New microbial resource: microbial diversity, function and dynamics in Chinese liquor starter

Yuhong Huang1,2,3, Zhuolin Yi2,3, Yanling Jin2,3, Yonggui Zhao2,3, Kaize He2,3, Dayu Liu1, Dong Zhao4, Hui He5, Huibo Luo6, Wenzhu Zhang7,8, Yang Fang2,3 & Hai Zhao1,2,3

Traditional Chinese liquor (Baijiu) solid state fermentation technology has lasted for several thousand years. The microbial communities that enrich in liquor starter are important for fermentation. However, the microbial communities are still under-characterized. In this study, 454 pyrosequencing technology was applied to comprehensively analyze the microbial diversity, function and dynamics of two most-consumed liquor starters (Jiang- and Nong-flavor) during production. In total, 315 and 83 bacterial genera and 72 and 47 fungal genera were identified in Jiang- and Nong-flavor liquor starter, respectively. The relatively high diversity was observed when the temperature increased to 70 and 62 °C for Jiang- and Nong-flavor liquor starter, respectively. Some thermophilic fungi have already been isolated. Microbial communities that might contribute to ethanol fermentation, saccharification and flavor development were identified and shown to be core communities in correlation-based network analysis. The predictively functional profile of bacterial communities showed significant difference in energy, carbohydrate and amino acid metabolism and the degradation of aromatic compounds between the two kinds of liquor starters. Here we report these liquor starters as a new functionally microbial resource, which can be used for discovering thermophilic and aerobic enzymes and for food and feed preservation.

Chinese liquor (Baijiu), approximately 40-60% alcohol by volume, is one of the world’s four most popular distilled spirits along with whisky, brandy and vodka. Chinese liquor production accounts for more than one-third of all spirits consumed in the world according to the International Wine and Spirit Research Group. The unique and traditional Chinese solid simultaneous saccharification and fermentation (SSF) technology have existed for more than 4000 years1,2. Even more intriguing, microbial communities associated with environmentally-enriched for Chinese liquor fermentation was significantly different from those of most western spirit productions3. Chinese liquor is classified primarily by its distinctive flavor and taste4. Nong-flavor (NF, also called thick flavor and strong flavor) liquor accounts for more than 70% of the Chinese liquor market4. Meanwhile, the consumption of Jiang-flavor (JF, also called sauce flavor) liquor has increased in recent years according to the China National Food Industry Association. These two most-consumed and popular liquors are highly valued in Chinese culinary culture4.

The liquor SSF process is mainly attributed to the interaction and metabolism of microbes in liquor starter (daqiu), zaopei and pit mud5. Among them, liquor starter is important for saccharification, fermentation and flavor-generation and is the most essential component for the production of Chinese liquor6. The production process for both NF and JF liquor starters usually takes approximately four months (Fig. 1a). Briefly, ground wheat is mixed with water and shaped into blocks, which are spontaneously fermented in a fermentation room for approximately one month and then dried in a storage room for another three months to mature7. During the one...
month period of fermentation, the room temperatures increase to 50 and 55 °C after three days, and quickly reach the highest values of 62 and 70 °C after around six and seven days for NF and JF liquor starters, respectively. This high temperature continues for approximately 7–8 days. After the high-temperature stage, the room temperature will finally decrease to approximately 25 °C. Both the NF and JF starters use wheat as feedstock, and no microorganisms are intentionally inoculated, as all the microbes are enriched from the feedstock, water, air and working environment. The main difference between the production processes of these two liquor starters is the changes of room temperature, which is mainly generated by microbial metabolic heat and is controlled by the placement of the starter blocks, turning over the blocks and ventilation. As NF liquor starter is characterized by moderately high temperature (62 °C), the blocks are usually tiled, spread and covered with a layer of straw in the fermentation room. The blocks must be turned every day to maintain the moderately high temperature for 8 days. However, JF is a high-temperature liquor starter (70 °C), so the blocks are heaped up and covered in thick straw. The blocks are only turned two or three times during the high-temperature period. Therefore, the dynamics of the microbial composition during the liquor starter production process are mainly determined by the working environment and the changes of room temperature during the liquor starter production process.

The liquor SSF process is a highly skilled and mature technology, and the microbial community in liquor starter is highly stable with a long history. Nonetheless, the microbial diversity, function and dynamics of liquor starter, which are related to the dynamic changes of room temperature during production, have remained under-characterized. Microbial communities have been analyzed by traditional and molecular methods, e.g., culture-dependent methods and culture-independent PCR-DGGE. However, many issues exist with these methods, such as missing most of the unculturable strains, time-consuming procedures, high workload, changing microbial composition, high cost, and, most importantly, a limited understanding of the microbial communities under study. Advances in high-throughput sequencing technology, such as 454 pyrosequencing, can comprehensively analyze the composition, diversity and richness of the microbial community for many samples in a single run. Based on pyrosequencing analysis, the Fen liquor starter (also called qing flavor (QF) and light

![Figure 1.](https://www.nature.com/scientificreports/)
Communities19–21, and this method has not yet been applied in liquor starters. Recently, the software 'Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PICRUSt)'22 was developed to predict the functional profiles of bacterial communities by bridging 16S rRNA gene information with a reference genomic database and the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. Currently, this method can only be applied to predict bacterial function, and the number of available fungal genomes is not sufficient to allow for such prediction. To date, PICRUSt has been successfully applied to reveal the functional profiles of bacterial communities in animals23,24, humans20,25, oceans26, crude oil sediments 27 and soil28,29.

In this study, the bacterial and fungal communities of two widely consumed liquor starters, NF and JF, were investigated using 454 pyrosequencing technology. Some thermophilic microbes were also isolated. Attempting to understand the comparative diversity, function and dynamics of microbial communities in liquor starter can help to improve Chinese liquor fermentation technology from “traditional experience-based model” to a “modern science-directed model”.

