




Metagenomic and metatranscriptomic analysis of the microbial community structure and metabolic potential of fermented soybean in Yunnan Province

Xiao-Feng LIU^{1,2#}, Chen-Jian LIU^{1#}, Xue-Qin ZENG¹, Hai-Yan ZHANG¹, Yi-Yong LUO¹, Xiao-Ran LI^{1*} 

Abstract

Traditional fermented soybean contains a variety of microflora. To obtain a comprehensive and accurate understanding of the microbial community structure and metabolic potential of fermented soybean, comparative metagenomics and metatranscriptomics were performed on a sample of fermented soybean in Yunnan Province. Metagenomic DNA and metatranscriptomic RNA were sequenced using Illumina HiSeq™2500, which yielded a total of 92,192,276 reads, with an average read length of 150 bp and 38,798,262 paired-end sequences, with an average length of 151 bp. The results show that *Providencia stuartii* was the most abundant species and the genes of carbohydrates (2296, 13.30%), protein metabolism (1530, 8.86%), and amino acids and amino acid derivatives (1423, 8.24%) were dominant. The expression levels of the genes belonging to amino acid transport and metabolism processes were the highest according to the reads per length of transcript in kilo-bases per million mapped reads (RPKM) values, followed by energy production and conversion and carbohydrate transport and metabolism. The metabolic pathways were primarily associated with carbohydrates, proteins and amino acids, which might be in accordance with the high levels of proteins and other nutrients in soybeans. Overall, these findings provide insights into the community structure and metabolic potential of the fermented soybean microbiome.

Keywords: fermented soybean; metagenomics; metatranscriptomics; microbial community.

Practical Application: This study analyzed the community structure and the metabolic pathway of fermented soybean in Yunnan province, which is helpful for the efficient detection of the microbiological functions.

1 Introduction

Fermented soybean is widely distributed throughout Yunnan Province as well as a majority of the regions in China. In Yunnan Province, fermented soybean is produced by spontaneous fermentation and no microorganisms are added to the fermentation process of fermented soybean. Until recently, most of the fermented soybean has been produced in traditional family workshops in China. In Yunnan province, the traditional processing method of fermented soybean is as follows: Firstly, the soybeans are washed and soaked in water for 24 h, then boiled for 1-2 h. After draining, the soybeans are tightly packed in a small bamboo basket layered with the leaves of bamboo or banana and the baskets are covered with the leaves of soybean plant to maintain an ambient temperature. Additionally about 12-15% (w/w) salt is added, and spices (such as sugar, Chinese prickly ash, fresh hot pepper paste, or dry hot pepper powder) are added after fermenting for 3 to 5 days. Finally, the mixture is packed in a tank for about one month. Spontaneous fermentation without the use of starter cultures or sterilization leads to the growth of various microorganisms during soybean preparation.

According to the different types of microorganisms present in fermentation, it is divided into two classes: bacterial-fermented soybean and mold-fermented soybean. The fermented soybean

in Yunnan Province is primarily bacterial-fermented (Liu et al., 2012). The microorganisms in fermented soybean could produce several types of enzymes (Sanjukta & Rai, 2016; Wang et al., 2013). These enzymes could convert the contents of fermented soybean into complex metabolic products, which contribute to the color and aroma of fermented soybean and increase the content of vitamins, essential amino acids, etc. The quality of fermented soybean depends on the microorganisms involved in its fermentation. Therefore, studying the microbial community structure and gene expression of fermented soybean is necessary and meaningful.

The initial approach to study the microbial community structure of fermented soybean is culture-based method. The number of species isolated by the culture method accounted for 1-5% of the estimated quantity (Amann et al., 1995). In some bacteria, culture-based approaches have limitations and challenges for their culture reproducibility. More recent culture-independent methods based on 16S rRNA gene sequencing have enhanced the ease of studying the microbial community structure (Devi et al., 2015; Matsui et al., 2013; Polka et al., 2015), but these methods are limited in taxonomic resolution and are subject to primer biases (Polz & Cavanaugh, 1998) and other system biases in

Received 13 Apr., 2018

Accepted 24 Aug., 2019

¹Faculty of Life Science and Technology, Kunming University of Science and Technology, Kunming, Yunnan, China

²Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao, China

*X.-F. Liu and C.-J. Liu contributed equally to this work

*Corresponding author: starkeyran@163.com

PCR (Carraway & Marinus, 1993). Thus, the precise microbial community profile has not been accurately and comprehensively studied using these methods. By directly analyzing the total DNA of the microbial community, metagenomics can reveal the species, number and proportion of different microbes in the community to the greatest extent and can reveal the metabolic potential of all microorganisms in the community. Although the bias of sequencing method cannot be eliminated, metagenomic approaches based on the directly random sequencing of environmental DNA are free of PCR biases compared with gene target PCR-based methods. In recent years, the application of metagenomics in fermented foods such as kimchi and kefir grains has increased (Jung et al., 2011; Nalbantoglu et al., 2014). Metagenomics approaches greatly extend the information on gene composition and metabolism but are limited in the recognition of gene activity, function, etc. The drawback of metagenomics can be overcome by metatranscriptomics. Metatranscriptomics can be applied to examine the gene function and metabolism of the active microorganisms, which is a promising approach for gaining unbiased insights into the functionality of a microbial community (Van Hijum et al., 2013). Results of analyzing lactic acid bacterial (LAB) gene expression during kimchi fermentation contributed to the knowledge of the active populations and gene expression in the LAB community are responsible for an important fermentation process (Jung et al., 2013).

