



Limosilactobacillus fermentum F-6 的遗传背景和功能基因组

吴琼, 李伟程, 李敏, 李瑜, 孙天松*

内蒙古农业大学乳品生物技术与工程教育部重点实验室, 农业部奶制品加工重点实验室, 内蒙古乳品生物技术与工程重点实验室, 内蒙古 呼和浩特 010018

吴琼, 李伟程, 李敏, 李瑜, 孙天松. *Limosilactobacillus fermentum* F-6 的遗传背景和功能基因组. 微生物学报, 2022, 62(4): 1438–1451.

Wu Qiong, Li Weicheng, Li Min, Li Yu, Sun Tiansong. Genetic background and functional genome of *Limosilactobacillus fermentum* F-6. *Acta Microbiologica Sinica*, 2022, 62(4): 1438–1451.

摘要:【目的】*Limosilactobacillus fermentum* 具有增强免疫力、产胞外多糖(exopolysaccharide, EPS)等多种功能特性, 广泛应用于食品领域, 具有较高经济价值。本文从群体遗传学角度, 解析 *L. fermentum* F-6 的遗传背景和功能基因特征, 为其开发利用提供遗传学基础。【方法】本研究对 NCBI 已公开的 23 株 *L. fermentum* 全基因组序列和 1 株模式菌株 ATCC 14931^T 的基因组序列进行比较基因组学分析。利用 Roary 软件识别核心基因集与泛基因集; 采用 rapid annotation using subsystem technology (RAST) 网站对基因组进行功能注释, 以探究 F-6 基因组特征。【结果】以识别到的 997 个核心基因构建系统发育树, 发现聚类趋势与分离源无关, 但 F-6 与 3 株食品分离株聚在同一分支。功能注释分析发现, 24 株 *L. fermentum* 中仅 F-6 含有参与支链氨基酸合成途径的基因(*ilvD*、*leuA* 等), 可为机体提供必需氨基酸。F-6 含有大量编码糖基转移酶和 UDP-葡萄糖 4-表异构酶的基因, 且含有 1 个完整的 *eps* 基因簇。与其他 *L. fermentum* 相比, F-6 携带的特有功能基因 *gadC* 与菌株耐酸能力有关。此外, F-6 具有与谷胱甘肽、抗菌肽和黏附分子合成相关的功能基因。【结论】本文从全基因组水平解析 *L. fermentum* F-6 遗传背景和功能基因特征, 认为其是一株具有耐酸基因、支链氨基酸合成基因和完整 *eps* 基因簇的潜在益生菌株, 为其应用研究提供了必要的理论依据。

关键词: *Limosilactobacillus fermentum* F-6; 比较基因组学; 遗传背景; 功能基因; *eps* 基因簇

基金项目: 财政部和农业农村部国家现代农业产业技术体系

Supported by the China Agriculture Research System of MOF and MARA

*Corresponding author. Tel: +86-471-4305357; E-mail: sts9940@sina.com

Received: 25 July 2021; Revised: 3 October 2021; Published online: 14 January 2022

Genetic background and functional genome of *Limosilactobacillus fermentum* F-6

WU Qiong, LI Weicheng, LI Min, LI Yu, SUN Tiansong*

Key Laboratory of Dairy Biotechnology and Engineering, Ministry of Education; Key Laboratory of Dairy Products Processing, Ministry of Agriculture and Rural Affairs; Inner Mongolia Key Laboratory of Dairy Biotechnology and Engineering, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China

Abstract: [Objective] Enhancing immunity and producing exopolysaccharide (EPS), *Limosilactobacillus fermentum* is widely used in food industry and is of huge economic value. Based on population genetics, this paper aims to analyze the genetic background and functional gene characteristics of *L. fermentum* F-6, which is expected to lay a genetic basis for the development and utilization of the strain. [Methods] Comparative genomics was used to analyze the whole-genome sequences of 23 *L. fermentum* strains and 1 model strain ATCC 14931^T from NCBI, Roary to identify core-pan genes, and rapid annotation using subsystem technology (RAST) to annotate the genomes. [Results] A phylogenetic tree was constructed based on the yielded 997 core genes and the analysis showed that the clustering had nothing to do with the source. However, strain F-6 was in the same clade with 3 food-derived strains. Genome annotation results revealed that only strain F-6 contained genes involved in branched chain amino acid synthesis pathway (*ilvD*, *leuA*, etc.) and thus can provide essential amino acids for the body. Strain F-6 harbored a large number of genes encoding glycosyltransferases and UDP-glucose 4-epimerase, and contained a complete *eps* cluster. Compared with other *L. fermentum* strains, strain F-6 had the unique functional gene *gadC* which is associated to acid resistance. In addition, potential probiotic genes such as genes related to the synthesis of glutathione, antimicrobial peptide, and adhesion molecules were also identified in the genome. [Conclusion] According to the genomic background and functional genome, F-6 is a potential probiotic strain with acid resistance genes, branched chain amino acid synthesis genes, and a complete *eps* gene cluster, which will contribute to our understanding of the genetic basis of *L. fermentum* F-6.

Keywords: *Limosilactobacillus fermentum* F-6; comparative genomics; genetic background; functional genes; *eps* cluster

Limosilactobacillus fermentum (*L. fermentum*) 是乳酸菌重要菌种之一, 属于兼性厌氧菌, 广泛存在于发酵食品、生殖道及人体肠道中^[1-2]。作为肠道固有菌种之一^[2], 大量研究证实 *L. fermentum* 具有多种益生特性。如 Do 等^[3] 从发酵竹笋中获得一株具有高产胞外多糖 (exopolysaccharide, EPS) 能力的 *L. fermentum* MC3 菌株; 研究还发现 *L. fermentum* HFD1 具有产抗菌肽的能力^[4]。此外, *L. fermentum* FUA

3321 可提高发酵高粱中儿茶素、没食子酸等活性物质含量^[5]。*L. fermentum* 作为潜在益生菌得到广泛开发与应用。

比较基因组学是对菌株基因组及功能基因等信息进行深入挖掘, 进而解析菌株间遗传进化关系及生境适应性机制的学科^[6-7]。目前, 对 *L. fermentum* 比较基因组学的研究主要集中在特有基因功能注释和遗传进化方面。如 Verce 等^[8] 利用比较基因组学分析不同来源的 *L. fermentum*

基因组特征及其功能,发现面团中的 *L. fermentum* 具有降解淀粉的功能基因,表明其遗传进化与生境相关。此外,有研究发现 *L. fermentum* 3872 存在与细菌聚集和抵抗胃肠道环境的相关基因,可在胃肠道存活并定殖,维持肠道菌群平衡^[9]。但对 *L. fermentum* 的功能基因特征方面的研究较少。因此, *L. fermentum* 功能基因特征分析对其功能开发与应用具有重要意义。

本团队前期已完成对 *L. fermentum* F-6 的全基因组测序工作^[10], 并发现 F-6 具有促进仔猪生长并提高其免疫的能力, 提高肉鸡生产性能^[11-12]等功能, 是潜在的益生菌株。但目前基于全基因组水平解析 *L. fermentum* F-6 功能基因的相关研究较匮乏。因此, 本研究对从 NCBI

Refseq 数据库下载的 23 株不同分离源 *L. fermentum* 全基因组序列(包括 *L. fermentum* F-6^[13])和模式菌株 ATCC 14931^T 的基因组序列进行分析, 从基因组水平解析 *L. fermentum* F-6 遗传背景及功能特性, 为其开发利用提供参考。

1 材料与方法

1.1 发酵乳杆菌基因组序列

截至 2020 年 11 月 9 日, 已将 NCBI (National Coalition Building Institute, <https://www.ncbi.nlm.nih.gov/>) Refseq 数据库中 23 株 *L. fermentum* 基因组完成图(包括 *L. fermentum* F-6^[13])和模式株 *L. fermentum* ATCC 14931^T 基因序列全部下载完成(表 1)。

