

基于比较基因组学研究粪肠球菌乳源分离株的安全性

沈婷婷, 马昕玮, 路镜达, 李伟程, 钟智*

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

摘要 粪肠球菌为同型发酵乳酸菌,在食品生产中应用广泛。然而,很多粪肠球菌为条件致病菌,使其在食品中的应用存在很大的争议。本文以粪肠球菌乳源分离株为研究对象,探究粪肠球菌乳源分离株的安全性。采用 Illumina HiSeq 高通量测序平台对 35 株粪肠球菌乳源分离株进行全基因组测序,与 NCBI 数据库中下载的 21 株粪肠球菌粪便分离株进行比较基因组学分析。结果表明,粪肠球菌整体基因组长度为(2.93±0.13)Mb,GC 含量在(37.43±0.13)%,转运 RNA(tRNA)有(56±4)个,编码序列(CDS)数量在(2768±146)个。粪肠球菌乳源分离株平均基因组长度与粪便分离株无显著差异,而平均 CDS 数量显著高于粪便分离株。系统发育树结果显示同分离源的菌株聚类在同一进化分支,该结果与平均核苷酸一致性(ANI)聚类结果基本一致。乳源分离株含有 5~11 种耐药基因,8~24 种毒力因子及 18~31 个可移动遗传元件,与粪便分离株相比较,乳源分离株平均每株菌少携带 3 个耐药基因和毒力因子,多 2.4 个可移动遗传元件。因此即使是在乳制品中分离得到的粪肠球菌,也存在致病风险,使用前需要经过严格的安全性评价。

关键词 粪肠球菌; 比较基因组学; 乳制品; 安全性评价

文章编号 1009-7848(2024)05-0359-14 **DOI:** 10.16429/j.1009-7848.2024.05.030

粪肠球菌(*Enterococcus faecalis*)广泛存在于血液、尿液、乳制品等环境^[1-2],也是人和动物肠道中固有菌群之一。粪肠球菌具有耐酸、耐热、肠道黏附能力强等特点,也可抑制病原菌、增强免疫力等^[3],在维持肠道菌群稳定中起着重要作用。个别粪肠球菌存在益生特性,在食品和微生物制剂领域^[4]发挥着重要作用。

作为条件致病菌的一种,粪肠球菌可引起菌血症、感染性心内膜炎、尿路感染等疾病^[5]。20 世纪 70 年代后,粪肠球菌作为条件致病菌成为医院获得性尿路感染疾病的主要原因。随着耐万古霉素肠球菌(*Vancomycin resistance Enterococcus*, VRE)的发现^[6],肠球菌临床感染率急剧增加。近年来,欧洲^[7-8]地区粪肠球菌引起的感染率不断攀升,为临床治疗带来很大的困难^[9-10]。由于人们滥用抗生素,细菌为抵御抗生素不断进化,使细菌出现耐药性,进而成为耐药菌^[11]。养殖业中,抗生素作为

添加剂在饲料中低剂量加入,促进养殖生物的生长,而这一方法加快了耐药菌的形成。有研究显示,将肠球菌加入仔猪饲料中,可降低仔猪的腹泻率^[12]。人类不经意间接触到动物粪便会使耐药细菌在人畜间传播,当人和畜牧使用同一种抗生素时,人体便对抗生素产生抗性,这为临床治疗带来不少困难。

粪肠球菌耐药性及其安全性已成为人们广泛关注的问题。随着测序技术的发展,利用全基因组测序技术结合生物信息学分析可预测不同分离源菌株中潜在的致病性和耐药性^[13-14]。以往粪肠球菌研究主要集中于临床分离株^[15],随着人们对食品安全问题的关注,近年来,人们逐渐热衷研究食品源粪肠球菌^[16],而关于食品来源粪肠球菌基因组研究报道甚少,食品源粪肠球菌是否安全成为首要研究问题。

本研究采用高通量测序技术对 35 株乳源分离株进行全基因组测序,并与美国国家生物技术信息中心(NCBI)GenBank 数据库中选取的 21 株粪便分离株进行耐药基因、毒力因子和可移动遗传元件比较分析。采用核心基因构建系统发育树,探讨不同来源菌株间的基因组差异和遗传进化关系,为评定粪肠球菌安全性提供依据。

收稿日期: 2023-05-10

基金项目: 内蒙古自然科学基金重点项目(2023ZD17);内蒙古自治区高等学校“青年科技英才支持项目”(NJYT24055)

第一作者: 沈婷婷,女,硕士生

通信作者: 钟智 E-mail: imu150zhongzhi@163.com

1 材料与方法

1.1 材料和试剂

1.1.1 试验菌株 本研究菌株共计 56 株粪肠球菌。其中人粪便分离株 (21 株) 基因组下载自 NCBI 的 Genbank 数据库 (<http://www.ncbi.nlm.nih.gov/genbank/>), 乳源分离株 (35 株) 由本实验室进行全基因组测序。

1.1.2 试剂与仪器 MRS Broth 培养基, 赛默飞世尔科技; DNA 提取试剂盒, 欧米伽生物技术公司; Nextera XT Library Prep Kit 文库制备试剂盒, HiSeq Reagent v3 600 cycles Kit, 美国 Illumina 公司。

HWS24-电热恒温水浴锅、DHP-9272-电热恒温培养箱、LRH-250-生化培养箱, 上海一恒科技仪器有限公司; ND-1000-超微量紫外分光光度计, 美国 NanoDrop 公司; KC-SX-700-高压蒸汽灭菌锅, 日本 TOMY KOGYO 公司; ZHJH-C1214C-超净工作台, 上海智城分析仪器制造有限公司; MUL9000-超纯水系统, 昆山总馨企业有限公司; AR2202CN-电子天平, 奥豪斯仪器上海有限公司; SP-650-全自动高压干热灭菌器, 日本三洋电器集团; BX50-光学显微镜, 日本奥林巴斯公司; Vortex-genie 2-漩涡振荡器, 美国 Scientific Industries 公司; Illumina HiSeq 测序仪, 美国 Illumina 公司。

1.2 试验方法

1.2.1 菌株活化及 DNA 提取 粪肠球菌在 MRS 液体培养基中 37 °C 有氧条件下培养 24 h, 传代培养至二代后进行镜检, 将二代菌以 2% 的接种量接种至 35 mL 的液体培养基中, 37 °C 有氧条件下扩培 24 h 后离心收集菌体。取收集的菌体使用 DNA 提取试剂盒提取菌株 DNA, 采用超量紫外分光光度计检验 DNA 的纯度和浓度, 将符合试验要求的 DNA ($OD_{260/280}=1.8\sim 2.0, >6 \mu\text{g}$) 置于冰箱 (-20 °C) 备用。

1.2.2 全基因组测序 使用 Nextera XT Library Prep Kit 试剂盒将质量良好的细菌基因组 DNA 进行文库构建, 使用 HiSeq Reagent v3 600 cycles Kit 试剂盒将文库进行测序前处理, 之后根据厂商提供的说明书运用 Illumina HiSeq 测序平台对菌体进行全基因组测序^[17]。

1.2.3 序列组装 首先使用 SOAPdenovo v1.06^[18] 软件将高质量的 Pair-end reads 数据进组装, 然后使用 GapCloser (<http://sourceforge.net/projects/soapdenovo2/files/GapCloser/>) 软件对组装结果进行补洞、单碱基校对。

1.2.4 计算平均核苷酸一致性 平均核苷酸相似度 (ANI) 分析是通过菌株基因组鉴定种内亲缘关系的方法, 该分析通过菌株之间同源性序列的比对获得, ANI 值大于 95% 即认定为同一种^[19]。参考 Goris 等^[20]报道的方法计算 56 株粪肠球菌间 ANI 值。

1.2.5 构建核心基因集和泛基因集 利用 Prokka v1.11^[21] 软件对 56 株粪肠球菌基因序列进行编码序列 (Coding sequences, CDS) 预测和注释, 采用 Roary v3.6.1^[22] 软件识别粪肠球菌核心基因集 (Core gene) 和泛基因集 (Pan gene), 以编码蛋白氨基酸相似性大于 95% 的原则识别基因家族^[23], 将菌株共同拥有的基因家族定义为核心基因集, 全部基因家族定义为泛基因集。最后利用 PanGP 软件绘制进行泛-核心基因积累曲线图绘制。