Here, we studied the microbial metagenomes and microbial metatranscriptomes from fermented soybean collected from the Prefecture of De Hong in Yunnan Province using Illumina HiSeq™2500. The aim of the present study was to investigate the microbial community structure and metabolic pathway of fermented soybean and analyze the metabolic processes of bioactive microbes.

2 Materials and methods

2.1 Sample collection

Based on previous research (data not shown) and unique environment of human geography of Prefecture of Dehong in Yunnan Province for China, moreover, the microbial community structure of fermented soybean in different places is very similar, so we chose Prefecture of Dehong for sampling. We repeated the sampling at the main vegetable market in the city, five parallel samples were collected in October 2014, which all were hand-made traditional fermented soybeans. The soybean samples were spontaneously fermented at ambient temperature for about 1 month by the local residents and did not contain any additives such as hot pepper, salt, etc. The fermented soybean samples were stored at -80 °C for total DNA and RNA extraction.

2.2 Total DNA extraction, determination and purification

Total DNA and RNA were extracted on the same day, and the five parallel samples were evenly mixed and ground for DNA and RNA extraction. DNA was extracted from 0.5 g of the samples in a 1.5 mL centrifuge tube by a CTAB-based method as previously described (Schmidt et al., 1991) with slight modifications. Briefly, the samples were resuspended in

450 µL of a lysis buffer (0.1 mol/L Tris/HCl, 0.1 mol/L EDTA, and 0.75 mol/L sucrose) treated with lysozyme (1 mg/mL) and incubated in a water bath at 37 °C for 30 min. Subsequently, the samples were treated with SDS (1%) and CTAB (1%), followed by treatment with phenol:chloroform:isoamyl alcohol (25:24:1, V/V) and chloroform:isoamyl alcohol (24:1, V/V) to remove impurities. The samples were resuspended in 50 µL of a TE buffer (0.01 mol/L Tris-HCl, pH 8.0, and 0.001 mol/L EDTA, pH 8.0) following treatment with ethanol. The DNA samples were stored at -20 °C until further analysis. Prior to sequencing, agarose gel electrophoresis was conducted to confirm the existence of the target band. Subsequently, the DNA samples were purified using the DNeasy plant Mini Kit. The purity and yield of the DNA samples were determined using the Thermo Scientific Nano Drop 1000 Spectrophotometer and Qubit fluorometer with DNA dye [picogreen (sigma, America)] to determine whether the samples were adequate and sufficient for sequencing and downstream reactions.

2.3 Metagenome sequencing, assembly and annotation

DNA was sequenced using a HiSeq 2500 sequencing platform (Illumina, America), according to the manufacturer's protocol. Paired-end sequencing technology was used and the read length was 300 bp. The obtained sequences were filtered. Quality control was performed using the following criteria: (1) short reads ligated with adapter sequences were considered as adapter contamination and removed; and (2) reads contained more than 50 bases with quality values lower than 2 were removed. Filtered clean data was used in a contig assembly by Metavelvet (Namiki et al., 2012) at 81 kmer. The contigs were further analyzed to discriminate lengths longer than 200 bp. Gene prediction was conducted using MetaGeneAnnotator. The obtained genes were blasted to the SEED database (Overbeek et al., 2005). Metagenomics analysis was conducted using Metagenome Rapid Annotation with Subsystem Technology (MG-RAST) (Meyer et al., 2008; Glass et al., 2010). The reads passing the MG-RAST quality filters were matched to the SILVA Small Subunit (SSU) rRNA database (Pruesse et al., 2007) and the functional and organism database M5NR (M5 non-redundant protein database), (Wilke et al., 2012) for analysing the diversity of species. To explore the metabolic potential of the fermented soybean microbiome. These data were compared with SEED subsystems using a maximum e-value of 1e-5, a minimum identity of 80%, and a minimum alignment length of 15 aa for protein and 15 bp (The default setting of MG-RAST) for RNA databases.

2.4 RNA extraction

The pretreatment was required prior to total RNA extraction. The samples were mixed with moderately sterile water, followed by incubation in a bed temperature incubator with shaking at 180 rpm and ambient temperature for 2h. The eluent was centrifuged at 5000 rpm, and then the sediment was collected for the extraction of total RNA. An RNA extraction kit (BioTeke, China) was used to isolate total RNA from fermented soybean according to the manufacturer's instructions. Finally, RNA was dissolved in 50 µL DEPC-H₂O, then 1 µL RNAase inhibitor

(40 u/μL) was added to the RNA solutions. Total RNA was quality tested using a 2100 Bio analyzer. The RNA Integrity Number (RIN) value was 6.1. To reduce DNA contamination in the RNA extract, the sample was treated with DNase I (Takara, Japan) for 30 min at 37 °C. The RNA treated with DNase I was used as a PCR template to amplify the 16S rRNA, which were analyzed in a 1% (w/v) agarose gel ensure for non-DNA contamination. Subsequently, rRNA was deleted from the RNA sample using the Ribo-Zero™ rRNA Removal Kit (Bacteria) (Epicenter, America) according to the manufacturer's instructions.

2.5 cDNA library construction and illumina sequencing

cDNA was synthesized using the SuperScript™ Double-Stranded cDNA Synthesis Kit (Invitrogen, America) according to the manufacturer's instructions with 100 ng of purified mRNA. The constructed library was assessed three times through quantification using a Qubit fluorometer, 2% agarose gel electrophoresis and additional quantification using a high-sensitivity DNA chip to ensure the quality of the library. The concentration of cDNA was determined with a NanoDrop 3300 (Thermo Fisher, USA) with Picogreen reagent. A total of 10 ng of cDNA library was needed, and cluster generation was realized using the TruSeq PE Cluster Kit (Illumina, America). Subsequently, the clustered cDNA was bidirectionally sequenced on an Illumina Hiseq™2500 platform.

