Reads assembly

Paired-end clean reads were merged as raw tags using FLSAH [3] (v 1.2.11) with a minimum overlap of 10 bp and mismatch error rates of 2%.

2.5.3 Raw tag filtering

Noisy sequences of raw tags were filtered by QIIME [4] (V1.9.1) pipeline under specific filtering conditions [5] to obtain high-quality clean tags.

2.5.4 Chimera checking and removal

Clean tags were searched against the reference database (http://drive5.com/uchime/uchime_download.html) to perform Reference-based chimera checking using UCHIME algorithm (http://www.drive5.com/usearch/manual/uchime_algo.html). All chimeric tags were removed and finally obtained effective tags for further analysis.

2.5.5 OTU cluster

The effective tags were clustered into operational taxonomic units (OTUs) of ≥ 97 % similarity using UPARSE [6] pipeline. The tag sequence with the highest abundance was selected as a reprehensive sequence within each cluster. Between groups, Venn analysis was performed in R to identify unique and common OTUs.

2.5.6 Taxonomy classification

The representative sequences were classified into organisms by a naive Bayesian model using RDP classifier [7] (Version 2.2) based on UNITE [8] Database

(https://unite.ut.ee/). The abundance statistics of each taxonomy and a phylogenetic tree was constructed in a Perl script and visualized using SVG. Biomarker features in each group were screened by Metastats and LEfSe software.

2.5.7 Alpha diversity analysis

Chao1, Simpson and all other alpha diversity indices were calculated in QIIME. OTU rarefaction curve and Rank abundance curves were plotted in QIIME. Statistics of between-group Alpha index comparison was calculated by a Welch's t-test and a Wilcoxon rank test in R. Alpha index comparisons among groups was computed by a Tukey's HSD test and a Kruskal-Wallis H test in R.

2.5.8 Beta diversity analysis

The weighted and unweighted unifrac distance matrix was generated by QIIME. Multivariate statistical techniques including PCA, principal coordinates analysis (PCoA) and NMDS of (Un)weighted unifrac distances were calculated and plotted in R. Statistics of Welch's t-test, Wilcoxon rank test Tukey's HSD test, Kruskal-Wallis H test, Adonis (also called Permanova) and Anosim test was calculated using R.

2.5.9 Functional prediction

The functional group (guild) of the OTUs was inferred using FUNGuild [9] (v1.0).

3. RESULTS AND DISCUSSION

3.1 Fungal diversity and richness in single rice samples and comparison of these indexes among the four regions

Total fungal ITS tags (106951, 108190, 111294, 105520, 108264, 113010, 115999, 104584, 108716, 123025, 108835 and 119401) were recovered from 12 (Rice 1, Rice 2, Rice 3, Rice 4, Rice 5, Rice 6, Rice 7, Rice 8, Rice 9, Rice 10, Rice 11, and Rice 12) samples, respectively. The library samples were then clustered into fungal Operational Taxonomic Units (OTUs) at 97% similarity (Table 1).

Table 1 Community richness, diversity and coverage indexes for the four regions*

Region	OTU	Chao 1	Ace	Good Coverage (%)	Shannon	Simpson
Wuchang	510±53^{a}	710±49^{a}	709±55^{a}	99.8±0.0^{a}	3.49±0.53^{a}	0.76±0.07^{a}
Jiamusi	379±45^{b}	531±72^{b}	541±77^{b}	99.8±0.0^{a}	3.24±0.34^{a}	0.75±0.07^{a}
Zhaoyuan	396±53^{a}	592±102^{b}	585±90^{b}	99.8±0.0^{a}	3.33±0.36^{a}	0.76±0.04^{a}
Tailai	383±64^{a}	636±87^{b}	537±87^{b}	99.8±0.0^{a}	3.43±0.30^{a}	0.78±0.06^{a}

Data represents mean±SD. Data followed by the same superscript letter in the same column are not significantly different.

The Chao and ACE are abundance indexes; the Simpson and Shannon are diversity indexes. Higher values of Chao (richness estimate) and ACE indicate more community richness. The Shannon (diversity index) and Simpson values indicate the community diversity, and higher Shannon and Simpson values indicate greater community diversity. The good coverage value indicates the depth of sequencing. The good sequencing coverage in all the four regions almost reached 99.8%, which indicated that almost all fungi have been detected. The number of OTUs determined

in the four regions showed that Wuchang got the maximum value, whereas Jiamusi obtained the minimum value. An OTU is usually recognized as a genus of microorganism. Consequently, a total of 507, 564, 459, 374, 398, 366, 440, 409, 338, 443, 316, and 390 fungal genera were identified in the 12 rice samples, respectively.

To compare fungal diversity and richness among the four regions, data was statistically analyzed and presented in Table 1. As seen in the table, no significant difference was found in the five indexes of the four regions. This is probably due to a close physical proximity among the four regions resulting in a lack of significant differences in environmental conditions. Since Jiamusi city is around 2° latitude north of the other three regions and greater than 1 °C of daily average minimum and maximum temperature lower than the other three regions, the observed numbers of fungal genera were the lowest as a result of cooler temperatures.