1.2.6 构建系统发育树 基于粪肠球菌核心基因核苷酸序列, 采用 TreeBeST 软件通过邻接法 (Neighbor-joining, NJ) 构建系统发育树。使用在线软件 iTol^[24] (<https://itol.embl.de/>) 对树文件进行可视化, Bootstrap 参数值设为 1 000。

1.2.7 粪肠球菌耐药基因鉴定 56 株粪肠球菌蛋白序列分别与耐药基因数据库 (Comprehensive Antibiotic Resistance Database, CARD, <http://arpcard.mcmaster.ca>) 进行 BLAST 比对, 将比对结果中相似度 >85%, E 值 < 1×10^{-5} 的蛋白序列确定为耐药基因^[25-26]。

1.2.8 粪肠球菌毒力基因鉴定 56 株粪肠球菌蛋白序列分别与毒力因子数据库 (Virulence Factor Database, VFDB, <http://www.mgc.ac.cn/VFs/main.htm>) 中的完整数据集中的蛋白质序列进行比对, 将比对结果中相似度 >95%, 匹配长度 >300, E 值 < 1×10^{-15} 的蛋白序列确定为毒力基因^[27]。

1.2.9 粪肠球菌可移动遗传元件分析 利用在线网站 IslandViewer 4^[28] (<https://www.pathogenomics.sfu.ca/islandviewer/>), ISfinder^[29] (<https://www-is.biotoul.fr/>), ICEfinder^[30] (<https://bioinfo-mml.sjtu.e->

du.cn/ICEfinder/)对 56 株菌进行基因岛(Gene island, GI),插入序列(Insert the sequence, IS)及整合性接合元件(Integrative and conjugative element, ICE)分析。统计株菌所包含的基因岛,插入序列及整合性接合位点数量。

1.2.10 数据作图 使用 R(V4.1.2)中 ggpubr 绘制箱型图, TBtools^[31]绘制热图, SPASS 26 处理数据分析。

1.2.11 序列上传 35 株粪肠球菌乳源分离株已测序数据评估组装, 其核酸序列提交至 GenBank 数据库, 序列号为 JALPRL000000000 ~ JALSCL000000000。

2 结果与分析

2.1 粪肠球菌基因组信息

35 株乳源分离株和 21 株粪便分离株信息统计如表 1 所示。粪肠球菌总体平均基因组长度在(2.93±0.13)Mb, GC 含量在(37.43±0.13)%, tRNA 数量在(56±4)个, CDS 数量在(2 768±146)个。发现 2 种分离源菌株在基因组长度和 GC 含量及 tRNA 数量未存在差异(图 1a、1b、1c)。乳源分离株平均 CDS 数量比粪便分离株平均 CDS 多 28 个(图 1d), 2 种分离源呈现显著差异($P<0.01$)。

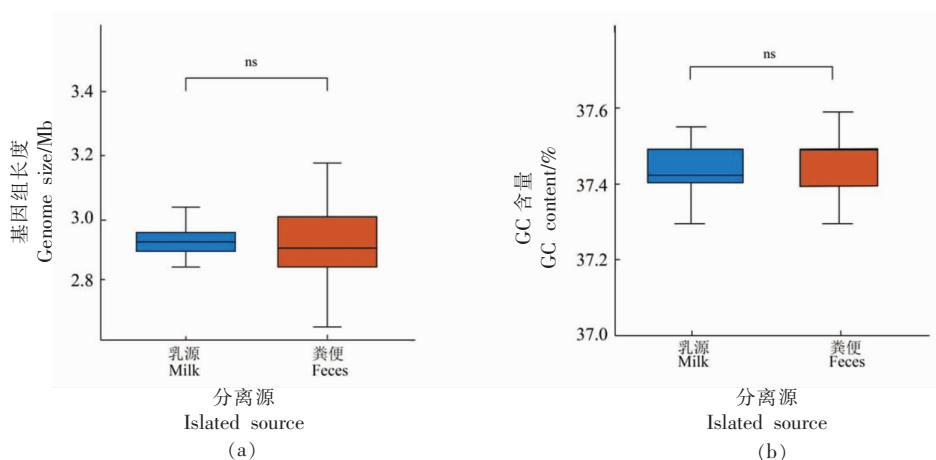
表 1 粪肠球菌菌株信息

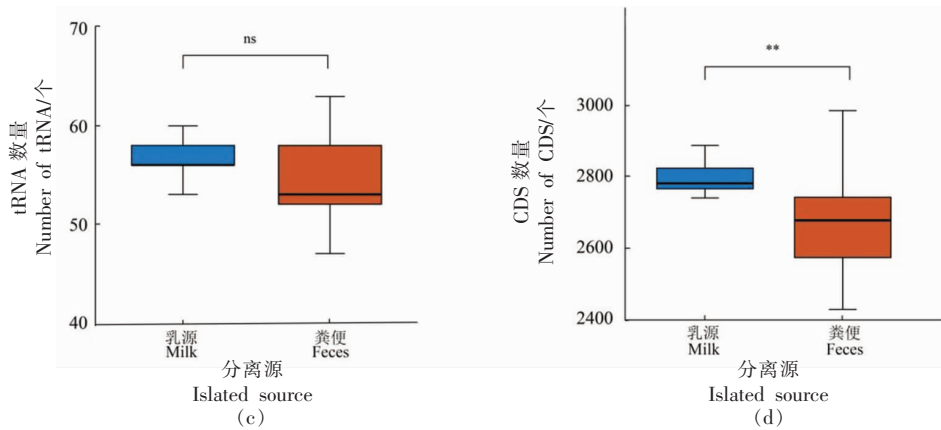
Table 1 List of the *Enterococcus faecalis* strains analyzed in this study

菌株名称	分离源	基因组长度/Mb	GC 含量/%	tRNA 数量/个	CDS 数量/个	NCBI 编号
1207/14	粪便	3.49	37.04	69	3 325	GCA_018986755.2
EGM183	粪便	2.77	37.72	60	2 545	GCA_014844235.1
HA-1	粪便	2.96	37.40	63	2 737	GCA_006349345.1
HUM-569	粪便	3.00	37.30	52	2 860	GCA_014853845.1
J107	粪便	2.93	37.50	52	2 712	GCA_007034685.1
J147	粪便	2.79	37.40	51	2 568	GCA_007034965.1
J188	粪便	2.87	37.50	53	2 647	GCA_007035245.1
J249	粪便	2.93	37.50	52	2 714	GCA_007035845.1
J303	粪便	2.87	37.50	53	2 649	GCA_007036085.1
JY32	粪便	2.93	37.50	61	2 702	GCA_010103655.1
K48_3	粪便	2.80	37.50	47	2 565	GCA_019731615.1
K52_2	粪便	2.90	37.40	56	2 673	GCA_019731515.1
K53_2	粪便	2.89	37.40	53	2 673	GCA_019731475.1
L3451	粪便	3.15	37.20	61	2 952	GCA_009832415.1
MCC499	粪便	2.86	37.50	58	2 660	GCA_018784985.1
OM08-2AT	粪便	2.67	37.60	57	2 432	GCA_003438055.1
S.22.19.Ef	粪便	2.84	37.50	53	2 661	GCA_020167115.1
S101	粪便	2.65	37.80	40	2 422	GCA_014353145.1
VRE33477	粪便	3.06	37.20	50	2 859	GCA_016391965.1
VRE33535	粪便	3.17	37.40	48	2 981	GCA_016394325.1
VRE69	粪便	3.10	37.30	52	2 944	GCA_003709925.1
BL1_16	奶酪	2.93	37.51	58	2 783	JALPRL000000000
BL13_1	乳饮料	2.92	37.41	53	2 761	JALSDS000000000
BL25_10	乳渣	2.96	37.41	63	2 811	JALS DR000000000
BL4_17	乳渣	2.89	37.50	56	2 750	JALS DQ000000000
FYM14_1	母乳	2.85	37.50	59	2 737	JALS DP000000000
FYM7_1	母乳	2.84	37.50	57	2 736	JALS DO000000000
FYM7_2	母乳	2.84	37.49	57	2 737	JALS DN000000000
FYM7_3	母乳	2.85	37.50	59	2 735	JALS DM000000000
FYM9_5	母乳	2.84	37.51	57	2 741	JALS DL000000000

