# MGI's DNBSEQ Sequencing Platform Facilitates the Identification of Functional Strains from the Pit Mud of Traditional Luzhou-flavor Baijiu "Wuliangye"

Baijiu can hardly be brewed without the pit mud. It is the unique functional microbes and microbial flora in the pit mud that give the brewed baijiu its unique flavor. A research team isolated, for the first time, a strain that can produce multiple beneficial flavor components of baijiu from the mud in an ancient pit where Wuliangye was brewed in the early Ming Dynasty[1]. After completing the whole genome sequencing based on the DNBSEQ sequencing platform, the team identified it as a new strain of *Clostridium* (No. WLY-B-L2$^T$), and officially named it *Clostridium aromativorans*[1].

Recommended application: Industrial Microbiology
Recommended model: DNBSEQ-G400RS, DNBSEQ-G99RS

- **High data detection accuracy**

  DNBSEQ technology, featured by high accuracy, low repeat sequence rate, and low index hopping rate, provides a guarantee for obtaining high-quality whole genome sequencing data.

- **Relatively high sequencing throughput**

  The DNBSEO-G400RS or DNBSEO-G99RS has arelatively high sequencing throughput.generating 55~1440 GB or 8~96 Gb of data in each run, which meets the sequencing needs forlarge number of samples..

- **Automatic operation compatible**

  The automated extraction and library preparation equipment of MGI significantly save labor costs in high-through-put sequencing and improve processing efficiency.

## Background

Technological innovations have facilitated the application of massively parallel sequencing (MPS) in food microbiome research. Compared to traditional methods, MPS has higher sensitivity to enable the identification of non-dominant microbiome that are essential in the sample. MPS can be divided into the targeted sequencing and shotgun sequencing based on different sequencing workflows.

Targeted sequencing can be applied in taxonomy study and microbial community identification by amplifying or capturing 16S ribosomal DNA (rDNA) of microbes. Meanwhile, shotgun sequencing is an economic and efficient whole genome sequencing (WGS) solution for the DNA sample extracted from enriched and cultured strains, which facilitates subsequent comparative genomics research[2,3]. Global researchers have carried out a number of studies on the microbiomes in fermented food, such as cheese, pickles, fermented wines and cocoa beans by shotgun metagenomics[2].

Baijiu, a traditional distilled white spirit fermented with grains, has become a symbol of Chinese traditional culture. There is a saying that good baijiu depends on old pits. Luzhou-flavor is one of the four basic types of baijiu flavors in China. Solid-state fermentation in a mixed distill and order way in a mud pit is a typical feature that distinguishes Luzhou-flavor baijiu from other baijiu, bringing the style of a sweet and pit flavor. Wuliangye, the superior Luzhou-flavor baijiu in China, was first produced in Wuliangye's ancient fermentation pits in 1368 (the first year of Hongwu in the early Ming Dynasty), with its active brewing technique continuing to the present day. These pit muds provide nutrient carriers and media for key microbiome in the baijiu brewing process, while the metabolism of these microbiome in these pit muds injects the "pit flavor" for the baijiu. The identification of relevant microbes is important for exploring the source of the sweet and mellow flavor of traditional Luzhou-flavor baijiu[4,5]. However, limitations of technical approaches forbade us to keep a good grip on the functions and features of these microbes in the past[6].

Recently, the "Wuliangye" team completed the WGS sequencing on the mud of an ancient pit in the Wuliangye brewing workshop in the early Ming Dynasty based on MGI's unique DNBSEQ sequencing technique, and isolated a strain that produced a variety of flavor components for the first time. The strain, WLY-B-L2$^T$, was identified as a new strain of *Clostridium*, and was named *Clostridium aromativorans*. This discovery is important to further understand the functional microbes in old pit muds and study the high quality and yield of the tradi-tional Wuliangye baijiu and the source of its sweet and mellow flavor.