2.6 Metatranscriptomic bioinformatics

The raw reads were obtained after sequencing and stored in the FASTQ file form. To obtain clean reads and guarantee the quality of information analysis, the raw reads were filtered using the FASTX-Toolkit. The expression results were gathered as reads per length of transcript, expressed as kilo-base per million mapped reads (RPKM) values (Ammar et al., 2012). This method is described using the following Formula 1:

$$RPKM = \text{total exon reads} / \text{mapped reads (millions)} \times \text{exon length (KB)} \quad (1)$$

2.7 Nucleotide sequence accession numbers

The metagenomics sequences and metatranscriptomics sequences from this study are available in the sequence read archive (SRA) under the accession number SRP083828.

3 Results

3.1 The information of metagenomics and metatranscriptomics sequences

The DNA and RNA of fermented soybean were sequenced on an Illumina Hiseq™2500 platform. For the metagenomics sequences, 92,192,276 raw reads were generated, averaging 13.83 Gb in size and 150 bp in length. For the metatranscriptomics sequences, 38,798,262 paired-end sequences were generated, averaging 11,717,075,124 bp, with an average length of 151 bp.

For the metagenomics data, the short Illumina reads were assembled into longer contigs that could be analyzed and annotated by standard methods. A total of 40,644 contigs were

obtained. The average length of the contigs was 1,264 bp, and the average guanine-cytosine content (GC content) was 41.93%. The assembled contigs were used to predict genes and then 80,514 genes were obtained.

3.2 The composition of the microbial communities in the fermented soybean

At the domain level, most of the contigs of the fermented soybean belonged to bacteria (95.22%, 64437) and a few of the contigs belonged to eukaryote (0.28%, 189) and viruses (0.58%, 392). In addition, classification information for the remaining contigs (3.92%, 2653) could not be determined. At the phylum level, the most abundant phylum was *Proteobacteria*, with 24,539 contigs. The second most abundant phylum was *Firmicutes*, containing 19,583 contigs.

The most abundant genus was *Providencia* of *Proteobacteria*, with 75.69% of all contigs. The second most abundant genus was *Enterobacter* of *Proteobacteria*, with 17.87% of all contigs. The sequence percentages of all other genera were less than 2%. Among these genera, some genera were associated with fermentation, such as *Bacillus* (0.5%), *Lactobacillus* (0.2%), *Streptococcus* (0.2%), *Lactococcus* (0.2%), etc. (Figure 1A).

The circular tree showed the relatedness of the metagenome selected for comparison, with color according to phylum (Figure 1B). The most abundant species was *Providencia stuartii* (67.10%), followed by *Enterobacter cloacae* (16.62%), *Providencia rettgeri* (7.36%) and *Escherichia coli* (1.88%). Notably, 0.3% of the contigs belonged to *Bacillus subtilis*. In addition, few contigs were assigned to *Bacillus thermoamylovorans* and *Bacillus licheniformis* (Figure 1B).

3.3 Comparative functional analysis of the fermented soybean microbiome

All sequences were analyzed using the publicly available SEED subsystem on the MG-RAST server in order to explore the metabolic potential of the fermented soybean microbiome. Based on analysis of the MG-RAST server, all genes were categorized into 28 functional subsystems, with the highest proportions of genes associated with clustering-based subsystems (2323, 13.46%), carbohydrates (2296, 13.30%), protein metabolism (1530, 8.86%), amino acids and their derivatives (1423, 8.24%), miscellaneous (1292, 7.49%), and cofactors, vitamins, prosthetic groups and pigments (984, 5.70%) (Figure 2A). All the genes associated with carbohydrates were categorized into 12 subsystems, with the greatest proportions of genes associated with central carbohydrate metabolism (725, 23.04%), monosaccharides (496, 15.76%), di- and oligosaccharides (355, 11.28%) and fermentation (243, 7.72%) (Figure 2B). In the categories of cofactors, vitamins, prosthetic groups and pigments, 589 genes belonged to folate and pterines (Figure 2C). In contrast, the other subsystems were represented at relatively low levels (<5%) in the metagenome (Figure 2A). In addition, 445 genes (2.58%) associated with virulence, disease and defense subsystems, in which 52.48% of the genes were assigned to antibiotic and toxic compound resistance subsystems (Figure 2A).

