

研究报告

基于比较基因组学解析耐酸乳杆菌 G10 的多碳源利用特征

林麟, 杜如冰, 吴群*, 徐岩

江南大学生物工程学院 工业生物技术教育部重点实验室 酿造微生物学与应用酶学研究中心,
江苏 无锡 214122

林麟, 杜如冰, 吴群, 徐岩. 基于比较基因组学解析耐酸乳杆菌 G10 的多碳源利用特征[J]. 微生物学通报, 2022, 49(8): 3279-3292

Lin Lin, Du Rubing, Wu Qun, Xu Yan. Comparative genomics-based analysis of *Lactobacillus acetotolerans* G10, a strain using multiple carbon sources[J]. Microbiology China, 2022, 49(8): 3279-3292

摘要:【背景】耐酸乳杆菌(*Lactobacillus acetotolerans*)是白酒发酵过程中的优势乳酸菌,对白酒发酵具有重要作用。*L. acetotolerans* G10 是分离自芝麻香型白酒发酵酒醅的一株能够利用多种碳源的菌株。【目的】基于全基因组测序,解析菌株 G10 多碳源利用机制。【方法】通过三代测序平台 Oxford Nanopore 完成菌株 G10 全基因组测序,分别利用 Circlator 和 Prodigal 对测序数据进行组装和基因预测;通过细菌基因组分析工具(bacterial pan genome analysis tool, BPGA)进行泛基因组分析。【结果】G10 能够利用 22 种糖类及糖类衍生物,其全基因组大小为 1 627 828 bp,含有 1 878 个编码基因;基于 Koyto Encyclopedia of Genes and Genomes (KEGG)数据库注释获得 292 个碳源代谢相关基因,基于 Carbohydrate-Active Enzymes (CAZy)数据库注释获得 44 个 CAZy 家族的编码基因。与其他发酵食品来源的耐酸乳杆菌相比, G10 基因组最小,但其总基因数量以及淀粉和蔗糖代谢相关基因数量均最多且含有 426 个独有基因;与发酵食品中植物乳杆菌 ATCC 14917^T、发酵乳杆菌 ATCC 14931^T、干酪乳杆菌 ATCC 393^T、短乳杆菌 ATCC 14869^T和布氏乳杆菌 ATCC 4005^T等其他主要代表性乳酸菌相比, G10 基因组最小,碳水化合物代谢相关基因数占比最高,而且具有 *glvA*、*malP* 和 *glvC* 等独有基因。【结论】*L. acetotolerans* G10 底物选择范围广,可利用多种碳源,能够适应碳源种类多变的发酵环境,对 *L. acetotolerans* G10 基因组信息的分析为进一步阐明 *L. acetotolerans* 的发酵性能提供了遗传学基础。

关键词: 耐酸乳杆菌; 碳源代谢; 全基因组测序; 比较基因组分析

基金项目: 国家自然科学基金(32172175)

Supported by: National Natural Science Foundation of China (32172175)

*Corresponding author: E-mail: wuq@jiangnan.edu.cn

Received: 2021-12-26; Accepted: 2022-03-30; Published online: 2022-06-02

Comparative genomics-based analysis of *Lactobacillus acetotolerans* G10, a strain using multiple carbon sources

LIN Lin, DU Rubing, WU Qun*, XU Yan

Laboratory of Brewing Microbiology and Applied Enzymology, Key Laboratory of Industrial Biotechnology, Ministry of Education; School of Biotechnology, Jiangnan University, Wuxi 214122, Jiangsu, China

Abstract: [Background] *Lactobacillus acetotolerans* is the dominant lactic acid bacteria species and plays an important role in Chinese liquor fermentation. *L. acetotolerans* G10, which was isolated from the fermented grains of sesame-flavor liquor, utilizes multiple carbon sources. [Objective] To analyze the mechanism for the multiple carbon source utilization of G10 based on whole genome sequencing. [Methods] The whole genome of G10 was sequenced by Oxford Nanopore Technologies, a third-generation platform for the sequencing of native DNA strands. Circlator was employed to circularize genome assemblies and Prodigal to predict genes and annotate protein-coding genes. bacterial pan genome analysis tool (BPGA) was used for pan-genome analysis. [Results] G10 was able to utilize 22 sugars and their derivatives. The genome size of G10 was 1 627 828 bp with 1 878 coding genes. G10 contained 292 genes related to carbohydrate metabolism, as annotated by Koyto Encyclopedia of Genes and Genomes (KEGG) annotation, and 44 genes related to Carbohydrate-Active EnZymes (CAZy) according to the CAZy annotation. Compared with other *L. acetotolerans* from food fermentations, G10 had the smallest genome with the highest numbers of total genes and genes related to starch and sucrose metabolism, and contained 426 unique genes. Compared with *Lactiplantibacillus plantarum* subsp. *plantarum* ATCC 14917^T, *Limosilactobacillus fermentum* ATCC 14931^T, *Lactocaseibacillus casei* ATCC 393^T, *Levilactobacillus brevis* ATCC 14869^T and *Lentilactobacillus buchneri* ATCC 4005^T from food fermentations. The genome of G10 was the smallest. The number of genes related to carbohydrate metabolism accounted for the highest proportion of total genes in G10. It had unique genes such as *glvA*, *malP* and *glvC*. [Conclusion] G10 can use a variety of carbon sources and adapt to various fermentation environments. The analysis of genomic information lays a genetic basis for further illustration of the fermentation performance of *L. acetotolerans*.

Keywords: *Lactobacillus acetotolerans*; carbon source metabolism; whole genome sequencing; comparative genomics-based analysis

耐酸乳杆菌(*Lactobacillus acetotolerans*)属于乳杆菌属,最早分离自米醋发酵液,能够耐受高浓度醋酸^[1]。因其耐酸、耐低氧及耐乙醇等特征,*L. acetotolerans*广泛存在于白酒^[2-4]、食醋^[5]、发酵蔬菜^[6-7]和酸面团^[8-10]等食品的发酵过程,与发酵食品的风味显著相关^[11]。

中国白酒是以谷物为原料通过多种微生物

边糖化边发酵制得的酿造食品^[12]。微生物通过分泌多种水解酶降解淀粉、纤维素和半纤维素等多糖,使发酵环境中存在麦芽糖、葡萄糖、果糖和阿拉伯糖等多种碳源^[13]。多碳源的发酵环境对于微生物生长具有影响,Liu等研究表明中国白酒发酵过程中的糖类物质能够影响微生物的群落结构^[14]。Lu等研究发现白酒发酵系统来源

的酿酒酵母 MT1 具有多碳源利用特征, 说明白酒发酵系统中特定微生物能够适应多碳源的发酵环境^[15]。*L. acetotolerans* 在中国白酒发酵中后期处于优势地位, 但对其碳源利用能力及其机制尚不明晰, 因此有必要对白酒发酵系统中 *L. acetotolerans* 的碳源利用能力及其基因组特征进行解析。

本研究从芝麻香型白酒酒醅中分离筛选 *L. acetotolerans* G10, 对菌株的碳源利用能力进行分析, 并利用三代测序技术对其全基因组进行测序, 同时对菌株碳源利用相关的关键基因进行挖掘, 以期为该菌株的进一步研究和应用提供参考。

1 材料与方法

1.1 主要试剂和仪器

MRS 培养基和快速测序试剂盒, Oxoid 公司; Biolog Gen III 微孔板和 IF-C 接种液, Biolog 公司; 细菌基因组 DNA 提取试剂盒, 生工生物工程(上海)股份有限公司。生化培养箱, 上海博迅实业有限公司; 荧光酶标仪, BioTek 公司; PCR 仪, Bio-Rad 公司; 核酸浓度检测仪, 赛默飞世尔科技(中国)有限公司。