3.2 Fungal community composition

For the 12 rice samples, fungal community compositions were detected at seven levels: Domain, Phylum, Class, Order, Family, Genus, and Species. Fig. 3 demonstrates the taxonomy stack distribution of the Phylum, Order, Genus, and Species of the identified fungi. Due to limited space, only the results of fungi community compositions at the Phylum, Order, Genus, and Species levels were presented in the Figure. Of the classifiable sequences, two Phyla were identified as seen in Fig. 3a: *Ascomycota* and *Basidiomycota*, in which *Ascomycota* accounts for absolute dominance. At Class level, *Dothideomycetes*, *Sordariomycetes*, *Tremellomycetes*, *Microbotryomycetes*, *Eurotiomycetes*, and *Saccharomycetes* were identified, where *Dothideomycetes* and *Sordariomycetes* account for absolute dominance. At Order level, *Capnodiales, Pleosporales, Hypocreales, Tremellales, Trichosphaeriales, Sporidiobolales*, and *Eurotiales* were determined, in which *Capnodiales, Hypocreales*, and *Pleosporales* were dominant (Fig. 3b). At Family level, *Cladosporiaceae, Pleosporaceae, Nectriaceae, Clavicipitaceae, Tremellaceae, Phaeosphaeriaceae, Trimorphomycetaceae, Sporidiobolaceae, Bionectriaceae,* and *Trichocomaceae* were detected, and *Cladosporiaceae, Pleosporaceae* as well as *Nectriaceae* are the dominant families (data not shown in Fig. 3). As seen in Fig. 3c, *Cladosporium* is the absolute dominant genus; followed by *Epicoccum*, *Alternaria, Gibberella,* and *Fusarium*; they are less abundant phylotypes at the genus level detected in these rice samples. From Fig. 3d, it can be observed that unclassified species account for large portions of the whole bars. This is due to limitations of the UNITE Database which prevents classification of large amounts of species. *Epicoccum_nigrum*, *Gibberella_zeae*, and *Fusarium_proliferatum* are the dominant fungal species found in the determined samples. High proportions of *Fusarium_proliferatum* were detected in two samples (ZY-4 and ZY-9). *Gibberella zeae*, usually known by the name of its anamorph *Fusarium graminearum*, is identified as a plant pathogen which causes *Fusarium* head blight and can produce toxins, particularly deoxynivalenol (DON). *Fusarium spp.* produces a diverse number of secondary metabolites, including some fatal mycotoxins [10], and they are attributed as the most important toxigenic fungi in the Northern temperate areas [11].

Among the *Fusarium* spp. isolated from rice, *F. proliferatum* and *F. verticillioides* were proven to be the most abundant Fumonisin producers [12]. *Fusarium proliferatum* can occur in a wide range of plants, including rice and produce mycotoxins such as fumonisin [12,13].

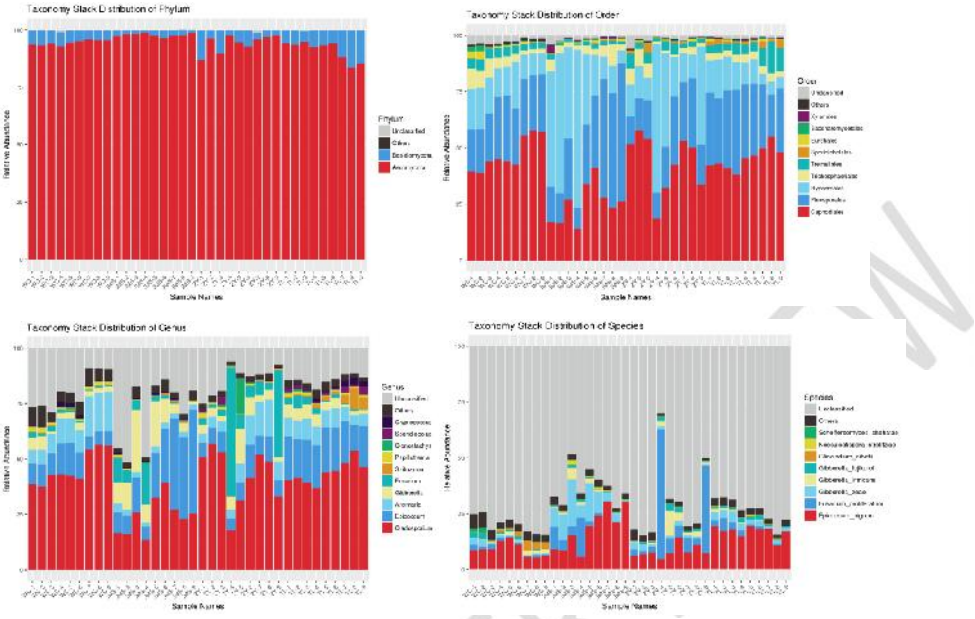


Fig.3 Taxonomy stack distribution of genus and species of the 12 rice samples in Heilongjiang province. a. At Phylum level; b. At Order level; c. At Genus level; d. At Species level.