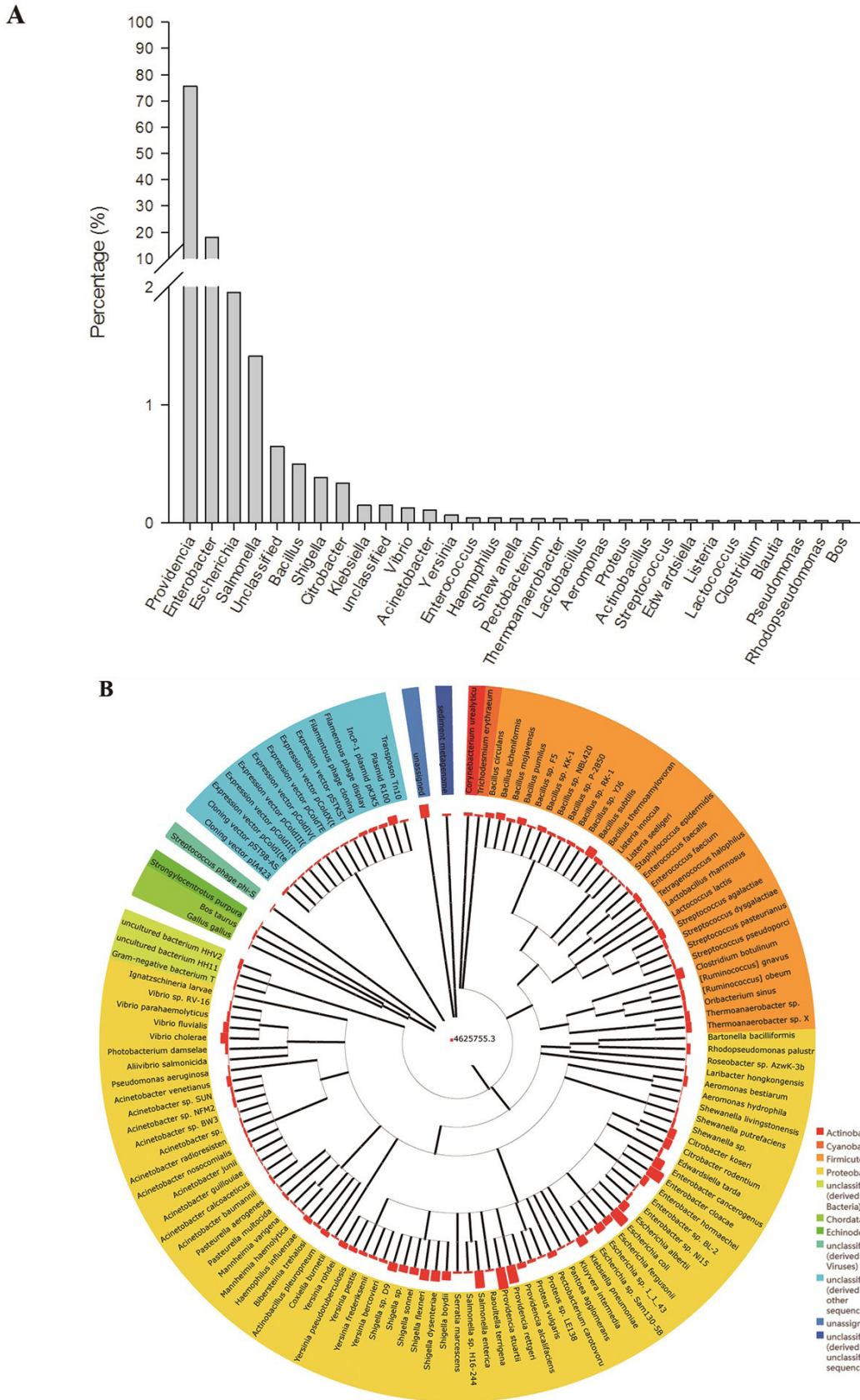


Figure 1. Microbial-classified information of the metagenomes of fermented soybean. (A) Classification at the genus level; (B) Classification at the species level.