1.2 菌株

L. acetotolerans G10 菌株于 2019 年 12 月从芝麻香型白酒酒醅中分离获得, 本实验室保存。

1.3 Biolog Gen III 板测定碳源利用能力

利用 Biolog Gen III 微孔板对 *L. acetotolerans* G10 的碳源代谢能力进行测定, 3 个生物学平行。实验操作过程: 将 *L. acetotolerans* G10 接种在 MRS 液体培养基中, 于 30 °C 厌氧条件下静置培养 3 d, 然后 9 000 r/min 离心 10 min, 在无菌条件下除去培养基, 再加入 5 mL 的无菌超纯水重悬, 将菌悬液转接至接种液 IF-C 中, 配成细

胞浓度为 90%–98% 的菌悬液。将菌悬液加到微孔板上, 每孔加入 100 μ L。将微孔板置于 30 °C 厌氧条件下静置培养 48 h。根据 OD_{595} 值确定 *L. acetotolerans* G10 碳源利用能力的大小。

1.4 菌株基因组 DNA 的提取和测序

利用快速细菌基因组 DNA 提取试剂盒提取菌株 G10 的基因组 DNA, 然后使用快速测序试剂盒制备测序库。通过三代测序平台 Oxford Nanopore 完成 *L. acetotolerans* G10 的全基因组测序。

1.5 比较基因组学分析

利用 Filtlong V0.2.0 根据大小和质量选择总读取的子集。用 Canu 将测序序列重新组装^[16]。基因组使用 Medaka 进行优化, 并用 Circlator-1.4.1 进行环化。*L. acetotolerans* G10 基因组序列已提交至 DNA data bank of Japan (DDBJ) 数据库, 序列号为 DRA013252。利用 Prodigal 进行基因预测并获得蛋白编码基因(protein-coding genes, CDS)^[17]。利用 BlastKOALA 对 *L. acetotolerans* G10 进行功能注释^[18]。使用 carbohydrate-active enzymes (CAZy) 数据库对细菌基因组进行快速注释, 并得到相应的注释信息。使用细菌泛基因组分析工具(bacterial pan genome analysis tool, BPGA)进行基因组比较分析^[19]。

用于比较基因组学分析的菌株及其在 NCBI 数据库中的基因组登录号分别为: *L. acetotolerans* NBRC 13120 (AP014808.1)、*L. acetotolerans* LA749 (CP044496.1)、*L. acetotolerans* CN247 (CP046528.1) 和 *L. acetotolerans* LJ49 (CP051649.1), 以及模式菌株植物乳杆菌 ATCC 14917^T (*Lactiplantibacillus plantarum* subsp. *plantarum* ATCC 14917^T, GCA_000143745.1)、发酵乳杆菌 ATCC 14931^T (*Limosilactobacillus fermentum* ATCC 14931^T, GCA_000159215.1)、干酪乳杆菌

ATCC 393^T (*Lactocaseibacillus casei* ATCC 393^T, GCA_000829055.1)、短乳杆菌 ATCC 14869^T (*Levilactobacillus brevis* ATCC 14869^T, GCA_000469365.1)和布氏乳杆菌 ATCC 4005^T (*Lentilactobacillus buchneri* ATCC 4005^T, GCA_004354745.1)。

2 结果与分析

2.1 *L. acetotolerans* G10 的碳源利用能力分析

利用 Biolog Gen III 微孔板分析 *L. acetotolerans* G10 的碳源利用能力。结果表明, *L. acetotolerans* G10 可利用的碳源物质有 67 种($OD_{595}>0.02$), 包括 22 种糖类及其衍生物、16 种酸类物质、5 种醇类物质、9 种氨基酸和 15 种其他碳源物质。其中, *L. acetotolerans* G10 可高效利用的碳源物质有 37 种($OD_{595}>0.1$), 表明该菌能够利用多种碳源物质。

L. acetotolerans G10 能够利用的糖类物质包括 7 种单糖、9 种二糖、1 种三糖、3 种低聚糖及 2 种磷酸己糖($OD_{595}>0.02$) (图 1A)。其中, *L. acetotolerans* G10 可高效利用 15 种糖类物质, 包括 D-果糖、D-甘露糖、D-麦芽糖、D-海藻糖、D-纤维二糖、龙胆二糖、蔗糖、D-松二糖、 α -D-乳糖、蜜二糖、D-水杨苷、水苏糖、糊精、棉子糖和 6-磷酸葡萄糖($OD_{595}>0.1$), 约占可高效利用碳源的 40.5%。

在以酸类物质为唯一碳源时, *L. acetotolerans* G10 可高效利用 6 种酸类物质, 包括 D-葡萄糖酸、奎宁酸、L-乳酸、柠檬酸、L-苹果酸和溴丁二酸($OD_{595}>0.1$) (图 1B), 分别参与磷酸戊糖途径、氨基酸生物合成、糖酵解与糖异生和柠檬酸循环等代谢途径。*L. acetotolerans* G10 能够高效利用 4 种醇类物质($OD_{595}>0.1$) (图 1C), 其中 D-山梨醇、D-甘露醇为果糖和甘露糖代谢途径中的

重要物质。除 L-丝氨酸外, *L. acetotolerans* G10 能够高效利用其余 8 种氨基酸($OD_{595}>0.1$), 表明其对氨基酸也具有较强的利用能力(图 1D)。同时, 该菌可高效利用的 N-乙酰-D-葡萄糖胺参与氨基糖和核苷酸糖的代谢(图 1E)。

L. acetotolerans G10 具有多碳源利用的能力, 而且其可高效利用的碳源能够参与多种代谢途径, 对于促进 *L. acetotolerans* G10 生长、提高 *L. acetotolerans* G10 适应能力具有重要意义。

2.2 *L. acetotolerans* G10 基因组信息

为分析 *L. acetotolerans* G10 多碳源利用特征的形成机制, 采用第 3 代 Oxford Nanopore 测序平台, 对 *L. acetotolerans* G10 的全基因组进行测序, 并完成了基因组圆形图谱(图 2)。*L. acetotolerans* G10 全基因组大小为 1 627 828 bp, 平均 GC 含量为 36.8%, 共编码基因 1 878 个。目前 NCBI 数据库中仅 4 株 *L. acetotolerans* 具有完整全基因组序列, 分别为 *L. acetotolerans* NBRC 13120 (分离自日本清酒)、*L. acetotolerans* LA749 (分离自食醋)、*L. acetotolerans* CN247 (分离自啤酒)和 *L. acetotolerans* LJ49 (分离自食醋)。将 *L. acetotolerans* G10 的全基因组信息与 NCBI 数据库中其他食品发酵系统来源的 *L. acetotolerans* 全基因组信息比较分析。如表 1 所示, *L. acetotolerans* G10 的基因组最小, 但 *L. acetotolerans* G10 的基因数却比其他 *L. acetotolerans* 多, 表明其代谢潜力较其他菌株更强, 在白酒发酵系统中能够更好地适应发酵环境, 提高转化效率。

2.3 *L. acetotolerans* G10 全基因组功能注释

将 *L. acetotolerans* G10 的全基因组信息与 Koyto Encyclopedia of Genes and Genomes (KEGG)数据库进行比对(图 3A), 共有 1 437 个基因在 KEGG 数据库中获得注释结果, 其中有