表 1 下载自 NCBI 的 24 株 *L. fermentum* 基本信息

Table 1 Basic information of 24 *L. fermentum* retrieved from NCBI

Strain name	Isolate source	Region location	NCBI Refseq assembly number
IFO 3956	Food	Unknown	GCA_000010145.1
ATCC 14931 ^T	Unknown	Unknown	GCA_000159215.1
CECT 5716	Milk from female	Unknown	GCA_000210515.1
F-6	Food	Unknown	GCA_000397165.1 ^[11]
3872	Milk from female	Others	GCA_000466785.3
NCC2970	Unknown	Others	GCA_001742205.1
SNUV175	Vaginal tract	Korea	GCA_001941785.1
FTDC 8312	Fecal	Malaysia	GCA_002119645.1
LAC FRN-92	Oral cavity	Others	GCA_002192435.1
SK152	Food	Korea	GCA_002242615.1
MTCC 25067	Food	Others	GCA_002356135.1
CBA7106	Fecal	Korea	GCA_003255875.1
LDTM 7301	Food	Korea	GCA_003346795.1
DR9	Food	Malaysia	GCA_003710225.1
LMT2-75	Food	Korea	GCA_003855655.1
YL-11	Food	China	GCA_003860425.1
SRCM 103285	Food	Korea	GCA_004063515.1
SRCM 103290	Food	Korea	GCA_004063635.1
2760	Food	Others	GCA_008802915.1
USM 8633	Food	Malaysia	GCA_009676625.1
AGR1487	Oral cavity	New Zealand	GCA_011032745.1
AGR1485	Oral cavity	New Zealand	GCA_011032765.1
HFD1	Fecal	Others	GCA_012273035.1
IMDO 130101	Food	Others	GCA_900205745.1

1.2 比较基因组分析

1.2.1 平均核苷酸一致性(average nucleotide identity, ANI)计算

本研究中, 24 株 *L. fermentum* ANI 值参考 Goris 等^[14]报道的方法, 用自制 Perl 脚本计算菌株间 ANI 值。

1.2.2 泛-核心基因集构建

利用 Prokka^[15]软件对菌株基因组进行基因预测后, 采用 Roary^[16]软件对核心基因集、泛基因集进行识别, 其中核心基因识别以编码蛋白氨基酸相似性大于 95% 的原则进行^[17]。泛基因组为 24 株 *L. fermentum* 所含有的全部基因, 核心基因为 24 株 *L. fermentum* 共有的基因, 而特异性基因只在某一株菌的基因组中出现。最后利用 R (v.4.0.4) 进行泛-核心基因趋势图绘制。

1.2.3 系统发育树构建

经 Roary^[16]软件分析得到的核心基因序列利用 TreeBeST 软件(<http://www.mybiosoftware.com/treebest>)中的邻接法(neighbor-joining, NJ)进行系统发育树构建。系统发育树可视化在 iTol 在线软件(<https://itol.embl.de/>)中进行。

1.2.4 基因功能注释

利用 rapid annotation using subsystem technology (RAST, <https://rast.nmpdr.org/>)数据库完成 24 株 *L. fermentum* 全基因组的功能注释。

1.3 CRISPR-Cas 系统分析

使用 CRISPR-Cas Finder^[18]在线分析软件对 24 株 *L. fermentum* 的 CRISPR 及 Cas 基因进行查找, 对 CRISPR 基因簇数量、Cas 基因簇分型及序列间隔区(spacer)进行统计。分析不同菌株间 CRISPR-Cas 系统基因簇类型及差异。

1.4 数据分析及绘图

利用 Perl 脚本统计基因组信息。利用

TBtools^[19]软件进行 EPS 基因簇热图绘制。利用 R 绘制 ANI 聚类热图。

2 结果与分析

2.1 核心基因构建系统发育树

系统发育树可直观反映菌株间群体结构和遗传进化关系^[20]。为研究 *L. fermentum* F-6 与其他 *L. fermentum* 种内遗传进化关系, 本文采用 NJ 法, 对 Roary 软件识别得到的 997 个核心基因构建系统发育树(图 1), bootstrap 值为 1 000。

系统发育树结果显示, 24 株 *L. fermentum* 在进化过程中由于遗传多样性使其被分为 4 大分支。Clade I、II 和 III 包含 5 株亲缘关系较远的 *L. fermentum*, 且不存在聚类现象。Clade IV 包含 19 株 *L. fermentum*, 可将其细分为 4 个亚分支(Clade A、Clade B、Clade C 和 Clade D)。Clade B 包括 2 株母乳分离株、2 株食品分离株和 3 株其他分离源菌株, 其中, 益生菌株 CECT5716 与产抗菌肽菌株 HFD1 在此分支中, 表明两者遗传关系较近, 但与菌株 F-6 遗传距离较远。特别地, 2 株马来西亚分离株聚类在此分支。*L. fermentum* F-6 与其他 5 株 *L. fermentum* 共同聚类于 Clade D 分支, 该分支包含食品分离株(4 株, 占 33.33%)、粪便分离株(1 株)和 1 株分离源未知菌株, 且菌株 F-6 与其他 3 株食品分离株聚类在一起。Clade C 所处分支的 2 株食品分离株在同一小分支中, 表明两者亲缘关系较近。

综上所述, 基于核心基因构建的系统发育树结果, 未发现分离源与分离地有明显聚类现象。该研究结果与 Illegheems 等^[21]研究结果一致。但食品分离株 F-6 与其他 3 株食品分离株共同聚类于 Clade D 分支, 遗传距离较近。

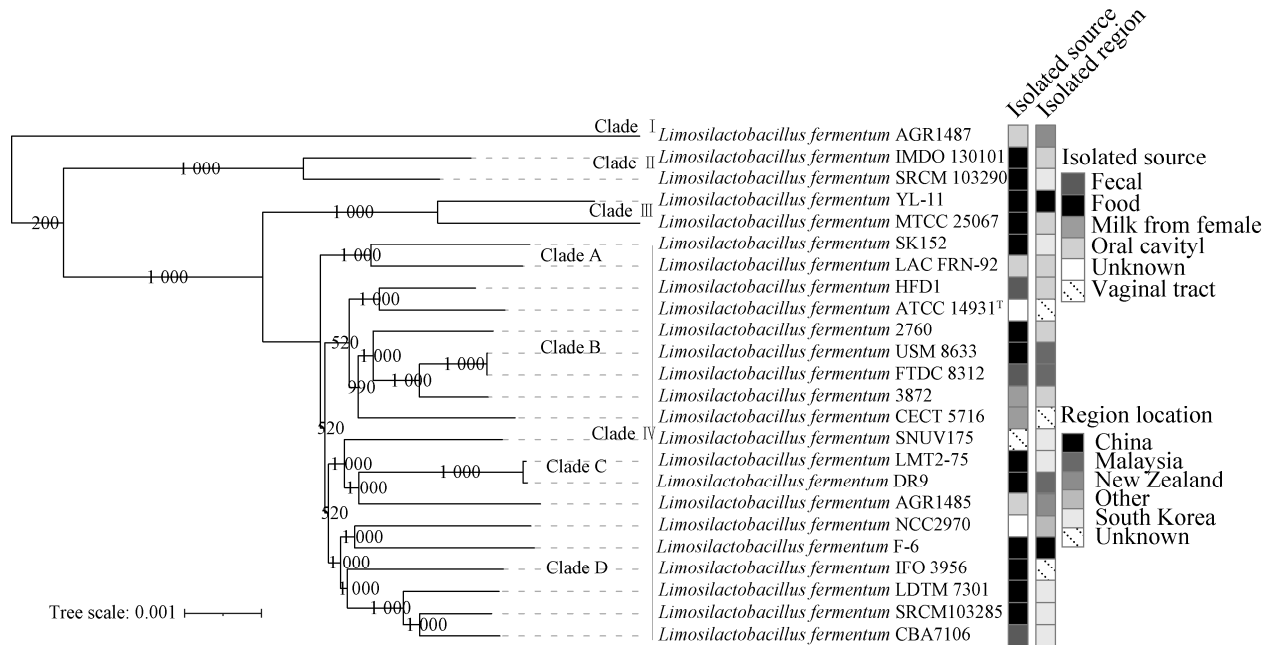


图 1 基于 24 株 *L. fermentum* 的 997 个核心基因构建的系统发育树

Figure 1 Phylogenetic tree based on 997 core genes in 24 *L. fermentum*. The numbers on the branch indicated that the reliability of this branch (bootstrap value); The ruler indicates the branch genetic variability of this length of 0.001 (the frequency of base substitution is 0.001 per locus in this genome).

