



副格氏乳杆菌 IMAU FB017 的遗传背景和益生相关基因分析

刘旭, 李伟程, 王可欣, 孙志宏, 孙天松*

- 1 内蒙古农业大学乳品生物技术与工程教育部重点实验室, 内蒙古 呼和浩特 010018
- 2 农业农村部奶制品加工重点实验室, 内蒙古 呼和浩特 010018
- 3 内蒙古自治区乳品生物技术与工程重点实验室, 内蒙古 呼和浩特 010018

刘旭, 李伟程, 王可欣, 孙志宏, 孙天松. 副格氏乳杆菌 IMAU FB017 的遗传背景和益生相关基因分析[J]. 微生物学报, 2023, 63(4): 1630-1644.

LIU Xu, LI Weicheng, WANG Kexin, SUN Zhihong, SUN Tiansong. Genetic background and analysis of probiotics related genes in *Lactobacillus paragasseri* IMAU FB017[J]. Acta Microbiologica Sinica, 2023, 63(4): 1630-1644.

摘要: 【目的】副格氏乳杆菌(*Lactobacillus paragasseri*)作为格氏乳杆菌(*Lactobacillus gasseri*)的近源物种,是肠道、生殖道重要菌种之一。本研究团队前期发现 *L. paragasseri* IMAU FB017 具有良好的胃酸和胆盐耐受性,拟从基因组水平解析 IMAU FB017 的遗传背景和功能基因特征,并挖掘潜在益生特性基因,为其开发利用奠定遗传学基础。【方法】采用 Nanopore 和 Illumina 两种测序技术完成了 IMAU FB017 全基因组测序及完成图组装,并结合 NCBI 已公开的 18 株 *L. paragasseri* 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 IMAU FB017 基因组特征。【结果】结果显示 IMAU FB017 基因组包含 1 条环状染色体和 1 个质粒,其中环状染色体大小为 1 880 023 bp, GC 含量为 34.90%, 包含 1 851 个蛋白质编码区(protein-coding sequence, CDS); 质粒大小 43 639 bp, GC 含量为 38.00%。基于 18 株 *L. paragasseri* 基因组序列识别到的 903 个核心基因构建系统发育树,发现 *L. paragasseri* 群体共形成 3 个分支。菌株 IMAU FB017 属于分支 I, 该分支包含菌株最多(13 株, 约占 68%), 且相同分离源的菌株无显著聚集趋势。功能注释分析发现, IMAU FB017 基因组中编码关于 N-乙酰半乳糖胺和半乳糖胺的利用相关基因以及草酸分解代谢相关益生基因, 并且包含 1 条完整的胞外多糖(exopolysaccharides, eps)基因簇和 3 条耐酸耐胆盐相关基因簇且与表型对应。【结论】本研究从基因组水平解析了 IMAU FB017 遗

资助项目: 财政部和农业农村部国家现代农业产业技术体系资助

This work was supported by the China Agriculture Research System of MOF and MARA.

*Corresponding author. E-mail: sts9940@sina.com

Received: 2022-09-15; Accepted: 2022-11-07; Published online: 2022-11-16

传背景, 并挖掘了潜在益生相关基因, 为其开发利用提供参考。

关键词: 副格氏乳杆菌; 比较基因组学; 胞外多糖; 耐酸耐胆盐

Genetic background and analysis of probiotics related genes in *Lactobacillus paragasseri* IMAU FB017

LIU Xu, LI Weicheng, WANG Kexin, SUN Zhihong, SUN Tiansong*

1 Key Laboratory of Dairy Biotechnology and Engineering, Ministry of Education, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China

2 Key Laboratory of Dairy Products Processing, Ministry of Agriculture and Rural Affairs, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China

3 Inner Mongolia Key Laboratory of Dairy Biotechnology and Engineering, Inner Mongolia Agricultural University, Hohhot 010018, Inner Mongolia, China

Abstract: [Objective] *Lactobacillus paragasseri* as a closely related species of *L. gasseri*, is one of the important bacteria in the intestinal and reproductive tract. Our research team previously found that *L. paragasseri* IMAU FB017 had good tolerance to stomach acid and bile salt. Therefore, we intended to analyze the genetic background of IMAU FB017 from the genome level, explore its functional gene characteristics, and dig into its potential probiotic genes, thereby laying a genetic foundation for its development and utilization. [Methods] In this study, the whole genome of IMAU FB017 was sequenced and assembled using Nanopore and Illumina sequencing technologies, and the genomic sequences of 18 *L. paragasseri* strains published by National Center for Biotechnology Information (NCBI) were combined for comparative genomic analysis. Roary software was used to identify core gene sets and pan gene sets. The Rapid Annotation using Subsystem Technology website was used for functional annotation of the genome to explore the genomic characteristics of IMAU FB017. [Results] The results showed that IMAU FB017 genome contained a ring chromosome and a plasmid, and the size of the ring chromosome was 1 880 023 bp, with 34.90% of the guanine/cytosine (G/C) content, containing 1 851 protein-coding sequences. The size of the plasmid was 43 639 bp, and the G/C content was 38.00%. A phylogenetic tree was constructed based on 903 core genes identified from 18 *L. paragasseri* genome sequences. It was found that *L. paragasseri* population formed three clades. Strain IMAU FB017 belonged to clades I, which contained the most strains (13 strains, about 68%), and the strains from the same source had no significant aggregation trend. Functional annotation analysis found that the IMAU FB017 genome encoded genes related to N-acetylgalactosamine and galactosamine utilization, as well as probiotic genes related to oxalate catabolism, and contained a complete exopolysaccharide (EPS) gene cluster and three acid and bile salt tolerance gene clusters corresponding to the phenotype. [Conclusion] In this paper, the genetic background of IMAU FB017 has been analyzed at the genomic level, and the potential probiotic genes have been mined to provide references for its development and utilization.

Keywords: *Lactobacillus paragasseri*; comparative genomics; exopolysaccharides; acid and bile salt resistance

格氏乳杆菌(*Lactobacillus gasseri*)于1980年首次被鉴定为一种同型发酵兼性厌氧菌^[1],广泛存在于人类口腔、胃肠道、阴道和母乳中,具有提高人体免疫、维持阴道菌群平衡等益生特性^[2]。2018年Tanizawa等通过平均核苷酸一致性(average nucleotide identity, ANI)、DNA指纹图谱以及表型特征分析发现,JCM 5344和JCM 1130是一个不同于*L. gasseri*的新物种,并将其命名为*L. paragasseri*^[3]。2020年Zhou等从中国粪便中分离出的92株*L. gasseri*中同样存在*L. paragasseri*^[4]。自*L. gasseri*分类学地位的重新划分以来,*L. paragasseri*得到了广泛研究。2020年Nilsen等发现*L. paragasseri* K7的无细胞上清液中存在3种活性肽,对*L. iners*均有抑制活性^[5]。2021年Mehra等对*L. paragasseri* UBLG-36(粪便)进行了测序和分析,发现*L. paragasseri* UBLG-36编码参与草酸盐分解代谢的基因,该基因被证明与肾草酸结石的治疗相关,并且该菌在体外草酸盐的降解实验中具有显著效果^[6]。*L. paragasseri*作为潜在益生菌具有广泛开发与应用前景。

比较基因组学是比较不同细菌的基因组特征和基因组结构,包括DNA序列、基因、调控序列等,从基因组的全部或大部分进行比较,去研究生物体之间的生物学相似和差异性以及进化关系^[7]。2006年Ventura等利用比较基因组学发现*L. johnsonii*和*L. gasseri*、*L. acidophilus*和*L. delbrueckii*包含的原噬菌体序列与宿主的进化并不相似,并且在种间和种内均表现出高度的多样性^[8]。Bukhari等通过比较基因组和泛基因组对*Akkermansia muciniphila*进行研究,确定了一种可以系统去除不需要基

