

基于基因功能分析的乳酶生与表飞鸣生产用菌株质量比对研究

丁 勃^{1,2,3,4,5}, 沈 振^{1,2,3,4}, 任丽宏^{1,2,3,4}, 徐晓洁^{1,2,3,4}, 邢 晟^{1,2,3,4}, 刘新泳⁵

(1 山东省食品药品检验研究院, 济南 250101; 2 山东省食品药品安全检测工程技术研究中心, 济南 250101; 3 国家药品监督管理局仿制药研究与评价重点实验室, 济南 250101; 4 山东省仿制药一致性评价工程技术研究中心, 济南 250101; 5 山东大学药学院, 济南 250012)

[摘要] **目的:** 利用高通量测序和生物信息学技术, 从基因组水平上评价、比较微生态活菌制品乳酶生(乳酶生株)和表飞鸣(表飞鸣株)生产用菌株的安全性和有效性, 为微生态活菌制品生产用菌株的质量研究、一致性评价等提供依据。**方法:** 以 DIAMOND 技术为核心, 结合个性化数据库构建检索技术, 通过序列比对寻找目标基因并分析其功能。**结果:** 乳酶生和表飞鸣生产用菌株均为屎肠球菌, 两株菌的基因组间共线性关系较好, 存在少量的插入、缺失、倒位和易位等基因重排事件。表飞鸣株在毒力基因、耐药性基因、细菌素基因、生物胺基因、原噬菌体以及移动遗传元件等 6 个方面的安全性与乳酶生株相似, 但是在包括应激反应基因、抗菌/拮抗活性基因、抗氧化活性基因以及附着力基因等益生性功能等方面, 表飞鸣株优于乳酶生株。**结论:** 本文初步建立了基于特定基因功能分析的微生态活菌制品生产用菌株质量评价方法, 进一步完善了微生态活菌制品的质量研究和质量评价体系。

[关键词] 乳酶生; 表飞鸣; 屎肠球菌; 生物信息学; 功能基因; 益生菌筛选; 药物一致性评价

[中图分类号] R975 **[文献标志码]** A **[文章编号]** 1003-3734(2023)06-0643-14

Quality comparison research of Lactasin strain and Biofermin strain based on gene function analysis

DING Bo^{1,2,3,4,5}, SHEN Zhen^{1,2,3,4}, REN Li-hong^{1,2,3,4}, XU Xiao-jie^{1,2,3,4}, XING Sheng^{1,2,3,4}, LIU Xin-yong⁵

(1 Shan Dong Institute for Food and Drug Control, Jinan 250101, China; 2 Shandong Food and Drug Safety Testing Engineering Technology Research Center, Jinan 250101, China; 3 NMPA Key Laboratory for Research and Evaluation of Generic Drugs, Jinan 250101, China; 4 Shandong Research Center of Engineering and Technology for Consistency Evaluation of Generic Drugs, Jinan 250101, China; 5 School of Pharmaceutical Sciences Shandong University, Jinan 250012, China)

[Abstract] **Objective:** To evaluate and compare the safety and efficacy of microecological live bacteria preparations Lactasin (Lactasin strain) and Biofermin (Biofermin strain) production strains at the genome level using high-throughput sequencing and bioinformatics technology, thus to provide basis for the quality research and consistency evaluation of strains used in the production of ecological live bacteria products. **Methods:** Combined with a personalized database, DIAMOND retrieval strategy was used to find target genes and analyze their functions through sequence comparison. **Results:** The strains used for Lactasin and Biofermin production were both *Enterococcus faecium*. The genomes of the 2 strains had good collinearity, but a few of genes had rearrangements such as insertions, deletions, inversions, and translocations. The safety of the Biofermin strain was similar to the Lactasin strain in 6 aspects, including virulence genes, drug resistance genes, bacteriocin genes, biogenic amine genes, prophages and mobile genetic elements. In the evaluation of probiotic function of stress response genes, antimicrobial/antagonistic