# Study Description

To explore how the microbes in the pit mud affect the flavor of the Luzhou-flavor baijiu Wuliangye, the "Wuliangye" research team isolated an anaerobic, Gram-positive strain from the mud of an ancient pit in the Wuliangye brewing workshop in the early Ming Dynasty in Yibin, Sichuan Province, and numbered it as WLY-B-L2$^T$. Based on the DNBSEQ sequencing platform, the team completed the WGS sequencing of the strain, and identified it as a new *Clostridium* strain through subsequent analyses. In addition, further physiological and metabolic analyses enabled the discovery of the strain's function of producing important flavor components such as butyric acid, ethyl valerate, ethyl acetate and 2-pentanone. The strain was officially named *Clostridium aromativorans*. The DNBSEQ sequencing platform provides a powerful tool for the identification and analysis of microbes in the food industry.

# Materials and Methods

## Sample collection

The pit mud sample collected from the 501# ancient pit in the Wuliangye brewing workshop in the early Ming Dynasty was anaerobically cultured for 5 days. The WLY-B-L2$^T$ strain was isolated by gradient dilution and streak plate technique, and single colonies were selected, purified and stored. Afterwards, DNA was extracted using the DNA extraction kit from Sangon Biotech. In the study, the closely related *C.luticellarii* FW 431$^T$ strain was selected for the comparative analysis of physiological and chemical features.

## Library preparation and sequencing

For the extracted DNA, the 16S rDNA was amplified using the universal primer pair 27F (5'-AGA GTT TGA TCC TGG CTCAG-3') and 1492R (5'-GGT TAC CTTGTT ACG ACTT-3'), the amplified product was then inserted into the pEASY-T1 cloning vector and subsequently transfected into DH5α *Escherichia coli* . Finally, the M13 universal primer was used for Sanger sequencing to identify the gene sequence of the strain.

To analyze WLY-B-L2$^T$ comprehensively, the study also prepared a WGS library for the strain, and completed whole genome sequencing on the DNBSEQ sequencing platform.

## Bioinformatic analysis

The 16S rDNA Sanger sequence was used to generate phylogenetic tree using MEGA (v7.0). The genetic relationship between this sequence and the genome of other similar species was analyzed on the EzTaxon server by neighbor-joining method, with a total of 17 selected species. The taxonomical evolutionary distance between species was analyzed by interior-branch and bootstrap methods. The genetic relationship between the 16S rDNA gene sequence of the strain WLY-B-L2$^T$ and its closely related species was explored. Blast was conducted on the sequence obtained by WGS sequencing on GeneBank. On the GGDC website, Formula2 was used to calculate the DNA hybridization value obtained based on genome information. Glimmer3.02 software was used to complete gene annotation; KEGG (Kyoto Encyclopedia of Genes and Genomes), COG (Cluster of Orthologous Groups) and GO (Gene Ontology) databases were used to complete gene function analysis and core gene prediction.

| Sample collection | Library preparation and sequencing | Bioinformatic analysis | Result analysis |
|---|---|---|---|
| Collect the mud in Wuliangye' s ancient pit, select single colonies and extract DNA |  DNBSEQ sequencing platform | MEGA;<br>GGDC website;<br>Glimmer3.02;<br>KEGG, COG and GO databases | Phylogenetic tree construction;<br>Genome distance calculation;<br>Gene annotation;<br>Core gene prediction;<br>Gene function analysis |

# Results

## Morphologic observation and biochemical metabolic characteristic analysis

This team observed the morphology and appearance of the isolated and cultured strain, and compared its growth conditions with those of *C. luticellarii* FW 431$^T$. The results showed that the strain was strictly anaerobic, Gram-positive, and straight or slightly curved rod-shaped cells with 0.5-0.7 μm width and 1.7-3.1 μm length were arranged alone or in pairs. The colony was round, pale yellow, bulged and opaque, with irregular edges. It grew at a temperature of 15~45°C, pH 5~8, with a NaCl concentration of 0~3% and an ethanol concentration of 0~6%. It could produce butyric acid, ethyl valerate, ethyl acetate, 2-pentanone and other important flavor components in baijiu, of which ethyl acetate and ethyl valerate had typical fruit flavors similar to the aroma of banana and apple and were important skeleton flavor components of Luzhou-flavor baijiu.