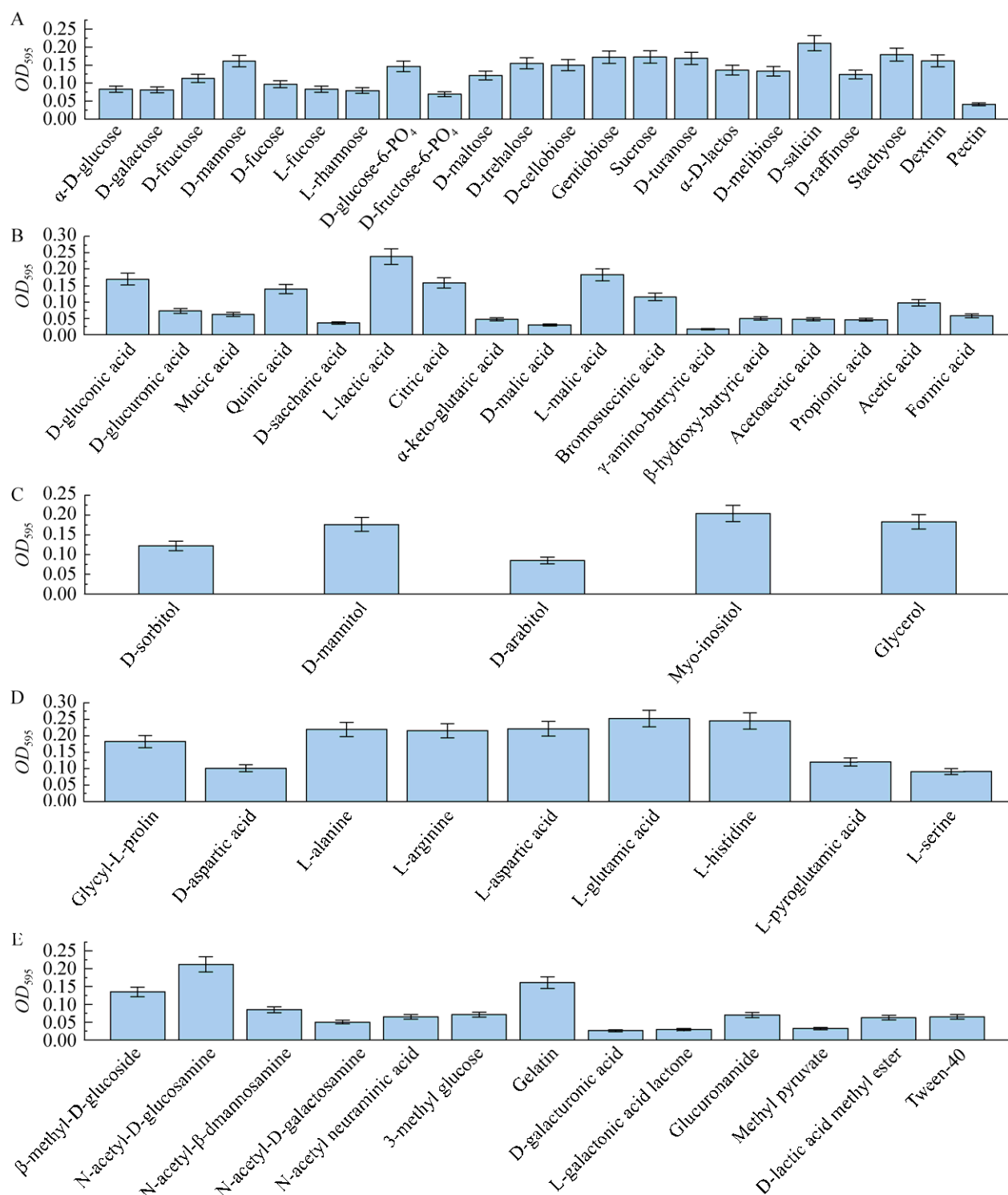


图 1 *L. acetotolerans* G10 在不同种类碳源上的生长情况($n=3$) A: 糖类及其衍生物; B: 酸类物质; C: 醇类物质; D: 氨基酸; E: 其他化合物

Figure 1 The growth of *L. acetotolerans* G10 with various carbon sources ($n=3$). A: Sugars and derivatives; B: Acids; C: Alcohols; D: Amino acids; E: Other compounds.

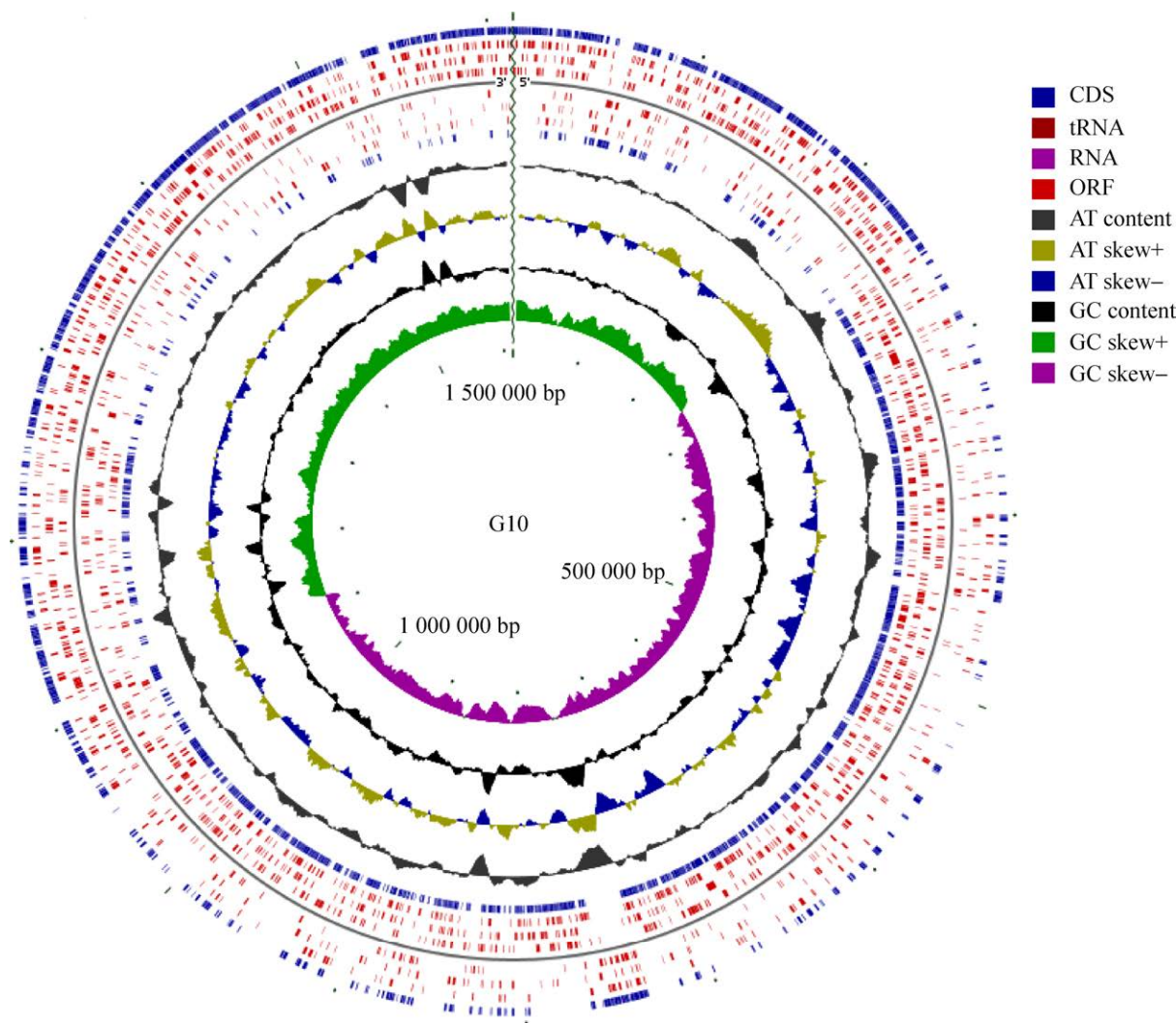


图 2 *L. acetotolerans* G10 基因组圆形图谱
Figure 2 Circular map of the *L. acetotolerans* G10.