2.2 ANI 与泛-核心基因分析

基于菌株之间同源序列的比对获得的 ANI 主要在全基因组水平对菌株亲缘关系进行鉴定^[22], 一般认为 ANI 值大于 95% 即为同一物种^[23]。本文对 24 株 *L. fermentum* 菌株进行 ANI 值计算, 并构建聚类热图(图 2A)。结果显示, 菌株两两间平均 ANI 为(98.68±0.49)%, 均大于 95%, 表明研究所用菌株均为同一物种即 *L. fermentum*。其中, AGR1487、SRCM103290 和 IMDO130101 与其他菌株间 ANI 值较低, 聚类于同一分支, 且其中 2 株食品分离株聚在同一分支, 与系统发育树结果一致。此外, 食品分离株两两间 ANI 为 98.46%, 粪便分离株两两间 ANI 为 99.13%, *L. fermentum* 食品分离株的基因组序列一致性比人类肠道/粪便分离株低, 表明 *L. fermentum* 基因组序列多样性可能与其分离源相关。

为进一步从泛基因水平了解 *L. fermentum* 差异, 利用 24 株 *L. fermentum* 泛基因组绘制基因存在缺失分布热图(图 2B), 每一行代表一个基因, 每一列代表不同菌株。其中, 黑色代表基因存在, 灰色代表基因缺失。由图 2B 可以看出, 分离自马来西亚地区的粪便分离株 FTDC 8312 与食品分离株 USM 8633 基因组成较为相似, 且两者 ANI 为 99.99%, 推测两者可能为同一菌株。此外, 食品分离株(DR9 和 LMT2-75; MTCC25067 和 YL-11)与韩国分离株(LDTM7301、CBA7106 和 SRCM103285)聚在同一分支, 与系统发育树结果一致。但菌株 F-6 与之前提到的 3 株食品分离株及其他菌株间基因组成均未呈现明显聚集现象。

上述结果发现, 菌株 F-6 与其他菌株基因组成存在较大差异。因此, 本文将以食品分离株 F-6 为主进行后续分析, 以揭示其基因组特征。

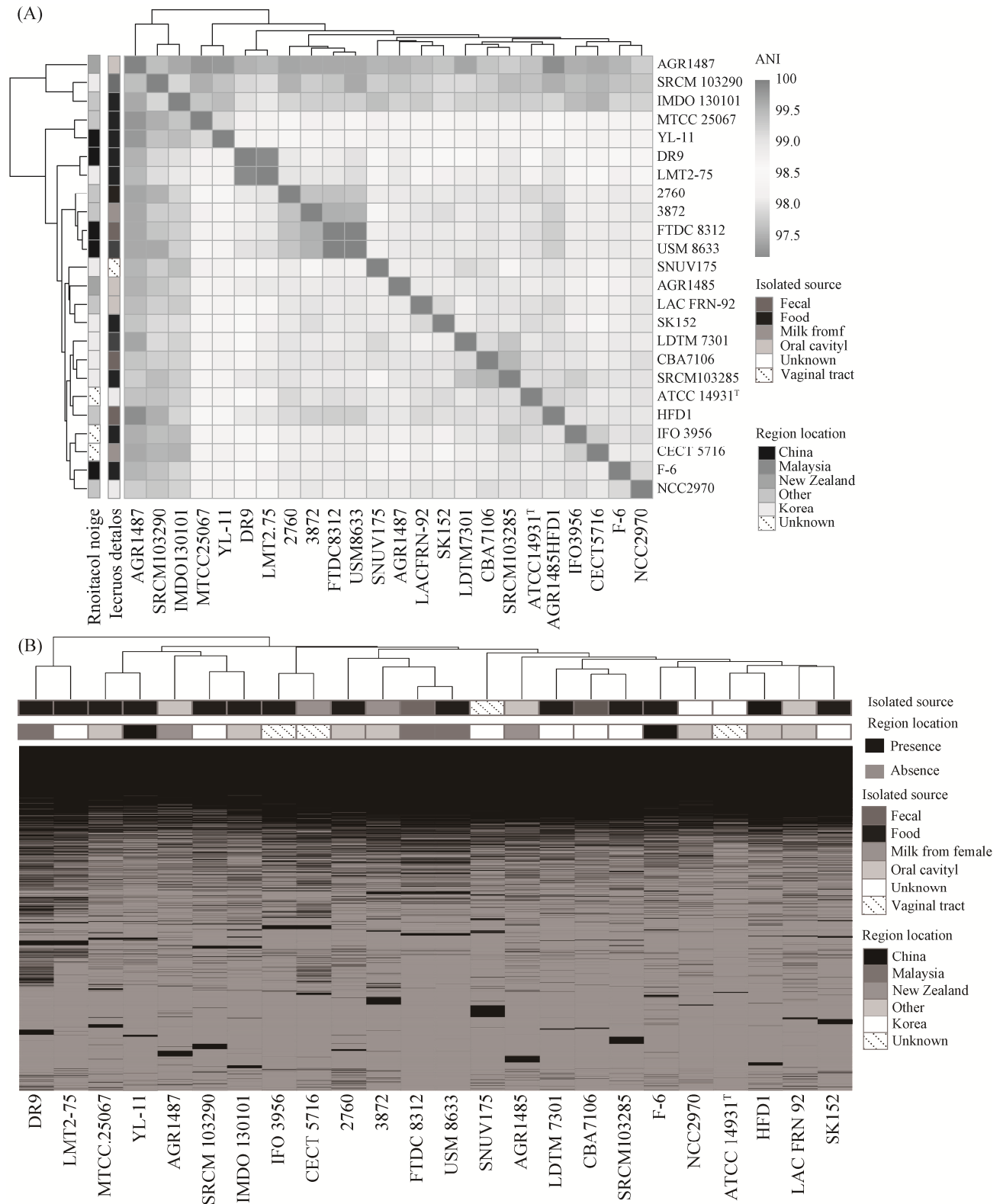


图 2 *L. fermentum* ANI 值聚类热图(A)与基因存在缺失分布图(B)

Figure 2 Heatmap of ANI (A) and the presence and absence of genes (B) based on the genome sequences of *L. fermentum*.