## Genetic relationship and genome analysis

The team further conducted the genome comparative analysis, demonstrating that the 16s rDNA sequence of the strain was 97.42% similar to the genome of *C. luticellarii* FW 431$^T$, and its genomic similarity to other *Clostridium* strains was less than 98.65%. According to the analysis of Ribosomal Database Project (RDP) (http://rdp.cme.msu.edu/classifier/classifier.jsp), WLY-B-L2$^T$ was a member of *Clostridium*, a new branch different from *C. luticellarii* FW431$^T$ (Figure 1).

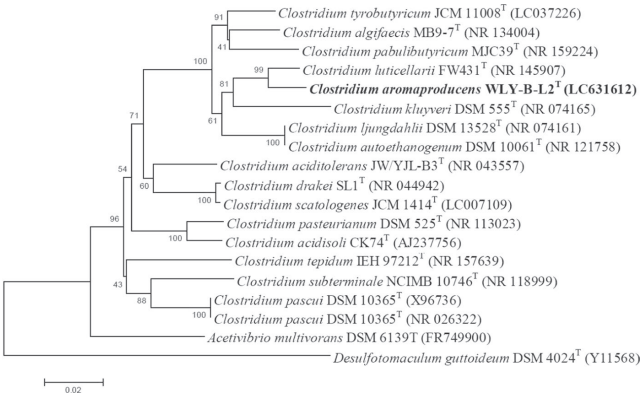


Figure 1. Phylogenomic tree shows the phylogenetic position of WLY-B-L2$^T$.

To explore the genomic information and function of the new strain (Figure 2), the whole genome sequencing analysis of WLY-B-L2$^T$ showed that its G+C% was 34.16 mol%, while that of *C. luticellarii* FW431$^T$ was 44.4 mol%, and the dDDH of WLY-B-L2$^T$ and *C. luticellarii* FW431$^T$ (NR_145907) was 28.1%, which was lower than the lower threshold of the distance between species. Therefore, this can further demonstrate that the WLY-B-L2$^T$ strain is a new strain.

## Function prediction and metabolic pathway analysis

The further core gene function prediction analysis showed that there were 218 genes involved in carbohydrate metabolism, 218 genes involved in amino acid metabolism, 154 genes involved in cofactor and vitamin metabolism, 159 genes involved in energy metabolism, and 87 genes involved in membrane transport on the WLY-B-L2$^T$ genome. The COG analysis showed that 156 genes were involved in unknown functions, and COG categories annotated most to metabolism included amino acid transport and metabolism (303 genes), inorganic ion transport and metabolism (172 genes), and coenzyme transport and metabolism (173 genes).

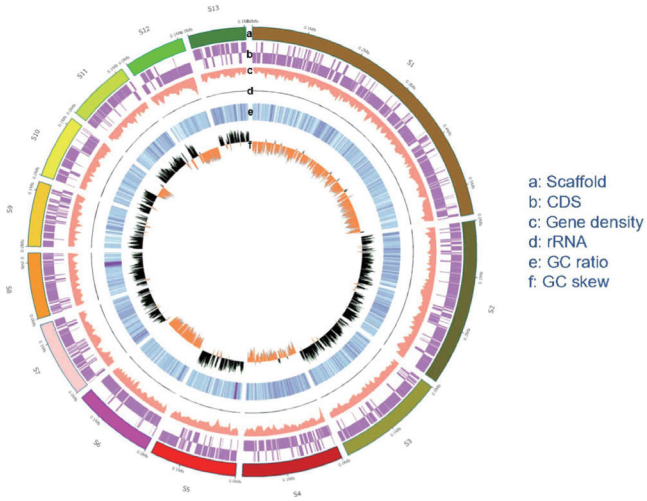


Figure 2. Genomic map of strain WLY-B-L2$^T$. The 16S rRNA gene sequence number of strain WLY-B-L2$^T$ uploaded to GenBank is LC631612, and the DDBJ/ENA/GenBank ID of the sketch of the genome is JAJJPB000000000.