Through a naive Bayesian model using RDP classifiers based on UNITE Database analysis of the assembled sequences, it was found that in the rice samples *Epicoccum nigrum* and *Fusarium proliferatum* were dominant, where *Epicoccum nigrum* is a plant pathogen and endophyte. *Fusarium proliferatum* is a fungal plant pathogen and usually infects asparagus. Huang et al [14,15]. isolated pathogens of rice spikelet rot disease from infected rice samples collected from Zhejiang province in Southern China and identified *Fusarium proliferatum* as one of the pathogens. Liu [16] confirmed that *Fusarium proliferatum* was the main pathogen which induced rice spikelet rot disease. Hou [17] demonstrated that *Fusarium proliferatum* was one of the five determined *Fusarium* and accounted for 63.4% of the total detected strains. Furthermore, they also determined that *Fusarium proliferatum* produced mycotoxins. Du et al [18]. detected *Penicillium*, *Aspergillus,* and *Fusarium* as the major fungal genus in Huaidao No. 5 rice freshly harvested in 2013 and indicated that *Penicillium* and *Aspergillus* are the dominant fungi genus. A great difference exists between their result and ours, likely because Huaidao No. 5 was planted in Jiangsu Province which is located on the east coast of China and has a climate type of subtropical monsoon climate to temperate monsoon climate, while Heilongjiang Province is located in northeastern China with a temperate continental monsoon climate. Consequently, the rice fungal communities in these two provinces are rather dissimilar.

Table 2 Fungal diversity and abundance (%) of rice samples at genus and species levels collected from the four regions*

Levels	Fungal Strains	Wuchang	Jiamusi	Zhaoyuan	Tailai
Genus	*Cladosporium*	45.83±7.58^{a}	24.40±8.36^{b}	42.80±12.83^{a}	43.91±8.18^{a}
	Epicoccum	9.65±3.32^{a}	23.16±18.15^{bc}	8.72±2.97^{a}	18.26±2.69^{ac}
	Alternaria	11.96±4.61^{a}	2.62±0.66^{b}	8.93±4.89^{ad}	7.48±1.74^{cd}
	Gibberella	3.25±1.49^{a}	10.97±8.52^{bc}	5.49±4.22^{ac}	4.11±1.93^{a}
	Fusarium	1.16±0.73^{a}	4.83±6.06^{a}	13.90±20.35^{a}	2.05±1.72^{a}
	Saitozyma	0.27±0.09^{a}	0.39±0.26^{a}	0.18±0.07^{a}	2.83±3.01^{b}
	Papiliotrema	1.31±0.77^{a}	0.32±0.25^{bc}	0.79±0.64^{ac}	0.93±0.33^{ac}
	Clonostachys	0.26±0.17^{a}	0.23±0.21^{a}	2.01±5.15^{a}	0.30±0.40^{a}
	Sporidiobolus	0.17±0.03^{a}	0.12±0.08^{a}	0.75±1.19^{ac}	1.61±1.17^{bc}
	Cryptococcus	0.74±0.82^{a}	0.47±0.24^{a}	0.18±0.11^{ab}	1.23±1.08^{ac}
	Unclassified	30.91±18.50^{a}	28.46±7.62^{a}	13.94±6.60^{a}	14.26±1.95^{a}
Species	*Epicoccum_nigrum*	9.22±3.09^{a}	18.26±9.18^{b}	8.22±2.92^{a}	17.05±2.66^{b}
	Fusarium_proliferatum	1.10±0.74^{a}	4.47±5.74^{a}	13.83±20.36^{a}	2.01±1.72^{a}
	Gibberella_zeae	1.25±1.16^{a}	9.04±6.68^{b}	1.68±0.92^{a}	2.60±1.31^{a}
	Gibberella_intricans	1.02±0.75^{a}	1.16±1.58^{a}	2.54±3.38^{a}	0.55±0.51^{a}
	Gibberella_fujikuroi	0.94±0.83^{a}	0.62±0.45^{a}	1.22±1.07^{a}	0.94±0.62^{a}
	Gliocladium_cibotii	1.68±1.50^{a}	0.10±0.09^{b}	0.12±0.09^{b}	0.11±0.03^{b}
	Unclassified	79.89±3.94^{a}	63.15±7.83^{bc}	69.74±18.42^{ac}	73.52±5.39^{ac}

* Values followed by the same superscript letter in the same row are not significantly different.

As seen in Table 2, for the four regions the dominant fungi at Genus level are *Cladosporium, Epicoccum, Alternaria, Gibberella*, and *Fusarium*, and almost no significant difference in their abundance was observed among the five Genera in the four regions. *Cladosporium* has been the most frequently found species in the four regions. However, the abundance of *Cladosporium* in Jiamusi was significantly lower than those of in the other three regions. *Cladosporium* is recognized as a psychrophile hence it is more adaptable to cool temperature condition. The cause of its low abundance in Jiamusi in comparison to the other three regions is still unclear. *Cladosporium* has been proven to be a potentially pathogenic mycotoxin-producing fungus frequently occurring in outdoor environments [19]. In addition, the proportion of *Epicoccum* in Jiamusi was greater than those in the other three regions. *Epicoccum* is a plant pathogen and widespread fungus which produces coloured pigments. Therefore, rice in Jiamusi region has a higher probability of contamination by coloured pigments which will in turn reduce rice quality.