表 1 *L. acetotolerans* G10 与其他 *L. acetotolerans* 的全基因组特征比较
Table 1 The genomic characteristics of the *L. acetotolerans* G10 genome and comparison with other *L. acetotolerans*

Strains	GenBank accession No.	Genome size (bp)	GC content (%)	Gene number
<i>L. acetotolerans</i> G10	DRA013252	1 627 828	36.8	1 878
<i>L. acetotolerans</i> NBRC 13120	AP014808.1	1 704 859	36.7	1 694
<i>L. acetotolerans</i> LA749	CP044496.1	1 683 905	36.5	1 643
<i>L. acetotolerans</i> CN247	CP046528.1	1 677 515	36.9	1 628
<i>L. acetotolerans</i> LJ49	CP051649.1	1 683 950	36.6	1 639

441 个基因未得到明确分类。*L. acetotolerans* G10 中已注释的功能基因分布在四类代谢途径中，包括细胞生长(62，4.3%)、环境信息处理(178，12.4%)、遗传信息处理(310，21.6%)和代谢(887，61.7%)。在代谢功能相关基因中碳水化合物代谢相关功能基因数量最多(存在 292 个基

因, 约占已注释功能基因总量的 20.3%)。

根据 CAZy 数据库对 *L. acetotolerans* G10 的基因组序列进行分析, 共有 44 个基因编码的蛋白质结构域属于 CAZy 家族, 包括碳水化合物结合模块(carbohydrate-binding modules, CBMs)、碳水化合物酯酶(carbohydrate esterases, CEs)、糖苷水解酶(glycoside hydrolases, GHs)和糖基转移酶(glycosyl transferases, GTs)。*L. acetotolerans* G10 有 26 个蛋白属于糖苷水解酶家族, 占总注释蛋白的 47.8%; 而且含有 13 个糖基转移酶、4 个碳水化合物结合模块和 1 个碳水化合物酯酶(图 3B)。*L. acetotolerans* G10 的全基因组功能注

释结果证实该菌株具有高效代谢碳源的潜能, 有助于其在白酒多碳源发酵环境中生长。

2.4 *L. acetotolerans* G10 碳源代谢途径构建

为明确 *L. acetotolerans* G10 全基因组特征与多碳源利用的关联, 以全基因组信息为基础, 对 *L. acetotolerans* G10 碳源代谢途径进行构建(图 4)。注释结果表明, *L. acetotolerans* G10 的碳源代谢基因主要分布在糖酵解与糖异生途径、淀粉和蔗糖代谢途径、氨基糖和核苷酸糖代谢及丙酮酸代谢途径, 所含基因数量分别为 46、54、42 和 37 个, 其中淀粉和蔗糖代谢途径中功能基因数量最多。

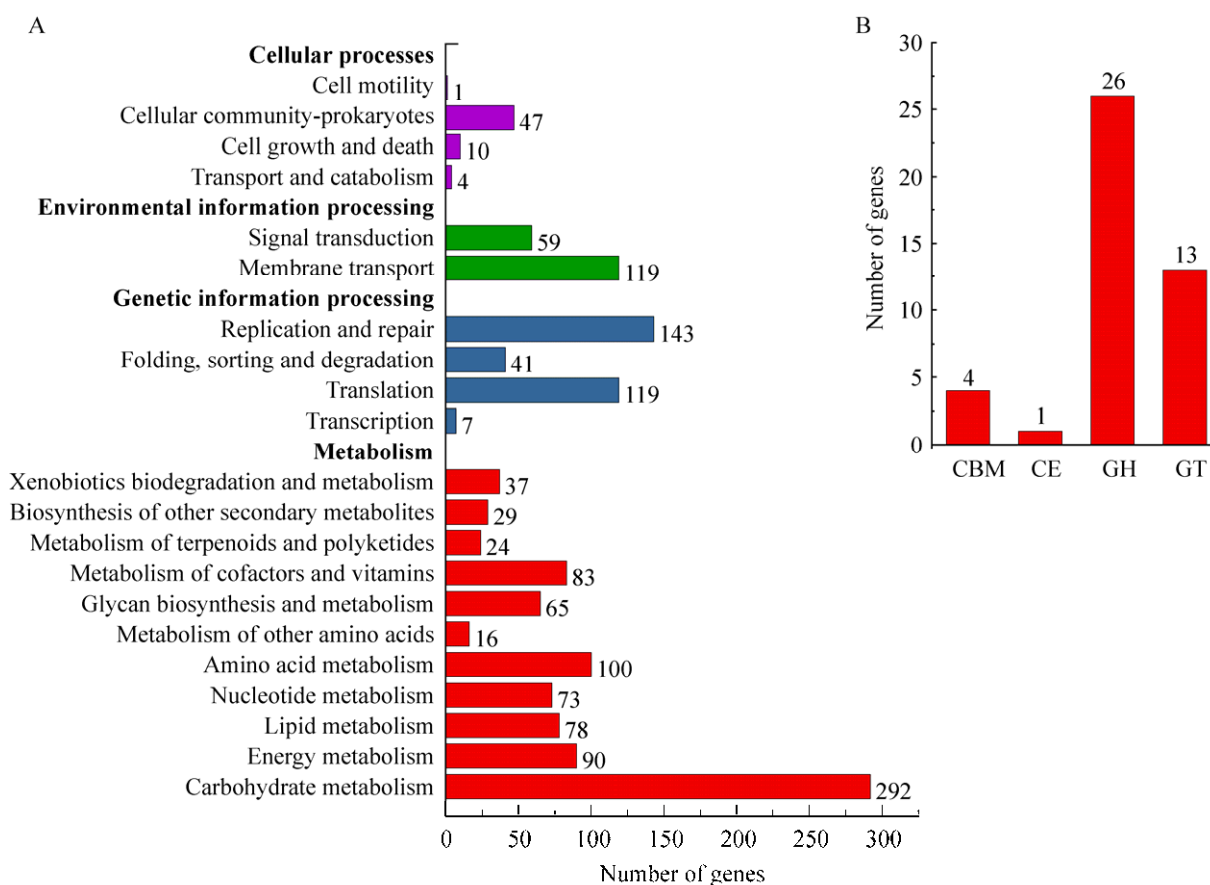


图 3 *L. acetotolerans* G10 基于 KEGG 功能注释(A)及 CAZy 功能注释(B) CBM: 碳水化合物结合模块; CE: 碳水化合物酯酶; GH: 糖苷水解酶; GT: 糖基转移酶

Figure 3 KEGG (A) and CAZy (B) functional classification map of *L. acetotolerans* G10. CBM: Carbohydrate-binding module; CE: Carbohydrate esterase; GH: Glycoside hydrolase; GT: Glycosyltransferase.