[基金项目] 国家药典委员会药品标准制修订研究课题: 微生态活菌制品分类及标准建立(2020S10)

[作者简介] 丁勃, 男, 主任药师, 硕士研究生, 主要从事药品微生物质量控制研究。联系电话: (0531)81216758, E-mail: 13791051023@126.com。

[通讯作者] 刘新泳, 男, 博士生导师, 教授, 主要从事基于靶点的抗病毒药物设计合成等研究。联系电话: (0531)88380270, E-mail: xinyongl@sdu.edu.cn。

activity, antioxidant activity genes and adhesion genes, Biofermin strain was better than Lactasin strains. **Conclusion:** In this paper, a quality evaluation method of strains for the production of microecological live bacteria products was initially established based on specific gene function analysis, and the quality research and quality evaluation system were also improved and perfected for microecological live bacteria products.

[Key words] lactasin; biofermin; *Enterococcus faecium*; bioinformatics; functional genes; probiotic screening; drug consistency evaluation

微生态活菌制品在我国已形成千亿级的产业市场,为保证行业正常发展,使微生态制剂安全有效地为人类健康服务,我国对相关产品的质量标准问题日益关注。与其他药品种类相比,微生态制剂的质量标准存在较大差距,主要表现为概念使用不当、核心菌株不明、功能性质难界定等诸多问题,致使行业发展存在潜在的隐患和风险。菌株是研究、生产和检定微生态活菌制品的物质基础,其质量直接影响产品的安全性和有效性^[1],完善的评价体系需要进一步阐明菌株的安全性、代谢力、适应性和功能机制等^[2]。目前市售的益生菌大多缺乏科学的分子生物学检查及鉴定手段,对菌株是否存在变异产生的遗传背景变化重视不够,也缺乏合适的方法进行分析评价。微生物在传代过程中随环境的改变而发生变异是毋庸置疑的,因此应在标准中考虑研究菌株的遗传背景变化,并建立适当方法^[3]快速筛查有害基因,以确保菌株的安全性和稳定性,同时对菌株进行科学合理的评价,并对其风险进行有效评估。

表飞鸣(Biofermin)为20世纪50年代日本开发生产的一种含有屎肠球菌(*Enterococcus faecium*)的活菌制剂,其在肠道内可分解糖类生成乳酸,从而有效抑制有害菌繁殖,同时防止蛋白质发酵,减少肠内产气,主要用于肠道菌群失调或肠内异常发酵引起的腹胀、腹泻及小儿饮食不当引起的腹泻等,该品种标准收载于日本药典。20世纪60年代前后,在表飞鸣之后,我国的同型药品乳酶生(Lactasin)研制成功,最早收载于《卫生部药品标准》第二部第六册。初期,表飞鸣与乳酶生标示菌株的名称相同,均为“*Streptococcus faecalis*”,之后随着细菌鉴定技术水平的提高以及菌种命名的规范化发展,乳酶生生产用菌种CMCC32901(原编号140623)的最初定名历经多次命名变化最终确定为屎肠球菌^[4]。虽然乳酶生(中国)和表飞鸣(日本)的生产菌株在标准中的命名不同,但是基于16S rRNA和同源基因系统发育树的分析结果均表明,两菌株均为屎肠球菌。

基于药品生产用菌种的菌株溯源、制剂工艺优化以及检定等方面考虑,进一步加强药品生产用菌种的质量控制,了解生产用菌株的分子特征和遗传背景就显得尤为重要。乳酶生作为国内特有品种,根据国家药品监督管理局289基本药物目录的评价建议,其仿制药疗效和质量一致性评价工作“无需开展临床有效性试验,豁免生物等效性试验”。然而由于该品种现行质量标准相对简单,如何在现有质量标准框架下,更有效地评价并提升产品质量,是生产企业和监管部门共同面临的问题。近年来,基于基因组学等方法形成的“益生菌组学”已成功应用于当前药用益生菌和候选益生菌有效性和安全性的研究和评价^[5]。高通量测序和生物信息学技术可实现菌株的全基因组测序、分析,对基因组中包含的抗性基因、毒力基因、基因穿梭和功能基因等进行详细的分析^[6],从基因组水平上全面评价菌株质量。本文采用基于基因组学的功能基因评价方法,对乳酶生及表飞鸣产品中的屎肠球菌菌株进行分析比较,探索新的可用于微生态制剂的评价及风险分析技术方法,希望有助于进一步完善相关产品的质量研究和评价体系。