The dominant fungi at the species level are *Epicoccum_nigrum, Fusarium_proliferatum, and Gibberella_zeae*. Like above, almost no significant difference in their abundance was found among the three species in the four regions. This is probably due to the relatively close geographical proximity of the four regions resulting in similar climatic conditions.

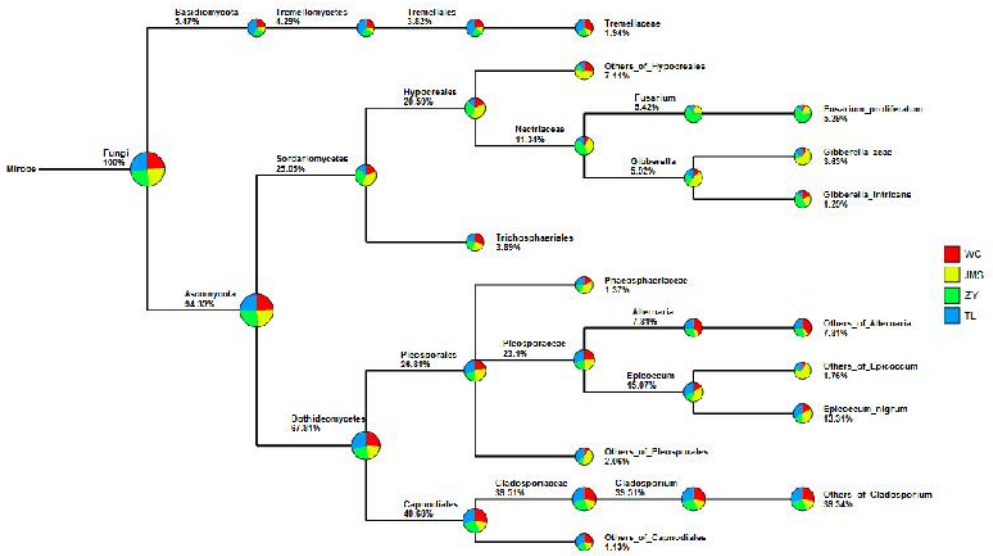


Fig. 4 Multi rice samples taxonomy analysis tree on the species level

As seen in Fig. 4, fungal community of the four regions (composed of 12 rice samples) was mainly composed of two Phylum, *Ascomycota* and *Basidiomycota,* which account for 94.33% and 5.47%, respectively. At the species level, others of *Cladosporium* accounts for 39.34% of the total species, followed by *Epicoccum nigrum* which accounts for *13.31%*, others of *Alternaria, Fusarium proliferatum*, *Gibberella zeae*, and others of *Epicoccum* account for 7.81%, 5.29%, 3.63%, and 1.76%, respectively. The least is *Gibberella intricans*, which accounts for 1.29%. For the mycotoxin-producing species, the proportion of *Fusarium proliferatum* accounts for absolute dominant in Zaoyuan region in comparison with the other three regions, and this probably is a hint that rice planted in Zhaoyuan has a greater potential to be contaminated by mycotoxins especially Fumonisin. The proportion of *Gibberella zeae* and others of *Epicoccum* in Jiamusi region account for absolute dominant compared with the other three regions. Since *Gibberella zeae* is the sexual stage of *Fusarium graminearum*, it has the possibility of producing deoxynivalenol (DON) and nivalenol (NIV) [20]. In addition, the proportion of *Gibberella intricans* in Zhaoyuan is the biggest in comparison with the other three regions. Although non-toxigenic fungi and yeasts themselves may only cause spoilage without safety issues, the damage they caused still not to be ignored.

3.3 Cluster analysis of the 12 rice samples

As seen in Fig. 5, the fungi of the 3 rice samples (three replicates for each sample) from Wuchang city were clustered into one group. This is probably a result of near geographical proximity among the three spots where the rice samples were collected resulting in a similar fungal community. However, not all the fungi from the same region can be clustered into one group. Many regions have rice fields with varying soil types, water resources, types of fertilization, rice varieties, and other

environmental factors which might increase the possibility of fungal diversity and make it difficult to cluster the fungi of rice samples from the same region into one group. Nevertheless, most of the fungi of rice samples from the same region can be clustered into the same group.