因的方法,实现了预测最小基因组的目标^[9]。2021年Zhao等对发酵粘液乳杆菌进行比较基因组分析发现,发酵粘液乳杆菌FXJCJ61和CECT5716基因组中的多个特定基因协同调节结肠中短链脂肪酸的产生,并且这2株菌比其他测试菌株更能显著减少DSS诱导的结肠炎^[10]。因此,对菌株进行比较基因组学分析对明确其遗传背景具有重要意义。

团队前期已完成了对IMAU FB017的体外耐酸耐胆盐实验,在模拟人工胃液3 h、8 h的存活率分别为89.45%、36.6%;肠液4 h存活率为61.19%;胆盐延滞期为0.51 h,发现其具有良好的耐受性^[11],具有成为益生菌的潜力。因此本研究将以IMAU FB017为研究对象,结合NCBI Refseq数据库下载的18株*L. paragasseri*和4株*L. gasseri*进行比较基因组学分析,从基因组水平解析IMAU FB017遗传背景,并挖掘其潜在的功能基因组,为其开发利用奠定基础。

1 材料与amp;方法

1.1 实验菌株

菌株IMAU FB017(accession number: CP102851-CP102852)于2011年从西藏自治区拉萨市城关区一位健康成人粪便中分离得到,现保藏在内蒙古农业大学乳酸菌菌种资源库(lactic acid bacteria collection center, LABCC)。

1.2 主要试剂和仪器

Wizard[®] Genomic DNA Purification Kit(普洛麦格生物技术有限公司);MRS培养基(Oxoid公司);SMRT 文库试剂盒DNA Template Prep Kit 1.0和测序试剂DNA/Polymerase Binding Kit P6 v2和DNA Sequencing Bundle 4.0(太平洋生物科

学公司); 电热恒温培养箱(北京一恒科技有限公司); 超微量紫外分光光度计(NanoDrop 公司); Qubit 2.0 荧光计(Life Technologies); Illumina NovaSeq 6 000 (Illumina 公司); Oxford Nanopore PromethION24 (Oxford Nanopore Technologies)。

1.3 副格式乳杆菌基因组序列

从 NCBI Refseq 数据库中进行 4 株 *L. gasseri* 和 18 株 *L. paragasseri* 基因组序列的下载 (JV-V03 和 JCM5343^T 仅使用完成图), 以及一株获得全基因组序列的实验室菌株 IMAU FB017 菌株基本信息见表 1。

1.4 菌株培养和基因组 DNA 的提取

采用 TIANGEN 试剂盒对 IMAU FB017 完成

基因组 DNA 提取, 具体方法参照试剂盒说明书。提取的 DNA 通过 1% 琼脂糖凝胶电泳进行完整度和纯度检测, 使用 Nanodrop 微量紫外分光光度计测定 DNA 的浓度及 A_{260}/A_{280} 和 A_{260}/A_{280} 比值。质量达到建库上机要求后, 分别利用 Illumina NovaSeq 6000 和 Oxford Nanopore PromethION24 平台完成二代和三代测序。

1.5 基因组测序和组装

基于 Nanopore 三代测序得到的原始 reads, 质控后获得高质量序列, 通过 NECAT 软件^[12]进行组装, 组装后利用 Pilon 软件^[13]进行二代数据对三代数据的校正, 并使用 Circlator (V1.5.5)^[14]软件将校正后的三代数据进行环化并最终得到 IMAU

表 1 四株 *Lactobacillus gasseri* 和 19 株 *Lactobacillus paragasseri* 的基本信息

Table 1 Basic information of 4 *Lactobacillus gasseri* strains and 19 *Lactobacillus paragasseri* strains

Organism name	Strain	Size (Mb)	GC (%)	CDS	NCBI Refseq assembly No.	Isolate source
<i>Lactobacillus paragasseri</i>	JV-V03	2.030 30	34.7	1 845	GCA_005886075.1	Vagina
	JCM 5343 ^T	1.937 63	34.9	1 721	GCA_003584685.1	Feces
	NCTC13720	1.982 16	35.0	1 782	GCA_900636895.1	Vagina
	NCK1347	1.944 20	35.0	1 709	GCA_009734365.1	Feces
	1001095IJ_161003_B4	2.018 49	34.5	1 826	GCA_015550525.1	Feces
	JCM 1130	2.027 61	34.9	1 855	GCA_003307255.1	Feces
	ADH	2.013 40	34.9	1 848	GCA_018588245.1	Feces
	J1101004_170508_E2	1.915 18	34.7	1 701	GCA_015552765.1	Feces
	JCM 11657	1.932 89	34.9	1 790	GCA_019972195.1	Miss
	UBLG-36	1.980 60	34.8	1 743	GCA_019459585.1	Vagina
	JCM 5344	1.969 16	34.7	1 825	GCA_003307295.1	Vagina
	UMB6975	2.108 39	34.8	1 932	GCA_007785845.1	Catheter
	D6t1_180914_C1	1.997 14	34.8	1 844	GCA_015560405.1	Feces
	Indica	1.932 64	34.7	1 653	GCA_010092655.1	Miss
	UBLG-36	1.941 91	34.9	1 766	GCA_014269295.2	Feces
	UMB1065	1.805 80	35.2	1 669	GCA_007785885.1	Catheter
	UMB0596	1.714 81	35.3	1 576	GCA_007785945.1	Catheter
	K7	1.994 16	34.8	1 830	GCA_000406345.2	Feces
	IMAU FB017	1.963 66	34.9	1 851	CP102851	Feces
IMAU FB017 (plasmid)	43 639	38.0	54	CP102852	Feces	
<i>Lactobacillus gasseri</i>	ATCC 33323 ^T	1.894 36	35.3	1 728	GCA_000014425.1	Miss
	HL20	2.176 46	35.0	2 035	GCA_017638885.1	Feces
	DSM14869	1.951 55	35.0	1 704	GCA_002287905.1	Vagina
	CECT5714	1.907 79	35.0	1 740	GCA_019192945.1	Breast milk

FB017 (登录号: CP102851–CP102852)的全基因组序列。

1.6 比较基因组分析

1.6.1 平均核苷酸一致性(ANI)计算

本研究中, 4株 *L. gasseri* 和 19株 *L. paragasseri* 的 ANI 值参考 Jain 等^[15]报道的方法计算菌株间 ANI 值。再利用 TBtools^[16]软件绘制 ANI 聚类热图。

1.6.2 系统发育树构建

利用 Prokka^[17]软件对菌株基因组进行基因预测后, 经 Roary^[18]软件分析得到的核心基因序列利用 TreeBest 软件(<http://www.mybiosoftware.com/treebest>)中邻接法(neighbor-joining, NJ)进行系统发育树构建。使用 iTol 在线软件对系统发育树进行可视化(<https://itol.embl.de/>)。

1.6.3 功能基因组预测和注释

将组装好的 IMAU FB017 和其他 *L. paragasseri* 的核酸序列文件分别上传至 Rapid Annotation using Subsystem Technology (RAST, <https://rast.nmpdr.org/>)数据库, 完成 *L. paragasseri* 基因组的功能注释。

1.6.4 耐药基因与毒力基因注释

将组装好的 IMAU FB017 和其他 *L. paragasseri* 的核酸序列文件分别上传至在线注释 CARD (<http://card.mcmaster.ca>)和 VFAnalyzer (<http://www.mgc.ac.cn/cgi-bin/VFs/v5/main.cgi?func=VFAnalyzer>)进行注释, 下载其注释文件。统计 *L. paragasseri* 注释结果(过滤一致性<75%的基因)。