材料与方法

1 试验菌株

乳酶生生产用屎肠球菌CMCC32901(原编号:140623)来源于中国食品药品检定研究院中国医学细菌保藏管理中心。

2 制剂

表飞鸣片(Biofermin Tablets)由日本ビオフェルミン制药株式会社生产,批号为7H600(1000片·盒⁻¹)。

3 主要仪器

日本SANYO的高压蒸汽灭菌锅、美国NUAIRE公司的生物安全柜、美国INFORS公司的恒温震荡培养箱、美国MEMMERT公司的恒温培养箱以及梅里埃诊断产品(上海)有限公司的VITEK compact 2

全自动微生物鉴定仪等。

4 主要试剂

碳酸钙、氯化钠、溴甲酚紫等购于天津科密欧试剂有限公司;纯化琼脂粉、乳糖、蛋白胨、牛肉浸膏等购于北京三药科技有限公司;胰酪大豆琼脂培养基购于青岛高科技工业园海博生物技术有限公司;试剂均为分析纯。

5 菌种分离鉴定

称取牛肉浸膏 3 g,氯化钠 5 g,蛋白胨 10 g,乳糖 20 g,碳酸钙 10 g,琼脂 15 g,溴甲酚紫 0.06 g,加至 1 000 mL 水中,加热搅拌均匀后,115 ℃ 灭菌 20 min,即得 1% 碳酸钙含糖牛肉琼脂培养基。取表飞鸣片 10 g,加 pH 7.0 无菌氯化钠-蛋白胨缓冲液至 100 mL,充分研磨混匀,作为 1:10 供试液,按 10 倍稀释法进行稀释,每个稀释度吸取 0.1 mL 稀释液,涂布于 1% 碳酸钙含糖牛肉琼脂培养基表面,置 36 ℃ 培养 3 d。选取平板上周围变黄的菌落接种至胰酪大豆琼脂培养基上,36 ℃ 培养 48 h 得到纯化后的菌种。使用 VITEK compact 2 全自动微生物鉴定仪进行鉴定,结果为屎肠球菌。

6 全基因组测序

将分离自表飞鸣的屎肠球菌(以下简称表飞鸣株,部分图表中以 Japan 标示)及购自中国食品药品检定研究院中国医学细菌保藏管理中心编号为 140623-200708 的屎肠球菌(以下简称乳酶生株,部分图表中以 China 标示)送至生工生物工程(上海)股份有限公司,使用美国 Illumina 公司的第二代 solexa 测序技术进行全基因组测序。原始数据通过 FastQC 进行质量评估。通过 Trimmomatic 对 Illumina 测序数据进行质量剪切,得到相对准确的有效数据。使用 SPAdes 拼接二代测序数据。采用 GapFiller 对拼接得到的 contig 补 GAP。利用 PrInSeS-G 进行序列矫正,修正拼接过程中的剪辑错误及小片段的插入缺失。采用 RepeatMasker 鉴定基因组上的重复序列。