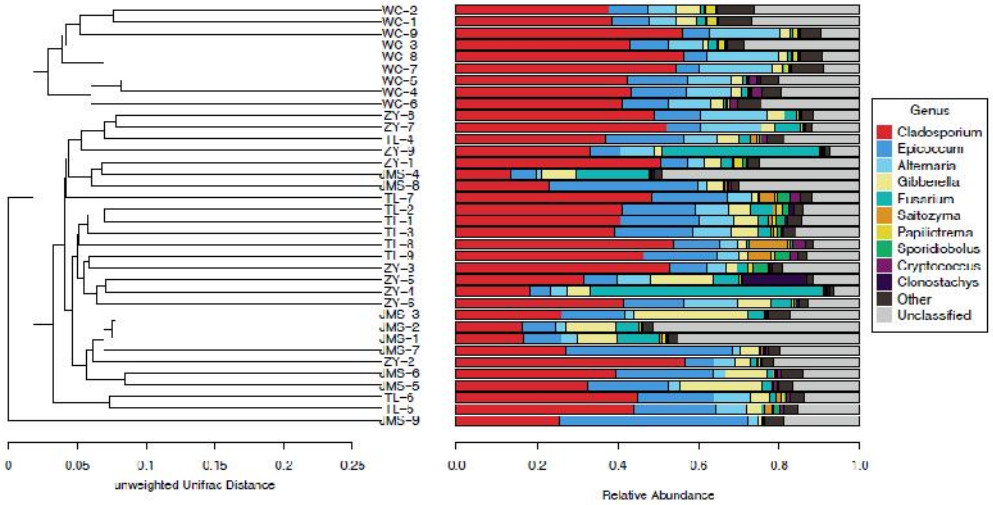


Fig. 5 Unweighted Pair-Group Method with Arithmetic Means (UPGMA) analysis of microbial community structure based on ITS gene amplicon sequencing data.

3.4 Fungal communities and functional guilds analysis

Fungal communities and functional guilds of the rice samples detected in the four regions are shown in Fig. 6. As seen in this figure, an open environment enables the rice to be a plant host to a wide range of environmental fungi. The most abundant phylotypes are seen to be plant pathogen, endophyte, fungal parasite, undefined saprotroph, wood saprotroph, soil saprotroph, as well as animal pathogen. For all the rice samples, plant pathogen, lichen parasite, soil saprotroph, wood saprotroph, and endophyte account for the largest proportions. For the mycotoxigenic fungi species, they are in the category of plant pathogen.

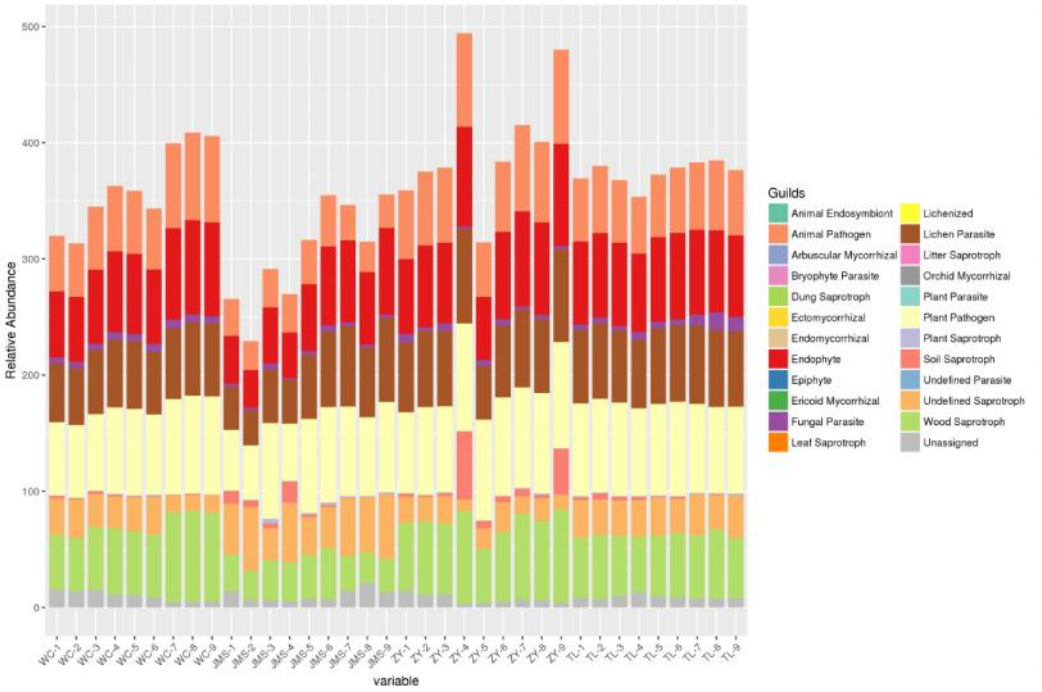
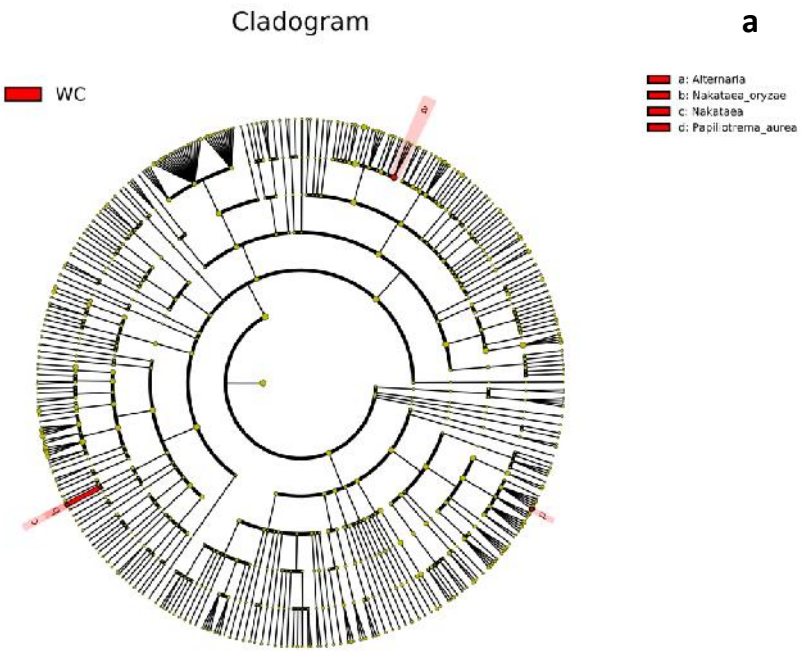