2 结果与讨论

2.1 *L. paragasseri* 基因组基本信息

对样品 IMAU FB017 测序数据评估组装, 拼接成全基因组圈图(图 1A)。IMAU FB017 基因组序列包含 1 条环状染色体和 1 个质粒, 其中环

状染色体大小为 1 880 023 bp, GC 含量为 34.9%, 包含 1 851 个蛋白质编码区 (protein-coding sequence, CDS); 质粒大小 43 639 bp, GC 含量为 38%。统计已注册的 19 株 *L. paragasseri* 的基因组信息, 结果见表 1, 19 株 *L. paragasseri* 的基因组大小为 1.71–2.11 Mb, 平均为 1.95 Mb; GC 含量为 34.5%–35.3%, 平均为 34.85%。通过比较 *L. paragasseri* 不同分离源基因组大小、GC 含量和 CDS 数量发现, 该菌分离源与其基因组大小(图 1B)、GC 含量(图 1C)和 CDS 数量(图 1D)均无相关性。

2.2 ANI 分析

ANI 是通过比对基因组的同源序列以鉴定菌株亲缘关系^[19]。在比较基因组学分析过程中, ANI 可用于评估基因组间多态性的程度, 也可判断基因组之间的相似性, 一般认为 ANI 值大于 95%即为同一物种^[20]。2018 年 Ciufu 等将 95%的阈值改为 96%, 并以此作为物种边界^[21]。

为解析 *L. paragasseri* 遗传多样性, 本研究对 18 株 *L. paragasseri*、4 株 *L. gasseri* 和 IMAU FB017 两两之间进行 ANI 计算, 并构建聚类热图。结果如图 2 所示, *L. gasseri* 和 *L. paragasseri* 各自聚为一分支。将 *L. paragasseri* 所在分支命名为副格氏乳杆菌分支; *L. gasseri* 所在分支命名为格氏乳杆菌分支。格氏乳杆菌分支中 *L. gasseri* 两两之间 ANI 值均大于 98%, 且平均 ANI 值为 98.99%; 副格氏乳杆菌分支中 *L. paragasseri* 两两之间 ANI 值均大于 97%, 且平均 ANI 值为 98.27%; 两分支间菌株 ANI 值均小于 95%。Zhou 等^[4]发现分离自中国人粪便的 92 株 *L. gasseri* 中也存在 *L. paragasseri*, 且 *L. paragasseri* 与 *L. gasseri* 存在较大差异, 本研究结果与该报道结果一致。菌株 IMAU FB017 位于 *L. paragasseri* 分支, 与模式菌株 JCM5343^T 的 ANI 值为 98.21%。

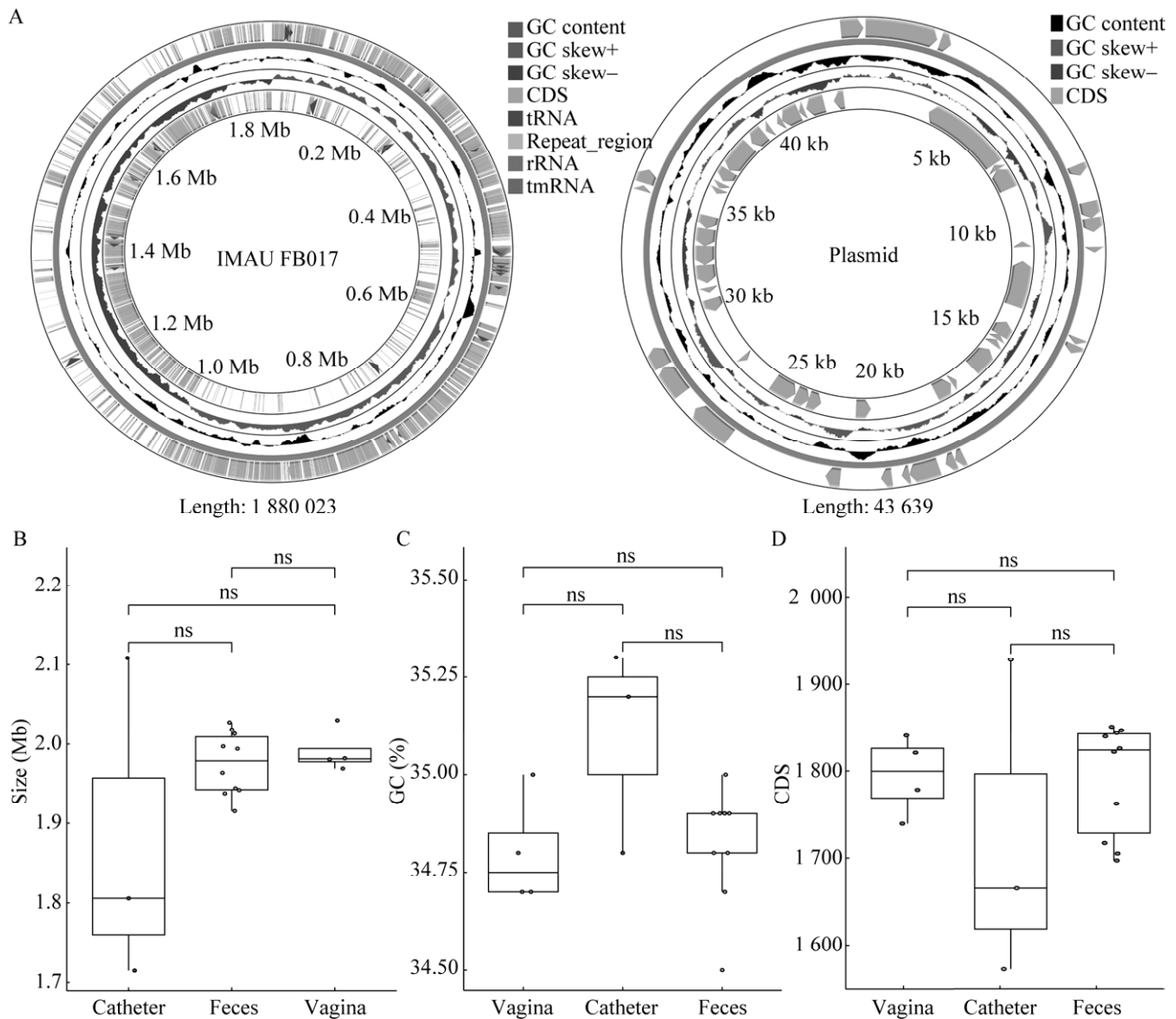


图 1 *Lactobacillus paragasseri* 基因组一般特征分析

Figure 1 *Lactobacillus paragasseri* genome general characteristics analysis. A: IMAU FB017 genome circle map of *L. paragasseri*. Isolation source and genome size (B), GC content (C), CDS (D) box map ns: No significance.