### Results

#### Microbial diversity and richness in the liquor starters.

The library samples were then clustered into bacterial and fungal Operational Taxonomic Units (OTUs) at 97% similarity (Table 1). After rarefied reads, the results showed that the most bacterial OTUs were found in the J3 and N2 samples (597 and 572 OTUs, respectively) and the highest number of fungal OTUs in the J4 and N3 samples (587 and 258 OTUs, respectively) (Table 1). The estimators of Shannon and Simpson also showed the similar results as the Chao and Ace indices discussed above.

| Sample ID | OTU  | Ace  | Chao  | Coverage (%) | Shannon | Simpson |
|-----------|------|------|-------|--------------|---------|---------|
| J1        | 415  | 777  | 704   | 93.9         | 6.05    | 0.04    |
| J2        | 319  | 696  | 746   | 94.7         | 5.04    | 0.1004  |
| J3        | 597  | 1119 | 1087  | 91.0         | 6.59    | 0.0548  |
| J4        | 223  | 381  | 381   | 97.0         | 4.36    | 0.1363  |
| N1        | 342  | 742  | 689   | 94.2         | 4.07    | 0.1766  |
| N2        | 572  | 1321 | 1258  | 90.0         | 6.41    | 0.0411  |
| N3        | 478  | 987  | 974   | 92.1         | 5.84    | 0.0668  |
| N4        | 274  | 649  | 621   | 95.2         | 3.43    | 0.3014  |
| J4        | 587  | 853  | 892   | 97.7         | 6.13    | 0.0525  |
| J3        | 173  | 278  | 268   | 99.2         | 1.66    | 0.6484  |
| J2        | 138  | 213  | 205   | 99.4         | 1.04    | 0.806   |
| J1        | 502  | 784  | 843   | 97.7         | 4.51    | 0.1555  |
| N4        | 258  | 388  | 372   | 98.9         | 2.96    | 0.2936  |
| N3        | 152  | 204  | 215   | 99.4         | 1.27    | 0.758   |
| N2        | 210  | 302  | 311   | 99.2         | 1.98    | 0.6328  |
| N1        | 282  | 405  | 426   | 99.0         | 3.87    | 0.2964  |
| J1        | 587  | 564  | 597   | 97.7         | 6.13    | 0.0525  |
| J2        | 173  | 278  | 268   | 99.2         | 1.66    | 0.6484  |
| J3        | 258  | 388  | 372   | 98.9         | 2.96    | 0.2936  |
| J4        | 152  | 204  | 215   | 99.4         | 1.27    | 0.758   |
| N1        | 502  | 784  | 843   | 97.7         | 4.51    | 0.1555  |
| N2        | 210  | 302  | 311   | 99.2         | 1.98    | 0.6328  |
| N3        | 282  | 405  | 426   | 99.0         | 3.87    | 0.2964  |
| N4        | 587  | 564  | 597   | 97.7         | 6.13    | 0.0525  |

Table 1. Community richness, diversity and coverage indices for Jiang-flavor (J) and Nong-flavor (N) liquor starter samples using 16S rRNA gene and ITS region rarefied to 3460 and 9622 sequences, respectively (OTU similarity was 97%). OTU: operational taxonomic unit. Higher values of Chao and Ace indicate more community richness. The Shannon and Simpson values indicate the community diversity. Higher Shannon values indicate greater community diversity, whereas a higher Simpson value indicates the opposite. The coverage value indicates the depth of sequencing. Bacteria (16S rRNA gene) and fungi (ITS) were analyzed in all the liquor starter samples. J1 and N1 were sampled at the beginning of liquor starter production; J2 and N2 were sampled after 3 days of liquor starter fermentation; J3 and N3 were sampled after 8 and 9 days liquor starter fermentation, respectively; J4 and N4 were sampled from the mature liquor starter. The temperatures of J1, J2, J3, J4, N1, N2, N3 and N4 were 30, 55, 70, 25, 30, 50, 62 and 25 °C, respectively.
Principal component analysis (PCA) was performed to comparatively analyze the relationships of the microbial communities in the JF and NF liquor starters. The bacterial community PCA scatterplot (Fig. 1b) revealed that the bacterial community in the JF liquor starter was distinct from that in the NF liquor starter. The bacterial community in the NF liquor starter was grouped to the right of the graph along PC1. However, the bacterial community in the JF liquor starter varied considerably. In contrast, the fungal community was similar in the JF and NF liquor starter samples during the early period (J1 and N1, J2 and N2) (Fig. 1b). However, the fungal community of liquor starter samples differed after the highest temperature period (J3, J4, N1, N2, N3 and N4).

Microbial community composition. Classification of the bacteria and fungi at the order level for all samples is shown in Fig. 2. Lactobacillales increased up to 92% and 61% during the first 3 days for both J2 and N2 samples, respectively. After the temperature reached its maximum, Lactobacillales decreased markedly. However, Bacillales increased quickly and their abundance reached 17% and 43% for samples J3 and N3, respectively. Meanwhile, a relatively high abundance of no-rank Actinobacteria (39%) was found in J3, while only few (0.2%) were found in N3. Surprisingly, some relatively low abundance bacterial communities (see supplementary Table 1) were only detected in JF liquor starter. When the room temperature decreased to approximately 25 °C, Bacillales increased, especially in the NF liquor starter, reaching up to 71% (N4). No-rank Actinobacteria in the JF liquor starter also increased markedly, reaching 78% (J4). For bacterial communities, it was further shown that there was no OTU in common among four JF liquor starter samples (Fig. S1A), and 36 OTUs were in common among four NF samples (Fig. S1B).

The total fungal diversity in both the JF and NF liquor starters was much lower than that in the bacterial communities (Fig. 2). After 3 days of incubation, Saccharomycetales were the dominant fungi, accounting for up to 96.3% and 99.7% in samples J2 and N2, respectively. During the highest temperature period, a large number of no-rank Eukaryota were detected in both liquor starters, comprising 55% and 41% of the relative abundance in samples J3 and N3, respectively. In sample J3, Eurotiales also increased up to 42% during this period. Interestingly, 48% of Saccharomycetales could tolerate 62 °C and exist in sample N3. After the highest temperature period, Eurotiales increased markedly and became the dominant fungal taxon with up to 85% and 96% of the observed fungi in samples J4 and N4, respectively. For fungi communities, there were 10 OTUs (Fig. S1C) in the JF liquor starter.
common across four JF liquor starters. And 3 OTUs (Fig. S1D) were in common across the NF starters. The NF liquor starter had *Pichia kudriavzevii* and uncultured *Saccharomyces* in common, while the JF liquor starter had *P. kudriavzevii*, *Thermoascus lanuginosus*, *Thermoaosus crustaceus*, *Byssochlamys spectabilis*, *Rhizomucor pusillus*, *Emericella rugulosa*, *Eurotium* spp. and uncultured compost fungi species in common.