2.3 RAST 功能注释

本研究利用 RAST 软件对 24 株 *L. fermentum* 全基因组序列进行功能注释,分析菌株 F-6 与其他菌株功能差异。24 株 *L. fermentum* 共注释到 25 个功能大类。由图 3 可知,24 株 *L. fermentum* 中参与氨基酸及其衍生物合成与降解(amino acids and derivatives)的基因数量最多(16.37%),其次是蛋白质代谢(protein metabolism, 14.16%)和碳水化合物代谢(carbohydrates metabolism, 14.20%)相关基因。芳香族化合物代谢(metabolism of aromatic compounds)及次级代谢(secondary metabolism)

相关功能编码基因仅在部分菌株中存在。前期研究发现,CECT5716 菌株具有多种益生功能,因此对菌株 CECT5716 与 F-6 进行比较分析。*L. fermentum* F-6 与 CECT5716 共预测到 23 个功能类别,且 CECT5716 含有的功能基因较多。与菌株 F-6 相比,CECT5716 含有大量氨基酸及其衍生物、蛋白质代谢和 DNA 代谢相关的功能基因。此外,两者都注释到编码与吸收硒酸盐与亚硒酸盐相关的功能基因(*dedA*)^[24]。研究发现,一定浓度的硒酸盐与亚硒酸盐会导致生物体病变^[25],因此,F-6 与 CECT 5716 菌株可在一定范围内降低病变的可能性。24 株 *L. fermentum*

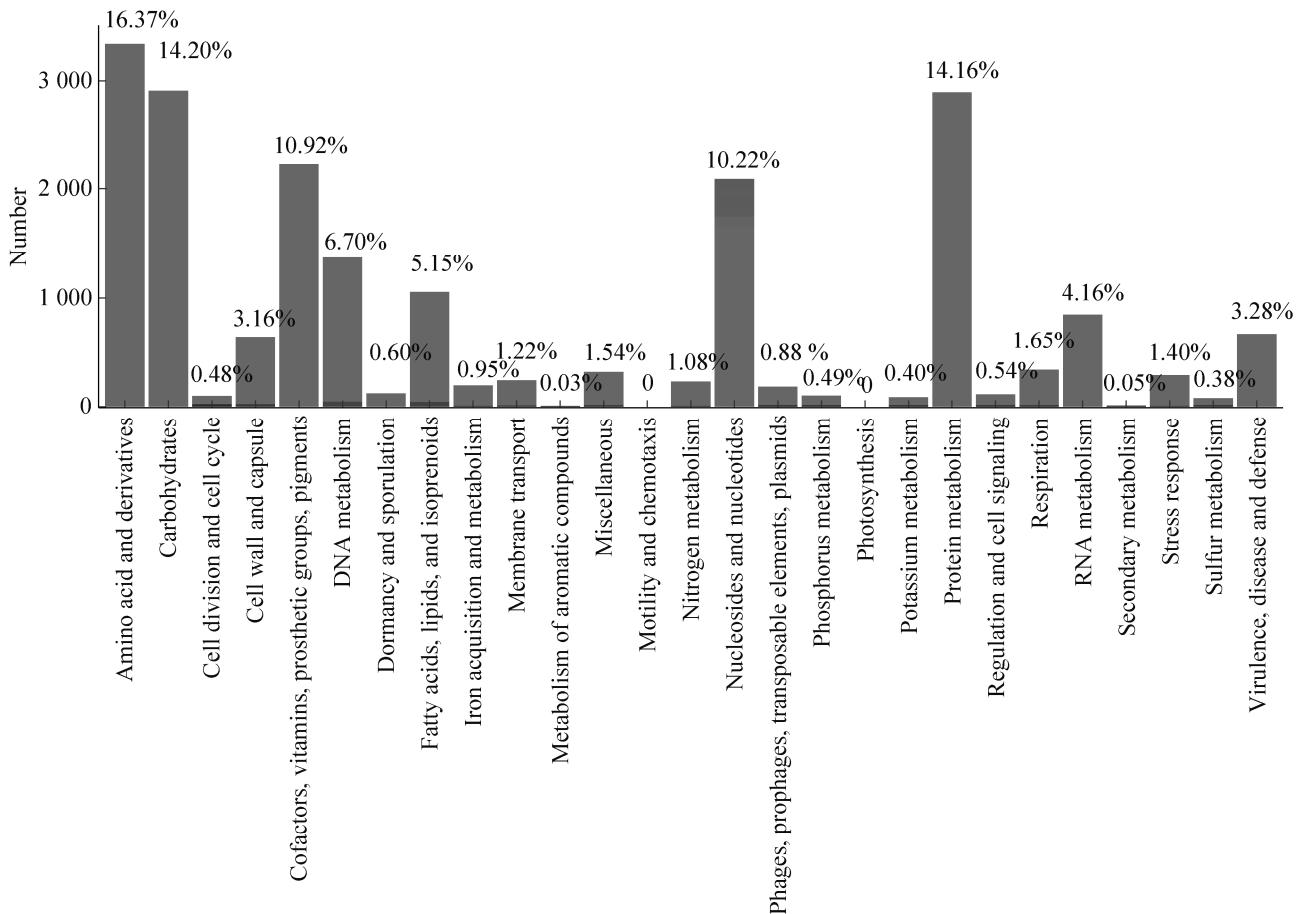


图 3 24 株 *L. fermentum* RAST 功能注释

Figure 3 Annotation of 24 *L. fermentum* strains based on the RAST database. The percentage of above the bar chart = (number of such functional genes/number of all functional genes) \times 100%. The ordinate number is the number of functional genes.

中仅菌株 F-6 注释到与支链氨基酸合成相关的基因, 可催化丙酮酸, 进而合成异亮氨酸、亮氨酸及缬氨酸, 促进人体对支链氨基酸的吸收。

功能注释结果分析发现, 菌株 F-6 注释到合成支链氨基酸相关的基因, 且其含有的 *dedA* 基因可吸收硒酸盐与亚硒酸盐, 可能具有抑制生物体病变并合成支链氨基酸的功能特性。

2.4 EPS 相关基因分析

研究发现, 乳酸菌产生的 EPS 具有多种对健康有益的功效。同时, EPS 作为天然的稳定剂、乳化剂及增稠剂广泛应用于发酵乳和干酪的生产。表型试验发现, *L. fermentum* 在培养期间会产生大量 EPS, 韦云路和李全宏^[26]发现 *L. fermentum* YL-11 胞外多糖能够清除 DPPH (2,2-diphenyl-1-picrylhydrazyl) 等自由基, 具有抗氧化活性。因此, 本文通过对 24 株 *L. fermentum* 基因组序列分析, 查找到与 EPS 合成相关的基因簇。图 4 为 EPS 合成相关基因簇的存在情况, 除食品分离株外, 其他不同分离源 *L. fermentum* *eps* 基因簇无聚集趋势。

研究证明, 糖基转移酶可能参与多糖的转运过程, 且其类型和数量决定了 EPS 中重复单元的范围, 有助于 EPS 的合成^[27-28]。仅有菌株 2760 不含糖基转移酶基因, 可能不产 EPS 或产 EPS 能力较弱(图 4)。另外, Degeest 等发现, UDP-葡萄糖 4-表异构酶的活性与 EPS 的产量有较高相关性^[29]。而与菌株 F-6 和 YL-11 处于同一簇的菌株均携带大量编码糖基转移酶和 UDP-葡萄糖 4-表异构酶的基因。其中, 菌株 DR9 与 F-6 还含有较多糖基转移酶基因, 且 F-6 与具有产 EPS 能力的 YL-11 含有的 EPS 基因种类数量较为一致。

为进一步确定 EPS 相关基因在全基因组中的分布情况, 对菌株 YL-11 和 F-6 EPS 相关基因分布情况进行分析, 发现在由 *epsJ* 基因编码

的糖基转移酶和 *yokT* 基因编码的糖基转移酶的上下游存在大量 EPS 合成相关基因。菌株 F-6 与 YL-11 各形成 2 个基因簇(图 5)。其中, F-6 的 *eps* 基因簇长度为 14 415 bp (A) 和 26 777 bp (C), YL-11 的 *eps* 基因簇长度为 13 210 bp (B) 和 29 945 bp (D)。4 个 *eps* 基因簇含有大量糖基转移酶基因和假定蛋白, 但主要集中在 A 和 B 基因簇。4 个基因簇均有一个多糖合成基因, 其中, C 和 D 携带的编码多糖合成基因 *cap8A* 可能还具有调节多糖链长的功能^[30]。另外, C 和 D 基因簇还包括与调节和转运相关基因, 可保证多糖的合成及其重复单元转移至胞外。综上所述, 与 A 和 B 基因簇相比, C 和 D 基因簇较为完整, 且 C 和 D 基因簇所含基因种类及基因簇长度一致。