2.3 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,其中NJ法对于短序列重建树准确度高、速度快是最有效的基于距离重建系统发育树的方法^[22]。因此为了进一步验证 ANI 的结果,并评估菌株之间的遗传距离,本研究采用 NJ 法,利用 Roary 软件,bootstrap 值为 1 000,构建 2 棵系

统发育树,一棵基于 19 株 *L. paragasseri*、4 株 *L. gasseri* 识别得到的 786 个核心基因构建(图 3A)(加入外群可以确定 *L. paragasseri* 的根部及其进化位置);另一棵去掉外群只用 19 株 *L. paragasseri* 识别得到的 903 个核心基因构建(图 3B)(外群的选择会让树形发生变化,因此判断 *L. paragasseri* 种内遗传关系,需将外群移除重新构建发育树)。

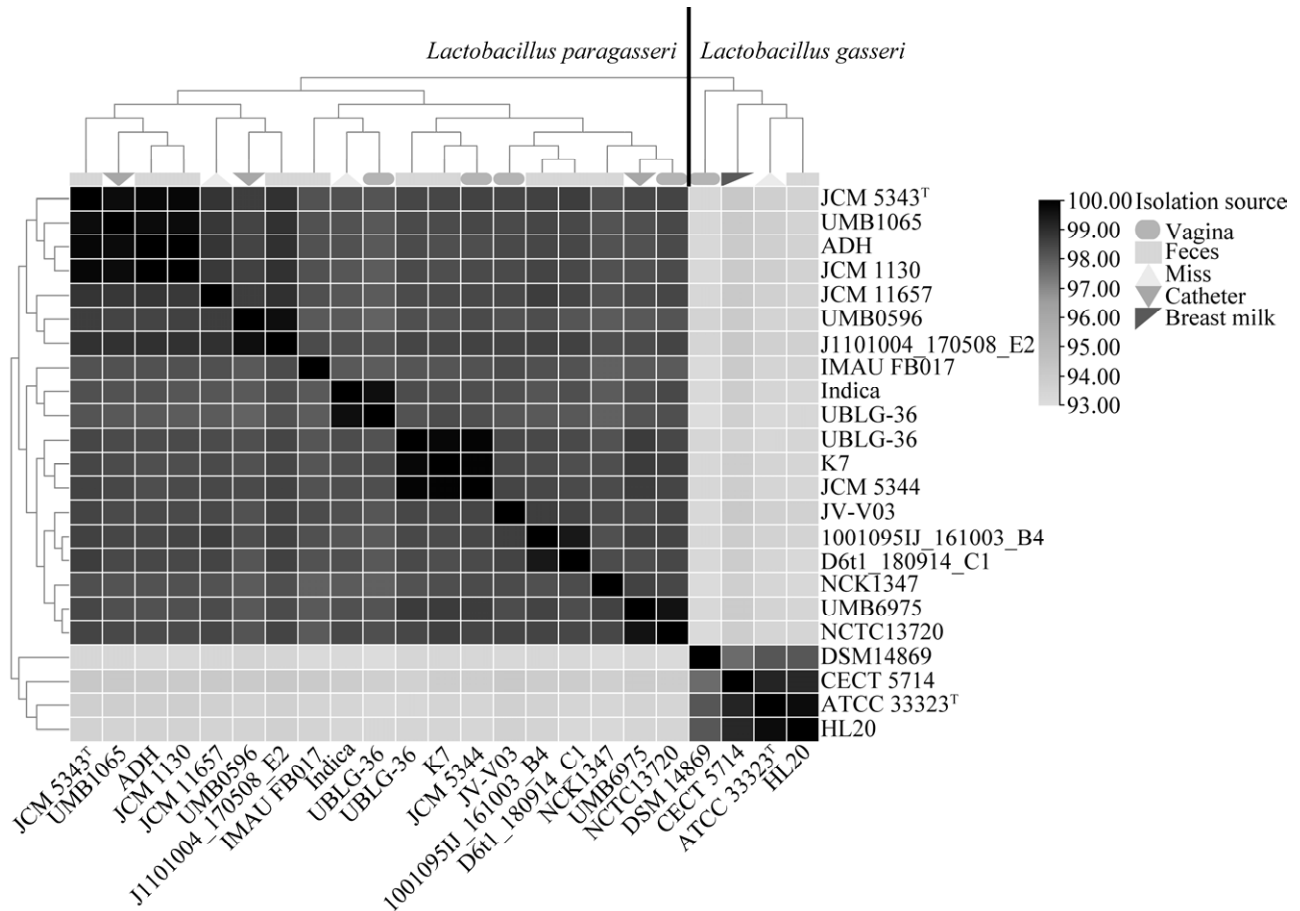


图 2 *Lactobacillus gasseri* 和 *Lactobacillus paragasseri* 的 ANI 值聚类热图

Figure 2 Heatmap of ANI based on the genome sequences of *Lactobacillus gasseri* and *Lactobacillus paragasseri*.

由图 3A 可知, 23 株菌主要分为两大分支。4 株 *L. gasseri* (DSM14869、CECT5714、HL20、DSM33323^T) 为一分支; 19 株 *L. paragasseri* 为一分支。上述结果发现 *L. paragasseri* 和 *L. gasseri* 遗传距离较远, 表明 2 个物种之间存在较大差异。

由图 3B 可知 19 株 *L. paragasseri* 菌群种内共分为 3 个大分支, 7 个亚分支, 其中分支 I 含 5 个亚分支, 包含菌株最多(约 68%)且模式菌株 JCM 5343^T 也在该分支, 平均每 2-3 个菌株就可形成一个新分支, 表明 *L. paragasseri* 存在较高的种内遗传多样性。

系统发育树显示 19 株 *L. paragasseri* 的遗传

关系与其分离地、分离源均无关, 与 Zhou 等对中国人粪便中分离出的 *L. paragasseri* 菌株分析发现一致^[4]。其中 IMAU FB017 (D) 与益生菌 UBLG-36 (粪便)^[6] (II) 处在不同分支, 遗传距离较远, 差异较大。

2.4 RAST 功能注释

本研究利用 RAST 网站对 19 株 *L. paragasseri* 基因组序列进行功能注释, 分析分离源与其功能之间的差异。将 19 株 *L. paragasseri* 按其分离源进行分组, 对其功能基因进行主成分分析(图 4A), 发现功能基因差异与其分离源不相关。

注释到的 *aga* 基因簇, 见表 2。*aga* 和 *gam* 基因簇均负责调控 N-乙酰半乳糖胺和半乳糖胺的利用^[23]。菌株 IMAU FB017 相比 UBLG-36 (粪便) 有更多关于 N-乙酰半乳糖胺和半乳糖胺的利用相关基因。同时两者均注释到草酸分解代谢相关酶的基因, 见表 2, 如甲酰辅酶 A 转移酶 (formyl-CoA transferase, Frc) 和甲酰辅酶 A 脱羧酶 (oxalyl coenzyme A decarboxylase, Oxc) 基因。有研究报道, 草酸盐代谢通路主要包括 3 种草酸降解酶参与草酸的代谢, 即草酸-甲转运体

(oxalate-formate antiporter, Oxlt)、Frc、Oxc^[24], 且该 3 种酶被证明与预防肾草酸结石的治疗相关^[6]。人体因为缺乏草酸代谢相关的代谢酶类, 因此维持机体正常运行, 只能通过持续摄入、排泄或者通过肠道菌群的降解来维持正常草酸水平。如果平衡被打乱, 草酸钙结石的形成风险就会明显增加^[25]。有实验证明益生菌 UBLG-36 (粪便) 在体外草酸盐的降解实验中具有显著效果^[6]。菌株 IMAU FB017 拥有与 UBLG-36 (粪便) 相同功效酶的基因。