**Pivotal yeast, filamentous fungi and bacteria communities in liquor starter.** In this study, five yeast genera, *Candida*, *Hyphopichia*, *Pichia*, *Saccharomyces* and *Saccharomyces* were found in two liquor starters (Table 2). *Saccharomyces fibuligerus* was detected in sample N1. However, *Saccharomyces* were not found in the JF liquor starter, and a low abundance of no-rank/unclassified *Saccharomyces* were present in the NF liquor starter. *Pichia* was the dominant genus and *P. kudriavzevii* was the most abundant species in samples J2 and N2. Recently, *P. kudriavzevii* and *Candida* spp. which can grow at temperatures up to 50 °C, have been isolated from NF liquor starter (N3). Surprisingly, *Hyphopichia*, which can tolerate higher temperature (62 °C), was relatively abundant (up to 40%) in the NF liquor starter (N3).

When comparing the important filamentous fungi in the JF and NF liquor starters (Fig. 3a), both had low relative abundance of the traditional genera *Aspergillus*, *Penicillium*, *Rhizomucor* and *Rhizopus*. Interestingly, both liquor starters had high relative abundances of *Thermoaosus* and *Thermoascus* at the highest temperature point (J3 and N3), and these increased even further when the temperature dropped to 25 °C (J4 and N4) (Fig. 3a). In the present work, some thermophilic fungi, *T. lanuginosus*, *T. crustaceus*, *R. pusillus*, and *R. microsporus* were successfully isolated from N3 and J3 samples.

For bacterial communities, similar abundances were found between *Acetobacter* and *Lactobacillus*, both of which increased quickly when the two liquor starters had incubated for 3 days (J2 and N2) and decreased when the temperature increased to its highest point (J3 and N3) (Fig. 3b). On the other hand, different abundances were shown between *Leucostoc* and *Weissella* communities in NF and JF liquor starters (Fig. 3b). Both liquor starters had a relatively abundant *Bacillus* community during the highest temperature point (Fig. 3b, J3 and N3). Meanwhile, highly abundant no-rank *Actinobacteria* were found in JF liquor starter (J3 and J4) (Fig. 2). In addition, relatively abundant *Thermoactinomyces* bacteria were detected in N3 and J4, with comparable values in J3 and few found in N4 samples (Fig. 3b).

**Correlation Network analysis.** A microbiome-wide network of microbial associations among 8 samples was performed to better understand the co-occurring relationship of the microbial community. In total, 339 and 239 pairs of nodes were significantly correlated in bacterial JF and NF samples, respectively (Table S2 and Table S3). Meanwhile, 69 and 68 pairs of nodes were strongly correlated in fungal JF and NF samples, respectively (Table S2 and Table S3).

In the bacterial network, hundreds of edges were found within each sample: 671 in N1, 2544 in N2, 1969 in N3, 289 in N4, 961 in J1, 544 in J2, 22983 in J3, and 438 in J4 (Fig. 4 and Table S3). Few edges were found among samples, however, except for N1-N2, N2-J1 and N3-J3, with 59, 71 and 90 edges, respectively (Fig. 4 and Table S2). Those associations among samples mainly resulted from co-occurring taxa. A large proportion of co-occurring taxa were found to be related to the genera *Weissella* (J1, N2), *Lactobacillus* (N1, N2), *Thermoactinomyces* (N3) and *Macroccocus* (J3) (Fig. 4 and Table S3).

Meanwhile, in the fungal network, the edge numbers within each sample were 153 (N1), 182(N2), 54 (N3), 28 (N4), 396 (J1), 90(J2), 26 (J3), and 88 (J4) (Fig. 4 and Table S4), which were all lower than those in the bacterial network. However, many associations were found among samples. Some associations even showed comparable edge numbers to those within samples, such as the 59 found in N1-N3, 24 in N2-N4, and 35 in N2-J2. Interestingly, the genus *Pichia* played a pivotal role in associations related to N1, N2, N3 and J2, accounting for 52% of nodes in N1, 92% of nodes in N2, 23% of nodes in N3 and 81% of nodes in J2, and contributing to 100% of the edges in N2-J2, N2-J3, N2-J4, N2-N4 and J2-J4 and 78% of edges in J2-N4 (Fig. 4, Table S4, Table S5). Additionally, the *Pichia* species were largely affiliated with *P. burtonii* (N1), *P. occidentalis* (N1) and *Candida inconspicua* (J2, N2 and N3). In addition to *Pichia*, the genera *Thermoaosus*, *Paecilomyces*, *Thermoascus* also played important roles in fungal correlation network, accounting for 75% of nodes in J3 and 50% of nodes in J4, 43% of nodes in J4, and 100% of nodes in N4. Moreover, together with the genus *Pichia*, *Trichosporon* (N1), *Candida* (N1), *Galactomyces* (N3) and *Thermoascus* (N3) made large contributions to association between N1 and N3.

**Table 2.** Relative abundance of the main microbial community with ethanol fermentation capability. J1 and N1 were sampled at the beginning of liquor starter production; J2 and N2 were sampled after 3 days of liquor starter fermentation; J3 and N3 were sampled after 8 and 9 days of liquor starter fermentation, respectively; J4 and N4 were from the mature liquor starter. The temperatures of J1, J2, J3, J4, N1, N2, N3 and N4 were 30, 55, 70, 25, 30, 50, 62 and 25 °C, respectively.

| Genus           | Relative abundance(%) |
|-----------------|------------------------|
|                 | J1 | J2 | J3 | J4 | N1 | N2 | N3 | N4 |
| *Hyphopichia*   | 0.09 | 0.11 | 0.01 | 0.00 | 83.03 | 0.11 | 40.40 | 0.00 |
| *Pichia*        | 18.70 | 96.13 | 0.20 | 5.35 | 1.38 | 98.92 | 4.82 | 0.24 |
| *Saccharomyces* | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| *Candida*       | 0.11 | 0.01 | 0.00 | 0.00 | 0.18 | 0.02 | 0.03 | 0.00 |
| *Saccharomyces* | 0.00 | 0.00 | 0.00 | 0.00 | 5.30 | 0.10 | 0.10 | 0.10 |
Predictive functional profiles of bacterial communities. Functional capacity of the bacterial communities during liquor starters making process was predicted by PICRUSt. Meanwhile the compositions of KEGG pathways were also analyzed for the 8 samples (Fig. S2). The abundant KEGG pathways were also analyzed between every pair of samples to investigate the differently functional profiles of bacterial communities among the 8 samples (Fig. 5). Notably, carbon metabolism was the most abundant pathways in 8 samples, especially for the microbes in the J3 sample. In addition to carbon metabolism, related carbohydrate and energy metabolism activities, such as oxidative phosphorylation, starch and sucrose metabolism, fructose and mannose metabolism, galactose metabolism, propanoate metabolism, sulfur metabolism, pyruvate metabolism, butanoate metabolism and glycolysis/gluconeogenesis, were highly represented in the top 50 pathways (Fig. S2). Interestingly, most of those carbohydrate and energy metabolism activities were in the top 20 significantly differential pathways in two-sample groups (Fig. 5), and some of them showed clear differences between two liquor starters. On the one hand, some pathways showed obviously higher differential significances compared to JF samples, e.g., starch and sucrose metabolism in J1-J2, J2-J3, J3-J4, N1-N2, N1-J1 and N2-J2, oxidative phosphorylation in J2-J3, J3-J4 and N3-J3, and carbon metabolism in J1-J2, J2-J3, J3-J4, N3-J3 and N4-J4 (Fig. 5). On the other hand, some pathways also showed higher differential significances compared to NF samples, e.g., N2-N3, N3-N4, N2-J2, N3-J3 and N4-J4 in galactose metabolism, and N1-N2, N1-J1 and N3-J3 in fructose and mannose metabolism (Fig. 5).