上述结果表明, 菌株 F-6 含有较多糖基转移酶和 UDP-葡萄糖 4-表异构酶基因, 与 YL-11 携带的 EPS 基因种类数量较一致, 且 F-6 具有 1 个较为完整的 *eps* 基因簇。

2.5 CRISPR-Cas 系统分析

CRISPR-Cas 系统是一种抵抗病毒感染的适应性系统^[31], 基于 CRISPR-Cas 系统的基因编辑可打破某些遗传限制, 从而改变某些菌株的生物学特性^[32]。Brandt 等^[33]研究 38 株 *L. fermentum* 的 CRISPR-Cas 系统多样性, 发现不同菌株之间 CRISPR-Cas 系统存在较高变异, 说明菌株之间具有遗传多样性。

对 24 株 *L. fermentum* 的 CRISPR 和 Cas 基因进行比较分析(图 6), 柱状图表示 *L. fermentum* CRISPR 及 spacer 数量, 热图表示 Cas 基因簇分型存在情况。24 株 *L. fermentum* 均预测到 CRISPR 基因簇, 且 LMT2-75 与 SK152 的 CRISPR 位点最多, 有 7 个; MTCC 25067、AGR1485 和 F-6 仅有 1 个 CRISPR 位点, 但 F-6 含有的 spacer 数量高于其余两株,

推测其具有较强抗噬菌体能力。另外，24株 *L. fermentum* 中 SK152 的 CRISPR 及

spacer 数量最多，可能具有较强的抗噬菌体能力。

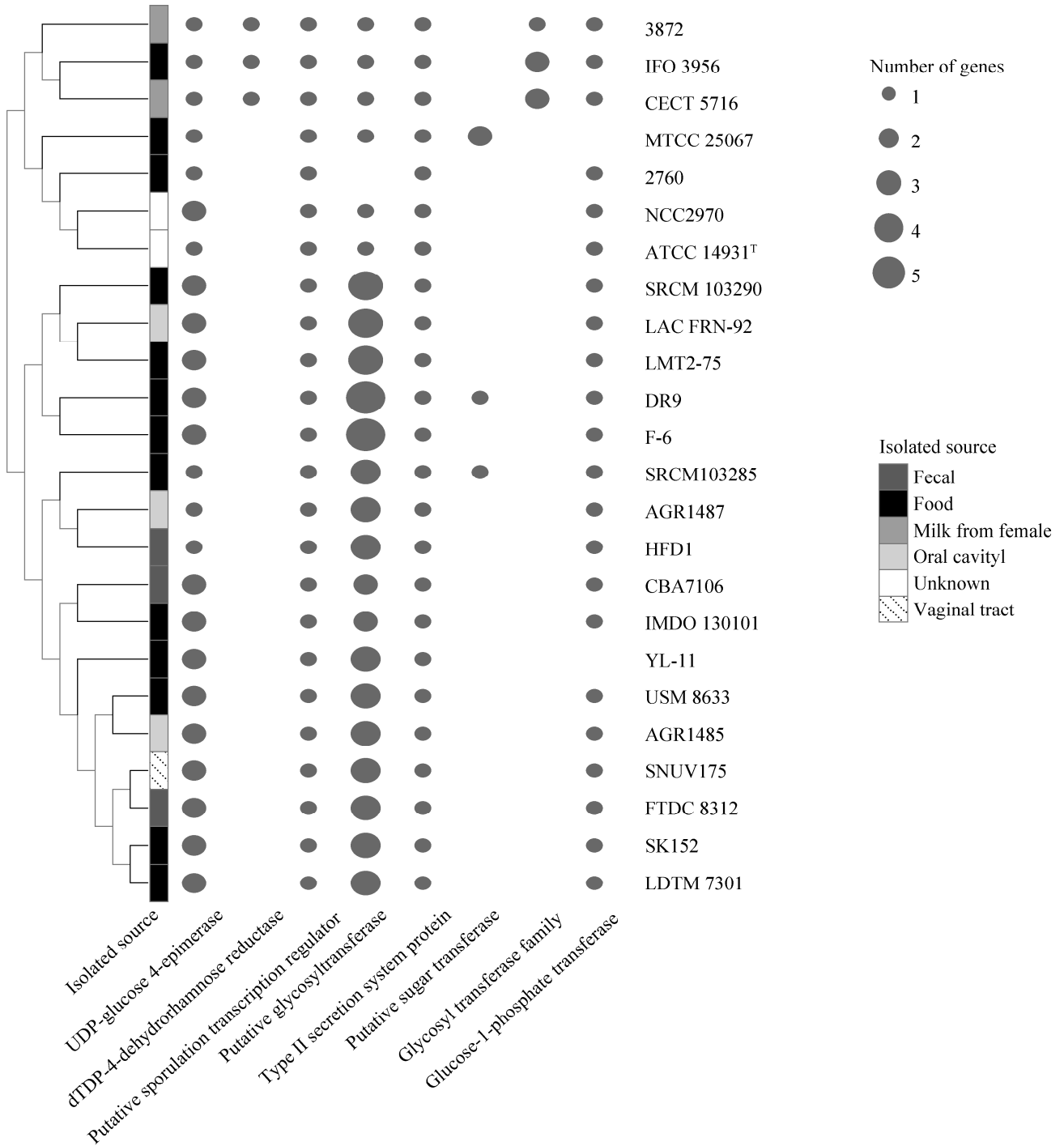


图 4 24 株 *L. fermentum eps* 基因簇分布情况

Figure 4 24 strains of *L. fermentum* dtribution of *eps* gene clusters.

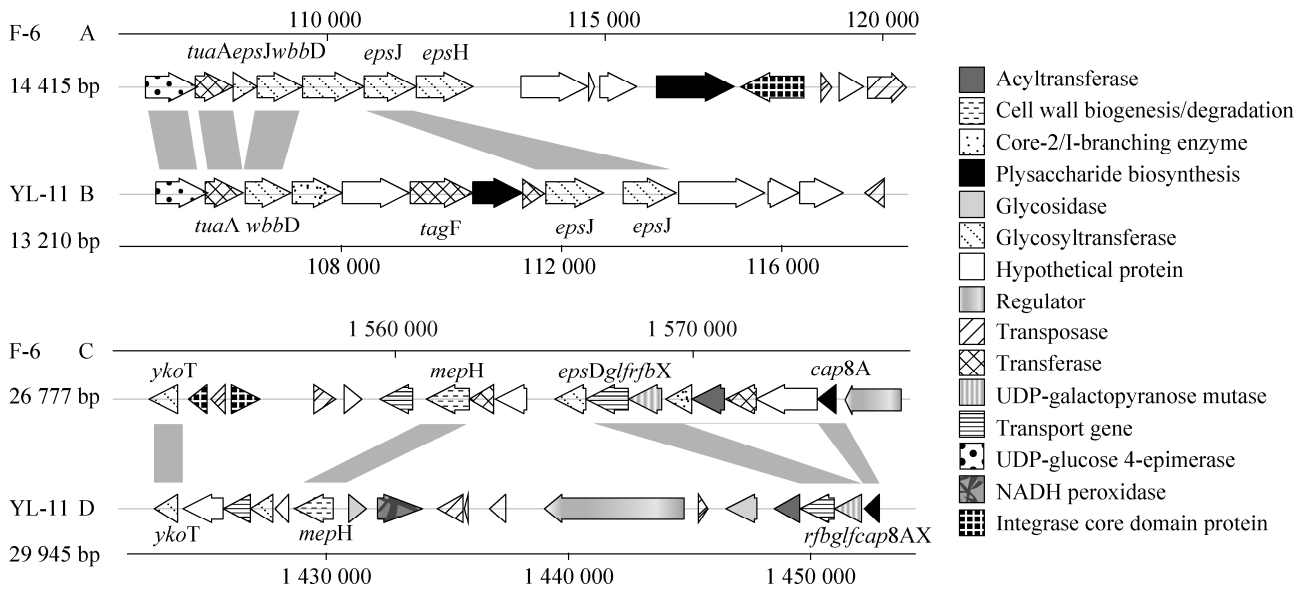


图 5 *L. fermentum* YL-11 与 F-6 的 *eps* 基因簇结构图

Figure 5 The *eps* gene clusters of *L. fermentum* YL-11 and F-6.