Metagenomics and metatranscriptomics of fermented soybean

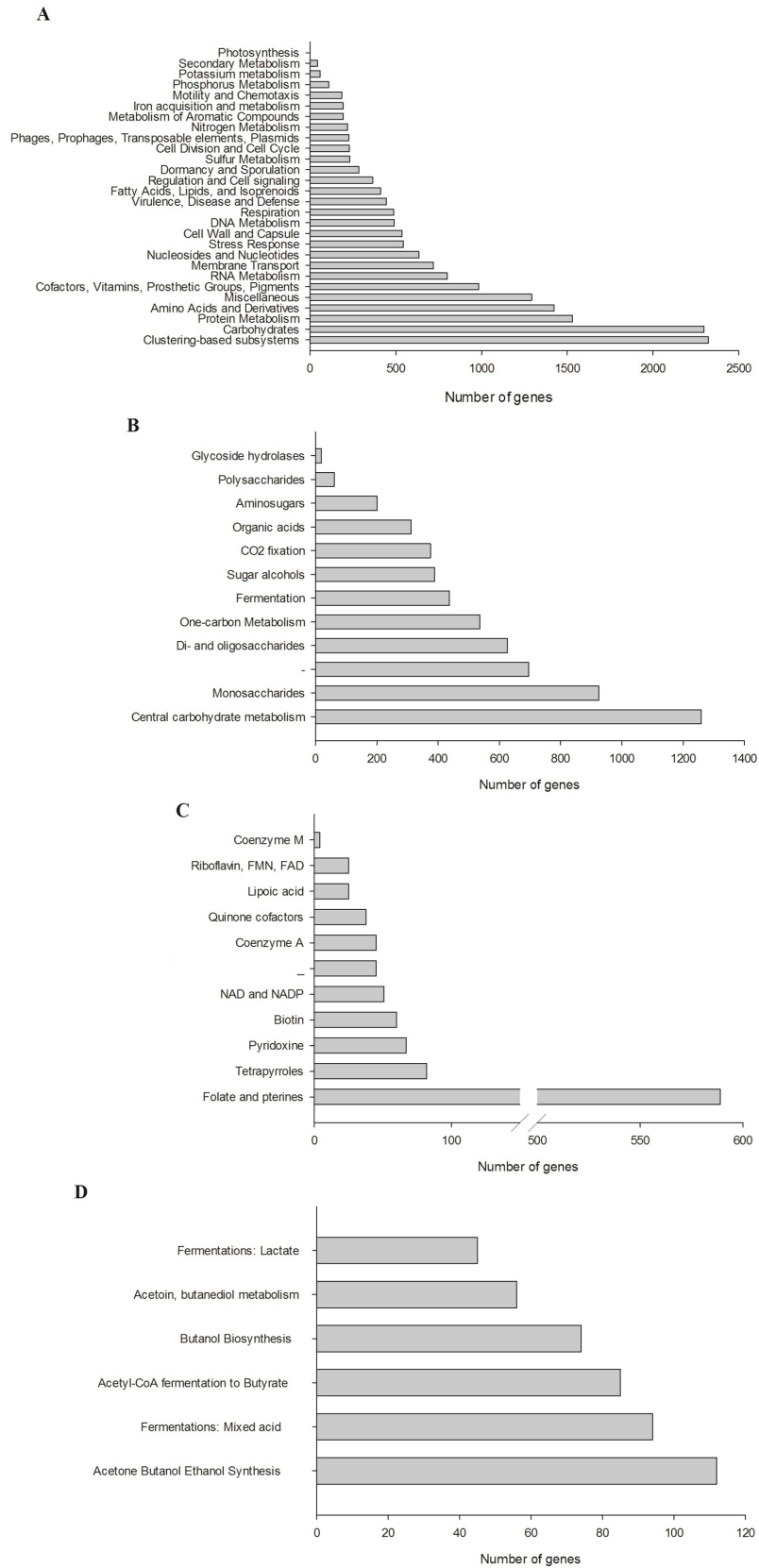


Figure 2. The distribution and classification of the gene function in the metagenomes of fermented soybean. (A) The general function classification of the genes in the metagenomes; (B) Detailed functional classification in subsystems of Carbohydrates of the genes in the metagenomes; (C) The detailed function classification in subsystems of Cofactors, vitamins, prosthetic groups and pigments of the genes in the metagenomes; (D) The fermentation metabolism subcategory).

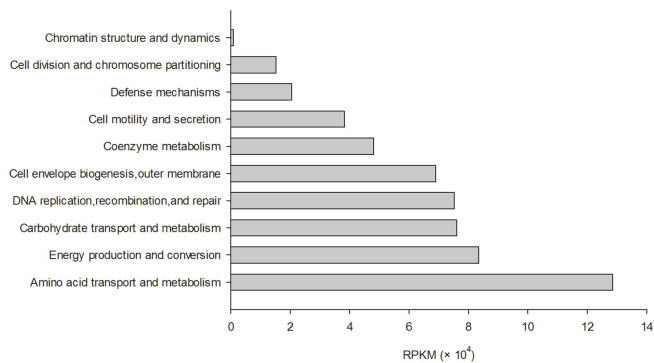


Figure 3. The gene expression levels at different functional classifications.

3.4 The metatranscriptomic analysis of the fermented soybean microbiome

To compare the gene expression level among clusters of orthologous group (COG) function clusters, we calculated the RPKM values of all predicted genes and the genes (total RPKM value > 5) were subsequently classified into the COG function clusters. The RPKM value for amino acid transport and metabolism was highest, followed by energy production and conversion, and carbohydrate transport and metabolism (Figure 3). The RPKM values of the genes related to proteases, sugar/maltose fermentation was 395.76 and 221.90.

4 Discussion

The incorporation of metagenomics and metatranscriptomic approaches provides an overview of both community taxonomies combined with the encoded and expressed functional diversity of these communities. The incorporation of various metagenomics and metatranscriptomic approaches has been applied to environment, human and animal samples to characterize the microbial and viral communities present in an ecosystem, thereby elucidating metabolic capabilities and native taxa, such as acid mine drainage systems (Chen et al., 2015b), cystic fibrosis patients (Lim et al., 2013) and hindgut paunch microbiota in wood- and dung-feeding higher termites (He et al., 2013). Currently, metagenomics has been applied to various fermented foods, such as Korean kimchi (Jung et al., 2011), pure tea (Lyu et al., 2013), Shaoxing rice wine (Xie et al., 2013) etc. However, the application of metagenomics and metatranscriptomics to study the microbial community structure and the metabolic characteristics of fermented foods is lacking. Fermented foods have a clear advantage of floras, and their microorganisms generally exhibit high activity. In the present study, we applied metagenomics and metatranscriptomics to investigate the microbial community structure and gene expression of fermented soybean, a traditional fermented food in China.

Although the metagenomics and metatranscriptomic data were large, the sequencing qualities were good. Thus, the results of the microbial community structure and gene expression were reliable. A minor amount of microorganisms, such as fungi and viruses, were detected, suggesting that the sequencing coverage was relatively high and that the results reflect the

comprehensive general picture of the microbial communities of fermented soybean.

Providencia of *Enterobacteriaceae* is an unusual pathogenic bacterium; however, its infection rate has increased, so the genus of *Providencia* has drawn increasing research attention. Relatively common bacterial species in the genus *Providencia* include *Providencia rettgeri* and *Providencia stuartii*. *Providencia rettgeri*, which has been detected in foods and have increased to large numbers during the process of food production, transport and sales; and their presence could contaminate foods and induce food poisoning. Currently, an increasing number of studies have shown that *Providencia rettgeri*, *Providencia stuartii*, *Escherichia coli* and *Enterobacter cloacae* could produce β -lactamases and extended-spectrum β -lactamases (Ferjani et al., 2015; Mahrouki et al., 2015; Osawa et al., 2015). Beta-lactam antibiotic is the most widely used antibiotic. In recent years, consistent with the development and widespread use of antibiotic medicines, the phenomenon of bacterial drug resistance has become increasingly serious. Bacterial drug resistance occurs because these bacteria could produce β -lactamases and extended-spectrum β -lactamases. Several studies have shown that the types of bacteria described above could participate in fermentation and that these species might be derived from the environment. Notably, there are many fermentation strains with low abundance in fermented soybean, such as *Bacillus subtilis*, *Bacillus thermoamylovorans* and *Bacillus licheniformis* and the genus of *Lactobacillus* and *Lactococcus*. A previous study revealed that fungi, yeasts, *Bacillus* and lactic acid bacteria are key players in fermented soybean fermentation, and that when identified probiotic microorganisms participate in fermentation, a higher quality of fermented soybean product was produced (Chen et al., 2015a). It has been speculated that fermented soybean is contaminated with pathogenic bacteria (*Providencia rettgeri*, *Providencia stuartii*, *Escherichia coli* and *Enterobacter cloacae*) during storage and that the pathogenic bacteria increased to large numbers that inhibit the growth of *Bacillus* and lactic acid bacteria. *Bacillus subtilis* widely exists in the natural environment and could produce proteases and esterases that degrade the protein and fat in soybeans, effectively promoting the absorption of nutrients in soybean for the human body (Kada et al., 2013; Li et al., 2014). *Bacillus subtilis* could also produce higher levels of isoflavone aglycones, which might enhance health benefits over traditional fermented natto (Wei et al., 2008). *Bacillus amyloliquefaciens* produces the broad-spectrum antifungal protein, baciamin, which represents one of the few bacterial antifungal proteins reported to date (Wong et al., 2008).