(续表 1)

菌株名称	分离源	基因组长度/Mb	GC 含量/%	tRNA 数量/个	CDS 数量/个	NCBI 编号
NM136_5	酸粥	2.94	37.43	58	2 796	JALSDJ000000000
NM144_4	酸牛奶	2.91	37.41	53	2 763	JALS DI000000000
NM146_4	酸牛奶	2.91	37.41	53	2 764	JALSDH000000000
QH4_4	酸牛奶	2.94	37.51	60	2 804	JALSDG000000000
QH5_5	酸牦牛奶	2.93	37.51	58	2 809	JALSCK000000000
WH16_S12	牦牛奶	2.98	37.34	56	2 853	JALSDF000000000
WH3_3	酸山羊奶	2.89	37.48	58	2 770	JALSDE000000000
WH8_1	酸牛奶	2.93	37.43	56	2 755	JALSDD000000000
WZ21_2	酸山羊奶	2.89	37.45	56	2 768	JALSDC000000000
WZ44_1	酸山羊奶	2.89	37.43	56	2 771	JALSDB000000000
WZ47_1	酸牛奶	3.03	37.30	54	2 918	JALS DA000000000
WZ57_1	奶油	2.89	37.46	56	2 771	JALSCZ000000000
WZ9_1	酸山羊奶	2.90	37.49	58	2 776	JALSCY000000000
XM12_1	酸牛奶	3.03	37.30	55	2 916	JALSCX000000000
XM13_1	酸马奶	2.95	37.43	58	2 840	JALSCW000000000
XM13_11	酸马奶	2.94	37.43	56	2 804	JALSCV000000000
XM13_5	嚼口	2.94	37.43	53	2 776	JALSCU000000000
XM13_6	酸马奶	2.86	37.56	59	2 750	JAL SCT000000000
XM16_2	酸马奶	2.88	37.45	56	2 774	JALSCS000000000
XM2_3	嚼口	2.94	37.43	56	2 803	JALSCR000000000
XM29_3	嚼口	2.90	37.44	59	2 819	JALSCQ000000000
XM30_1	酸马奶	2.95	37.38	56	2 856	JALSCP000000000
XM5_1	嚼口	2.98	37.38	55	2 883	JALSCO000000000
XM5_5	嚼口	3.28	37.02	58	3 161	JALSCN000000000
XM5_7	嚼口	3.05	37.30	59	2 925	JALSCM000000000
YM39_5	嚼口	2.95	37.34	54	2 800	JAL SCL000000000





注:采用 R 软件中 ggpubr 绘制箱型图,利用 wilcox 检验计算各分组差异,ns 代表无显著性差异 ($P>0.05$),** 代表差异极显著 ($P<0.01$)。

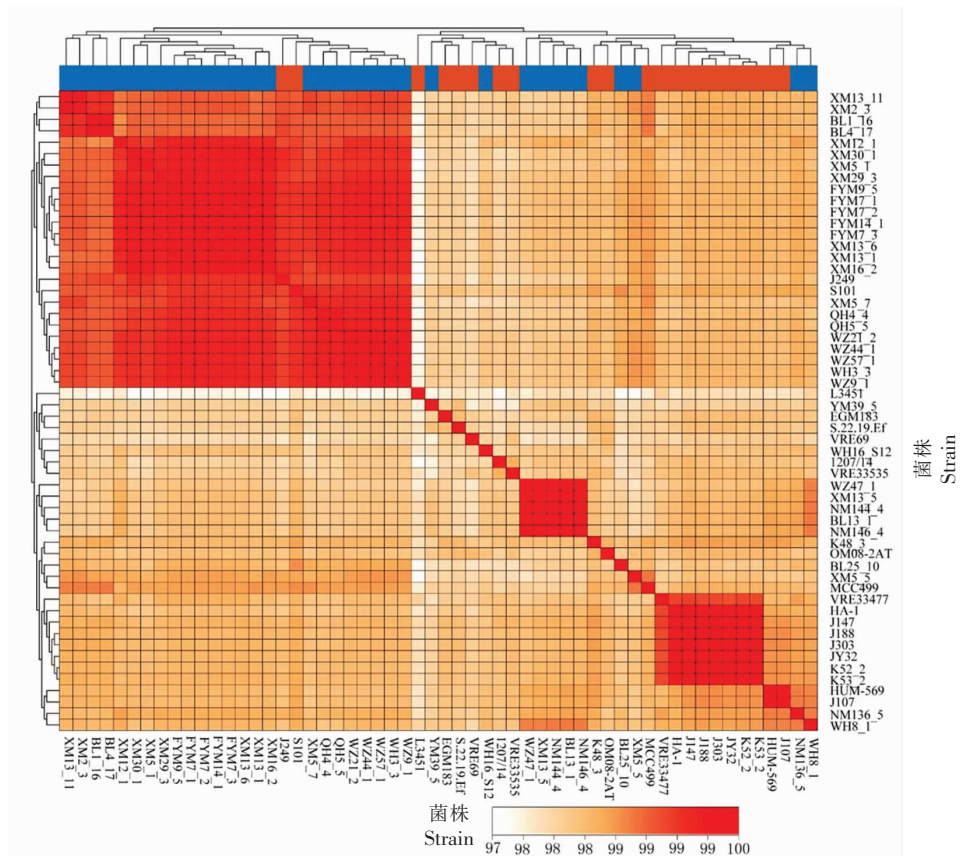
图 1 粪肠球菌基因组特征

Fig.1 Genomic characteristics of *Enterococcus faecalis*

2.2 粪肠球菌 ANI 值分析

本研究对 56 株粪肠球菌进行 ANI 值计算并构建聚类热图 (图 2)。结果显示,56 株粪肠球菌 ANI 值在 97.80%~99.99%,菌株两两间 ANI 值均大于 95%,表明本研究中测序的 35 株粪肠球菌乳

源分离株均为粪肠球菌。此外,粪肠球菌乳源分离株平均 ANI 值为 $(99.17\pm 0.29)\%$,粪便分离株平均 ANI 值为 $(98.76\pm 0.23)\%$,乳源分离株的序列一致性高于粪便分离株。



注:图 2 为菌株两两间平均 ANI 值所绘制的热图,采用 TTools 软件中 HeatMap 绘制。上方聚类中蓝色代表乳源分离株,橙色代表粪便分离株。

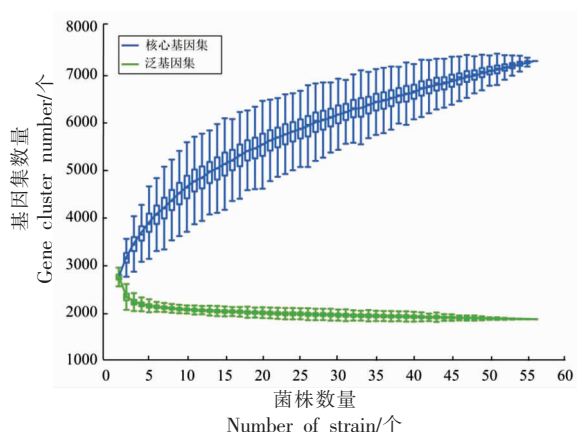
图 2 平均核苷酸一致性

Fig.2 Average nucleotide identity

2.3 粪肠球菌核心基因集和泛基因集

利用 Roary 软件计算到 56 株粪肠球菌构建了 1 884 个核心基因集,发现核心基因随着菌株数量的增加趋于稳定,而泛基因集中基因数量随着菌株数量的增加,呈现开放状态(图 3)。利用 56