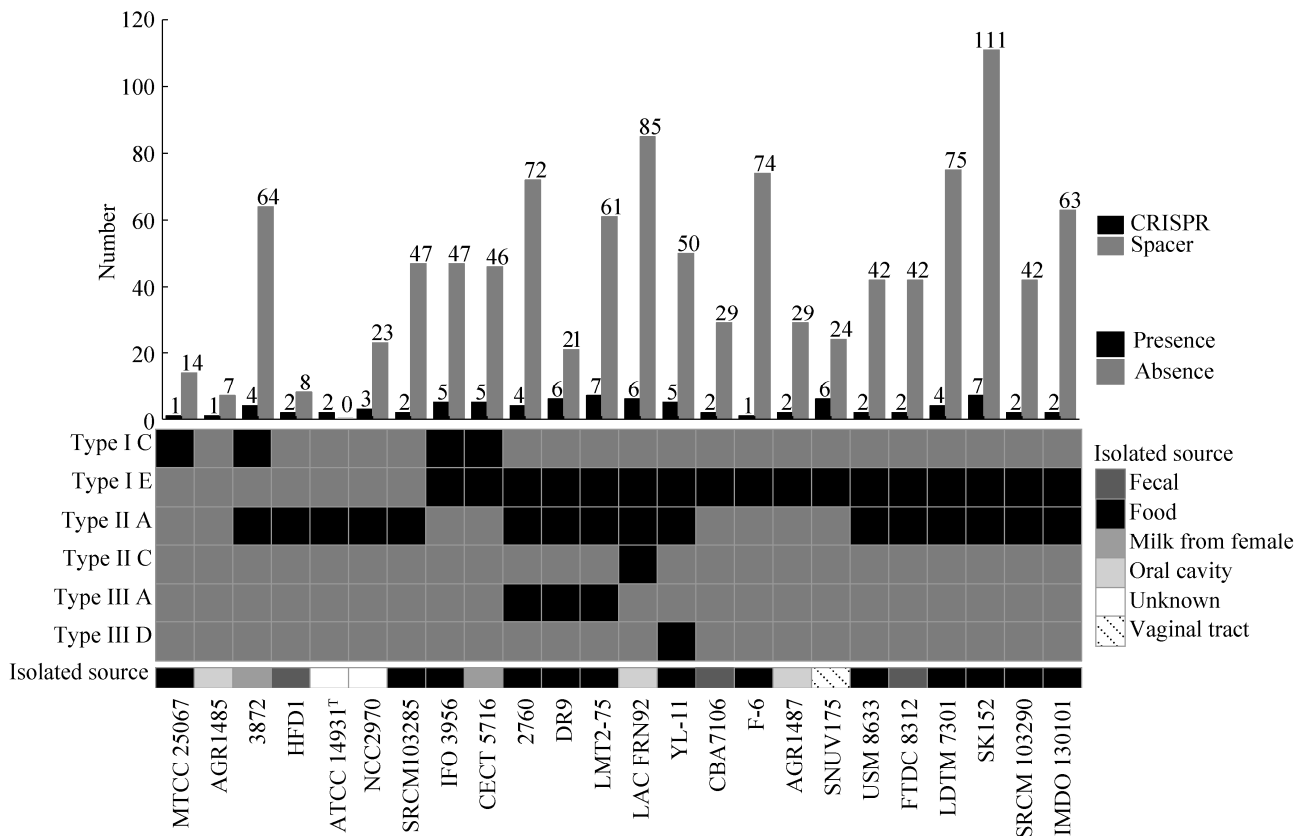


图 6 24 株 *L. fermentum* 中 CRISPR 及 Cas 基因簇存在情况

Figure 6 The number of CRISPR and type of Cas genes in 24 *L. fermentum* strains. The bar graph shows the number of CRISPR and spacer in 24 *L. fermentum* strains. The heatmap indicates the presence or absence of the Cas gene cluster.

由 Cas 基因簇分型热图可知,除 AGR1485 外,其余菌株均具有 Cas 基因簇。其中 17 株 (70.8%) 含有 I E 型 CRISPR-Cas 系统,16 株 (66.7%) 含有 II A 型 CRISPR-Cas 系统,两者出现频率较高,但 II C 型与 III D 型仅在 1 株 *L. fermentum* (4.2%) 中存在。有趣的是, Cas 蛋白作为识别噬菌体的结构,在 AGR1485 中未识别到,但其携带有 spacer 序列,可能是在进化过程中 Cas 基因丢失造成。另外,食品分离株的 Cas 基因簇分型具有明显聚类现象,且大部分食品分离株中存在 I E 型与 II A 型 Cas 基因簇。其中,菌株 F-6 仅有 I E 型 CRISPR-Cas 系统。基于 CRISPR-Cas 系统分析结果显示,不同 *L. fermentum* 的 CRISPR 基因、Cas 基因簇类型及 spacer 数量均存在差异,但食品分离株所含的 Cas 基因簇类型较为相似。

2.6 特有功能基因及益生相关基因分析

为深入解析 *L. fermentum* 菌株间基因组差异,挖掘 F-6 功能基因,解析其潜在益生功能基因。本研究将以益生菌株 F-6 和处于不同分支的益生菌株 CECT 5716 的特有功能基因为例进行差异分析。24 株 *L. fermentum* 共识别到 2 694 个特有基因,特有基因数量较多。菌株 F-6 与 CECT 5716 分别存在 103 和 151 个特有基因,除去一些假定蛋白编码基因后,菌株 F-6 有 4 个特有功能基因,CECT 5716 有 7 个特有功能基因。CECT 5716 特有功能基因主要参与细胞活动调节(*gppA*、*pacS*)、生物合成途径(*aroF*、*rluC*、*pyrF*、*mobB*)及其他代谢活动(*mccB*)(表 2)。*mccB* 催化胱硫醚向半胱氨酸转化,并刺激高半胱氨酸向硫化物的转化。研究发现,高半胱氨酸通过 progranulin/EphA2 途径损伤内皮细胞屏障功能和血管生成潜能,易引发心血管疾病^[34]。但对其在心血管疾病方面的研究较少,具体预防机制还有待挖掘。

菌株 F-6 特有功能基因主要参与还原反应(*zaoB*)、氨基酸代谢(*gadC*、*kynB*)和免疫调节(*bca*)途径(表 2)。其中,*bca* 基因编码一种细胞表面抗原蛋白,可能诱发免疫反应^[35]。另外,F-6 含有的 *gadC* 基因编码谷氨酸/γ-氨基丁酸转运蛋白,该蛋白可使胞外谷氨酰胺转运到细胞内生成 L-谷氨酸和游离氨,游离氨中和 H⁺,使胞内 pH 升高,赋予菌株对极端酸性环境的抗性^[36]。因此,该基因可使菌株在极端酸性环境保持活性,发挥益生功能。