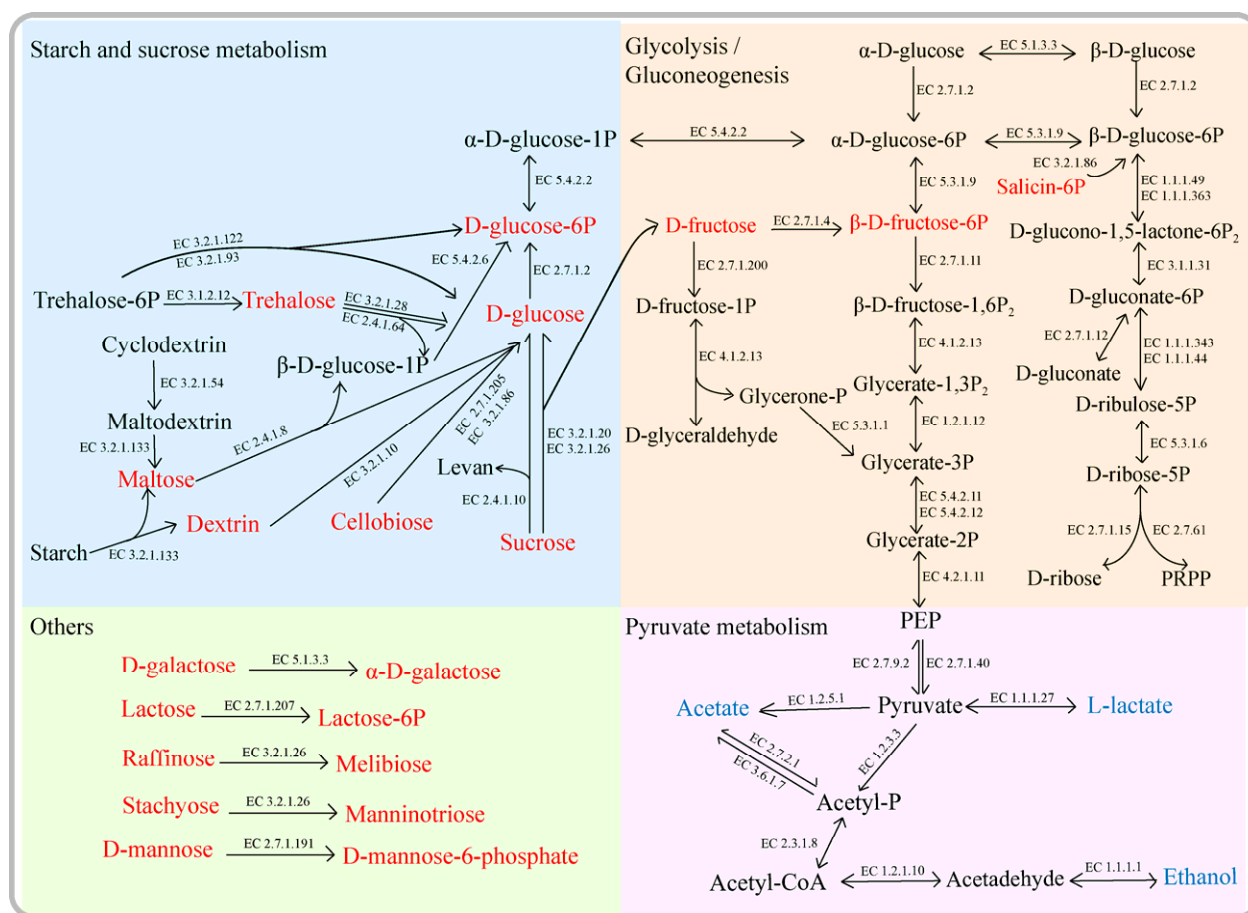


图 4 *L. acetotolerans* G10 碳源代谢途径预测

Figure 4 Prediction of carbohydrate metabolism pathway in *L. acetotolerans* G10.

在淀粉和蔗糖代谢途径中, *L. acetotolerans* G10 具有可以利用 D-葡萄糖、6-磷酸葡萄糖、D-海藻糖、D-麦芽糖、D-纤维二糖、蔗糖和糊精相关的基因, 这些糖类物质均为 *L. acetotolerans* G10 可高效利用的碳源($OD_{595} > 0.1$) (图 1A)。在糖酵解与糖异生途径中, *L. acetotolerans* G10 具有可利用 D-果糖、6-磷酸果糖和 6-磷酸水杨苷相关的基因。同时, D-果糖为 *L. acetotolerans* G10 可高效利用碳源之一($OD_{595} > 0.1$)。在其他碳源代谢途径中, *L. acetotolerans* G10 具有能够利用 D-甘露糖、α-D-乳糖、水苏糖和棉子糖相关的基因, 进一步证实了菌株具有代谢多碳源的潜能。

2.5 *L. acetotolerans* 碳源代谢潜能及泛基因组分析

利用 CAZy 数据库, 对 NCBI 数据库中下载的 4 株 *L. acetotolerans* 的全基因组序列进行注释, 并将注释结果与 *L. acetotolerans* G10 进行比较。结果如表 2 所示, *L. acetotolerans* G10 中能够编码 CBM 和 GH 相关基因的含量最高, 分别为 4 个和 26 个。

将 *L. acetotolerans* G10、*L. acetotolerans* NBRC 13120、*L. acetotolerans* LA749、*L. acetotolerans* CN247 和 *L. acetotolerans* LJ49 的全基因组信息与 KEGG 数据库进行比对(表 3)。KEGG 富集结果表明, 在糖酵解/糖异生通路

表 2 五株 *L. acetotolerans* 菌株的 CAZy 功能注释
Table 2 CAZy functional classification map of 5 *L. acetotolerans* strains

Strains	Gene number			
	CBM	CE	GH	GT
<i>L. acetotolerans</i> G10	4	1	26	13
<i>L. acetotolerans</i> NBRC 13120	1	2	22	19
<i>L. acetotolerans</i> LA749	1	2	23	20
<i>L. acetotolerans</i> CN247	1	1	20	15
<i>L. acetotolerans</i> LJ49	0	1	20	17

上 *L. acetotolerans* G10 有 31 种基因, 高于其他 4 株菌。同时, *L. acetotolerans* G10 比 *L. acetotolerans* LA749、*L. acetotolerans* CN247 和 *L. acetotolerans* LJ49 具有更多淀粉和蔗糖代谢相关基因, 这表

明 *L. acetotolerans* G10 在碳源利用方面具有更大潜能。

为分析不同发酵系统中 *L. acetotolerans* 菌株的差异, 利用 BPGA 对 5 株 *L. acetotolerans* 的全基因组序列进行比较分析。以氨基酸序列相似度大于 50% 为标准, 统计得到包含 1 913 个基因的泛基因组, 其中核心基因 1 188 个(图 5)。与其余 4 株菌相比, *L. acetotolerans* G10 含有 426 个独有基因。基于核心基因序列构建的系统发育树见图 5, 5 株 *L. acetotolerans* 主要分为 2 个分支, *L. acetotolerans* G10 与 *L. acetotolerans* CN247 遗传距离最近。

表 3 五株 *L. acetotolerans* 的碳水化合物代谢途径及其相关基因数量

Table 3 Carbohydrate metabolism pathways and gene number of 5 *L. acetotolerans* strains

KO No.	Pathway name	G10	NBRC 13120	LA749	CN247	LJ49
00010	Glycolysis/Gluconeogenesis	31	23	26	23	25
00020	Citrate cycle (TCA cycle)	2	2	2	2	3
00030	Pentose phosphate pathway	16	17	17	15	15
00040	Pentose and glucuronate interconversions	3	4	4	3	3
00051	Fructose and mannose metabolism	18	20	23	13	24
00052	Galactose metabolism	10	12	14	9	7
00500	Starch and sucrose metabolism	31	31	25	24	19
00520	Amino sugar and nucleotide sugar metabolism	26	19	18	19	20
00620	Pyruvate metabolism	26	23	21	22	24
00630	Glyoxylate and dicarboxylate metabolism	5	5	5	5	5
00640	Propanoate metabolism	11	13	10	10	11
00650	Butanoate metabolism	6	7	5	5	7
00562	Inositol phosphate metabolism	2	2	4	2	3

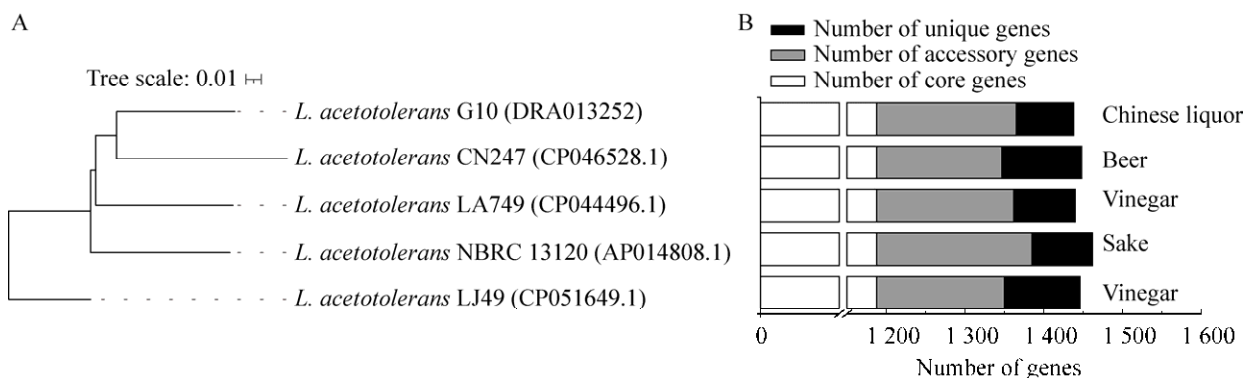


图 5 *L. acetotolerans* 基于核心基因的系统发育树(A)及泛基因组分析(B)

Figure 5 The phylogenetic tree based on core gene sequences of *L. acetotolerans* (A) and the analyze of panggenome (B).