7 软件及数据库

生物信息学分析软件为 CLC Genomics Workbench Version 22.0.2。

数据库包括:KEGG^[7] 数据库(<http://www.kegg.jp>)、NCBI NR 数据库(<http://ncbi.nlm.nih.gov/>)、GO^[8] 数据库(<http://www.geneontology.org/>)、NCBI COG/KOG^[9] 数据库(<https://www.ncbi.nlm.nih.gov/COG/>)、Uniprot^[10] 数据库(<http://www.uniprot.org/>)、PFAM^[11] 数据库(<http://pfam.xfam.org/>)、CARD^[12] 数据库(<https://card.mcmaster.ca/>)、CAZy^[13] 数据库(<http://www.cazy.org/>)、PHI-base^[14] 数据库(<http://www.phi-base.org/>)、VFDB^[15] 数据库(<http://www.mgc.ac.cn/VFs/>)、PHASTER^[16-17] 数据库(<http://phaster.ca/>)、BACTIBASE^[18] 数据库(<http://bactibase.hammamilab.org/main.php>)、CRISPRCasFinder^[19] 数据库(<https://crisprcas.i2bc.paris-saclay.fr/CrisprCasFinder/Index/>)、UHGP^[20] 数据库(http://ftp.ebi.ac.uk/pub/databases/metagenomics/mgnify_genomes/)以及 CENSOR^[21] 数据库(<https://www.girinst.org/rebase/update/browse.php>)。可视化数据分析绘图软件为 OriginPro[®] 2022。

可视化数据分析绘图软件为 OriginPro[®] 2022。

8 数据库检索

基因检索过程使用的 NR (Non-Redundant Protein Sequence Database, 非冗余蛋白库) 数据为非冗余蛋白序列,对于所有已知或可能的编码序列, NR 都给出了相应的氨基酸序列,因此,该数据库容量高达 100 GB 以上。采用常规 BLAST 检索策略将导致巨大的计算资源和时间消耗,因此本文采用 DIAMOND 检索策略,该技术的检索速度较 BLAST 最高可提高 20 000 倍,且获得相对一致的检索结果^[22]。

目标菌株的功能基因分析主要分为安全性评估和益生性评估,安全性评估包括毒力基因、耐药性基因、细菌素基因、生物胺基因、原噬菌体以及移动遗传元件等 6 个方面,益生性评估包括适应性基因、抗菌/拮抗活性基因、抗氧化活性基因以及附着力基因等 4 个方面。其中毒力基因^[15]、耐药性基因^[12]、细菌素基因^[18]、原噬菌体^[16] 以及移动遗传元件^[19] 等采用成熟数据库进行分析。其他功能基因的分析通过本项目组汇总 NR 及 Uniprot^[10] 数据库相关联基因序列信息^[5-6] 所构建的相关数据库,经由 CLC Genomics Workbench Version 22.0.2 建立 DIAMOND Index,对目标菌株基因组进行基因信息检索和分析。

结 果

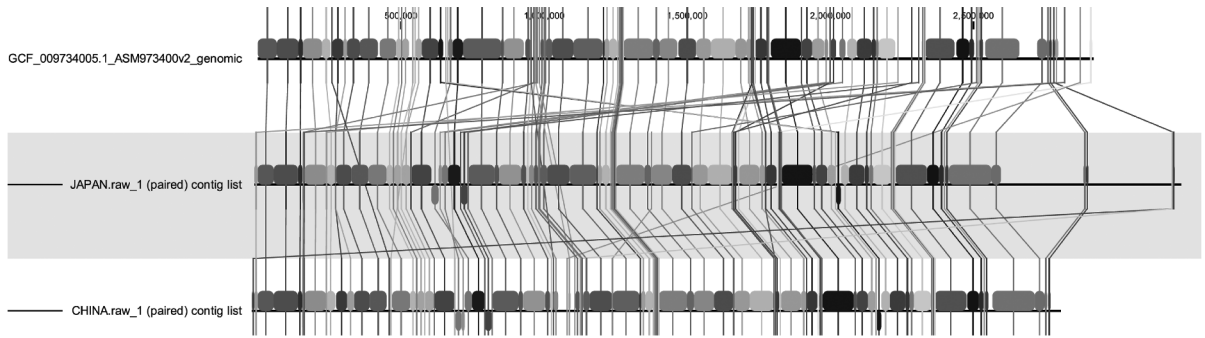
1 基因组结构分析

乳酶生株基因组序列全长为 2 821 945 bp,含有 2 810 个编码基因,编码区域占整个基因组的 87.64%,平均对照基因(carated gene, GC)含量为 38.36%,基因平均大小为 880.12 bp。表飞鸣株基因组序列全长为 2 690 331 bp,含有 2 679 个编码基因,编码区域占整个基因组的 87.89%,平均 GC 含量为