## Summary

Whole genome sequencing was conducted on the strain isolated from the pit mud based on the DNBSEQ sequencing platform. The unique DNBSEQ technique, combined with Patterned Array, cPAS (combinatorial Probe Anchor Synthesis) and other techniques, has the advantages of low index hopping and low error accumulation. Its high accuracy provides a guarantee for obtaining high-quality whole genome data of the strain and a basis for the subsequent functional gene research.

The MGI DNBSEQ sequencing platform provides a powerful tool for the identification of microbes in the food industry, facilitating the research on the source of the flavor of traditional Luzhou-flavor baijiu.


Genetic Sequencer DNBSEQ-G400RS


Genetic Sequencer DNBSEQ-G99RS

## References

1. Luo, Q., Zheng, J., Zhao, D. & Liu, D., Clostridium aromativorans sp. nov., isolated from pit mud used for producing Wuliangye baijiu. *ANTON LEEUW INT J G* 116 739 (2023).
2. De Filippis F, Parente E, Ercolini D. Recent Past, Present, and Future of the Food Microbiome. *Annu Rev Food Sci Technol*. 2018 Mar 25;9:589-608. doi: 10.1146/annurev-food-030117-012312. Epub 2018 Jan 18. PMID: 29350561.
3. Yao Z, Zhu Y, Wu Q, Xu Y. Challenges and perspectives of quantitative microbiome profiling in food fermentations. *Crit Rev Food Sci Nutr*. 2024;64(15):4995-5015. doi: 10.1080/10408398.2022.2147899. Epub 2022 Nov 22. PMID: 36412251.
4. 蔡鹏飞,邵传贞,姚庆乐, 等.窖泥在酿酒中的两个作用[J].酿酒, 2011,38(5):59-61.
5. 梁欢, 许长峰, 朱立宁, 等.浓香型白酒地域性及窖泥微生物研究进展[J].中国酿造, 2020, 39(11): 5-8.
6. 骆红波, 孙优兰, 范奇高, 等.组学技术在白酒酿造微生物中的应用进展[J].中国酿造, 2023, 42(12): 13-21.

## Recommended Ordering Information

| Category | Product | Cat. NO. |
|---|---|---|
| Instruments | DNBSEQ-G400RS Genetic Sequencer | 900-000170-00 |
| | DNBSEQ-G99ARS Genetic Sequencer | 900-000609-00 |
| | MGISP-100RS Automated Sample Preparation System | 900-000206-00 |
| | MGISP-960RS Automated Sample Preparation System | 900-000146-00 |
| Software | MegaBOLT Bioinformatics Analysis accelerator (workstation server) | 900-000555-00 |
| Library Prep Reagents | MGIEasy FS DNA Library Prep Set V2.1 (16 RXN) | 1000006987 |
| Sequencing Reagents | DNBSEQ-G400RS High-throughput Sequencing Set (FCL PE100) | 1000016950 |
| | DNBSEQ-G99RS High-throughput Sequencing Set (G99 SM FCL PE150) | 940-000410-00 |


# MGI Tech Co.,Ltd

Building 11, Beishan Industrial Zone, Yantian District, Shenzhen, CHINA, 518083





Version: August 2024



+86-4000-688-114

en.mgi-tech.com

MGI-service@mgi-tech.com



Authors: Tongtong Ge

Editor-in-Charge: Qiwei Wang

Reviewer: Yao Jiang