基于 KEGG 数据库对全基因组的比较结果进行功能注释, 结果表明 5 株 *L. acetotolerans* 的差异基因主要是碳水化合物代谢和膜转运相关功能基因, 主要与糖酵解和糖质异生、半乳糖代谢、淀粉和蔗糖代谢、氨基酸糖和核苷酸糖代谢、糖胺聚糖降解及磷酸转移酶系统等相关(图 6)。

对 *L. acetotolerans* G10 含有的 426 个独有基因进行功能注释, 结果如表 4 所示, 完全独有基因中能够注释到功能基因的有 10 个, 分别为 *cps1C*、*glf*、*glvC*、*glvR*、*glvA*、*nagZ*、*ptbA*、*repA*、*rlmH* 和 *rgpA*, 主要与糖类物质转运、转录调控相关。*cps1C* 编码合成多糖转运体, *glvC* 编码合成 α -葡萄糖苷 PTS 系统的 EIICB 成分,

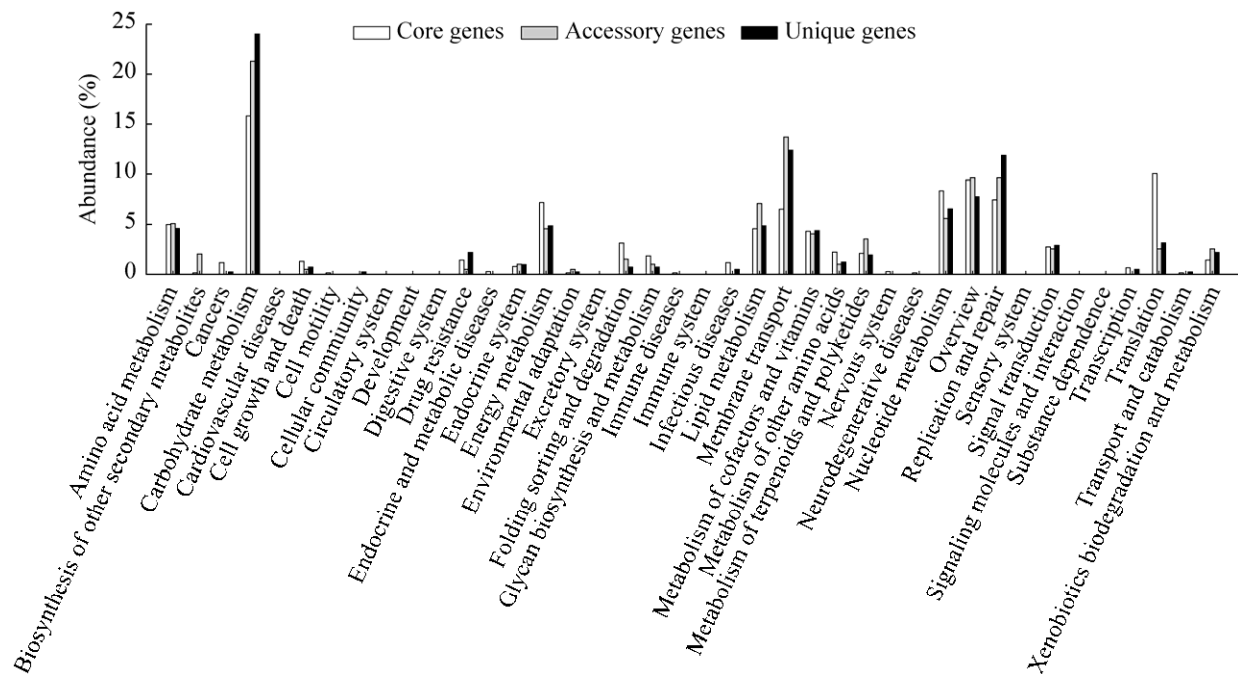


图 6 *L. acetotolerans* 基于泛基因组分析的 KEGG 功能注释图

Figure 6 KEGG functional classification map of *L. acetotolerans* G10 based on the analyze of core-pan genome.

表 4 *L. acetotolerans* G10 与其他 *L. acetotolerans* 菌株相比完全独有基因的 KEGG 功能注释

Table 4 KEGG annotations of unique genes in *L. acetotolerans* G10 compared with other *L. acetotolerans* strains

ID	Gene name	Description
K03328	<i>cps1C</i>	Polysaccharide transporter
K01854	<i>glf</i>	UDP-galactopyranose mutase
K02749	<i>glvC</i>	Alpha-glucoside PTS system EIICB component
K03481	<i>glvR</i>	RpiR family transcriptional regulator
K01232	<i>glvA</i>	Maltose-6'-phosphate glucosidase
K01207	<i>nagZ</i>	Beta-N-acetylhexosaminidase
K20108	<i>malT</i>	Maltose PTS system EIICB or EIICBA component
—	<i>repA</i>	Replication initiator protein A
K00783	<i>rlmH</i>	23S rRNA (pseudouridine1915-N3)-methyltransferase
K12996	<i>rgpA</i>	Rhamnosyltransferase

Note: —: No ID.

malT 编码合成麦芽糖 PTS 系统 EIICB 或 EIICBA 组分, 均能将环境中的大分子糖类物质运输到细胞内供菌株利用。这表明 *L. acetotolerans* G10 与其他 *L. acetotolerans* 菌株相比其碳源代谢能力更强。

2.6 发酵食品中常见乳酸菌碳源代谢潜能比较分析

乳酸菌在发酵食品中广泛分布, 其利用碳源的能力往往会影响其在发酵食品中的分布^[20]。发酵食品中常见的乳酸菌包括植物乳杆菌、发酵乳杆菌、干酪乳杆菌、短乳杆菌和布氏乳杆菌等^[21]。为分析 *L. acetotolerans* G10 与其他种类乳酸菌碳源代谢能力的差异, 选择了上述 5 种乳酸菌的模式菌株作为参照, 包括植物乳杆菌 ATCC 14917^T (*Lactiplantibacillus plantarum* subsp. *plantarum* ATCC 14917^T)、发酵乳杆菌 ATCC 14931^T (*Limosilactobacillus fermentum* ATCC 14931^T)、干酪乳杆菌 ATCC 393^T (*Lacticaeibacillus casei* ATCC 393^T)、短乳杆菌 ATCC 14869^T (*Levilactobacillus brevis* ATCC 14869^T) 和布氏乳杆菌 ATCC 4005^T (*Lentilactobacillus buchneri* ATCC 4005^T), 将 *L. acetotolerans* G10 与这些常见乳酸菌进行比较基因组分析(表 5)。

根据 KEGG 数据库的注释结果, *L. acetotolerans* G10 菌株基因组长度小于其他 5 株乳酸菌, 但 *L. acetotolerans* G10 中与碳水化合物代谢相

关基因数量仅次于 *L. plantarum* ATCC 14917^T (表 5); 而且 *L. acetotolerans* G10 中碳水化合物代谢相关基因数量占基因总数量的 15.5%, 高于 *L. plantarum* ATCC 14917^T (11.7%)、*L. fermentum* ATCC 14931^T (10.3%)、*L. casei* ATCC 393^T (11.1%)、*L. brevis* ATCC 14869^T (9.5%) 和 *L. buchneri* ATCC 4005^T (9.8%)。表明 *L. acetotolerans* G10 的基因组结构精简, 碳源代谢潜能高于其他 5 株乳酸菌。

KEGG 注释结果表明, *L. acetotolerans* G10 的碳水化合物代谢基因中与糖酵解/糖异生、淀粉和蔗糖代谢相关的基因数量均占碳水化合物代谢基因总量的 16.6%, 其次是氨基糖和核苷酸糖代谢(13.9%)、丙酮酸代谢(13.9%) (图 7)。*L. acetotolerans* G10 中糖酵解/糖异生、淀粉和蔗糖代谢及氨基糖和核苷酸糖代谢相关的基因数量占碳水化合物代谢基因总量的比例均高于其他 5 种乳酸菌。