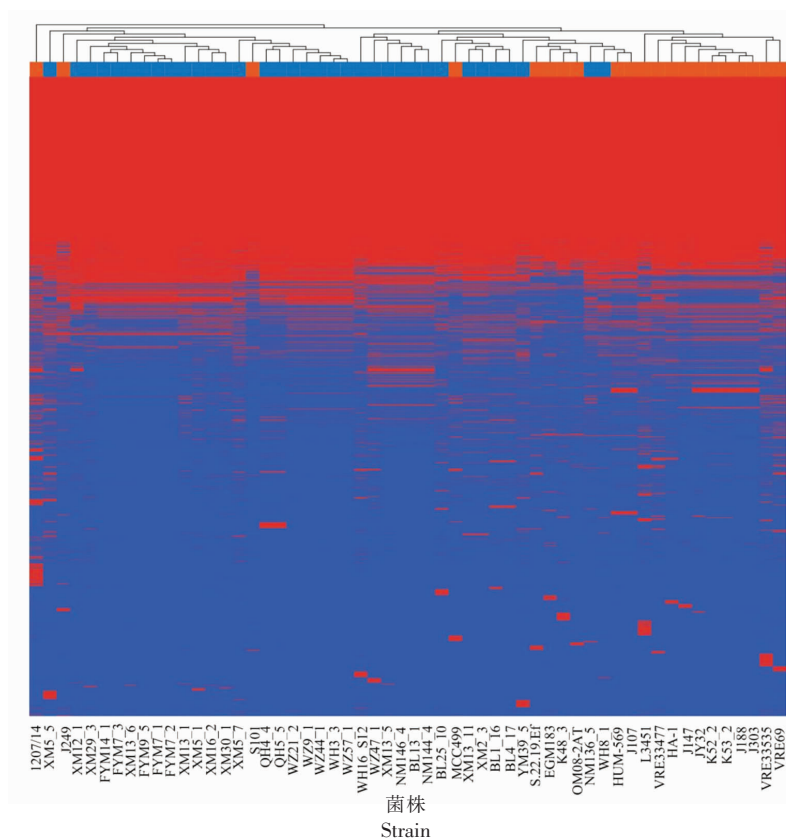
株粪肠球菌构建的 7 311 个泛基因集绘制基因存在与缺失分布热图,通过图 4 可知,粪肠球菌不同分离源菌株有聚类现象,表明菌株基因组差异与分离源相关。



注:图 3 由 PanGP 软件所绘制的趋势图。图中绿色代表核心基因,蓝色代表泛基因。

图 3 粪肠球菌泛基因-核心基因趋势图

Fig.3 Trend map of *Enterococcus faecalis* pan gene-core genes



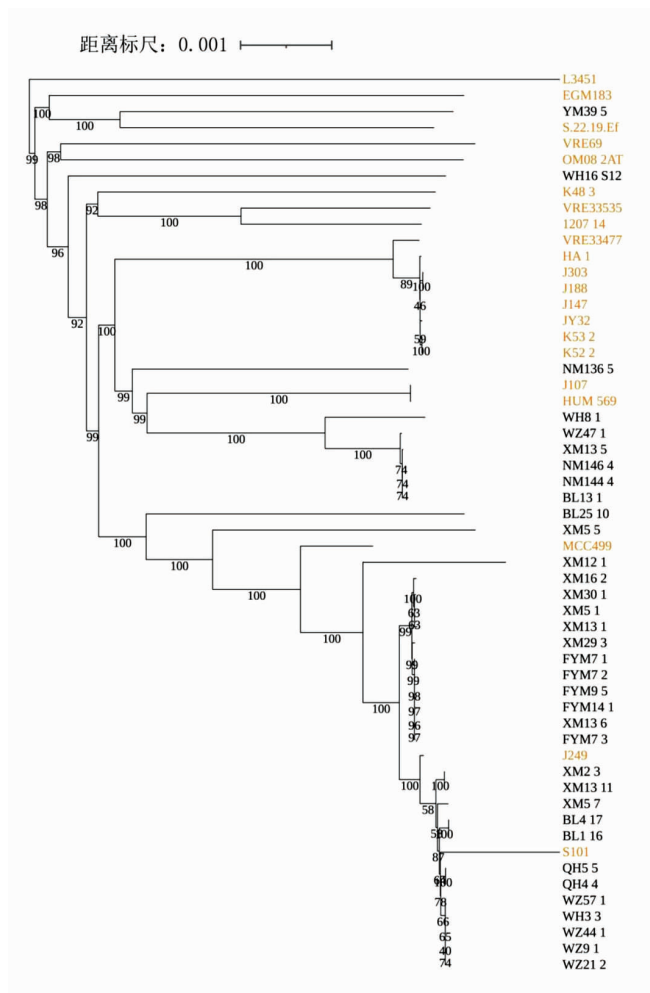
注:该图由 0-1 矩阵绘制,其中存在该基因标为 1 即红色,不存在该基因标位 0 即为蓝色;图上方聚类中标为蓝色的为乳源分离株,橙色的为粪便分离株。

图 4 粪肠球菌基因存在缺失分布图

Fig.4 Gene deletion distribution map of *Enterococcus faecalis*

2.4 系统发育树

为研究粪肠球菌种内不同分离源的遗传关系,利用 56 株粪肠球菌的 1 884 个核心基因构建系统发育树。粪肠球菌有明显的分离源聚集现象,由此认为粪肠球菌系统发育关系与分离源有一定相关性。如图 5 所示,乳源分离株 WH16_S12 和 YM39_5 分布在同一分支中,其次与两者关系最近的菌株为 NM163_5(乳源分离株)。粪肠球菌乳源分离株所聚集的小分支中粪便分离株 J249 和 S101 所聚类,表明这 2 株菌与乳源分离株遗传关系较近,这一现象与 ANI,基因存在缺失分析结果一致。



注:菌株名称黑色代表乳源分离株,黄色代表粪便分离株。此图由 iToll 在线网站绘制。

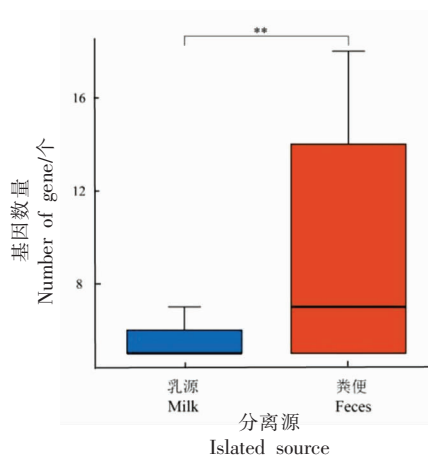
图 5 基于粪肠球菌核心基因构建的系统发育树

Fig.5 Phylogenetic tree based on core genes of *Enterococcus faecalis*

2.5 耐药基因

通过耐药基因数据库对 56 株粪肠球菌基因组进行 BLAST 比对,结果显示粪肠球菌中携带大量的耐药基因。56 株粪肠球菌共鉴定到 35 种潜在的耐药基因,2 种分离源耐药基因数量呈现显著差异(图 6, $P < 0.05$)。粪便分离株耐药基因数量在 5~18 种,乳源分离株耐药基因 5~11 种。为了解菌株耐药基因分布是否与分离源相关,利用耐药基因存在与缺失绘制热图(图 7)。结果显示 *emeA* 基因, *dfrE* 基因和 *lsaA* 基因在所有菌株中均存在。*EmeA* 基因^[32]是粪肠球菌耐药外排泵,其耐药机制为外排系统表达量增加时,细菌对抗生素外排量增加,使得生物膜通透性差、抗生素外排量过大从而产生耐药性, *emeA* 基因可导致肠球菌产生多药耐药。*DfrE*^[33]基因为二氢叶酸还原酶,使得粪肠球菌对甲氧苄氨嘧啶产生耐药性。*LsaA* 基因^[34]表达 ABC 外排泵,该基因会使粪肠球菌对克林霉素、奎奴普汀-达福普汀产生耐药性。