In addition, amino acid metabolism, terpenoid and polyketide metabolism were also found in the top 20 significantly differential pathways (Fig. 5). Among these, cysteine and methionine metabolism, lysine degradation and terpenoid backbone biosynthesis were in the top 30 pathways for 8 samples (Fig. S2). Importantly, cysteine and methionine metabolism (J1-J2, J2-J3, J3-J4, N1-N2, N1-J2, N2-J2 and N3-J3) and terpenoid backbone biosynthesis (J1-J2, J2-J3, J3-J4, N1-N2, N1-J2, N2-J2 and N3-J3) presented relatively high differential significances compared to JF samples. Furthermore, the number of genes involved in the degradation of aromatic compounds was much higher in J3 and J4 samples (Fig. S2) and showed highly differential significances in J2-J3, J3-N3 and J4-N4 (Fig. 5).

Discussion
Understanding the diversity, function and dynamics of liquor starter microbial communities is important for improving liquor quality and yield. Although the unique and traditional Chinese solid state SSF and liquor brewing technology has continued for over several thousand years, the structure of the NF and JF liquor starter microbial communities is still not well characterized. This is the first study to comprehensively analyze the microbial composition of Chinese NF and JF liquor starters during the production process. A total of 315 and 83 bacterial genera and 72 and 47 fungal genera were detected in the JF and NF liquor starters, respectively, indicating that both liquor starters have complex and abundant bacterial and fungal communities. This study obtained
a relatively larger number of microbes and higher coverage than studies of Fen liquor starter (QF) and Maotai liquor starter (JF) which used traditional and molecular methods; fewer than 30 species was detected in one starter. Moreover, our study found a surprising result that bacterial and fungal communities had the highest diversity and richness at the higher temperature stage in NF liquor starter. And in JF liquor starter, the highest diversity and richness of bacterial communities were also detected at the highest temperature stage of J3, while the highest diversity and richness of fungal communities were shown at the mature stage of J4.

The metabolic products of bacteria are important to develop unique flavors for liquor. In this study, Weissella, Lactobacillus and Thermoactinomyces were identified as the most abundant bacterial taxa and made large contributions in the bacterial correlation network, so they could constitute the core genera in bacterial communities for these two liquor starters. The genera Lactobacillus and Acetobacter, belongs to the group of acetic acid bacteria. Lactobacillus is mainly applied to ferment traditional food and beverages, and Acetobacter is the main functional bacteria for producing vinegar. The metabolic products (e.g., acetic acid, lactic acid) of Acetobacter and Lactobacillus can be further synthesized by esterase into esters, such as ethyl acetate, ethyl lactate, isoamyl lactate, hexyl acetate, 2-phenylethyl acetate, 3-methylbutyl acetate and benzenacetate, which make large contributions to the special flavor of NF and JF liquors. It has been reported that Weissella and Leuconostoc are

Figure 4. A correlation network of co-occurring microbial OTUs in the Jiang-flavor and Nong-flavor liquor starter samples. Each node represents a bacterial genus or fungal species, and color corresponds to different stages of liquor starter, with red for Nong-flavor (N1, N2, N3 and N4) and blue for Jiang-flavor (J1, J2, J3 and J4) samples. Square nodes stand for Jiang-flavor samples, triangle nodes stand for Nong-flavor samples, and circle nodes stand for co-occurring samples. Edges between nodes represent strong correlations (Spearman's $\rho > 0.7$ for bacteria and $\rho > 0.6$ for fungi), including both positive and negative correlations.
usually present in fermented food along with lactic acid bacteria and are active as a good dietary source of vitamins. Therefore, the *Lactobacillus*, *Acetobacter*, *Weissella* and *Leuconostoc* identified in both liquor starters may be important for flavor contribution. This study was the first discovery that JF liquor starter had a relatively high abundance of thermophilic Actinobacteria, which are an important and functionally diverse group of microbes with potential for biodegradation, carbon cycling and nutrient transformation. Meanwhile, the highly abundant no-rank Actinobacteria in JF liquor starter (J3 and J4) could produce volatile organic compounds; the secondary metabolic products of *Actinomyces* could also contribute to flavor development.

*Thermoactinomyces* can produce plenty of thermostable extracellular proteolytic enzymes, e.g., alkaline proteinase, glutamyl endopeptidase and carboxypeptidase. This study further found that Bacillales and no-rank Actinobacteria were the dominant communities when the temperature increased to the highest point. *Bacillus* is well known to play an important role in liquor flavor generation. It has been reported that *B. licheniformis* in JF liquor starter can produce specific flavor compounds such as 2,3-butanediol, 3-hydroxy-2-butanone, 2-methylpropionic acid and 3-methylbutanoic acid. Interestingly, among bacterial community with relatively low abundant in JF liquor starter, *Burkholderiales* has a great ability to degrade a vast array of aromatic compounds, and *Clostridiales* is an expert at reducing a wide range of cinnamic acid derivatives to their corresponding 3-phenylpropionic acid derivatives. Additionally, *Fusobacteriales* and *Clostridiales* work as versatile bacteria and have unique abilities to ferment amino acids. Furthermore, *Cytophagales* was reported to have a high capacity of degrading cellulose. Surprisingly, *Lactobacillus*, *Leuconostoc*, *Lactococcus* and *Thermoactinomyces* were also dominant in the mature status of another NF liquor starter.