与其他模式菌株相比, *L. acetotolerans* G10 有 17 种特有碳水化合物代谢基因在至少一种乳酸菌中缺失(表 6); 而且 *L. acetotolerans* G10 中含有 3 种完全独有基因 *glvA*、*malP* 和 *glvC*。*glvA* 编码麦芽糖-6'-磷酸葡萄糖苷酶, 可以水解多种含有 6-磷酸- α -D-葡萄糖苷的碳水化合物; *malP* 和 *glvC* 均能编码 α -葡萄糖苷 PTS 系统 EIICB 组

表 5 发酵食品中常见乳酸菌的碳水化合物代谢相关基因 KEGG 功能注释

Table 5 Carbohydrate metabolism genes of the common lactic acid bacteria in fermented food from KEGG database

Strains	Accession No.	Genome size (bp)	Number of genes	
			Total	Carbohydrate metabolism
<i>L. acetotolerans</i> G10	DRA013252	1 627 828	1 878	292
<i>L. plantarum</i> ATCC 14917 ^T	GCA_000143745.1	3 212 261	2 948	346
<i>L. fermentum</i> ATCC 14931 ^T	GCA_000159215.1	1 867 005	1 693	174
<i>L. casei</i> ATCC 393 ^T	GCA_000829055.1	2 952 961	2 514	278
<i>L. brevis</i> ATCC 14869 ^T	GCA_000469365.1	2 473 047	2 370	225
<i>L. buchneri</i> ATCC 4005 ^T	GCA_004354745.1	2 551 787	2 268	223

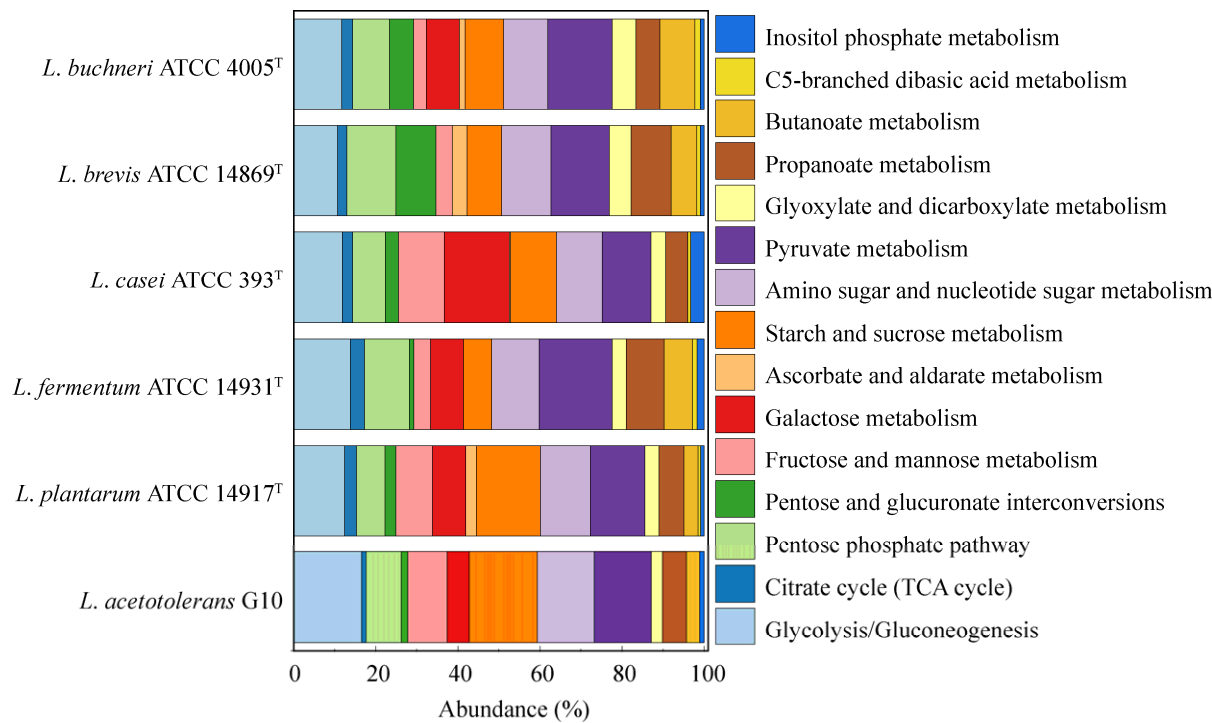


图 7 碳源代谢相关基因 KEGG 富集
Figure 7 The KEGG enrichment of carbohydrate metabolism genes.

表 6 *L. acetotolerans* G10 中特有碳水化合物代谢基因在其他乳酸菌中的存在情况
Table 6 The presence of carbohydrate metabolism genes of *L. acetotolerans* G10 in other lactic acid bacteria

Gene name	Description	G10	ATCC 14917 ^T	ATCC 14931 ^T	ATCC 393 ^T	ATCC 14869 ^T	ATCC 4005 ^T
<i>pgi</i>	Glucose-6-phosphate isomerase	+	+	+	—	+	+
<i>pfkA</i>	6-phosphofructokinase	+	+	—	+	—	—
<i>bglA</i>	6-phospho-β-glucosidase	+	+	—	+	+	—
<i>fbaA</i>	Fructose-bisphosphate aldolase	+	+	—	+	—	—
<i>ppsA</i>	Pyruvate, water dikinase	+	+	—	—	—	—
<i>sacA</i>	β-fructofuranosidase	+	+	+	+	—	—
<i>glvA</i>	Maltose-6'-phosphate glucosidase	+	—	—	—	—	—
<i>treC</i>	Trehalose-6-phosphate hydrolase	+	+	—	+	—	—
<i>treP</i>	α,α-trehalose phosphorylase	+	+	—	—	+	+
<i>malP, glvC</i>	α-glucoside PTS system EIICB component	+	—	—	—	—	—
<i>nplT</i>	Cyclomaltodextrinase	+	+	+	+	—	—
<i>fruA</i>	Fructose PTS system EIIA component	+	+	—	+	—	—
<i>fruK</i>	1-phosphofructokinase	+	+	—	+	—	—
<i>glcD</i>	Glycolate oxidase	+	+	—	+	—	—
<i>frdA</i>	Succinate dehydrogenase	+	+	+	—	—	+
<i>glf</i>	UDP-galactopyranose mutase	+	+	+	—	—	—

注：—：无编码基因；+：存在编码基因
Note: —: No gene; +: Have genes.

分,能够将大分子糖类物质从周质或细胞外空间运输到细胞质并使其磷酸化(表 6)。因此,相较于其他 5 种乳酸菌的模式菌株, *L. acetotolerans* G10 具有更强的水解及转运碳水化合物化合物的潜能,有利于提高碳源利用效率。

3 讨论与结论

在食品发酵过程中,微生物生长代谢消耗大量营养物质,不同物种利用底物能力的差异会影响其生态位^[15]。*L. acetotolerans* 在谷物类发酵食品中分布广泛,而且在发酵中后期保持优势地位,这不仅与 *L. acetotolerans* 耐酸、耐醇的特征有关,更与物种本身的代谢特征密切相关。本研究对 *L. acetotolerans* G10 的碳源利用能力进行分析,发现 *L. acetotolerans* G10 能够利用的碳源种类高达 67 种,其中 22 种为糖类及糖的衍生物。同时基于全基因组测序和比较基因组学方法,揭示了中国白酒发酵体系中 *L. acetotolerans* G10 与其他食品发酵体系菌株之间的差异;*L. acetotolerans* G10 具有更多糖类物质转运相关基因,具有能够合成多种转运酶的潜力,多碳源利用能力有助于菌株能够更好地适应中国白酒发酵体系。

来自不同发酵系统的 *L. acetotolerans* 菌株除在基因组大小、基因含量上存在差异外,基因功能上也有区别。形成微生物物种内差异的原因除了基因突变、基因漂移外,还受到环境带来的选择性压力^[22]。白酒发酵系统中微生物群落结构组成较其他发酵系统更复杂,不同微生物之间存在着复杂的相互作用,而且微生物的代谢产物能够影响其他微生物的生长^[23]。中国白酒在边糖化边发酵的生产过程中,微生物降解原料产生的碳源种类丰富且数量多^[14]。*L. acetotolerans* G10 具有更多的糖类转运相关基因,能够更好地从环境中获取营养物质、适应发酵环境,在发酵