特别地,基于前期对菌株 F-6 益生功能的研究^[10-12],本文注释到与益生特性相关的基因(表 3)。F-6 含有谷胱甘肽合成(*gshAB*)及核黄素合成(*ribBA*、*ribE*)相关基因。且上述基因在乳源益生菌 *L. fermentum* CECT 5716^[37]中同样存在。研究表明,谷胱甘肽参与肠黏膜的抗氧化机制,保护肠道免受氧化剂诱导的组织损伤。另外,谷胱甘肽可提高菌株在胃肠道的存活率^[38],有助于发挥其益生功能。研究还发现,F-6 存在

表 2 24 株 *L. fermentum* 中 F-6 和 CECT 5716 的特有功能基因

Table 2 The unique genes of strains F-6 and CECT 5716 in 24 *L. fermentum* strains

Strains	Genes	Annotation
F-6	<i>gadC</i>	Glutamate/gamma-aminobutyrate antiporter
	<i>azoB</i>	NAD(P)H azoreductase (EC 1.7.-. -)
	<i>bca</i>	C protein alpha-antigen precursor
	<i>kynB</i>	Kynurenine formamidase (EC 3.5.1.9)
CECT 5716	<i>gppA</i>	3-dehydroquinate dehydratase
	<i>aroF</i>	Shikimate dehydrogenase
	<i>rluC</i>	HTH-type transcriptional regulator
	<i>pyrF</i>	Inosose isomerase
	<i>pacS</i>	Putative copper-transporting ATPase
	<i>mobB</i>	Molybdopterin-guanine dinucleotide biosynthesis adapter protein
	<i>mccB</i>	Cystathionine gamma-lyase

表 3 *L. fermentum* F-6 中与益生特性相关的基因Table 3 Genes related to probiotic properties of *L. fermentum* F-6

Genes	Annotation	Probiotic properties
<i>gshAB, garB</i>	Glutathione biosynthesis bifunctional protein glutathione amide reductase	Glutathione biosynthesis
<i>lmrA</i>	Multidrug resistance ABC transporter ATP-binding and permease protein	Production of bioactive peptides
<i>LuxS</i>	S-ribosylhomocysteine lyase (EC 4.4.1.21)	
<i>tagE</i>	Putative poly (glycerol-phosphate) alpha-glucosyltransferase (EC 2.4.1.52)	Improving host metabolism
<i>atpC</i>	ATP synthase epsilon chain	Production of adhesion molecules
<i>ribBA</i>	Riboflavin biosynthesis protein	Riboflavin synthase
<i>ribF</i>	Riboflavin biosynthesis protein	
<i>ribE</i>	Riboflavin synthase (EC 2.5.1.9)	

提高宿主代谢能力(*tagE*)、与乳酸(*ldh*)、黏附分子 (*atpC*)和生物活性肽(*lmrA*、*LuxS*)产生相关的基因,可提高菌株益生功能。另外, *L. fermentum* HFD1 是一株具有产抗菌肽功能的菌株,对其基因组研究发现含有 4 个负责合成抗菌肽的基因^[4]。而 F-6 基因组中存在 2 个相同的抗菌肽基因,推测 F-6 可能具有一定抑菌能力。

特有功能基因分析发现,菌株 F-6 在 24 株 *L. fermentum* 中含有的特有功能基因 *bca* 与耐酸能力有关,可能较 CECT 5716 等其他菌株耐酸性强。对益生功能相关基因分析发现, F-6 含有与抗氧化能力、抑菌能力和黏附能力等相关的益生功能基因。因此,本文基于比较基因组学从基因水平揭示 F-6 具有益生特性相关基因,认为其是一株具有潜在益生功能的菌株。

3 结论

本文利用 NCBI Refseq 数据库中 23 株 *L. fermentum* 全基因组序列和 1 株模式菌株基因组进行比较基因组学分析。系统发育树分析发现分离源不存在明显聚集趋势,与之前研究结果一致。功能注释分析发现,仅 F-6 含有与支链氨基酸合成相关的基因,可为机体提供必需氨基酸。*eps* 基因簇研究发现, F-6 与 YL-11 均含

有大量糖基转移酶和 UDP-葡萄糖 4-表异构酶的基因,且含有 1 个完整的 *eps* 基因簇。此外, F-6 携带有与谷胱甘肽、抗菌肽和黏附分子合成相关的基因,且其含有的特有基因 *gadC* 为耐酸相关基因。综上所述,本文从基因组水平解析到菌株 F-6 是一株具有完整 *eps* 基因簇、耐酸相关基因及与合成必需氨基酸、谷胱甘肽和抗菌肽等相关功能基因的潜在益生菌株,为后续菌株 F-6 基因组研究及益生功能开发奠定了遗传学基础。

参考文献

- [1] 韩丽媛, 李琳, 安新, 王晓平. 阴道正常菌群分布调查分析. 华北理工大学学报: 医学版, 2020, 22(6): 457-461.
Han LY, Li L, An X, Wang XP. Investigation and analysis of normal vaginal flora. *Journal of North China University of Science and Technology: Health Sciences Edition*, 2020, 22(6): 457-461. (in Chinese)
- [2] Goldin BR, Gorbach SL, Saxelin M, Barakat S, Gualtieri L, Salminen S. Survival of *Lactobacillus* species (strain GG) in human gastrointestinal tract. *Digestive Diseases and Sciences*, 1992, 37(1): 121-128.
- [3] Do TBT, Tran TAL, Tran T, Le TH, Jayasena V, Nguyen THC, Nguyen CC, Kim SY, Le QV. Novel exopolysaccharide produced from fermented bamboo shoot-isolated *Lactobacillus fermentum*. *Polymers*, 2020, 12(7): 1531.
- [4] Ozhegov GD, Pavlova AS, Zhuravleva DE, Gogoleva