Moreover, the predicted functional profiles of bacterial communities have performed for both liquor starters. The liquor starter samples in this study showed that the Nearest Sequenced Taxon Index (NSTI) values (0.007 to 0.062) (Table S6) was lower than the reported similar sample (soil) (0.17) in Langille’s work, suggesting sufficient accuracy for metagenome predictions by PICRUSt. The prediction result indicated that both liquor starter samples had the most active energy metabolism, especially, J3 had the highest number of genes involved in almost all of the pathways (Fig. S2). This result also showed that the microbial community in sample J3 possibly had the most metabolic activity, which might be consistent with its highest temperature (70 °C). Those genes involved in carbohydrate digestion metabolisms were also high in both liquor starters. The two liquor starters might have preferences in carbohydrate digestion, e.g., starch and sucrose for JF samples, galactose, fructose and mannose...
for NF samples. Amino acid, terpenoid and polyketide metabolisms possibly were also active with differential significances between the two liquor starters (e.g., cysteine and methionine metabolism and terpenoid backbone biosynthesis in JF samples). Therefore, intermediates from those pathways (e.g., butanoate, propanoate, terpenoids, polyketides, aliphatic amino acids and branched-chain amino acids) might act as precursors or stimulators of the aromatic compounds in Chinese NF and JF liquor. Additionally, the relatively high differential significance of cysteine and methionine metabolism in JF samples might be consistent with reports that cysteine and its metabolisms play an important role in the formation of JF compounds. Furthermore, the degradation of aromatic compounds showed high differential significances in JF samples, and more genes related to this pathway were found in J3 and J4 when compared to NF samples and other JF samples by PICRUSt prediction, which is not surprising that many aromatic compounds and their derivatives have been found in JF liquor. To some extent, the difference of those pathways between two liquor starters might be consistent with their different flavors. This observation might suggest a confirmation for difference in the flavor-development of those bacteria mentioned above, e.g., *Actinomyces* and *B. licheniformis* in JF liquor starter. In addition, the capacity of aromatic compounds degradation by *Balkholderiales* in JF liquor starter might be consistent with the PICRUSt predicted result regarding the pathways of degradation of aromatic compounds. In sum, these bacteria in these liquor starters could be significant for developing their special and well-characterized flavor of this liquor.

Yeasts constitute the most important microbial community in the ethanol production process. As described above, the genus *Pichia* was the most abundant yeast and made large contribution in fungal correlation network, thus it could be one of the core genera in fungal communities in these two liquor starter. In Fen liquor starter (QF), *P. kudriavzevii* has been found to be the most abundant species, which is similar to our study. The surprising discovery is that *Hyphopichia* could tolerate 62 °C. *Hyphopichia* has a broad habitat and unique ecological niches such as starch, rumen contents, food, insects and insect frass. However, little is known about the thermostability of *Hyphopichia*. In addition to their function in the ethanol production, yeasts may also play a pivotal role in liquor flavor generation. *P. anomala* has been reported to increase acetate-ester hydrolyzing esterase activity when mixed with *S. cerevisiae* in liquor fermentation and regulate several important volatile compounds in Chinese liquor, such as ethyl lactate, octanoic acid, and ethyl tetradecanoate. When co-fermented with *S. cerevisiae*, *P. kluveri* increased varietal thiol concentrations in Sauvignon Blanc wine. Therefore, the yeast communities that are found in the NF and JF liquor starters may contribute to both ethanol fermentation and favor generation.

Filamentous fungi are significant during the production process of liquor starters for their saccharification function by secreting a wide spectrum of carbohydrate degrading enzymes. *Aspergillus*, *Rhizomucor* and *Penicillium* have been well studied for their good lignocellulose hydrolysis performance. Traditionally, *Rhizopus* and *Aspergillus* have been widely involved in fermenting food and beverages, and considered as the standards in saccharification. Here, both *Rhizopus* and *Aspergillus* did not show high abundance throughout the production process. However, *Thermoascus* and *Thermomyces* were the most abundant fungi. Therefore, the important thermophilic fungi that were isolated from the NF and JF liquor starters were completely different from the fungi which were found by traditional notion and experience during Chinese liquor fermentation. In this study, *Thermoascus*, *Paecilomyces* and *Thermomyces* were identified as abundant fungal taxa and played a pivotal role in the fungal correlation network, so they could be the core genera in fungal communities. Furthermore, *T. anomala*, *T. crustaceus* and *T. lanuginosus* have all been reported to be promising sources of thermophilic enzymes for high-level carbohydrate degradation. Currently, the microbiome in the giant panda gut and wood-feeding higher termites and cow rumen have been closely studied, as they specialize in the efficient degradation of biomass. All these studies have underscored the importance of microbial communities in ecology. However, most of the dominant microorganisms are bacteria that live in anaerobic, high water content or animal temperature conditions. It has been reported that filamentous fungi are the most potent biomass degraders, as they produce a high number and a broad variety of enzymes. Many studies on cellulase production have focused on aerobic species because the anaerobes show limited growth and do not produce enzymes in high quantities. Furthermore, thermophilic enzymes have received significant attention for biomass degradation in industrial processes. In this study, some thermophilic fungi have been further isolated and identified, e.g., *T. lanuginosus*, *T. crustaceus*, *R. pusillus* and *R. microspores*. Besides these isolated thermophilic fungi, more functional fungi will be used in future work from the promising microbial resource of liquor starter. These isolated fungi will be used for further discovering novel thermophilic enzymes.

This study provides insight into the diversity, function and dynamics of liquor starter microbial communities and sets the stage for understanding liquor starter microbial composition, paving the way for further active microbial community studies by metatranscriptomics and metaproteomics technologies and functional microorganism isolation, which may finally optimize the liquor production process, increasing alcohol yield and developing flavor generation. This new microbial resource can also be used for other industrial applications, such as bioenergy production, thermophilic and aerobic enzymes discovery, and fermented products used in food or feed preservation.