Fig. 6 Stacks of Guilds of the 12 rice samples from Heilongjiang province

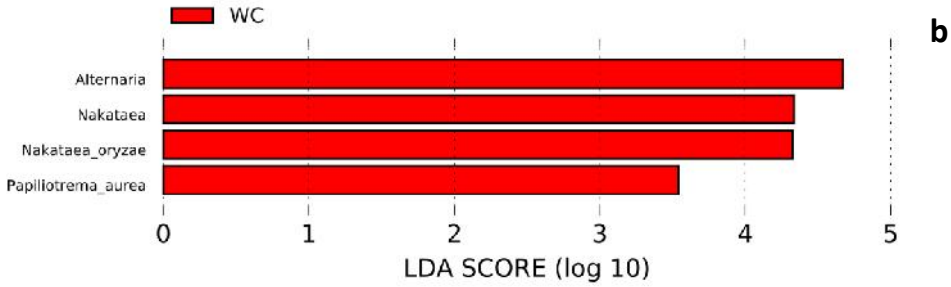
Around 70% of all major crop diseases were induced by fungal plant pathogens. Furthermore, 15% of global agricultural production was destroyed through yield losses and mycotoxin contamination [21]. Plant pathogens, especially mycotoxigenic fungi are considered to be the most harmful class of plant pathogens by far. As a cosmopolitan genus of filamentous ascomycete fungi, *Fusarium* includes a number of toxin-producing plant pathogens of agricultural importance [22]. For the rice freshly harvested in Heilongjiang province, the *Fusarium_proliferatum* determined likely includes mycotoxigenic species, although a fungi toxicity test has not been conducted yet.

3.5 LEfSe analysis

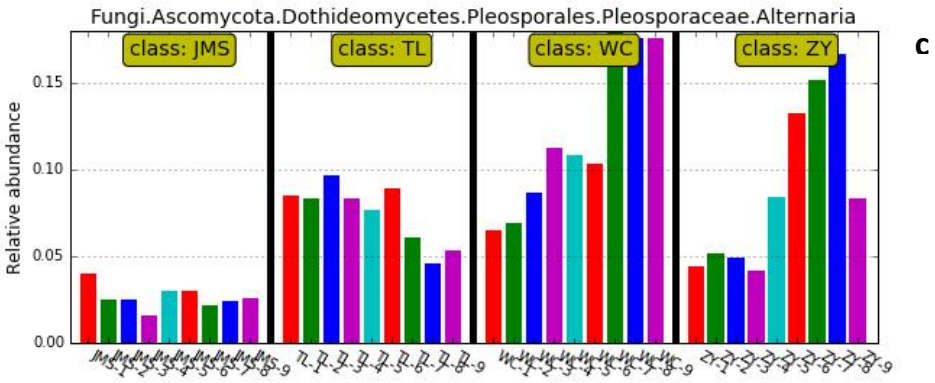
Key phylotypes of rice fungi microbiota representing the four regions identified using linear discriminant analysis (LDA) effect size (LEfSe) are shown in Fig. 7. As seen in Fig 7a, the Cladogram indicates that the numbers of four fungi genera and species in Wuchang city are significantly greater than those of in the other three regions; they are *Alternaria, Nakataea, Nakataea_oryzae,* and *Papiliotrema_aurea*. Their LDA scores are greater than 3 (Fig. 7b) and they might be considered as specific fungi associated with Wuchang region. Fig. 7c, d, e, and f illustrate the relative abundance of the four fungi given above in the four regions. Consequently, it might be possible to develop biomarkers using the four fungi given above to distinguish rice from Wuchang region.