- NE, Shagimardanova EI, Markelova MI, Yarullina DR, Kayumov AR. Whole genome sequence data of *Lactobacillus fermentum* HFD1, the producer of antibacterial peptides. *Data in Brief*, 2020, 32: 106105.
- [5] Adebo OA, Njobeh PB, Kayitesi E. Fermentation by *Lactobacillus fermentum* strains (singly and in combination) enhances the properties of ting from two whole grain *Sorghum* types. *Journal of Cereal Science*, 2018, 82: 49–56.
- [6] Kellis M, Patterson N, Endrizzi M, Birren B, Lander ES. Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature*, 2003, 423(6937): 241–254.
- [7] Cliften P, Sudarsanam P, Desikan A, Fulton L, Fulton B, Majors J, Waterston R, Cohen BA, Johnston M. Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science*, 2003, 301(5629): 71–76.
- [8] Verce M, De Vuyst L, Weckx S. Comparative genomics of *Lactobacillus fermentum* suggests a free-living lifestyle of this lactic acid bacterial species. *Food Microbiology*, 2020, 89: 103448.
- [9] Lehri B, Seddon AM, Karlyshev AV. Potential probiotic-associated traits revealed from completed high quality genome sequence of *Lactobacillus fermentum* 3872. *Standards in Genomic Sciences*, 2017, 12: 19.
- [10] Sun ZH, Zhang WY, Bilige M, Zhang HP. Complete genome sequence of the probiotic *Lactobacillus fermentum* F-6 isolated from raw milk. *Journal of Biotechnology*, 2015, 194: 110–111.
- [11] 郭元晟, 闫素梅, 史彬林, 倪丽丽, 张和平. 发酵乳杆菌 *L. fermentum* F6 对肉鸡免疫功能的影响. 第六次全国饲料营养学术研讨会论文集. 杨凌, 2010: 354.
- [12] 付果花. 发酵乳杆菌-*Lactobacillus fermentum* F-6 对肉鸡生长性能及消化功能的影响. 内蒙古农业大学硕士学位论文, 2010.
- [13] Bao Y, Zhang YC, Zhang Y, Liu Y, Wang SQ, Dong XM, Wang YY, Zhang HP. Screening of potential probiotic properties of *Lactobacillus fermentum* isolated from traditional dairy products. *Food Control*, 2010, 21(5): 695–701.
- [14] Goris J, Konstantinidis KT, Klappenbach JA, Coenye T, Vandamme P, Tiedje JM. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. *International Journal of Systematic and Evolutionary Microbiology*, 2007, 57(1): 81–91.
- [15] Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 2014, 30(14): 2068–2069.
- [16] Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics: Oxford, England*, 2015, 31(22): 3691–3693.
- [17] 刘文君, 吕瑞瑞, 李伟程, 孙志宏, 陈永福, 张和平. 基于比较基因组学揭示不同植物乳杆菌的遗传特征及菌株差异——以 *Lactobacillus plantarum* P9 和 Lp-6 研究为例. *微生物学报*, 2021, 61(8): 2370–2381. Liu WJ, Lü RR, Li WC, Sun ZH, Chen YF, Zhang HP. Comparative genomics revealed genetic characteristics of different *Lactobacillus plantarum* strains: using P9 and Lp-6 strains as examples. *Acta Microbiologica Sinica*, 2021, 61(8): 2370–2381. (in Chinese)
- [18] Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. *BMC Bioinformatics*, 2007, 8: 172.
- [19] Chen CJ, Chen H, Zhang Y, Thomas HR, Frank MH, He YH, Xia R. TBtools: an integrative toolkit developed for interactive analyses of big biological data. *Molecular Plant*, 2020, 13(8): 1194–1202.
- [20] Morita H, Toh H, Fukuda S, Horikawa H, Oshima K, Suzuki T, Murakami M, Hisamatsu S, Kato Y, Takizawa T, Fukuoka H, Yoshimura T, Itoh K, O’Sullivan DJ, McKay LL, Ohno H, Kikuchi J, Masaoka T, Hattori M. Comparative genome analysis of *Lactobacillus reuteri* and *Lactobacillus fermentum* reveal a genomic island for reuterin and cobalamin production. *DNA Research*, 2008, 15(3): 151–161.
- [21] Illeghems K, De Vuyst L, Weckx S. Comparative genome analysis of the candidate functional starter culture strains *Lactobacillus fermentum* 222 and *Lactobacillus plantarum* 80 for controlled cocoa bean fermentation processes. *BMC Genomics*, 2015, 16: 766.
- [22] Sun Z, Harris HM, McCann A, Guo C, Argimón S, Zhang W, Yang X, Jeffery IB, Cooney JC, Kagawa TF, Liu W, Song Y, Salvetti E, Wrobel A, Rasinkangas P, Parkhill J, Rea MC, O’Sullivan O, Ritari J, Douillard FP, Paul Ross R, Yang R, Briner AE, Felis GE, De Vos WM, Barrangou R, Klaenhammer TR, Caufield PW, Cui Y, Zhang H, O’Toole PW. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated

- genera. *Nature Communications*, 2015, 6: 8322.
- [23] Arahal DR. Whole-genome analyses: average nucleotide identity. *Methods in Microbiology*, 2014, 41: 103–122.
- [24] Doerrler WT, Sikdar R, Kumar S, Boughner LA. New functions for the ancient DedA membrane protein family. *Journal of Bacteriology*, 2013, 195(1): 3–11.
- [25] 卿艳, 郭亚杰, 陈锦瑶, 张立实. 体内外微核试验评价亚硒酸钠的致突变性和抗突变性. *预防医学情报杂志*, 2013, 29(3): 234–237.
- Qing Y, Guo YJ, Chen JY, Zhang LS. Detection of mutagenicity and antimutagenicity of sodium selenite *in vivo* and *in vitro* micronucleus test. *Journal of Preventive Medicine Information*, 2013, 29(3): 234–237. (in Chinese)
- [26] 韦云路, 李全宏. 发酵乳杆菌胞外多糖的结构表征及其生物活性研究. 中国食品科学技术学会第十七届年会摘要集. 西安, 2020: 331–332.
- [27] 扎娃. 乳酸菌胞外多糖合成条件的优化及 *eps* 基因簇分析. 哈尔滨工业大学硕士学位论文, 2020.
- [28] Cui YH, Jiang X, Hao MY, Qu XJ, Hu T. New advances in exopolysaccharides production of *Streptococcus thermophilus*. *Archives of Microbiology*, 2017, 199(6): 799–809.
- [29] Degeest B, De Vuyst L. Correlation of activities of the enzymes alpha-phosphoglucomutase, UDP-galactose 4-epimerase, and UDP-glucose pyrophosphorylase with exopolysaccharide biosynthesis by *Streptococcus thermophilus* LY03. *Applied and Environmental Microbiology*, 2000, 66(8): 3519–3527.
- [30] Sau S, Bhasin N, Wann ER, Lee JC, Foster TJ, Lee CY. The *Staphylococcus aureus* allelic genetic loci for serotype 5 and 8 capsule expression contain the type-specific genes flanked by common genes. *Microbiology: Reading, England*, 1997, 143 (Pt 7): 2395–2405.
- [31] Peran L, Camuesco D, Comalada M, Nieto A, Concha A, Adrio JL, Olivares M, Xaus J, Zarzuelo A, Galvez J. *Lactobacillus fermentum*, a probiotic capable to release glutathione, prevents colonic inflammation in the TNBS model of rat colitis. *International Journal of Colorectal Disease*, 2006, 21(8): 737–746.
- [32] Hidalgo-Cantabrana C, O’Flaherty S, Barrangou R. CRISPR-based engineering of next-generation lactic acid bacteria. *Current Opinion in Microbiology*, 2017, 37: 79–87.
- [33] Brandt K, Nethery MA, O’Flaherty S, Barrangou R. Genomic characterization of *Lactobacillus fermentum* DSM 20052. *BMC Genomics*, 2020, 21(1): 328.
- [34] Tian D, Qin Q, Li MF, Li XY, Xu Q, Lv QZ. Homocysteine impairs endothelial cell barrier function and angiogenic potential via the progranulin/EphA2 pathway. *Frontiers in Pharmacology*, 2021, 11: 614760.
- [35] Michel JL, Madoff LC, Olson K, Kling DE, Kasper DL, Ausubel FM. Large, identical, tandem repeating units in the C protein alpha antigen gene, *bca*, of group B streptococci. *PNAS*, 1992, 89(21): 10060–10064.
- [36] Hersh BM, Farooq FT, Barstad DN, Blankenhorn DL, Slonczewski JL. A glutamate-dependent acid resistance gene in *Escherichia coli*. *Journal of Bacteriology*, 1996, 178(13): 3978–3981.
- [37] Mustafa MI, Makhawi AM. SHERLOCK and DETECTR: CRISPR-cas systems as potential rapid diagnostic tools for emerging infectious diseases. *Journal of Clinical Microbiology*, 2021, 59(3): e0074520.
- [38] Cárdenas N, Laiño JE, Delgado S, Jiménez E, Juárez Del Valle M, Savoy De Giori G, Sesma F, Mayo B, Fernández L, LeBlanc JG, Rodríguez JM. Relationships between the genome and some phenotypical properties of *Lactobacillus fermentum* CECT 5716, a probiotic strain isolated from human milk. *Applied Microbiology and Biotechnology*, 2015, 99(10): 4343–4353.

(本文责编 李磊)