**Methods**

**Liquor starter sampling.** NF liquor starter was sampled from a fermentation workshop of Yibin Hongloumeng Distillery Group Co., Ltd in Yibing, Sichuan, China in July 2013. JF liquor starter was sampled from a fermentation workshop of Kweichow Hanwang Group Co., Ltd. in Renhuai, Guizhou, China in July 2013. Both liquor starters were sampled at different time points. Samples were harvested from three locations in the fermentation room at each time point. Samples J1 and N1 were collected at the beginning of liquor starter production; J2 and N2 were collected after 3 days of liquor starter fermentation; J3 and N3 were collected after 8 and 9 days of liquor starter fermentation (the second day after the temperature reached to the highest point), respectively; J4 and N4 were collected from the mature liquor starter. The temperatures of J1, J2, J3, J4, N1, N2,
Table 3. Accession numbers of ITS sequence for the isolated thermophilic fungi.

| Source | Microorganism         | Sequence ID     | Accession number |
|--------|------------------------|-----------------|------------------|
| N3     | *Rhizomucor pusillus*  | BankIt1856062 seq 1 | KT737214         |
| N3     | *Rhizopus microsporus* 1 | BankIt1856062 seq 2 | KT737215         |
| N3     | *Rhizopus microsporus* 2 | BankIt1856062 seq 3 | KT737216         |
| N3     | *Thermomyces lanuginosus* | BankIt1856062 seq 4 | KT737217         |
| N3     | *Thermoausus crustaceus* | BankIt1856062 seq 5 | KT737218         |
| N3     | *Pichia kudriavzevii* | BankIt1856062 seq 6 | KT737219         |
| N3     | *Candida spp.* | BankIt1856062 seq 7 | KT737220         |
| J3     | *Thermoausus crustaceus* | BankIt1856062 seq 8 | KT737221         |
| J3     | *Thermomyces lanuginosus* | BankIt1856062 seq 9 | KT737222         |
| J3     | *Rhizomucor pusillus* 1 | BankIt1856062 seq 10 | KT737223        |
| J3     | *Rhizomucor pusillus* 2 | BankIt1856062 seq 11 | KT737224        |

The raw sequences were trimmed of the adaptor, barcode, forward and reverse primer using SEQCLN (14-03-20) (http://sourceforge.net/projects/seqclean)\(^9\). The resulting sequences were then quality-filtered by MOTHUR software package (v.1.33.0)\(^9\) to remove low quality sequences (shorter than 200 bp, with quality scores < 25, containing any ambiguous bases (N) or homopolymers longer than 6 nucleotides). The valid sequences were reduced to a subset of unique sequences in order to eliminate redundancy and reduce computation time. The reduced sequences were then clustered into OTUs at the 97% similarity level.

In detail, for bacterial OTU clustering, chimeric sequences in reduced sequences were identified and removed by UCHIME (4.2.40) (http://drive5.com/uchime)\(^8\). Uncorrected pairwise distances were calculated according to sequences’ base differences. Then, sequences having 97% similarity were clustered into OTUs using the furthest neighbor method in MOTHUR.

Meanwhile, for fungal OTU clustering, the reduced sequences were blasted against the reference Silva eukarya database (silva SSU111) (http://www.arb-silva.de)\(^9\) to check the reference-based chimera by using the script identify_chimeric_seq.py (QIIME 1.7.0, Usearch61) (http://qiime.org/scripts/identify_chimeric_seq.html)\(^9\), and all chimeric tags were removed. Then sequences having 97% similarity were clustered into OTUs using the script pick_otus.py (QIIME 1.7.0, Usearch61) (http://qiime.org/scripts/pick_otus.html)\(^9\).
The Ribosome Database Project (RDP) Naive Bayesian classifier (11.2) (http://rdp.cme.msu.edu/)\(^{91}\) in MOTHUR was used to analyze bacterial and fungal clustered OTUs against reference SSU rRNA (small subunit ribosomal RNA) database in Silva (silva SSU111)\(^ {96}\) for taxonomical classification. To fairly compare all samples at the same sequencing depth, the number of the sequences was normalized by randomly extracting 3460 reads for 16S rRNA gene data and 9622 reads for ITS sequence data from each sample using daisychopper.pl\(^ {92}\) before alpha diversity indices analysis. Alpha diversity indices of Chao1, ACE, Shannon, Simpson and sample coverage (Good’s coverage) were calculated using QIME 1.7.0\(^ {96}\). PCA was constructed using R package to quantify differences in community composition\(^ {93}\). A Venn diagram was created with the online tool Venny (2.1) (http://bioinfogp.cnb.csic.es/tools/venny/). The sequencing data was deposited in the NCBI Sequence Read Archive database under the study accession number PRJNA283746.

**Correlation Network Analysis.** All microbial OTUs in liquor starter samples were used for network analysis. Spearman’s rank correlations between selected OTUs were calculated using R package\(^ {95}\). A valid co-occurrence was selected as a strong correlation if the Spearman’s correlation coefficient (\(p\)) was >0.6 and >0.7 for fungi and bacteria, respectively\(^ {96}\). Correlation networks were constructed with the open source platform Cytoscape (3.3.0)\(^ {95}\).

**Predictive functional profiles of bacterial communities involved in metabolic pathways.** PICRUSt (1.0.0)\(^ {22}\) and KEGG\(^ {96}\) were applied to predict functional differences of bacterial communities in the liquor starter samples based on the raw 16S rRNA gene data. Briefly, the closed-reference OTU was chosen to blast against the Greengenes database (13.5)\(^ {97}\), then the metagenomes of selected OTUs were predictively analyzed against the KEGG database. The prediction of PICRUSt was evaluated by the Nearest Sequenced Taxon Index (NSTI)\(^ {22}\). Differentially abundant KEGG pathways were calculated using the program Metastats\(^ {98}\) (using an FDR correction, \(p<0.05\), \(\text{Log}_2(\text{FC})|>1\)), and pathways with significant difference were used for drawing bubble profile.

**References**

1. Xu, M. L., Yu, Y., Ramaswamy, H. S. & Zhu, S. M. Characterization of Chinese liquor aroma components during aging process and liquor age discrimination using gas chromatography combined with multivariable statistics. *Sci Rep* 7, 59671, https://doi.org/10.1038/srep39671 (2017).

2. Yao, F. et al. Chemical Analysis of the Chinese Liquor LuzhouLaojiao by Comprehensive Two-Dimensional Gas Chromatography/Time-of-Flight Mass Spectrometry. *Sci Rep* 5, 9553, https://doi.org/10.1038/srep09553 (2015).

3. Farias, M. E., Sosa, O. A., Mendoza, L. M. & Fernández, P. A. Microbial Interaction in Fermented Beverages 65 (Nova Science pub Inc, New York, 2011).

4. Xu, Y., Wang, D., Fan, W. L., Mu, X. Q. & Chen, J. in *Biotecnologia en China II: Chemicals, Energy and Environment* (eds G. T. Tsao, Pingkai Ouyang, & Jian Chen) 189–233 (Springer Berlin Heidelberg, 2010).

5. Chen, B., Wu, Q. & Xu, Y. Filamentous fungal diversity and community structure associated with the solid state fermentation of Chinese Maotai-flavor liquor. *Int J Food Microbiol* 179, 80–84, https://doi.org/10.1016/j.ijfoodmicro.2014.03.011 (2014).