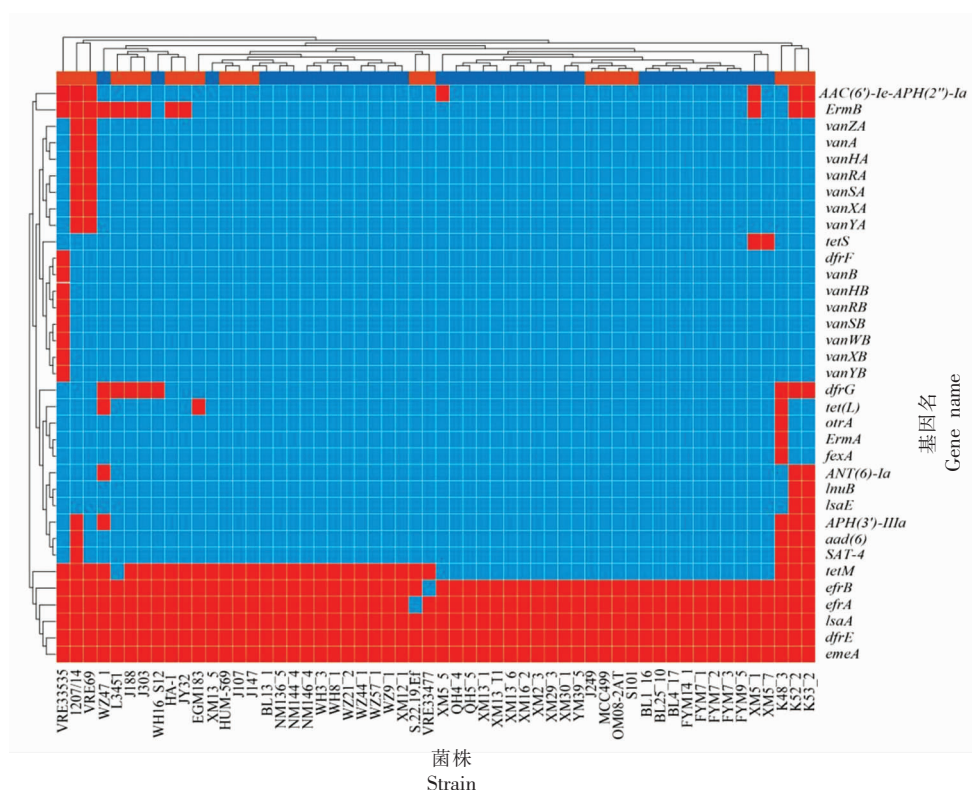
EfrA、*efrB* 基因只在粪便分离株 VRE33477 中未被分析到。在 31 株菌中检测到 *tetM* 基因,而 *tetS* 耐药基因仅存在于乳源分离株 XM5_1 和 XM5_7。*TetM* 基因是一种核糖体保护蛋白,Neela 等^[35]报道菌株虽携带 *tetM* 基因,但仍对四环素敏感,表明 *tetM* 基因在菌株中的检出频率并不代表抗性水平,仍需要从基因表达的角度去分析该



注:此图采用 R 软件中 ggpubr 绘制箱型图,利用 wilcox 检验计算两组差异,** 代表差异极显著 ($P < 0.01$)。

图 6 2 种分离源耐药基因数量比较

Fig.6 Comparison of the number of drug resistance genes between two isolated sources



注:利用0-1矩阵通过TBtools软件绘制此热图,红色即为1,表明菌株存在此基因,蓝色为0,表示菌株中不存在此基因;聚类中蓝色代表乳源分离株,橙色代表粪便分离株。

图7 粪肠球菌中耐药基因分布热图

Fig.7 Heat map of drug resistance genes in *Enterococcus faecalis*

基因。

以万古霉素耐药基因在菌株中的分布,将万古霉素耐药基因簇分为 *vanA* 基因簇、*vanB* 基因簇和 *van* 缺失基因簇。粪便分离株 1207_14 和 VRE69 2 株菌拥有 *vanA* 基因簇。该基因簇以 6 个万古霉素耐药基因组成,分别为 *vanHA*、*vanRA*、*vanSA*、*vanXA*、*vanYA*、*vanZA*。粪便分离株 VRE33535 含有 *vanB* 基因簇,该基因簇是由 *vanHB*、*vanRB*、*vanSB*、*vanWB*、*vanXB*、*vanYB* 组成。在其它粪便分离株(18 株)及乳源分离株(35 株)中没有检测到万古霉素耐药基因。

2.6 毒力因子

通过毒力因子数据库对 56 株粪肠球菌基因组进行 BLAST 比对,结果显示粪肠球菌中携带大量的毒力因子。56 株粪肠球菌共鉴定到 5 类编码基因功能的 43 种假定的毒力因子(表 2)。统计学结果显示 2 种分离源中含有的毒力因子数量呈现

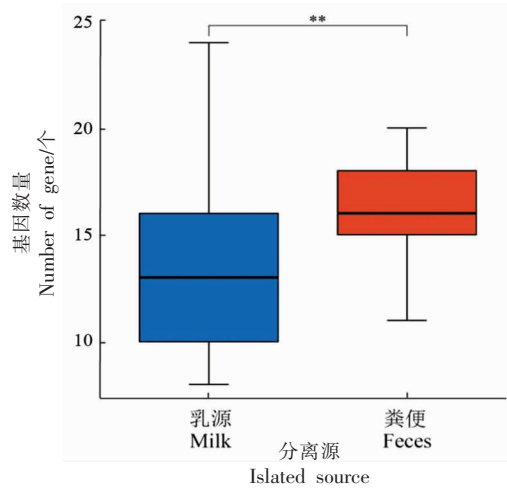
显著差异($P < 0.01$, 图 8)。乳源分离株平均毒力因子数量为 13 个,相比较于粪便分离株平均毒力因子数量少 3 个基因。含有最多毒力因子菌株分别为乳源分离株 XM5_5 和粪便分离株 VRE33535、1207_14,有 24 个基因。乳源分离株 BL25_10 和粪便分离株 S101 中毒力因子数量最少,为 8 个基因。

通过对 43 种毒力因子数量进行统计(图 9),发现 *fss1*、*efaA* 基因存在于所有菌株中。*EfaA* 基因^[34]与粪肠球菌对宿主产生心内膜炎相关,在 35 株乳源分离株中均存在该基因。*Fss1* 基因在 35 株乳源分离株中均存在,该基因是纤维蛋白原结合蛋白,可增强菌株表面的黏附能力。*bee1*、*bee3*、*srt1* 和 *srt2* 基因仅存在于 4 株乳源分离株(BL1_16、BL4_17、QH4_4、QH5_5),而 *bopD* 和 *esp* 基因仅在粪便分离株中分析出。