38.53%, 基因平均大小为 882.61 bp。

以屎肠球菌 SRR24 为模式菌株的全基因组比对结果(见图 1)显示,乳酶生株与表飞鸣株的基因组间共线关系较好,存在少量的插入、缺失、倒位和

易位等基因重排事件。相较于模式菌株,乳酶生株检测到单一位点突变 111 606 处,表飞鸣株检测到单一位点突变 109 842 处,且突变分布广泛。综合而言,2 株菌株在全基因组结构层面上无显著差异。



线性连接的相同颜色模块表示共线区域,白色表示低相似性区域,上方模块为正向链,下方模块为反向链

图 1 全基因组比对

2 基因组功能分析

COG 数据库注释结果显示,乳酶生株共有 2 028 个基因,表飞鸣株共有 1 973 个基因。与碳水化合物代谢相关的基因最多,分别为 297 和 277 个,各占基因总

数约 15%,其他基因主要涉及转录、翻译及复制、重组和修复等功能。另外乳酶生株和表飞鸣株分别有 231 和 224 个基因功能未定(见图 2)。除碳水化合物代谢相关基因外,COG 聚类结果未显示其他显著差异。

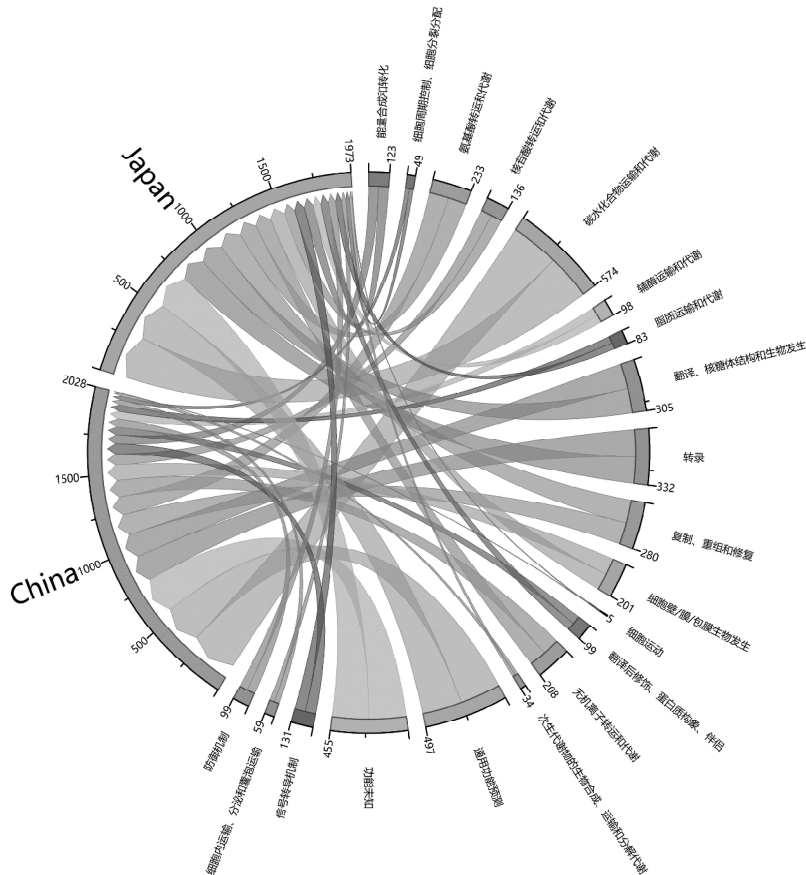


图 2 COG 聚类分析

虽然乳酶生株碳水化合物代谢相关基因总数大于表飞鸣株,但 CAZY 数据库分析结果显示,乳酶生株的 *GH23*, *GT4* 和 *GT35* 等家族基因数量明显少于表飞鸣株(见表 1)。根据序列以及结构相似性,糖苷水解酶(glycoside hydrolase, GH)分为 173 个家族,其中 *GH23* 家族基因主要包括几丁质酶^[23]、G 型溶菌酶和肽聚糖裂解酶^[24],这些酶能够有效降解细菌细胞壁并干扰微生物生长,*GH23* 家族基因丰度高的微生物在复杂微生态环境下的生存能力强。糖苷转移酶(glycosyl transferase, GT)分为 115 个家族,其