过程中处于优势,并且该特征对于提高白酒发酵的资源利用率具有重要的作用。

L. acetotolerans 作为中国白酒发酵中后期的主要乳酸菌,对白酒风味及品质具有重要影响。在白酒发酵过程中, *L. acetotolerans* 含量减少时发酵会有明显滞后,酯的含量也有明显降低^[24]。在构建白酒可再生风味代谢的合成微生物群中, *L. acetotolerans* 是核心物种之一^[25]。*L. acetotolerans* G10 具有更多的糖类物质转运相关基因,能够广泛利用环境中的营养物质,其代谢产物对于白酒品质具有一定贡献。本研究通过全基因组测序揭示白酒发酵过程中优势乳酸菌的功能潜力及不同菌株之间的差异,为 *L. acetotolerans* G10 的应用提供了理论基础。

REFERENCES

- [1] Entani E, Masai H, Suzuki KI. *Lactobacillus acetotolerans*, a new species from fermented vinegar broth[J]. International Journal of Systematic Bacteriology, 1986, 36(4): 544-549
- [2] Toh H, Morita H, Tsuji H, Iwashita K, Goto N, Nakayama J, Sekine M, Kato Y, Suzuki KI, Fujita N. Complete genome sequence of *Lactobacillus acetotolerans* RIB 9124 (NBRC 13120) isolated from putrefied (hiochi) Japanese sake[J]. Journal of Biotechnology, 2015, 214: 214-215
- [3] Goto S, Motomura A, Kawahara A, Shiratsuchi H, Tanaka K, Matsusaki H. Cloning and heterologous expression of lactate dehydrogenase genes from acid-tolerant *Lactobacillus acetotolerans* HT[J]. Food Science and Technology Research, 2018, 24(5): 861-868
- [4] 张霞, 郑佳, 赵东, 乔宗伟, 安明哲, 杨康卓, 罗青春. 多粮浓香型白酒中特征酵母菌与耐酸乳杆菌的关系[J]. 微生物学通报, 2019, 46(7): 1571-1581
Zhang X, Zheng J, Zhao D, Qiao ZW, An MZ, Yang KZ, Luo QC. Relationship between typical yeasts and *Lactobacillus acetotolerans* in multi-grain strong aroma Baijiu[J]. Microbiology China, 2019, 46(7): 1571-1581 (in Chinese)
- [5] Chai LJ, Shen MN, Sun J, Deng YJ, Lu ZM, Zhang XJ, Shi JS, Xu ZH. Deciphering the D-/L-lactate-producing microbiota and manipulating their accumulation during

- solid-state fermentation of cereal vinegar[J]. Food Microbiology, 2020, 92: 103559
- [6] 颜娜, 王玉荣, 廖华, 赵慧君, 张振东, 郭壮. 恩施地区酸萝卜中细菌多样性研究[J]. 中国食品添加剂, 2019, 30(1): 69-76
- Yan N, Wang YR, Liao H, Zhao HJ, Zhang ZD, Guo Z. Study on bacteria diversity of sour radish in Enshi[J]. China Food Additives, 2019, 30(1): 69-76 (in Chinese)
- [7] Guan QQ, Zheng WD, Huang T, Xiao YS, Liu ZG, Peng Z, Gong DM, Xie MY, Xiong T. Comparison of microbial communities and physiochemical characteristics of two traditionally fermented vegetables[J]. Food Research International, 2020, 128: 108755
- [8] Gül H, Özçelik S, Sağdıç O, Certel M. Sourdough bread production with *Lactobacilli* and *S. cerevisiae* isolated from sourdoughs[J]. Process Biochemistry, 2005, 40(2): 691-697
- [9] Vera A, Ly-Chatain MH, Rigobello V, Demarigny Y. Description of a French natural wheat sourdough over 10 consecutive days focussing on the lactobacilli present in the microbiota[J]. Antonie Van Leeuwenhoek, 2012, 101(2): 369-377
- [10] Yagmur G, Tanguer H, Leventdurur S, Elmaci S, Turhan E, Francesca N, Settanni L, Moschetti G, Erten H. Identification of predominant lactic acid bacteria and yeasts of Turkish sourdoughs and selection of starter cultures for liquid sourdough production using different flours and dough yields[J]. Polish Journal of Food and Nutrition Sciences, 2016, 66(2): 99-107
- [11] Xiao YS, Huang T, Huang C, Hardie J, Peng Z, Xie MY, Xiong T. The microbial communities and flavour compounds of Jiangxi yancai, Sichuan paocai and Dongbei suancai: three major types of traditional Chinese fermented vegetables[J]. LWT, 2020, 121: 108865
- [12] Wang BW, Wu Q, Xu Y, Sun BG. Synergistic effect of multiple saccharifying enzymes on alcoholic fermentation for Chinese baijiu production[J]. Applied and Environmental Microbiology, 2020, 86(8): e00013-e00020
- [13] Wang Z, Ji XA, Wang SL, Wu Q, Xu Y. Sugar profile regulates the microbial metabolic diversity in Chinese Baijiu fermentation[J]. International Journal of Food Microbiology, 2021, 359: 109426
- [14] Liu CC, Feng SB, Wu Q, Huang HQ, Chen ZX, Li SW, Xu Y. Raw material regulates flavor formation via driving microbiota in Chinese liquor fermentation[J]. Frontiers in Microbiology, 2019, 10: 1520
- [15] Lu XW, Wu Q, Zhang Y, Xu Y. Genomic and transcriptomic analyses of the Chinese Maotai-flavored liquor yeast MT1 revealed its unique multi-carbon co-utilization[J]. BMC Genomics, 2015, 16: 1064
- [16] Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation[J]. Genome Research, 2017, 27(5): 722-736
- [17] Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification[J]. BMC Bioinformatics, 2010, 11: 119
- [18] Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences[J]. Journal of Molecular Biology, 2016, 428(4): 726-731
- [19] Chaudhari NM, Gupta VK, Dutta C. BPGA- an ultra-fast pan-genome analysis pipeline[J]. Scientific Reports, 2016, 6: 24373
- [20] 邢敏钰. 芝麻香型白酒发酵过程中乳酸菌群结构及功能[D]. 无锡: 江南大学硕士学位论文, 2017
- Xing MY. Community struction and function of lactic acid bacteria during fermentation process for sesame-flavor liquor making[D]. Wuxi: Master's Thesis of Jiangnan University, 2017 (in Chinese)
- [21] Du RB, Wu Q, Xu Y. Chinese liquor fermentation: identification of key flavor-producing *Lactobacillus* spp. by quantitative profiling with indigenous internal standards[J]. Applied and Environmental Microbiology, 2020, 86(12): e00456-e00420
- [22] Van Rossum T, Ferretti P, Maistrenko OM, Bork P. Diversity within species: interpreting strains in microbiomes[J]. Nature Reviews Microbiology, 2020, 18(9): 491-506
- [23] Du RB, Liu J, Jiang J, Wang YQ, Ji XA, Yang N, Wu Q, Xu Y. Construction of a synthetic microbial community for the biosynthesis of volatile sulfur compound by multi-module division of labor[J]. Food Chemistry, 2021, 347: 129036
- [24] Pang XN, Han BZ, Huang XN, Zhang X, Hou LF, Cao M, Gao LJ, Hu GH, Chen JY. Effect of the environment microbiota on the flavour of light-flavour Baijiu during spontaneous fermentation[J]. Scientific Reports, 2018, 8: 3396
- [25] Zhao TF, Ni DR, Hu GY, Wang L, Chen S, Xu Y. 6-(2-Formyl-5-methyl-1 H-pyrrol-1-yl)hexanoic acid as a novel retronasal burnt aroma compound in soy sauce aroma-type Chinese baijiu[J]. Journal of Agricultural and Food Chemistry, 2019, 67(28): 7916-7925