表 2 毒力因子分布信息

Table 2 Virulence factor distribution information

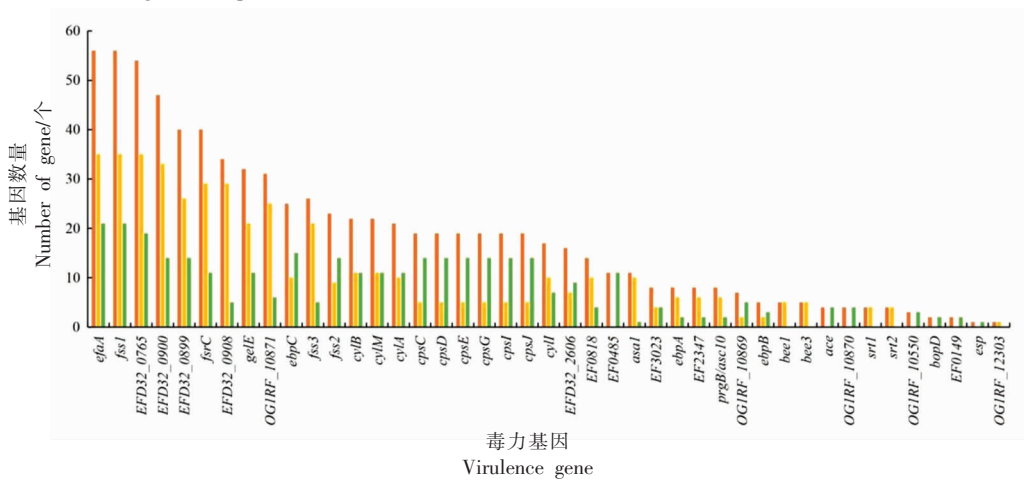
编码基因功能	毒力基因
黏附素 (Adhesins)	<i>Ace</i> 、 <i>asa1</i> 、 <i>ebpA</i> 、 <i>ebpB</i> 、 <i>ebpC</i> 、 <i>EF0149</i> 、 <i>EF2347</i> 、 <i>efaA</i> 、 <i>EFD32_0899</i> 、 <i>EFD32_0900</i> 、 <i>EFD32_0908</i> 、 <i>OG1RF_10869</i> 、 <i>OG1RF_10870</i> 、 <i>OG1RF_10871</i> 、 <i>prgB/asc10</i>
表面蛋白 (Surface protein)	<i>bee1</i> 、 <i>bee3</i> 、 <i>bopD</i> 、 <i>EFD32_0765</i> 、 <i>esp</i> 、 <i>fss1</i> 、 <i>fss2</i> 、 <i>fss3</i> 、 <i>srt1</i> 、 <i>srt2</i> 、 <i>EF0485</i>
溶血素 (Cytolysin)	<i>cylA</i> 、 <i>cylB</i> 、 <i>cylI</i> 、 <i>cylM</i>
荚膜多糖 (Capsule)	<i>cpsC</i> 、 <i>cpsD</i> 、 <i>cpsE</i> 、 <i>cpsG</i> 、 <i>cpsI</i> 、 <i>cpsJ</i>
产胞外酶 (Protease)	<i>EF0818</i> 、 <i>EF3023</i> 、 <i>EFD32_2606</i> 、 <i>fsrc</i> 、 <i>gelE</i> 、 <i>OG1RF_10550</i> 、 <i>OG1RF_12303</i>



注: ** 代表差异极显著 ($P < 0.01$)。

图 8 2 种分离源毒力因子数量比较

Fig.8 Comparison of the number of virulence factors in two isolated sources



注: 橙色为 56 株粪肠球菌中毒力因子共计数量, 黄色为乳源分离株中的数量, 绿色为粪便分离株中的数量。

图 9 毒力因子数量分布

Fig.9 Quantity distribution of virulence factors

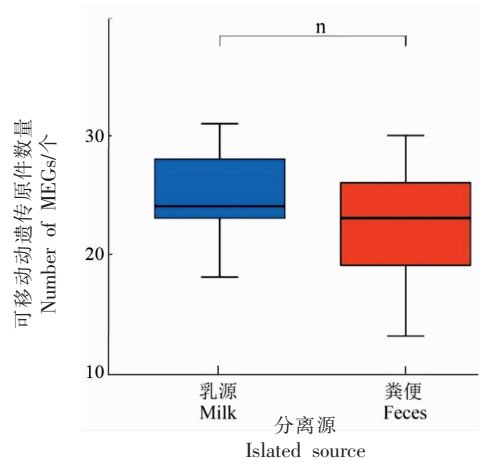
2.7 可移动遗传元件

通过分析发现 35 株乳源分离株与 21 株粪便分离株移动元件数量呈现显著差异(图 10), 粪肠

球菌乳源分离株比粪便分离株平均多 2 个可移动遗传元件。粪肠球菌乳源分离株整合性接合元件平均数量多粪便分离株 0.5 倍, 乳源分离株比粪

便分离株平均多 2 个插入序列，而乳源分离株平均基因岛数量少粪便分离株 0.9 个。图 11a 显示，18 株菌中未检测到整合性接合元件，其中 8 株为乳源分株。乳源分离株平均插入序列数量为 11.42，比粪便分离株平均多 2.57 个插入序列(图

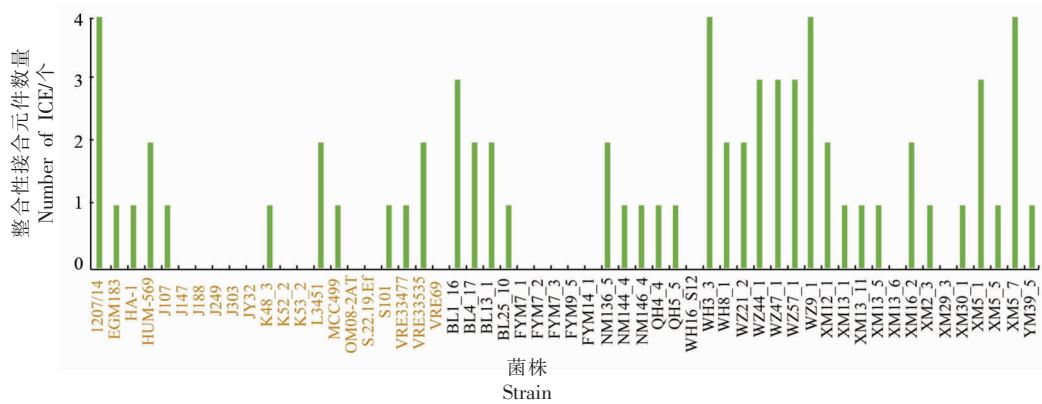
11b)。图 11c 中发现，粪肠球菌粪便分离株基因岛数量波动范围较大，平均基因岛数量为(12.81±4.53)个，而乳源分离株平均基因岛数量为(11.94±1.94)个。



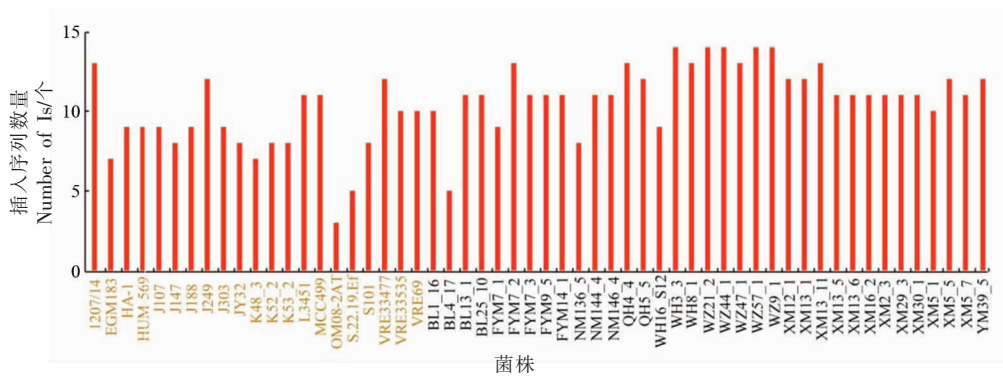
注：图中 n 代表无显著性差异。

图 10 2 种分离源可移动遗传元件(MEGs)数量比较

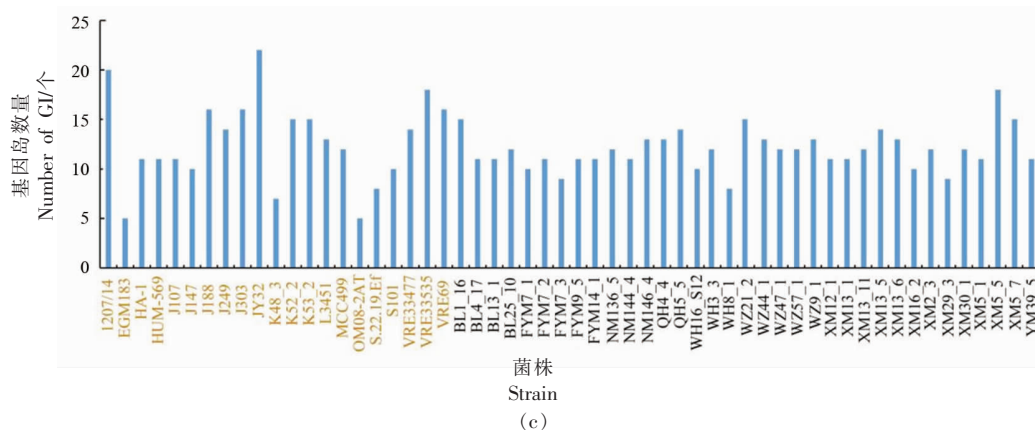
Fig.10 Comparison of the number of mobile genetic elements in two isolated sources