中 *GT4*^[25] 和 *GT35*^[26] 具有广泛的糖苷转移及磷酸化的能力,可有效提高微生物的定殖及黏附能力。在本项目组前期的研究过程中发现,表飞鸣株的菌种活力较强,计数培养仅需 24 h,培养基中的菌落形态较大,菌落周围的黄色光圈明显,且不易污染杂菌,而乳酶生株一般需培养 48 h 以上才能出现上述现象,且易污染杂菌,结合以上分析推测表飞鸣株在抵抗外源微生物侵入以及有效定殖方面优于乳酶生株,该特性具有提高表飞鸣株在肠道内功能发挥的作用,相关结果可通过转录组等方法进一步验证。

表 1 菌株的 Cazy 基因差异汇总

序号	家族基因 No.	乳酶生株	表飞鸣株	序号	家族基因 No.	乳酶生株	表飞鸣株	序号	家族基因 No.	乳酶生株	表飞鸣株
1	3.2.1.14	1	0	44	GH13_30	2	2	87	GH72	1	1
2	3.2.1.23	1	1	45	GH13_31	4	5	88	GH73	15	15
3	3.2.1.83	1	1	46	GH130	3	3	89	GH76	2	1
4	AA10	2	2	47	GH133	0	1	90	GH78	1	1
5	AA17	0	1	48	GH140	1	0	91	GH84	1	1
6	AA2	1	1	49	GH152	1	2	92	GH88	1	2
7	AA5	1	1	50	GH154	1	1	93	GH92	3	3
8	CBM0	1	1	51	GH170	6	6	94	GT0	4	3
9	CBM13	2	3	52	GH171	1	1	95	GT1	3	3
10	CBM16	1	1	53	GH18	4	4	96	GT106	1	1
11	CBM23	1	1	54	GH19	1	1	97	GT19	3	3
12	CBM26	1	1	55	GH2	7	7	98	GT2	56	56
13	CBM43	2	2	56	GH20	2	2	99	GT20	2	2
14	CBM47	1	1	57	GH23	11	15	100	GT22	1	1
15	CBM48	2	2	58	GH24	3	3	101	GT26	1	1
16	CBM50	21	20	59	GH25	5	4	102	GT27	1	1
17	CBM6	2	1	60	GH27	1	1	103	GT28	10	10
18	CBM66	1	0	61	GH28	5	5	104	GT29	1	1
19	CBM9	2	2	62	GH3	8	8	105	GT30	1	1
20	CE0	2	3	63	GH32	2	2	106	GT31	1	1
21	CE11	2	2	64	GH33	1	1	107	GT32	1	1
22	CE12	3	4	65	GH35	1	2	108	GT34	1	1
23	CE14	1	1	66	GH36	3	4	109	GT35	0	2
24	CE15	1	1	67	GH38	4	5	110	GT4	38	43
25	CE4	5	4	68	GH39	4	4	111	GT41	1	0
26	CE8	1	2	69	GH4	7	7	112	GT47	1	1
27	CE9	5	7	70	GH42	3	3	113	GT5	3	3
28	GH0	8	8	71	GH43_11	4	4	114	GT51	10	10
29	GH1	24	26	72	GH43_13	1	1	115	GT58	1	1
30	GH100	1	1	73	GH43_18	1	1	116	GT62	1	1
31	GH103	1	2	74	GH43_26	1	1	117	GT66	1	1
32	GH105	3	3	75	GH43_28	1	1	118	GT68	1	1
33	GH12	1	1	76	GH43_3	1	1	119	GT7	2	2
34	GH123	1	1	77	GH43_9	1	1	120	GT77	1	1
35	GH125	2	2	78	GH5_11	1	1	121	GT8	4	4
36	GH128	1	1	79	GH5_27	1	1	122	GT83	1	1
37	GH13	4	5	80	GH5_7	1	1	123	GT9	3	3
38	GH13_11	2	2	81	GH51	1	1	124	GT96	1	1
39	GH13_18	1	1	82	GH53	0	1	125	PL1_1	2	2
40	GH13_20	1	1	83	GH55	1	1	126	PL8	1	1
41	GH13_23	1	1	84	GH57	1	1	127	PL9_1	1	1
42	GH13_29	4	4	85	GH6	1	1				
43	GH13_3	2	2	86	GH65	8	9				