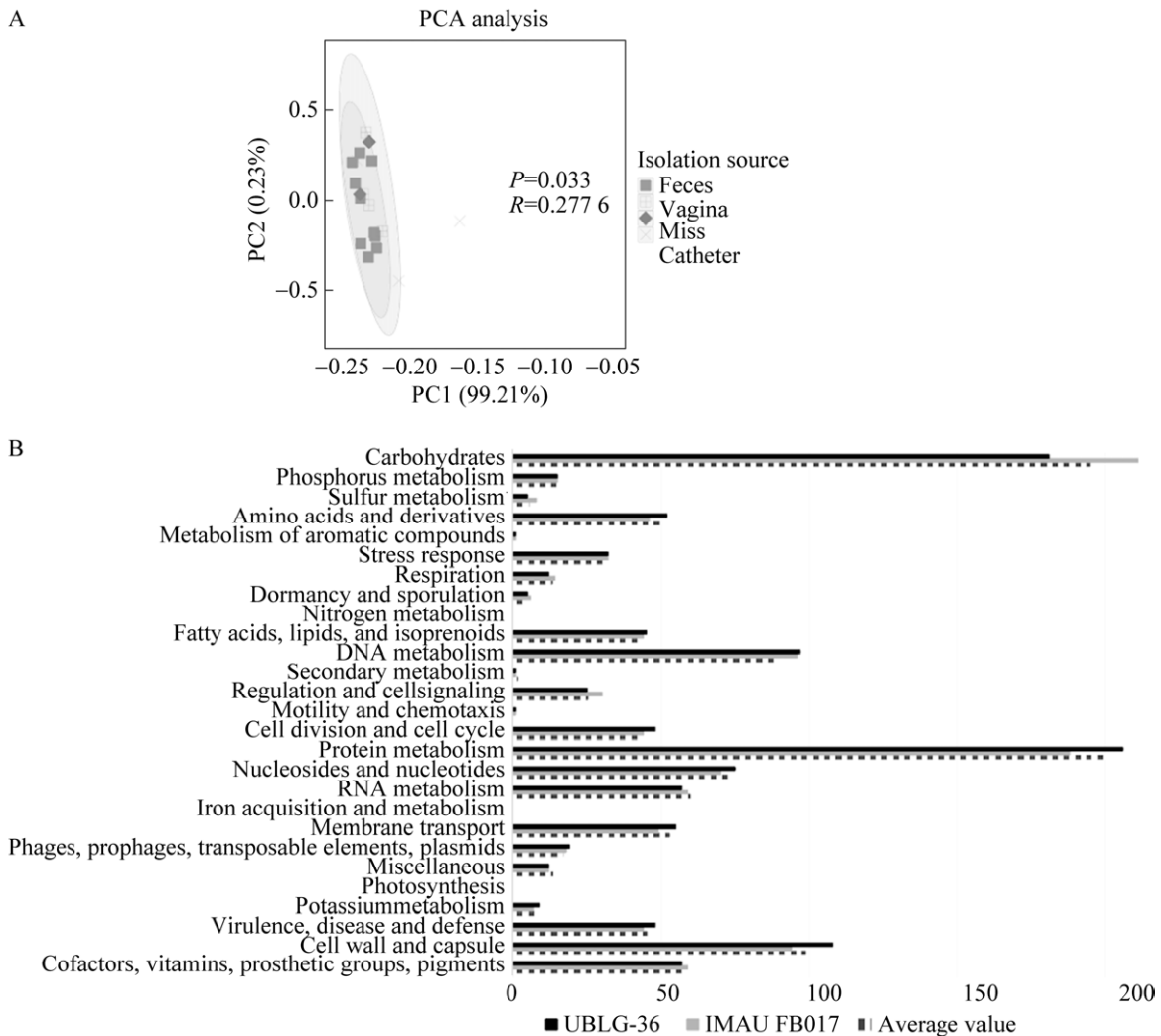


图 4 *Lactobacillus paragasseri* 功能基因分析

Figure 4 Functional gene analysis of *Lactobacillus paragasseri*. Principal component analysis of functional genes (A) and RAST functional annotation (B).

表 2 IMAU FB017 和 UBLG-36 (粪便) 编码的 N-乙酰半乳糖胺和半乳糖胺的利用及草酸代谢相关基因
Table 2 Genes encoded by N-acetyl-galactosamine and galactosamine utilization and oxalate catabolism of IMAU FB017 and UBLG-36 (feces)

Subsystem feature counts	Gene	IMAU FB017	UBLG-36 (feces)
N-acetyl-galactosamine and galactosamine utilization	<i>aga</i>	2	0
	<i>gam</i>	3	1
Oxalate catabolism	<i>frc</i>	1	1
	<i>oxc</i>	1	1
	<i>oxlt</i>	0	0

Numbers represent gene copy numbers.

2.5 胞外多糖(EPS)相关基因分析

研究发现, 乳酸菌产生的胞外多糖(exopolysaccharides, EPS)具有多种对健康有益的功效, 如调节机体免疫功能、抗氧化效应等, 且其独特的物理化学和生物活性使其成为许多工业应用的候选物, 例如食品工业中的稳定剂、增黏剂、乳化剂或胶凝剂^[26-27]。2018年 Rani 等发现 *L. gasseri* FR4 产生的 *eps* 可抑制多种食源性致病菌的生长, 也可作为抗氧化剂用于食品工业^[28]。有研究报道 *L. gasseri* 和 *L. paragasseri* 作为近源种之间存在着大量的水平转移^[4]。因此对 19 株 *L. paragasseri* 基因组序列分析, 查找与 *eps* 合成相关的基因簇。图 5A 为 *eps* 合成相关基因的存在情况。分析发现, *L. paragasseri* 存在 *eps* 基因簇与分离源无关。

有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^[29-30]。另外, *galE* 编码的 UDP-葡萄糖 4-表异构酶的活性与 *eps* 的产量有关^[31]; *epsJ*、*epsF*、*epsA*、*epsC* 和 *epsD* 编码的糖基转移酶起聚合和输出作用^[32]。由图 6A 可知, 所有 *L. paragasseri* 均携带编码糖基转移酶和 UDP-葡萄糖 4-表异构酶的基因, 其中 IMAU FB017 和 NCK1347 相较其他 *L. paragasseri* 拥有更多糖基转移酶。

为进一步确定 IMAU FB017 的 *eps* 相关基因

在基因组中的分布情况, 对其基因组进行注释发现存在一条完整的 *eps* 基因簇与 *L. gasseri* ATCC33323^T 的 *eps* 基因簇相似, 相似度 97.66% (图 5B), 发现在 *epsA*、*epsC* 和 *epsD* 这 3 个基因的上下游存在大量 *eps* 合成相关基因。IMAU FB017 的 *eps* 基因簇长度为 20 544 bp, ATCC33323^T 的 *eps* 基因簇长度为 17 821 bp。由图 5B 可知, 2 株菌的 *eps* 基因簇均含有大量糖基转移酶基因; 2 个基因簇均含有多个多糖合成基因, 多糖合成基因 *cap8A*、*cpsM* 和 *cpsF* 编码的荚膜多糖生物合成蛋白可能控制多糖链的长度和多糖的输出^[33]。2 个基因簇均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。

上述结果表明, IMAU FB017 与已知产胞外多糖的菌株 *L. gasseri* ATCC33323^T 携带的 *eps* 基因簇较为一致, 同时 IMAU FB017 在培养过程中菌液较为黏稠浑浊, 结合以上信息推测其浑浊物可能为胞外多糖, 关于 IMAU FB017 的 EPS 分析仅为初步推测。

2.6 耐酸耐胆盐相关基因

益生菌能否耐受胃肠道环境是其在肠道发挥生理效益的决定性因素^[34]。2021年 Pan 等^[35]对一株具有胆盐耐受性的唾液乳杆菌进行基因敲除实验发现, 菌株胆盐耐受性可能与分子伴侣、PTS 和肽聚糖的合成基因有关。因此对 18 株