(a)



(b)



注:黄色字体的为粪便分离株,黑色字体的为乳源分离株。

图 11 可移动遗传元件数量分布

Fig.11 Number distribution of mobile genetic elements

3 讨论

粪肠球菌作为常用乳酸菌,在食品生产过程中应用十分广泛,而作为主要的医院获得性感染病原体,粪肠球菌对人类健康安全产生危险性,因此粪肠球菌一直存在很大的争议^[30]。随着技术的发展,基于生物信息学工具研究菌株功能基因成为重要的一种手段。本研究对 35 株粪肠球菌乳源分离株进行全基因组测序,结合公共数据库中 21 株粪肠球菌粪便分离株进行功能基因解析,发现 56 株粪肠球菌泛基因组呈现开放性,这可能由于粪肠球菌为适应新环境而与该环境中其它的细菌进行基因交换或基因插入所导致的现象。

细菌在进化及适应环境过程中通过可移动元件出现基因水平转移 (Horizontal gene transfer, HGT) 从而获得外源性耐药基因或毒力因子^[35-37]。因此可移动遗传元件 (Mobile genetic elements, MGEs) 成为细菌基因组进化和适应特定环境压力的重要机制^[38-39]。粪肠球菌乳源分离株所生存的环境营养物质丰富,与其共同生存的多为其它种乳酸菌,相比较于粪肠球菌粪便分离株其生存压力较小。乳源分离株在系统进化中被环境所影响出现基因水平转移的风险进而成为致病菌。通过比较 2 种分离源可移动遗传元件数量,发现乳源分离源数量多于粪便分离株,这表明乳源分离株中存在大量的特性基因,这进一步可推测乳源分离株在系统进化过程中容易通过基因插入或接合新基因以便适应新环境。

近年来,粪肠球菌可对免疫低下的宿主造成致命性感染。粪肠球菌引起感染或致死的主要原因是由多个毒力因子协同作用所引起的。本研究对不同分离源粪肠球菌进行毒力因子分析,发现 2 种分离源中毒力因子具有多样性,这些毒力因子多数参与编码荚膜多糖、表面蛋白、溶血素、黏附素和产胞外酶。Inoğlu 等^[40]对土耳其图卢姆奶酪中的粪肠球菌进行毒力因子检测,共检测肠球菌 14 个毒力基因。其中肠球菌毒力基因以 *ccf*、*efaA* 和 *gelE* 最为常见。Nasiri 等^[41]对巴氏奶中肠球菌进行分析,表明大部分粪肠球菌菌株中均携带 *gelE*、*hyl*、*asal*、*esp*、*cylA*、*ace* 和 *efaA* 毒力因子。孙淑玥等^[42]也发现食品源的粪肠球菌中存在大量的毒力因子。不同生态环境中的粪肠球菌毒力因子多样性高且无明显差别,进一步表明粪肠球菌乳源分离株与其它环境分离株同样存在致病风险。

4 结论

本研究在基因组层面分析了 56 株不同分离源粪肠球菌中含有的耐药基因,毒力因子及可移动遗传元件。发现粪肠球菌乳源分离株在耐药基因及毒力因子数量上与粪便分离株有显著差异,可移动遗传元件上比粪便分离株多出 0.5 倍。由此推断,即使是在乳制品中分离得到的粪肠球菌,同样存在致病风险,使用前需要经过严格的安全性评价。

参 考 文 献

- [1] PAULSEN I T, BANERJEE L, MYERS G S A, et al. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*[J]. Science, 2003, 299(5615): 2071–2074.
- [2] DAPKEVICIUS M L E, SGARDIOLI B, CAMARA S P A, et al. Current trends of *Enterococci* in dairy products: A comprehensive review of their multiple roles[J]. Foods, 2021, 10(4): 821.
- [3] RAZA T, ULLAH S R, MEHMOOD K, et al. Vancomycin resistant *Enterococci*: A brief review[J]. J Pak Med Assoc, 2018, 68(5): 768–772.
- [4] GIRAFFA G. Functionality of *Enterococci* in dairy products[J]. International Journal of Food Microbiology, 2003, 88(2/3): 215–222.
- [5] FIORE E, VAN TYNE D, GILMORE M S. Pathogenicity of *Enterococci*[J]. Microbiology Spectrum, 2019, 7(4): 1–23.
- [6] KAO P H N, KLINE K A. Dr. Jekyll and Mr. Hide: How *Enterococcus faecalis* subverts the host immune response to cause infection[J]. J Mol Biol, 2019, 431(16): 2932–2945.
- [7] MANNU L, PABA A, DAGA E, et al. Comparison of the incidence of virulence determinants and antibiotic resistance between *Enterococcus faecium* strains of dairy, animal and clinical origin[J]. Int J Food Microbiol, 2003, 88(2/3): 291–304.
- [8] GOOSSENS H. Spread of vancomycin-resistant *Enterococci*: Differences between the United States and Europe[J]. Infect Control Hosp Epidemiol, 1998, 19(8): 546–551.
- [9] WEINER L M, WEBB A K, LIMBAGO B, et al. Antimicrobial-resistant pathogens associated with healthcare associated infections: Summary of data reported to the national healthcare safety network at the centers for disease control and prevention, 2011–2014[J]. Infect Control Hosp Epidemiol, 2016, 37(11): 1288–1301.
- [10] STROMPFOVA V, MARCINAKOVA M, SIMONOVA M, et al. *Enterococcus faecium* EK13 – An enterocin A-producing strain with probiotic character and its effect in piglets[J]. Anaerobe, 2006, 12(5/6): 242–248.
- [11] KERRY J, COYNE R, GILROY D, et al. Spatial distribution of oxytetracycline and elevated frequencies of oxytetracycline resistance in sediments beneath a marine salmon farm following oxytetracycline therapy[J]. Aquaculture, 1996, 145(1/2/3/4): 31–39.
- [12] WERNER G, COQUE T M, HAMMERUM A M, et al. Emergence and spread of vancomycin resistance among *Enterococci* in Europe[J]. Eurosurveillance, 2008, 13(47): 19046.
- [13] PATRO J N, RAMACHANDRAN P, BARNABA T, et al. Culture-independent metagenomic surveillance of commercially available probiotics with high-throughput next-generation sequencing[J]. mSphere, 2016, 1(2): 57–72.
- [14] GHATTARGI V C, GAIKWAD M A, METI B S, et al. Comparative genome analysis reveals key genetic factors associated with probiotic property in *Enterococcus faecium* strains[J]. BMC Genom, 2018, 19(1): 1–16.
- [15] LIESE J, SCHÜLE L, OBERHETTINGER P, et al. Expansion of vancomycin-resistant *Enterococcus faecium* in an academic tertiary hospital in Southwest Germany: A large-scale whole-genome-based outbreak investigation[J]. Antimicrob Agents Chemother, 2019, 63(5): 1978–1990.
- [16] HE Q W, HOU Q C, WANG Y J, et al. Comparative genomic analysis of *Enterococcus faecalis*: Insights into their environmental adaptations[J]. BMC Genomics, 2018, 19(1): 1–12.
- [17] LEPAGE E, BRINSTER S, CARON C, et al. Comparative genomic hybridization analysis of *Enterococcus faecalis*: Identification of genes absent from food strains[J]. Journal of Bacteriology, 2006, 188(19): 6858–6868.
- [18] LUO R B, LIU B H, XIE Y L, et al. SOAPdenovo2: An empirically improved memory-efficient short-read de novo assembler[J]. Gigascience, 2012, 1(1): 1–18.
- [19] ARAHAL D R. Whole-genome analyses: Average nucleotide identity[M]// GOODFELLOW M, SUTCLIFFE I, CHUN J, ed. Methods in microbiology. Amsterdam: Academic Press, 2014, 41: 103–122.
- [20] GORIS J, KONSTANTINIDIS K T, KLAPPENBACH J A, et al. DNA–DNA hybridization values and their relationship to whole-genome sequence similarities[J]. International Journal of Systematic and Evolutionary Microbiology, 2007, 57(1): 81–91.