3 安全性分析

3.1 毒力因子 (virulence factors) 毒力因子对微生物相关疾病的发病机制具有重要的影响,目前已知的毒力因子包括酶、毒素、分泌效应物或细胞相关产物等^[27]。毒力因子分析采用 VFDB 数据库,选取相似度 $\geq 30\%$,覆盖度 $\geq 30\%$ 的分析结果,比对结果显示两株菌的毒力因子完全一致(见表2)。

表2 潜在毒力因子分析结果

序号	毒力因子分类	功能
1	ClpP	ATP 依赖性 Clp 蛋白酶亚基
2	ClpC	内肽酶 Clp ATP 结合链 C
3	ClpE	ATP 依赖性蛋白酶
4	Hyaluronic acid capsule	UDP-葡萄糖脱氢酶
5	Hyaluronic acid capsule	UTP-葡萄糖-1-磷酸尿苷酰转移酶
6	Capsule	8 型荚膜多糖合成蛋白 Cap8D
7	Capsule	荚膜多糖合成蛋白 Cps4E
8	Hsp60	Hsp60,60K 热休克蛋白 HtpB
9	LpeA	脂蛋白促进细胞侵袭
10	LplA1	脂蛋白连接酶
11	Lsp	信号肽酶 II
12	BSH	胆盐水解酶
13	EfaA	心内膜炎特异性抗原
14	Capsule	外膜 β 桶蛋白
15	Capsule	磷脂酰胞苷酰转移酶
16	Hyaluronidase	多糖裂解酶,家族 8
17	BopD	糖结合转录调节因子, LacI 家族
18	PI-1	分选酶
19	Lap	李斯特菌黏附蛋白 LapB
20	LPS	(3R)-羧基肉豆蔻酰 ACP 脱水酶
21	Ebp pili	分选酶
22	Acm	胶原黏附素前体 Acm
23	Scm	胶原黏附蛋白 Scm
24	Capsule	UDP-葡萄糖 6-脱氢酶
25	Capsule	NADP 依赖性磷酸葡萄糖酸脱氢酶

发现的潜在毒力因子基因主要是细胞表面因子,主要与宿主或表面黏附和促进生物膜形成有关,这些毒力因子被认为是诱发疾病的关键因素^[28]。同时也应该注意到,这些毒力因子在微生物细胞中本身并不具有毒力,在特定环境及代谢变化等多重因素影响下才可能具有毒力影响。如信号肽酶基因,只有当存在外源蛋白进入微生物细胞并且该蛋白为分泌性蛋白的情况下,信号肽酶基因才可能被触发并产生毒力作用,因此毒力因子的存在仅表明微生物具有安全性风险,需具体考察其功能转化条件才能确认其是否具有真实的毒力作用。