L. paragasseri 基因组进行 RAST 功能注释, 并使用 roary 软件识别并统计分子伴侣、PTS、肽聚糖和耐酸耐胆盐相关基因, 结果如图 6A 所示。所有菌株均注释到分子伴侣、PTS 和肽聚糖的合成相关基因, 但是相关基因在菌株基因组中携带数量不同, 其中 NCTC13720、UBLG-36、IMAU FB017 相关基因携带数最多, 并且益生菌 UBLG-36 (粪便)^[36]和 IMAU FB017 已被证明具

有良好耐酸耐胆盐特性^[11]。耐酸耐胆盐涉及机制、相关基因较多, 存在菌株差异, 因此我们对所有菌株进行耐酸耐胆盐基因簇预测, 发现所有菌株均存在耐酸耐胆盐相关基因簇, 但是除菌株 UBLG-36 (粪便)、D6t1_180914_C1、1001095IJ_161003_B4、JCM11657、JCM5343、UMB1065、K7 和 NCTC13720 外, 其他菌株注释到的基因簇与已知基因簇相似度较低(<45%)。

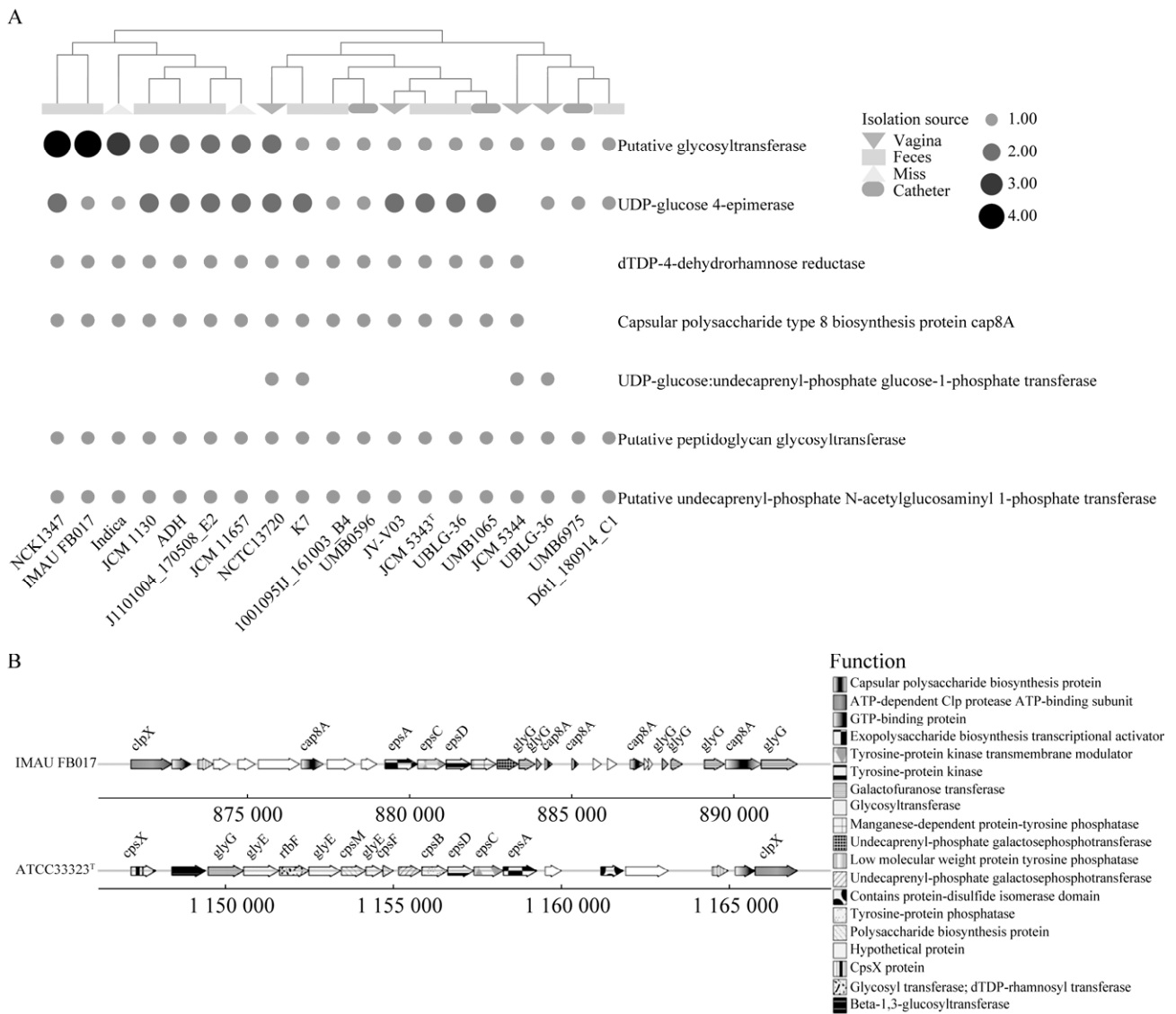


图 5 19 株 *Lactobacillus paragasseri* eps 基因簇分析

Figure 5 *Lactobacillus paragasseri* eps gene cluster analysis. Distribution of eps gene clusters (A) and structure of eps gene cluster (B).