- [21] SEEMANN T, Prokka; Rapid prokaryotic genome annotation[J]. *Bioinformatics*, 2014, 30(14): 2068–2069.
- [22] PAGE A J, CUMMINS C A, HUNT M, et al. Roary: Rapid large-scale prokaryote pan genome analysis[J]. *Bioinformatics*, 2015, 31(22): 3591–3593.
- [23] 吕瑞瑞, 李伟程, 康小红, 等. 副干酪乳杆菌 PC-01 全基因组测序及不同副干酪乳杆菌菌株比较基因组学分析[J]. *微生物学通报*, 2021, 48(9): 3025–3038.
- LÜ R R, LI W C, KANG X H, et al. Whole genome sequencing of *Lactobacillus paracasei* PC-01 and comparative genomics analysis about *Lactobacillus paracasei* strains[J]. *Microbiology China*, 2021, 48(9): 3025–3038.
- [24] LETUNIC I, BORK P. Interactive tree of life (iTOL) v4: Recent updates and new developments[J]. *Nucleic Acids Research*, 2019, 47(W1): W256–W259.
- [25] MCARTHUR A G, WAGLECHNER N, NIZAM F, et al. The comprehensive antibiotic resistance database[J]. *Antimicrob Agents Chemother*, 2013, 56(7): 3348–3357.
- [26] TANG F, WANG J F, LI D Z, et al. Comparative genomic analysis of 127 *Escherichia coli* strains isolated from domestic animals with diarrhea in China[J]. *BMC Genomics*, 2019, 20(1): 1–11.
- [27] CHEN L, YANG J, YU J, et al. VFDB: A reference database for bacterial virulence factors[J]. *Nucleic Acids Research*, 2005, 33(1): 325–328.
- [28] BERTELLI C, LAIRD M R, WILLIAMS K P, et al. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets[J]. *Nucleic Acids Research*, 2017, 45(W1): W30–W35.
- [29] SIGUIER P, PÉROCHON J, LESTRADE L, et al. ISfinder: The reference centre for bacterial insertion sequences[J]. *Nucleic Acids Research*, 2006, 34(suppl_1): D32–D36.
- [30] LIU M, LI X, XIE Y, et al. ICEberg 2.0: An updated database of bacterial integrative and conjugative elements[J]. *Nucleic Acids Research*, 2019, 47(1): 660–665.
- [31] CHEN C J, CHEN H, ZHANG Y, et al. TBtools: An integrative toolkit developed for interactive analyses of big biological data[J]. *Molecular Plant*, 2020, 13(8): 1194–1202.
- [32] CASALTA E, MONTEL M C. Safety assessment of dairy microorganisms: The *Lactococcus* genus[J]. *International Journal of Food Microbiology*, 2008, 126(3): 278–285.
- [33] LEE E W, CHEN J, HUDA M N, et al. Functional cloning and expression of *emeA*, and characterization of EmeA, a multidrug efflux pump from *Enterococcus faecalis*[J]. *Biological & Pharmaceutical Bulletin*, 2003, 26(2): 266–270.
- [34] 王彦杰. 粪肠球菌乳源分离株与其它环境分离株基因组比较分析[D]. 呼和浩特: 内蒙古农业大学, 2018.
- WANG Y J. Comparative genomic analysis of *Enterococcus faecalis* isolated from dairy products and other environments[D]. Hohhot: Inner Mongolia Agricultural University, 2018.
- [35] NEELA F A, NONAKA L, SUZUKI S. The diversity of multi-drug resistance profiles in tetracycline-resistant *Vibrio* species isolated from coastal sediments and seawater[J]. *Journal of Microbiology*, 2007, 45(1): 64–68.
- [36] 郑晓辉. 肠球菌 *efaA* 基因阳性转化株的构建及其对生物膜形成影响的初步探讨[D]. 福州: 福建医科大学, 2010.
- ZHENG X H. The construction of gene *efaA*+ *Enterococcus* transformant and preliminary study of its influence on biofilm formation[D]. Fuzhou: Fujian Medical University, 2010.
- [37] 邓瑞, 陈亮. 细菌基因岛研究进展[J]. *重庆医学*, 2018, 47(6): 846–848.
- DENG R, CHEN L. Research progress of bacterial gene island[J]. *Chongqing Medicine*, 2018, 47(6): 846–848.
- [38] 黄元铭, 李臻鹏, 程倩, 等. 河弧菌可移动遗传元件的预测与分析[J]. *疾病监测*, 2022, 37(5): 591–597.
- HUANG Y M, LI Z P, CHENG Q, et al. Prediction and analysis of mobile gene elements in *Vibrio fluvialis* [J]. *Disease Surveillance*, 2022, 37(5): 591–597.
- [39] 别路垚. 细菌中耐药性传播相关的可移动遗传元件和抗生素响应策略的研究[D]. 济南: 山东大学, 2019.
- BIE L Y. Study on mobile genetic elements and antibiotic response strategies associated with resistance transmission in bacteria[D]. Jinan: Shandong Univer-

- sity, 2019.
- [40] İNOĞLU Z N, TUNCER Y. Safety assessment of *Enterococcus faecium* and *Enterococcus faecalis* strains isolated from Turkish Tulum cheese[J]. Journal of Food Safety, 2013, 33(3): 369–377.
- [41] NASIRI M, HANIFIAN S. *Enterococcus faecalis* and *Enterococcus faecium* in pasteurized milk: Prevalence, genotyping, and characterization of virulence traits[J]. LWT, 2022, 153: 112452.
- [42] 孙淑玥, 王珊珊, 崔生辉, 等. 不同来源粪肠球菌的比较基因组学分析[J]. 中国病原生物学杂志, 2022, 17(7): 784–789.
- SUN S Y, WANG S S, CUI S H, et al. Different sources of comparative genomics in enterococcus analysis[J]. Chinese Journal of Pathogenic Biology, 2022, 17(7): 784–789.

Safety Evaluation of Dairy-derived *Enterococcus faecalis* Isolates Based on Comparative Genomic Analysis

Shen Tingting, Ma Xinwei, Lu Jingda, Li Weicheng, Zhong Zhi*

(Inner Mongolia Agricultural University, 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, Hohhot 010018)

Abstract *Enterococcus faecalis* (*E. faecalis*) is a homofermentative lactic acid bacterium widely used in food production. However, many *E. faecalis* strains are opportunistic pathogens, leading to significant controversy regarding their use in foods. This study focuses on dairy-derived isolates of *E. faecalis* to explore their safety. Using the Illumina HiSeq high-throughput sequencing platform, we performed whole-genome sequencing on 35 dairy-derived *E. faecalis* isolates and conducted comparative genomics analysis with 21 fecal isolates downloaded from the NCBI database. The results showed that the overall genome length of *E. faecalis* was (2.93 ± 0.13) Mb, GC content was $(37.43\pm 0.13)\%$, the number of transfer RNA (tRNA) was 56 ± 4 , and the number of coding sequences (CDS) was $2\,768\pm 146$. There was no significant difference in the average genome length of *E. faecalis* milk isolates from fecal isolates, but the average number of CDS was significantly higher than that of fecal isolates. The phylogenetic tree results showed that the strains of the same segregation were clustered in the same evolutionary branch, which was basically consistent with the average nucleotide identity (ANI) clustering results. Compared with fecal isolates, each strain contained 3 less drug resistance genes and 2.4 more mobile genetic elements. Therefore, even *E. faecalis* isolated from dairy products has a risk of pathogenic disease and needs to undergo strict safety evaluation before use.

Keywords *Enterococcus faecalis*; comparative genomics; dairy products; safety evaluation