6. Zhang, C. et al. Characterization of volatile compounds from Daqu-a traditional Chinese liquor fermentation starter. *Int J Food Sci Technol* 46, 1591–1599, https://doi.org/10.1111/1365-2621.12660 (2011).

7. Zheng, X.-W. & Han, B.-Z. Baijiu (白酒), Chinese liquor: History, classification and manufacture. *J Ethn Foods* 3, 19–25, https://doi.org/10.1016/j.jef.2016.03.001 (2016).

8. Shi, J. H. et al. Analyses of microbial consortia in the starter of Fen Liquor. *Let Appl Microbiol* 48, 478–485, https://doi.org/10.1111/j.1472-765X.2009.02554.x (2009).

9. Yan, Z. et al. Monitoring the Ecology of Bacillus During Daqu Incubation, a Fermentation Starter, Using Culture-Dependent and Culture-Independent Methods. *J Microbiol Biotechnol* 23, 614–622 (2013).

10. Zheng, X.-W. et al. Complex microbiota of a Chinese “Fen” liquor fermentation starter (Fen-Daqu), revealed by culture-dependent and culture-independent methods. *Food Microbio* 51, 293–300, https://doi.org/10.1016/j.fm.2012.03.008 (2012).

11. Liu, X., Guo, K. & Zhang, H. Determination of microbial diversity in Daqu, a fermentation starter culture of Maotai liquor, using nested PCR-denaturing gradient gel electrophoresis. *World J Microbiol Biotechnol* 28, 2375–2381, https://doi.org/10.1007/s11274-012-1045-y (2012).

12. Wang, H. Y. & Xu, Y. C. Effect of temperature on microbial composition of starter culture for Chinese light aroma style liquor fermentation. *Let Appl Microbiol* 60, 85–91, https://doi.org/10.1111/lam.12344 (2015).

13. Zhang, L., Wu, C., Ding, X., Zheng, J. & Zhou, R. Characterisation of microbial communities in Chinese liquor fermentation starters of Daqu using nested PCR-DGGE. *World J Microbiol Biotechnol* 30, 3055–3063, https://doi.org/10.1007/s11274-014-1723-y (2014).

14. Hamady, M., Walker, J. J., Harris, J. K., Gold, N. J. & Knight, R. Error-correcting barcoded primers allow hundreds of samples to be pyrosequenced in multiplex. *Nat Methods* 5, 235–237, https://doi.org/10.1038/nmeth1184 (2008).

15. Li, X.-R. et al. Bacterial and fungal diversity in the starter production process of Fen liquor, a traditional Chinese liquor. *J Microbiol Meth* 91, 430–438, https://doi.org/10.1016/j.jmib.2012.02.003 (2013).

16. Li, X.-R. et al. Bacterial and fungal diversity in the traditional Chinese liquor fermentation process. *Int J Food Microbiol* 146, 31–37, https://doi.org/10.1016/j.ijfoodmicro.2011.01.030 (2011).

17. Zhang, X., Zhao, J. & Du, X. C. Barcoded pyrosequencing analysis of the bacterial community of Daqu for light-flavour Chinese liquor. *Let Appl Microbiol* 58, 549–555, https://doi.org/10.1111/lam.12225 (2014).

18. Wang, X., Du, H. & Xu, Y. Source tracking of prokaryotic communities in fermented grain of Chinese strong-flavor liquor. *Int J Food Microbiol* 244, 27–35, https://doi.org/10.1016/j.ijfoodmicro.2016.12.018 (2017).

19. Faust, K. et al. Microbial Co-occurrence Relationships in the Human Microbiome. *PLoS Comput Biol* 8, e1002606, https://doi.org/10.1371/journal.pcbi.1002606 (2012).

20. Govers, D. et al. The Treatment-Naive Microbiome in New-Onset Crohn’s Disease. *Cell Host Microbe* 15, 382–392, https://doi.org/10.1016/j.chom.2014.02.005 (2014).

21. Rui, J. et al. The core populations and co-occurrence patterns of prokaryotic communities in household biogas digesters. *Biotechnol Biofuels* 8, 158, https://doi.org/10.1186/s13068-015-0339-3 (2015).

22. Langille, M. G. I. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat Biotechnol* 31, 814–821, https://doi.org/10.1038/nbt.2676 (2013).