Unexpectedly, lactic acid bacteria were not the predominant floras, suggesting that fermented soybean is contaminated with pathogenic bacteria during storage. Nevertheless, carbohydrate categories, including the metabolism of mono-, di-, and oligosaccharides and fermentation, were key categories for the fermented soybean microbiome, and the RPKM value of the carbohydrate transport and metabolism was higher. Soybeans are rich in protein, so the genes participating in protein metabolism are relatively larger to degrade the protein in soybeans, and the RPKM value of the amino acid transport and metabolism was the biggest. Among the many gene categories in the fermented soybean metagenome, the carbohydrate categories were primarily focused on carbohydrate metabolism, yielding an average of

13.30% of all matches (Figure 2A). The fraction of the kimchi microbiome and pure tea reads in the carbohydrate category was 14.49% (Jung et al., 2011) and 12.89% (Lyu et al., 2013), respectively. Therefore, the genes in the carbohydrate categories were generally higher than those of microbiomes derived from the termite gut (12.89%), acid mine drainage biofilm (12.17%), and Sargasso Sea (12.23%). However, the metagenomics read fractions of monosaccharide (15.76%), di- and oligosaccharide (11.28%), and fermentation (7.72%) within the carbohydrate category in the fermented soybean microbiome were lower than those in the microbiomes derived from kimchi (28.70%, 28.38%, 14.38%), which might be affected by pathogenic bacteria. Kimchi fermentation was predominantly accomplished by heterofermentative lactic acid bacteria, thus the fermentation metabolism subcategory was primarily associated with various lactate fermentations (65.93%) (Jung et al., 2011). In the present study, the dominant microorganisms were replaced with pathogenic bacteria, thus the genes of lactate fermentation were far less (9.66%) in the fermentation metabolism subcategory (Figure 2D). However, the genes were primarily associated with acetone butanol ethanol synthesis (24.03%) and mixed acid fermentations (20.17%) and acetyl-CoA fermentation to butyrate (18.24%), which might be closely associated with fermented soybean fermentation. In addition, the most predominant subsystem of cofactors, vitamins, prosthetic groups and pigments was folate and pterines metabolism (57.13%). Folate is a type of essential nutrient element, particularly for pregnant women and children. Two strains of *Lactobacillus sakei* and a strain of *Lactobacillus plantarum* have been confirmed to produced high levels of extracellular folate (Masuda et al., 2012).

5 Conclusion

In conclusion, the present results revealed the microbial community structure and metabolic potential of fermented soybean in Yunnan Province. *Providencia stuartii* was the predominant species and not the normal LAB, possibly because the fermented soybean was rotten and deteriorated or polluted during storage. Nevertheless, the gene expression of the amino acid transport and metabolism was the highest, thus the protein and amino acid metabolism was still most active.

Acknowledgements

This work were financially supported through the National Natural Science Foundation of China (31260397), the Fundamental Application Research Foundation of Yunnan Province [2017FE468 (-247)] and Yunnan Province Innovation Team of Intestinal Microecology Related Disease Research and Technological Transformation .