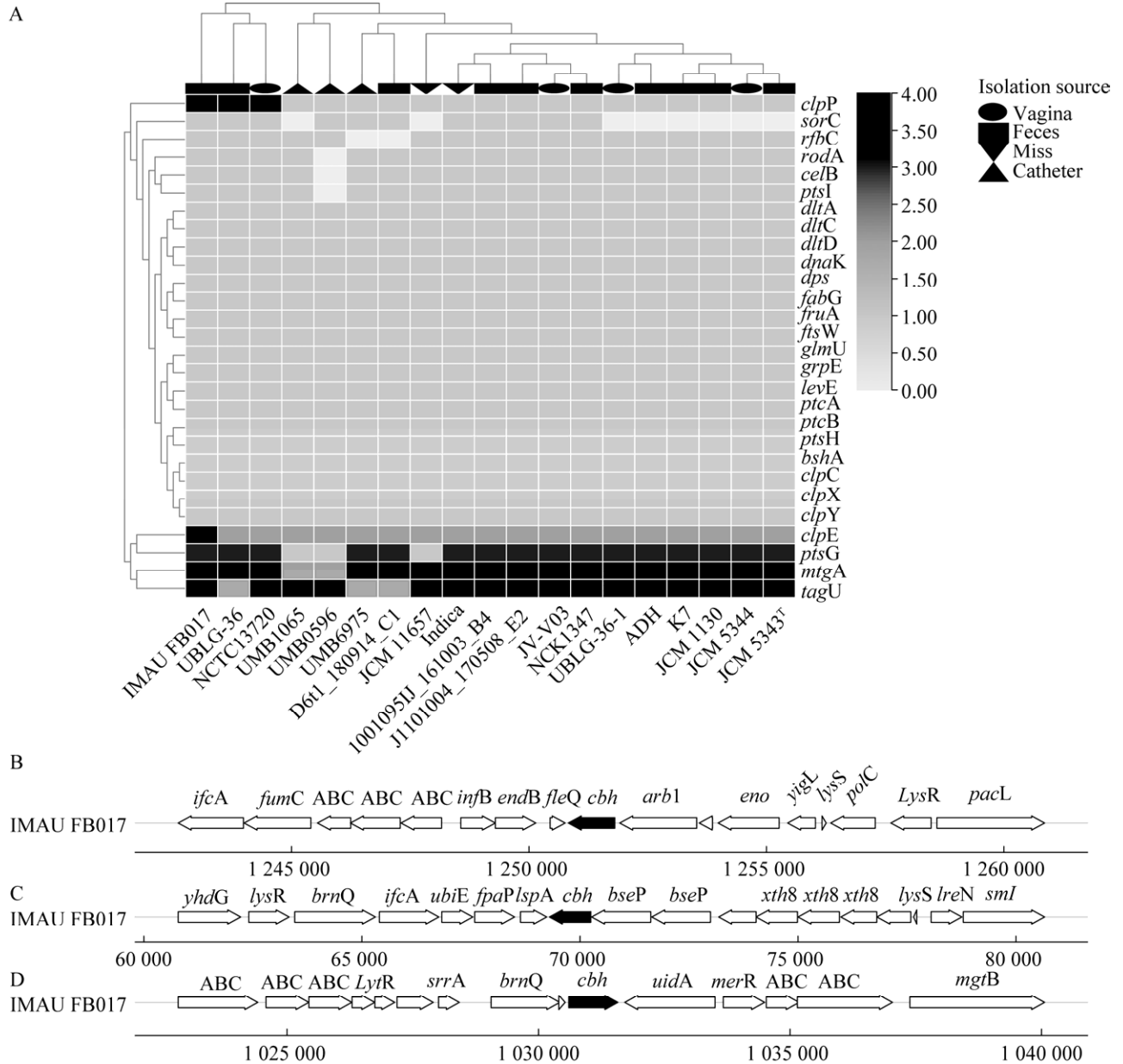


图 6 耐酸耐胆盐分析

Figure 6 Acid and bile salt resistance analysis. Acid and bile salt tolerance related genes (A) and gene cluster distribution (B, C, D).

IMAU FB017的基因组中注释到的3条与耐酸耐胆盐相关的基因簇分别如图 6B-6D 所示, 其与 *Lactobacillus reuteri* JCM 1112 耐酸耐胆盐基因簇^[37]相似, 相似度分别为 55.83%、55.83%、72.52%。截至目前已报道具有耐酸耐胆盐菌株

有 UBLG-36 (粪便)、JCM5343、IMAU FB017。

综上所述,我们发现基因组中携带关于分子伴侣、PTS、肽聚糖合成相关基因数量与完整的耐酸耐胆盐基因簇共同影响菌株的耐酸耐胆盐特性。

2.7 耐药基因与毒力基因

细菌耐药性是主要的公共卫生威胁,而细菌的耐药性通常是由编码基因确定,因此菌株是否携带耐药基因至关重要。通常我们认为2个基因相似性<75%则不具备等同功能^[38]。

19株 *L. paragasseri* 的耐药基因和毒力基因注释分析发现均不携带耐药基因和毒力基因。

3 结论

本研究采用 Nanopore 和 Illumina 两种测序技术完成了 IMAU FB017 全基因组测序及完成图组装。结果显示 IMAU FB017 基因组包含1条环状染色体和1个质粒。结合 NCBI 已公开的18株 *L. paragasseri* 基因组序列,基于903个核心基因构建系统发育树,发现系统发育树存在多个亚分支,且相同分离源的菌株无显著聚集趋势。功能注释分析发现,IMAU FB017 编码1条完整的 *eps* 基因簇、3条与耐酸耐胆盐相关基因簇、N-乙酰半乳糖胺和半乳糖胺的利用相关基因以及草酸分解代谢相关基因,且未注释到潜在耐药基因与毒力基因。本研究从基因组水平揭示了 IMAU FB017 是一株具有完整 *eps* 基因簇、耐酸耐胆盐基因簇和耐酸相关基因的潜在益生菌株,为后续 IMAU FB017 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] LAUER E, KANDLER O. *Lactobacillus gasseri* sp. nov., a new species of the subgenus *Thermobacterium*[J]. Zentralblatt Für Bakteriologie: I Abt Originale C: Allgemeine, Angewandte Und Ökologische Mikrobiologie, 1980, 1(1): 75-78.
- [2] MALDONADO-BARRAGÁN A, CABALLERO-GUERRERO B, MARTÍN V, RUIZ-BARBA JL, RODRÍGUEZ JM. Purification and genetic characterization of gassericin E, a novel co-culture inducible bacteriocin from *Lactobacillus gasseri* EV1461 isolated from the vagina of a healthy woman[J]. BMC Microbiology, 2016, 16: 37.
- [3] TANIZAWA Y, TADA I, KOBAYASHI H, ENDO A, MAENO S, TOYODA A, ARITA M, NAKAMURA Y, SAKAMOTO M, OHKUMA M, TOHNO M. *Lactobacillus paragasseri* sp. nov., a sister taxon of *Lactobacillus gasseri*, based on whole-genome sequence analyses[J]. International Journal of Systematic and Evolutionary Microbiology, 2018, 68(11): 3512-3517.
- [4] ZHOU XY, YANG B, STANTON C, ROSS RP, ZHAO JX, ZHANG H, CHEN W. Comparative analysis of *Lactobacillus gasseri* from Chinese subjects reveals a new species-level taxa[J]. BMC Genomics, 2020, 21(1): 119.
- [5] NILSEN T, SWEDEK I, LAGENAUR LA, PARKS TP. Novel selective inhibition of *Lactobacillus iners* by *Lactobacillus*-derived bacteriocins[J]. Applied and Environmental Microbiology, 2020, 86(20): e01594-e01520.
- [6] MEHRA Y, VISWANATHAN P. High-quality whole-genome sequence analysis of *Lactobacillus paragasseri* UBLG-36 reveals oxalate-degrading potential of the strain[J]. PLoS One, 2021, 16(11): e0260116.
- [7] 张少存. 两株溶血不动杆菌 TJS01 和 TJR01 基因组学研究[D]. 天津: 天津大学硕士学位论文, 2017. ZHANG SC. Genomics research of *Acinetobacter haemolyticus* TJS01 and TJR01[D]. Tianjin: Master's Thesis of Tianjin University, 2017 (in Chinese).
- [8] VENTURA M, CANCHAYA C, BERNINI V, ALTERMANN E, BARRANGOU R, MCGRATH S, CLAESSON MJ, LI Y, LEAHY S, WALKER CD, ZINK R, NEVIANI E, STEELE J, BROADBENT J, KLAENHAMMER TR, FITZGERALD GF, O'TOOLE PW, van SINDEREN D. Comparative genomics and transcriptional analysis of prophages identified in the genomes of *Lactobacillus gasseri*, *Lactobacillus salivarius*, and *Lactobacillus casei*[J]. Applied and Environmental Microbiology, 2006, 72(5): 3130-3146.
- [9] BUKHARI SAR, IRFAN M, AHMAD I, CHEN LJ. Comparative genomics and pan-genome driven prediction of a reduced genome of *Akkermansia muciniphila*[J]. Microorganisms, 2022, 10(7): 1350.
- [10] ZHAO Y, ZHANG CC, YU LL, TIAN FW, ZHAO JX, ZHANG H, CHEN W, ZHAI QX. Phylogenetic and comparative genomic analysis of *Lactobacillus fermentum* strains and the key genes related to their intestinal anti-inflammatory effects[J]. Engineering, 2021.