Cladogram
a
WC
a: Alternaria
b: Nakataea_oryzae
c: Nakataea
d: Papiliotrema_aurea

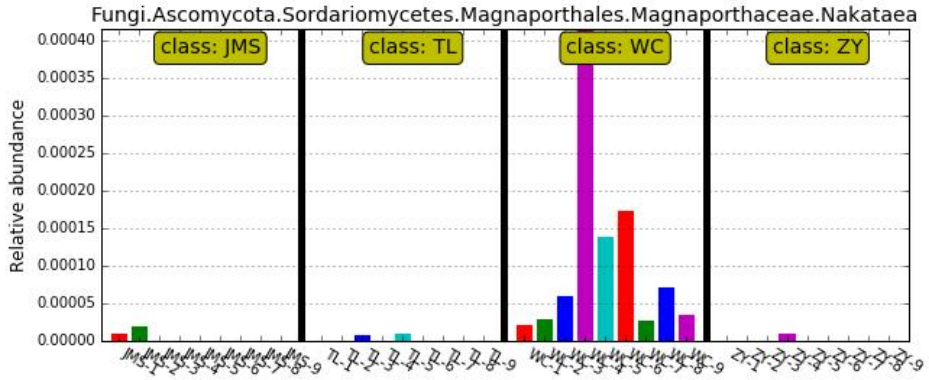
350


WC
b
Alternaria
Nakataea
Nakataea_oryzae
Papiliotrema_aurea
0
1
2
3
4
5
LDA SCORE (log 10)

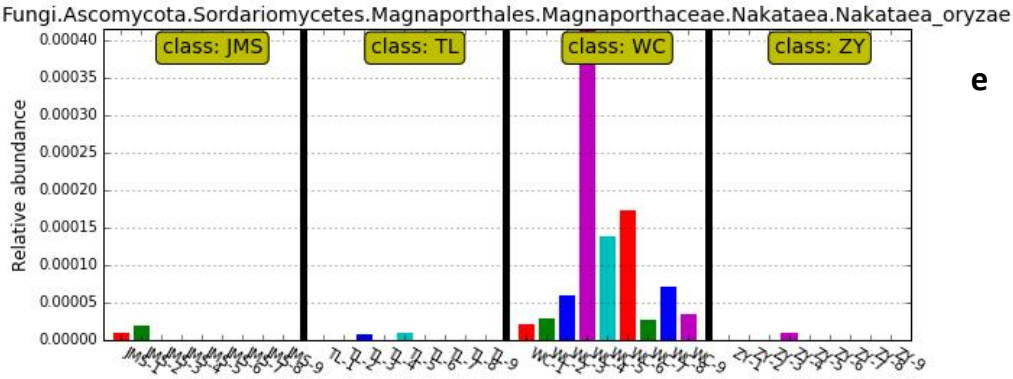
351


Fungi.Ascomycota.Dothideomycetes.Pleosporales.Pleosporaceae.Alternaria
c
class: JMS
class: TL
class: WC
class: ZY
Relative abundance
0.15
0.10
0.05
0.00

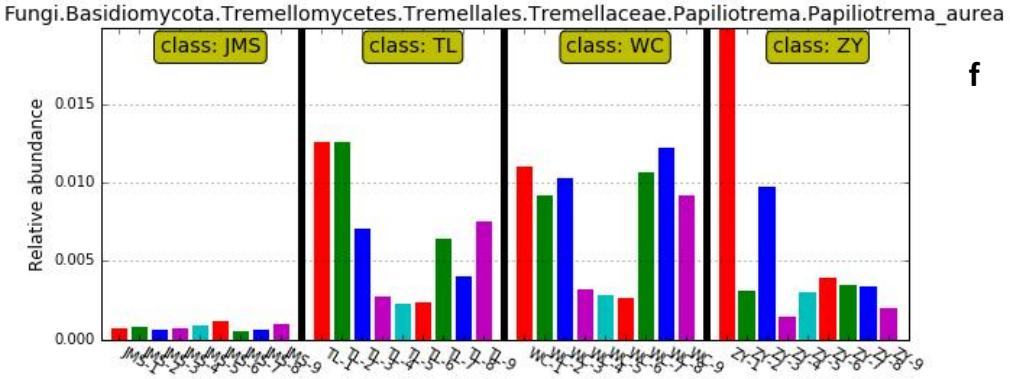
352



353



354



355

356 Fig. 7 Cladogram, LDA score, and relative abundance of fungi of rice samples from the four
357 regions. a. Cladogram; b. LDA score; c. relative abundance of *Alternaria* of rice samples from
358 the four regions; d. relative abundance of *Alternaria* of rice samples from the four regions; e.
359 a relative abundance of *Nakataea_oryzae* of rice samples from the four regions; f. a relative
360 abundance of *Papiliotrema_aurea* of rice samples from the four regions

361

362 **4. CONCLUSION**

363 To explore the potential of fungi contamination as well as mycotoxin production,
364 it is necessary to investigate field fungal diversity in rice in Heilongjiang province
365 through high throughput sequencing of freshly harvested rice samples. Our results
366 indicate that *Cladosporium* accounts for an absolute dominant at the genus level and

Epicoccum_nigrum, Fusarium_proliferatum, and *Gibberella_zeae* are relatively abundant fungi species, in which *Fusarium_proliferatum* has the potential to produce mycotoxins such as fumonisin. Rice planted in Zhaoyuan has the greatest potential to produce fumonisin whereas rice grown in Jiamusi is most likely to be contaminated by DON and NIV in comparison with the other three regions. Consequently, it is necessary to take adequate measures to prevent mycotoxin production during rice storage, as well as related damage induced by non-mycotoxins-producing fungus growth and reproduction. In addition, *Alternaria, Nakataea, Nakataea_oryzae,* and *Papiliotrema_aurea* are the specific fungi genera and species which can distinguish rice planted in Wuchang from the other three regions.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1 Food and Agriculture Organization of the United Nations (FAO). http://www.fao.org/faostat/en/#data. Accessed Dec.26, 2018.