References

- Amann, R., Ludwig, W., & Schleifer, K. (1995). Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiological Reviews*, 59(1), 143-169. PMID:7535888.
- Ammar, A., Elouedi, Z., & Lingras, P. (2012). RPKM: the rough possibilistic K-modes. In L. Chen, A. Felfernig, J. Liu & Z. W. Raš (Eds.), *Foundations of intelligent systems* (Lecture Notes in Computer Science, Vol. 7661, pp. 81-86). Switzerland: Springer. Retrieved from https://link.springer.com/chapter/10.1007/978-3-642-34624-8_9
- Carraway, M., & Marinus, M. G. (1993). Repair of heteroduplex DNA molecules with multibase loops in *Escherichia coli*. *Journal of Bacteriology*, 175(13), 3972-3980. <http://dx.doi.org/10.1128/jb.175.13.3972-3980.1993>. PMID:8320213.
- Chen, L., Hu, M., Huang, L., Hua, Z., Kuang, J., Li, S., & Shu, W. (2015a). Identification of key microorganisms involved in Douchi fermentation by statistical analysis and their use in an experimental fermentation. *Journal of Applied Microbiology*, 119(5), 1324-1334. <http://dx.doi.org/10.1111/jam.12917>. PMID:26251195.
- Chen, L., Hu, M., Huang, L., Hua, Z., Kuang, J., Li, S., & Shu, W. (2015b). Comparative metagenomic and metatranscriptomic analyses of microbial communities in acid mine drainage. *The ISME Journal*, 9(7), 1579-1592. <http://dx.doi.org/10.1038/ismej.2014.245>. PMID:25535937.
- Devi, K. R., Deka, M., & Jeyaram, K. (2015). Bacterial dynamics during yearlong spontaneous fermentation for production of ngari, a dry fermented fish product of Northeast India. *International Journal of Food Microbiology*, 199, 62-71. <http://dx.doi.org/10.1016/j.ijfoodmicro.2015.01.004>. PMID:25637876.
- Ferjani, S., Saidani, M., Amine, F. S., & Boutiba-Ben Boubaker, I. (2015). Prevalence and characterization of plasmid-mediated quinolone resistance genes in extended-spectrum beta-lactamase-producing *Enterobacteriaceae* in a Tunisian hospital. *Microbial Drug Resistance*, 21(2), 158-166. <http://dx.doi.org/10.1089/mdr.2014.0053>. PMID:25247633.
- Glass, E. M., Wilkening, J., Wilke, A., Antonopoulos, D., & Meyer, F. (2010). Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes. *Cold Spring Harbor Protocols*, 2010(1), pdb.prot5368. <http://dx.doi.org/10.1101/pdb.prot5368>. PMID:20150127.
- He, S., Ivanova, N., Kirton, E., Allgaier, M., Bergin, C., Scheffrahn, R. H., Kyrpides, N. C., Warnecke, F., Tringe, S. G., & Hugenholtz, P. (2013). Comparative metagenomic and metatranscriptomic analysis of hindgut paunch microbiota in wood- and dung-feeding higher termites. *PLoS One*, 8(4), e61126. <http://dx.doi.org/10.1371/journal.pone.0061126>. PMID:23593407.
- Jung, J. Y., Lee, S. H., Jin, H. M., Hahn, Y., Madsen, E. L., & Jeon, C. O. (2013). Metatranscriptomic analysis of lactic acid bacterial gene expression during kimchi fermentation. *International Journal of Food Microbiology*, 163(2-3), 171-179. <http://dx.doi.org/10.1016/j.ijfoodmicro.2013.02.022>. PMID:23558201.
- Jung, J. Y., Lee, S. H., Kim, J. M., Park, M. S., Bae, J.-W., Hahn, Y., Madsen, E. L., & Jeon, C. O. (2011). Metagenomic analysis of kimchi, a traditional korean fermented food. *Applied and Environmental Microbiology*, 77(7), 2264-2274. <http://dx.doi.org/10.1128/AEM.02157-10>. PMID:21317261.
- Kada, S., Ishikawa, A., Ohshima, Y., & Yoshida, K.-i. (2013). Alkaline serine protease AprE plays an essential role in poly-γ-glutamate production during Natto fermentation. *Bioscience, Biotechnology, and Biochemistry*, 77(4), 802-809. <http://dx.doi.org/10.1271/bbb.120965>. PMID:23563567.
- Li, M., Yang, L. R., Xu, G., & Wu, J. P. (2014). Identification of organic solvent-tolerant lipases from organic solvent-sensitive microorganisms. *Journal of Molecular Catalysis B, Enzymatic*, 99, 96-101. <http://dx.doi.org/10.1016/j.molcatb.2013.11.011>.
- Lim, Y. W., Schmieder, R., Haynes, M., Willner, D., Furlan, M., Youle, M., Abbott, K., Edwards, R., Evangelista, J., Conrad, D., & Rohwer, F. (2013). Metagenomics and metatranscriptomics: Windows on CF-associated viral and microbial communities. *Journal of Cystic*