- [11] 郭艳荣. 格氏乳杆菌体外筛选及高密度培养工艺研究[D]. 呼和浩特: 内蒙古农业大学硕士学位论文, 2020.
GUO YR. The *in vitro* screening of *Lactobacillus gasseri* and the research of high cell density culture[D]. Hohhot: Master's Thesis of Inner Mongolia Agricultural University, 2020 (in Chinese).
- [12] CHEN Y, NIE F, XIE SQ, ZHENG YF, DAI Q, BRAY T, WANG YX, XING JF, HUANG ZJ, WANG DP, HE LJ, LUO F, WANG JX, LIU YZ, XIAO CL. Efficient assembly of nanopore reads via highly accurate and intact error correction[J]. *Nature Communications*, 2021, 12(1): 60.
- [13] WALKER BJ, ABEEL T, SHEA T, PRIEST M, ABOUELLIEL A, SAKTHIKUMAR S, CUOMO CA, ZENG QD, WORTMAN J, YOUNG SK, EARL AM. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement[J]. *PLoS One*, 2014, 9(11): e112963.
- [14] HUNT M, SILVA ND, OTTO TD, PARKHILL J, KEANE JA, HARRIS SR. Circlator: automated circularization of genome assemblies using long sequencing reads[J]. *Genome Biology*, 2015, 16: 294.
- [15] JAIN C, RODRIGUEZ-R LM, PHILLIPPY AM, KONSTANTINIDIS KT, ALURU S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries[J]. *Nature Communications*, 2018, 9(1): 5114.
- [16] 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[J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [17] SEEMANN T. Prokka: rapid prokaryotic genome annotation[J]. *Bioinformatics*, 2014, 30(14): 2068-2069.
- [18] 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[J]. *Bioinformatics*: Oxford, England, 2015, 31(22): 3691-3693.
- [19] SUN ZH, HARRIS HMB, MCCANN A, GUO CY, ARGIMÓN S, ZHANG WY, YANG XW, JEFFERY IB, COONEY JC, KAGAWA TF, LIU WJ, SONG YQ, SALVETTI E, WROBEL A, RASINKANGAS P, PARKHILL J, REA MC, O'SULLIVAN O, RITARI J, DOUILLARD FP, et al. Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera[J]. *Nature Communications*, 2015, 6: 8322.
- [20] ARAHAL DR. Whole-genome analyses[A]//Methods in Microbiology[M]. Amsterdam: Elsevier, 2014: 103-122.
- [21] CIUFO S, KANNAN S, SHARMA S, BADRETDIN A, CLARK K, TURNER S, BROVER S, SCHOCH CL, KIMCHI A, DICUCCIO M. Using average nucleotide identity to improve taxonomic assignments in prokaryotic genomes at the NCBI[J]. *International Journal of Systematic and Evolutionary Microbiology*, 2018, 68(7): 2386-2392.
- [22] 宋宇琴. 德氏乳杆菌保加利亚亚种的群体遗传学和功能基因组学研究[D]. 呼和浩特: 内蒙古农业大学博士学位论文, 2018.
SONG YQ. Study on population genetics and functional genomics of *Lactobacillus delbrueckii* subsp. *bulgaricus*[D]. Hohhot: Doctoral Dissertation of Inner Mongolia Agricultural University, 2018 (in Chinese).
- [23] BRINKKÖTTER A, KLÖSS H, ALPERT C, LENGELER JW. Pathways for the utilization of N-acetyl-galactosamine and galactosamine in *Escherichia coli*[J]. *Molecular Microbiology*, 2000, 37(1): 125-135.
- [24] JIANG T, CHEN WW, CAO LS, HE YF, ZHOU HL, MAO HP. Abundance, functional, and evolutionary analysis of oxalyl-coenzyme A decarboxylase in human microbiota[J]. *Frontiers in Microbiology*, 2020, 11: 672.
- [25] 邱瑾, 刘剑新, 钟慧. 肠道菌群和抗生素与肾结石的形成[J]. *临床泌尿外科杂志*, 2020, 35(12): 1011-1014.
QIU J, LIU JX, ZHONG Y. Intestinal flora, antibiotics and formation of kidney stones[J]. *Journal of Clinical Urology*, 2020, 35(12): 1011-1014 (in Chinese).
- [26] WANG J, ZHAO X, TIAN Z, YANG YW, YANG ZN. Characterization of an exopolysaccharide produced by *Lactobacillus plantarum* YW11 isolated from Tibet Kefir[J]. *Carbohydrate Polymers*, 2015, 125: 16-25.
- [27] HUANG KH, CHEN BY, SHEN FT, YOUNG CC. Optimization of exopolysaccharide production and diesel oil emulsifying properties in root nodulating bacteria[J]. *World Journal of Microbiology and Biotechnology*, 2012, 28(4): 1367-1373.
- [28] RANI RP, ANANDHARAJ M, DAVID RAVINDRAN A. Characterization of a novel exopolysaccharide produced by *Lactobacillus gasseri* FR4 and demonstration of its *in vitro* biological properties[J]. *International Journal of Biological Macromolecules*, 2018, 109: 772-783.

- [29] 扎娃. 乳酸菌胞外多糖合成条件的优化及 *eps* 基因簇分析[D]. 哈尔滨: 哈尔滨工业大学硕士学位论文, 2020.
ZHA W. Optimization of extracellular polysaccharide synthesis conditions of lactic acid bacteria and analysis of *eps* gene cluster[D]. Harbin: Master's Thesis of Harbin Institute of Technology, 2020 (in Chinese).
- [30] CUI YH, JIANG X, HAO MY, QU XJ, HU T. New advances in exopolysaccharides production of *Streptococcus thermophilus*[J]. Archives of Microbiology, 2017, 199(6): 799-809.
- [31] ALMIRÓN-ROIG E, MULHOLLAND F, GASSON MJ, GRIFFIN AM. The complete *cps* gene cluster from *Streptococcus thermophilus* NCFB 2393 involved in the biosynthesis of a new exopolysaccharide[J]. Microbiology: Reading, England, 2000, 146 (Pt 11): 2793-2802.
- [32] 赵洁. 自然发酵乳中嗜热链球菌群体遗传学和功能基因组学研究[D]. 呼和浩特: 内蒙古农业大学博士学位论文, 2018.
ZHAO J. Population genetics and functional genomics of *Streptococcus thermophilus* isolated from natural fermented dairy products[D]. Hohhot: Doctoral Dissertation of Inner Mongolia Agricultural University, 2018 (in Chinese).
- [33] 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[J]. Microbiology: Reading, England, 1997, 143 (Pt 7): 2395-2405.
- [34] 刘之园, 贾俊霞, 姜昊蔚, 李志芸, 郭静, 刘志坤, 高秀珍. 耐酸耐胆盐益生菌的筛选及其益生特性研究[J]. 中国酿造, 2020, 39(11): 103-108.
LIU ZY, JIA JX, JIANG HW, LI ZY, GUO J, LIU ZK, GAO XZ. Screening of probiotics resistant to acid and bile salt and their probiotic properties[J]. China Brewing, 2020, 39(11): 103-108 (in Chinese).
- [35] PAN QQ, SHEN XD, YU LL, TIAN FW, ZHAO JX, ZHANG H, CHEN W, ZHAI QX. Comparative genomic analysis determines the functional genes related to bile salt resistance in *Lactobacillus salivarius*[J]. Microorganisms, 2021, 9(10): 2038.
- [36] MEHRA Y, RAJESH NG, VISWANATHAN P. Analysis and characterization of *Lactobacillus paragasseri* and *Lacticaseibacillus paracasei*: two probiotic bacteria that can degrade intestinal oxalate in hyperoxaluric rats[J]. Probiotics and Antimicrobial Proteins, 2022, 14(5): 854-872.
- [37] BEGLEY M, SLEATOR RD, GAHAN CGM, HILL C. Contribution of three bile-associated loci, *bsh*, *pva*, and *btLB*, to gastrointestinal persistence and bile tolerance of *Listeria monocytogenes*[J]. Infection and Immunity, 2005, 73(2): 894-904.
- [38] BOOLCHANDANI M, D'SOUZA AW, DANTAS G. Sequencing-based methods and resources to study antimicrobial resistance[J]. Nature Reviews Genetics, 2019, 20(6): 356-370.