23. Corrigan, A., de Leeuw, M., Penaud-Frézet, S., Dimova, D. & Murphy, R. A. Phylogenetic and Functional Alterations in Bacterial Community Compositions in Broiler Ceca as a Result of Mannan Oligosaccharide Supplementation. *Appl Environ Microbiol* 81, 3460–3470, https://doi.org/10.1128/aem.04194-14 (2015).
64. Liu, G. et al. Genomic and Secretome Analyses Reveal Unique Features of the Lignocellulolytic Enzyme System of Penicillium decumbens. *PloS ONE* 8, e51585, https://doi.org/10.1371/journal.pone.0051585 (2013).
65. Krisch, J., Bencsik, O., Papp, T., Vágóvölgyi, C. & Tkó, M. Characterization of a β-glucosidase with transgalactosylation capacity from thezymocyte Rhizomucor miehei. *Biocell* 114, 555–560, https://doi.org/10.1016/j.biotech.2012.02.117 (2011).
66. de Souza, W. R. et al. Transcriptome analysis of Aspergillus niger grown on sugarcane bagasse. *Biotechnol Biofuels* 4, 40–40, https://doi.org/10.1186/1754-6834-4-40 (2011).
67. Blandino, A., Al-Aserri, M. E., Pandella, S. S., Cantero, D. & Webb, C. Cereal-based fermented foods and beverages. *Food Res Int* 36, 527–543 (2003).
68. Caplice, E. & Fitzgerald, G. F. Food fermentations: role of microorganisms in food production and preservation. *Int J Food Microbiol.* 50, 131–149, https://doi.org/10.1016/S0168-1609(99)00082-3 (1999).
69. Dashuban, M., Schratt, H. & Qin, W. Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives. *Int J Biol Sci* 5, 578–595 (2009).
70. McClendon, S. D. et al. Thermoascom auranticus is a promising source of enzymes for biomass deconstruction under thermophilic conditions. *Biotechnol Biofuels* 5, 54–54, https://doi.org/10.1186/1754-6834-5-54 (2012).
71. Basaveswara Rao, V. et al. Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives. *Int J Biol Sci* 5, 578–595 (2009).
72. Herrera Bravo de Laguna, I., Toledo Marante, F. J. & Mioso, R. C. Enzymes and bioproducts produced by the ascomycete fungus *Paecilomyces variotii*. *Appl Microbiol* 119, 1455–1466, https://doi.org/10.1111/j.1365-2958.2014.12570.x (2015).
73. Zhu, L., Wu, Q., Dai, J., Zhang, S. & Wei, F. Evidence of cellulose metabolism by the giant panda gut microbiome. *Proc Natl Acad Sci USA* 108, 17714–17719, https://doi.org/10.1073/pnas.1107561108 (2011).
74. Schloss, P. D. et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–565, https://doi.org/10.1038/nature06269 (2007).
75. He, S. et al. Comparative Metagenomic and Metatranscriptomic Analysis of Hindgut Paunch Microbiota in Wood- and Dung-Feeding Higher Termites. *PloS ONE* 11, e016126, https://doi.org/10.1371/journal.pone.0016126 (2013).
76. Warnecke, F. et al. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* 450, 560–565 (2007).
77. Hess, M. et al. Metagenomic Discovery of Biomass-Degrading Genes and Genomes from Cow Rumen. *Science* 331, 463–467, https://doi.org/10.1126/science.1200387 (2011).
78. Frandsen, R. J. N., Andersson, J. A., Kristensen, M. B. & Giese, H. Efficient four fragment cloning for the construction of vectors for targeted gene replacement in filamentous fungi. *BMC Mol Biol* 9, 78–70, https://doi.org/10.1186/1471-2199-7-10 (2006).
79. Herrera Bravo de Laguna, I., Toledo Marante, F. J. & Mioso, R. C. Enzymes and bioproducts produced by the ascomycete fungus *Paecilomyces variotii*. *Appl Microbiol* 119, 1455–1466, https://doi.org/10.1111/j.1365-2958.2014.12570.x (2015).
80. Zhu, L., Wu, Q., Dai, J., Zhang, S. & Wei, F. Evidence of cellulose metabolism by the giant panda gut microbiome. *Proc Natl Acad Sci USA* 108, 17714–17719, https://doi.org/10.1073/pnas.1107561108 (2011).
81. Martin, K. J. & Rygiewicz, P. T. Fungal-specific PCR primers developed for analysis of the ITS region of environmental DNA extracts. *BMC Microbiol* 5, 28–28, https://doi.org/10.1186/1471-2180-5-28 (2005).
82. Prosdocimi, F., de Carvalho, D. C., de Almeida, R. N. & Beheragary, L. B. The complete mitochondrial genome of two recently derived species of the fish genus Nanosperma (Perciformes, Percichthyidae). *Mol Bio Rep* 39, 2767–2772, https://doi.org/10.1007/s11033-011-1034-5 (2012).
83. Schloss, P. D. et al. Introducing mothur: Open-Source, Platform-Independent, Community-Supported Software for Describing and Comparing Microbial Communities. *Appl Environ Microbiol* 75, 7537–7541, https://doi.org/10.1128/aem.01541-09 (2009).
84. Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. & Knight, R. UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200 (2011).
85. Quest, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids res* 41, D590–D596, https://doi.org/10.1093/nar/gks1219 (2012).
86. Caporaso, J. et al. QIIME allows analysis of high-throughput community sequencing data. *Nat Meth* 7, 335–336, https://doi.org/10.1038/nmeth.f.303 (2010).
87. Wang, Q., Garrity, G. M., Tiedje, J. M. & Cole, J. R. Naive Bayesian classifier for rapid assignment of rRNA sequences into the new *Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB.* *Appl Environ Microbiol* 75, 5261–5267, https://doi.org/10.1128/AEM.00062-07 (2007).
88. Gilbert, J. J. et al. The Taxonomic and Functional Diversity of Microbes at a Temperate Coastal Site: A ‘Multi-Omic’ Study of Seasonal and Diel Temporal Variation. *PloS ONE* 5, e15545, https://doi.org/10.1371/journal.pone.0015545 (2010).
89. Ithaka, R. & Gentleman, R. R. A Language for Data Analysis and Graphics. *J Comput Graph Stat* 5, 299–314, https://doi.org/10.1080/1061860.1996.10474713 (1996).
90. Barbérian, A., Bates, S. T., Casamayor, E. O. & Fierer, N. Using network analysis to explore co-occurrence patterns in soil microbial communities. *ISME* J 6, 343–351, https://doi.org/10.1038/ismej.2011.118 (2012).
91. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498–2504, https://doi.org/10.1101/gr.123903.0 (2003).
92. Kanehisa, M., Furumichi, M., Tanabe, M., Sato, Y. & Morishima, K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 45, D353–D361, https://doi.org/10.1093/nar/gkw1092 (2017).
93. DeSantis, T. Z. et al. Greengenes, a Chimera-Checked 16S rRNA Gene Database and Workbench Compatible with ARB. *Appl Environ Microbiol* 72, 5089–5092, https://doi.org/10.1128/aem.03095-08 (2006).
94. Paulson, J. N., Pop, M. & Bravo, H. C. Metastats: an improved statistical method for analysis of metagenetic data. *Genome Biol* 12(Suppl 1), P17, https://doi.org/10.1186/gb-2011-12-s1-p17 (2011).

**Acknowledgements**

This study was supported by the National Key Technology R&D Program of China (2015BAD16B01), China Agriculture Research System (No. CARS-10-B22), the National Natural Science for Youth Foundation of China (13106085), the National Natural Science for General Foundation of China (31770395), Science & Technology Program of Sichuan Province (2017NZ018 and 2017HH0077), and the Open Fund of Key Laboratory of Environmental and Applied Microbiology (KLEMACAS201501, KLCAS-2014-05 and KLCAS-2016-02) of the Chinese Academy of Sciences. We thank Professors Zhongyan Wang, Mingyang Zhuang and Yanyong Wu for providing valuable suggestions concerning liquor starter research.
Author Contributions
H.Z., Y.H., D.L., H.L., Y.J. and Y.F. designed the experiment; Y.H. and Z.Y. performed the experiment and analyzed data; Y.H., D.Z., W.Z. and H.H. collected samples and communicated with the liquor factory; Y.H. and Z.Y. wrote the main manuscript, Y.H. prepared Figs 1–3 and Z.Y. prepared Figs 4–5; Y.H., Z.Y., K.H., Y.F. and H.Z. revised the manuscript. All authors reviewed the manuscript.

Additional Information
Supplementary information accompanies this paper at https://doi.org/10.1038/s41598-017-14968-8.

Competing Interests: The authors declare that they have no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2017