- Fibrosis*, 12(2), 154-164. <http://dx.doi.org/10.1016/j.jcf.2012.07.009>. PMID:22951208.
- Liu, C., Gong, F., Li, X., Li, H., Zhang, Z., Feng, Y., & Nagano, H. (2012). Natural populations of lactic acid bacteria in douchi from Yunnan Province, China. *Journal of Zhejiang University, Science. B.*, 13(4), 298-306. <http://dx.doi.org/10.1631/jzus.B1100221>. PMID:22467371.
- Lyu, C., Chen, C., Ge, F., Liu, D., Zhao, S., & Chen, D. (2013). A preliminary metagenomic study of puer tea during pile fermentation. *Journal of the Science of Food and Agriculture*, 93(13), 3165-3174. <http://dx.doi.org/10.1002/jsfa.6149>. PMID:23553377.
- Mahrouki, S., Chihi, H., Bourouis, A., Ayari, K., Ferjani, M., Moussa, M. B., & Belhadj, O. (2015). Nosocomial dissemination of plasmids carrying *bla*(TEM-24), *bla*(DHA-1), *aac* (6')-Ib-cr, and *qnrA6* in *Providencia* spp. strains isolated from a Tunisian hospital. *Diagnostic Microbiology and Infectious Disease*, 81(1), 50-52. <http://dx.doi.org/10.1016/j.diagmicrobio.2014.09.021>. PMID:25315769.
- Masuda, M., Ide, M., Utsumi, H., Niuro, T., Shimamura, Y., & Murata, M. (2012). Production potency of folate, vitamin B-12, and thiamine by lactic acid bacteria isolated from Japanese pickles. *Bioscience, Biotechnology, and Biochemistry*, 76(11), 2061-2067. <http://dx.doi.org/10.1271/bbb.120414>. PMID:23132566.
- Matsui, H., Tsuchiya, R., Isobe, Y., & Narita, M. (2013). Analysis of bacterial community structure in Saba-Narezushi (Narezushi of Mackerel) by 16S rRNA gene clone library. *Journal of Food Science and Technology*, 50(4), 791-796. <http://dx.doi.org/10.1007/s13197-011-0382-4>. PMID:24425983.
- Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J., & Edwards, R. A. (2008). The metagenomics RAST server: a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatics*, 9(1), 386. <http://dx.doi.org/10.1186/1471-2105-9-386>. PMID:18803844.
- Nalbantoglu, U., Cakar, A., Dogan, H., Abaci, N., Ustek, D., Sayood, K., & Can, H. (2014). Metagenomic analysis of the microbial community in kefir grains. *Food Microbiology*, 41, 42-51. <http://dx.doi.org/10.1016/j.fm.2014.01.014>. PMID:24750812.
- Namiki, T., Hachiya, T., Tanaka, H., & Sakakibara, Y. (2012). MetaVelvet: an extension of Velvet assembler to de novo metagenome assembly from short sequence reads. *Nucleic Acids Research*, 40(20), e155. <http://dx.doi.org/10.1093/nar/gks678>. PMID:22821567.
- Osawa, K., Shigemura, K., Shimizu, R., Kato, A., Kusuki, M., Jikimoto, T., Nakamura, T., Yoshida, H., Arakawa, S., Fujisawa, M., & Shirakawa, T. (2015). Molecular characteristics of extended-spectrum beta-lactamase-producing *Escherichia coli* in a university teaching hospital. *Microbial Drug Resistance*, 21(2), 130-139. <http://dx.doi.org/10.1089/mdr.2014.0083>. PMID:25361040.
- Overbeek, R., Begley, T., Butler, R. M., Choudhuri, J. V., Chuang, H. Y., Cohoon, M., de Crécy-Lagard, V., Diaz, N., Disz, T., Edwards, R., Fonstein, M., Frank, E. D., Gerdes, S., Glass, E. M., Goesmann, A., Hanson, A., Iwata-Reuyl, D., Jensen, R., Jamshidi, N., Krause, L., Kubal, M., Larsen, N., Linke, B., McHardy, A. C., Meyer, F., Neuweger, H., Olsen, G., Olson, R., Osterman, A., Portnoy, V., Pusch, G. D., Rodionov, D. A., Rückert, C., Steiner, J., Stevens, R., Thiele, I., Vassieva, O., Ye, Y., Zagnitko, O., & Vonstein, V. (2005). The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Research*, 33(17), 5691-5702. <http://dx.doi.org/10.1093/nar/gki866>. PMID:16214803.
- Polka, J., Rebecchi, A., Pisacane, V., Morelli, L., & Puglisi, E. (2015). Bacterial diversity in typical Italian salami at different ripening stages as revealed by high-throughput sequencing of 16S rRNA amplicons. *Food Microbiology*, 46, 342-356. <http://dx.doi.org/10.1016/j.fm.2014.08.023>. PMID:25475305.
- Polz, M. F., & Cavanaugh, C. M. (1998). Bias in template-to-product ratios in multitemplate PCR. *Applied and Environmental Microbiology*, 64(10), 3724-3730. PMID:9758791.
- Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W. G., Peplies, J., & Glockner, F. O. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Research*, 35(21), 7188-7196. <http://dx.doi.org/10.1093/nar/gkm864>. PMID:17947321.
- Sanjukta, S., & Rai, A. K. (2016). Production of bioactive peptides during soybean fermentation and their potential health benefits. *Trends in Food Science & Technology*, 50, 1-10. <http://dx.doi.org/10.1016/j.tifs.2016.01.010>.
- Schmidt, T., DeLong, E., & Pace, N. (1991). Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *Journal of Bacteriology*, 173(14), 4371-4378. <http://dx.doi.org/10.1128/jb.173.14.4371-4378.1991>. PMID:2066334.
- Van Hijum, S., Vaughan, E., & Vogel, R. (2013). Application of state-of-art sequencing technologies to indigenous food fermentations. *Current Opinion in Biotechnology*, 24(2), 178-186. <http://dx.doi.org/10.1016/j.copbio.2012.08.004>. PMID:22960050.
- Wang, H., Zhang, S., Sun, Y., & Dai, Y. (2013). ACE-inhibitory peptide isolated from fermented soybean meal as functional food. *International Journal of Food Engineering*, 9(1), 1-7. <http://dx.doi.org/10.1515/ijfe-2012-0207>.
- Wei, Q.-K., Chen, T.-R., & Chen, J.-T. (2008). Use of *Bacillus subtilis* to enrich isoflavone aglycones in fermented natto. *Journal of the Science of Food and Agriculture*, 88(6), 1007-1011. <http://dx.doi.org/10.1002/jsfa.3181>.
- Wilke, A., Harrison, T., Wilkening, J., Field, D., Glass, E. M., Kyrpides, N., Mavrommatis, K., & Meyer, F. (2012). The M5nr: a novel non-redundant database containing protein sequences and annotations from multiple sources and associated tools. *BMC Bioinformatics*, 13(1), 141. <http://dx.doi.org/10.1186/1471-2105-13-141>. PMID:22720753.
- Wong, J., Hao, J., Cao, Z., Qiao, M., Xu, H., Bai, Y., & Ng, T. (2008). An antifungal protein from *Bacillus amyloliquefaciens*. *Journal of Applied Microbiology*, 105(6), 1888-1898. <http://dx.doi.org/10.1111/j.1365-2672.2008.03917.x>. PMID:19120637.
- Xie, G., Wang, L., Gao, Q., Yu, W., Hong, X., Zhao, L., & Zou, H. (2013). Microbial community structure in fermentation process of Shaoxing rice wine by Illumina-based metagenomic sequencing. *Journal of the Science of Food and Agriculture*, 93(12), 3121-3125. <http://dx.doi.org/10.1002/jsfa.6058>. PMID:23553745.