2 Heilongjiang Daily Newspaper (HLJD). http://epaper.hljnews.cn/hljrb/20181010/384488.html. Accessed Dec. 26, 2018.

3 Magoč T and Salzberg SL, FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**:2957-2963 (2011).

4 Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK et al., QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* **7**:335-336 (2010).

5 Bokulich NA, Subramanian S, Faith JJ, Gevers D, Gordon JI et al., Quality-filtering vastly improves diversity estimates from Illumina amplicon sequencing. *Nat Methods* **10**: 57-59 (2013).

6 Edgar RC, UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat methods* **10**:996-998 (2013).

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

8 Kõljalg U, Larsson KH, Abarenkov K, Nilsson RH, Alexander IJ, Eberhardt U, Erland S, Høiland K, Kjøller R, Larsson E, Pennanen T, Sen R, Taylor AFS, Tedersoo L and Vrålstad T, UNITE: a database providing web ‐ based methods for the molecular identification of ectomycorrhizal fungi. *New Phytol* **166**:1063-1068 (2005).

9 Nguyen NH, Song ZW, Scott T, Branco BS, Tedersoo L, Menke J, Schilling JS and Kennedy PG, FUNGuild: an open annotation tool for parsing fungal community datasets by ecological guild. *Fungal Ecol* **20**:241-248 (2016).

10 Desjardins AE and Proctor RH, Molecular biology of Fusarium mycotoxins. *Int J*

411 *Food Microbiol* **119**:47-50 (2007).
412 11 Gutleb AC, Morrison E and Murk AJ, Cytotoxicity assays for mycotoxins
413 produced by *Fusarium* strains: a review. *Environ Toxicol and Phar* **11**: 309-320
414 (2002).
415 12 Matić S, Spadaro D, Prelle A, Gullino ML and Garibaldi A, Light affects
416 fumonisin production in strains of Fusarium fujikuroi, Fusarium proliferatum,
417 and Fusarium verticillioides isolated from rice, *Int J Food Microbiol* **166**:515-
418 523 (2013).
419 13 Desjardins AE, Manhanadhar HK, Plattner RD, Manandhar GG, Poling SM and
420 Maragos CM, Fusarium species from Nepalese rice and production of
421 mycotoxins and gibberellic acid by selected species. *Appl Environ Microbiol*
422 **66**:1020−1025 (2000).
423 14 Huang SW, Wang L, Liu LM, Tang SQ, Zhu DF and Savary S, Rice spikelet rot
424 disease in China: 1. Characterization of fungi associated with the disease. *Crop*
425 *Prot* **30**:1-9 (2011).
426 15 Huang SW, Wang L, Liu LM, Liu EY, Hou EQ, Xiao DF and Fan ZL, Isolation,
427 identification and biological characters of pathogens of rice spikelet rot disease.
428 *Chinese J Rice Sci* **26**:341-350 (2012).
429 16 Liu, E. Y. Main infection sources identification and control of rice spikelet rot
430 disease. Master Thesis, Nanning, Guangxi University (2011).
431 17 Hou EQ, Studies on biological characteristics and toxin of rice spikelet rot disease
432 pathogens (RSRD) *Fusarium spp*. Master Thesis, Nanning, Guangxi University
433 (2013).
434 18 Du LH, He XY, Liu LP, Yuan J and Ju XR, Fungal diversity of Huaidao No. 5 rice
435 and the dominant culturable fungal strains during storage. *Sci Agri Sinica*
436 **49**:1371-1381 (2016).
437 19 Alwatban MA, Hadi S and Moslem MA, Mycotoxin production in Cladosporium
438 species influenced by temperature regimes. *J Pure Appl Microbiol* **8**:4061-4069
439 (2014).
440 20 Lee SH, Lee J, Nam YJ, Lee S, Ryu JG and Lee T, Population structure of
441 *Fusarium graminearum* from maize and rice in 2009 in Korea. *Plant Pathol J*
442 **26**: 321-327 (2010).
443 21 Prado S, Nay B and Kunz C, *Paraconiothyrium variabile*, an ascomycete
444 endophyte, suppresses mycotoxin production in the plant pathogen *Fusarium*
445 *oxysporum*. *J Mycol Méd* **25**:e96-e97 (2015).
446 22 Ma LJ, Geiser DM, Proctor RH, Rooney AP, O'Donnell K, Trail F, Gardiner DM,
447 Manners JM and Kazan K, *Fusarium* pathogenomics. *Annu Rev*
448 *Microbiol* **67**:399-416 (2013).
449