3.2 抗性因子 (antibiotic resistance) 微生物可作为抗生素抗性基因的储存库,抗生素抗性基因会转移给部分条件致病菌,因此了解相关的抗性机制、遗传性质以及有助于抗性传播的因素非常重要^[29]。抗生素抗性因子利用 CARD 数据库进行分析^[12],分析结果显示两菌株的抗生素抗性因子相同,有 13 种共 6 类抗生素抗性基因,分别是通过分子旁路赋予抗生素抗性的蛋白质编码基因、改变细胞壁电荷的基因、抗生素修饰失活基因、抗生素靶点保护蛋白、抗生素靶标替代蛋白以及外排泵复合物或亚基等(见图3,因两菌株具有完全相同的抗性因子,因此不进行区分)。这些基因可以通过分子修饰、细胞运输等机制改变微生物对抗生素的耐受性,当益生菌存在移动遗传元件时,可导致条件致病菌产生抗药性或通过其他途径影响人体健康。与毒力因子类似,抗生素抗性因子的存在仅表明微生物具有安全性风险,需考察其功能转化条件才能确认其是否具有真实作用。

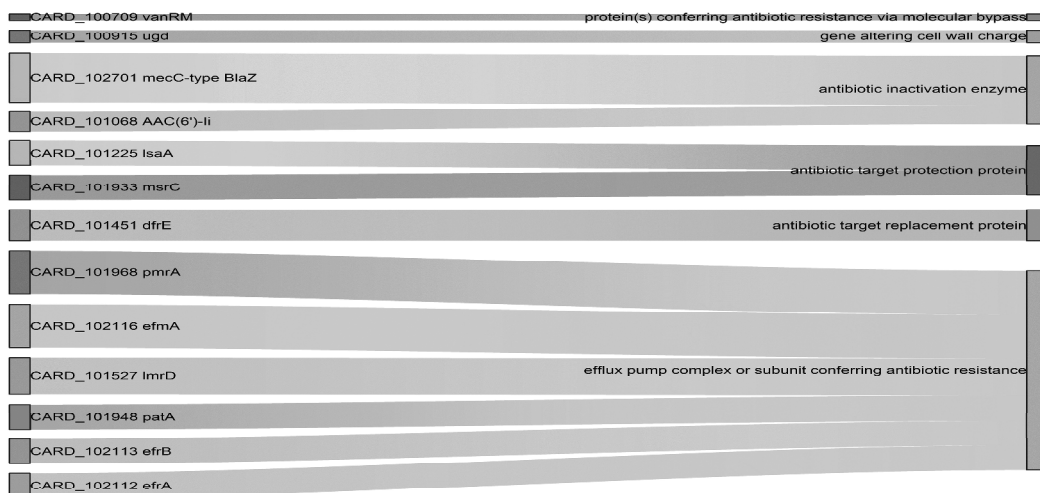


图3 抗生素抗性因子分析结果

3.3 细菌素基因 (bacteriocins) 细菌素是特定细菌在代谢过程中通过核糖体合成机制产生的一类具有抑菌活性的多肽或前体多肽,细菌素对同种近缘菌株呈现狭窄的抑制谱,通过细胞穿孔、抑制肽聚糖合成、抑制蛋白质合成、降解靶细胞 DNA 等机制起到抑菌或杀菌作用^[30]。细菌素编码基因可能与毒素基因相关,因此有必要探索其性质、作用机制和受体分子^[31]。通过检索 BACTIBASE 数据库^[18],在目标菌株中发现了多种细菌素,其中,乳酶生株发现了 Enterolysin A (EMBL: AAG29099.1) 的存在,表飞鸣株发现了包括 Enterocin P (EMBL: AAC45870.1), Zoocin A (EMBL: AAC46072.1), Gassericin T (EMBL: BAA82353.1) 以及 Lactacin-F (EMBL: AAS08589.1) 在内的 4 种细菌素。细菌素是细菌用于调控菌群结构的一种有力武器,在有利于细菌素生产菌侵入其他稳定微生物生态的同时,也可以作为细菌素生产菌的一种防御武器,防止自身建立的稳定生态受到其他细菌的侵犯。综合评价,表飞鸣株的细菌素种类较多,在肠道平衡维持能力方面存在较多的选择,但同时也意味着更高的肠道紊乱风险。

3.4 胺代谢基因 (biogenic amines) 肠道微生物合成的生物胺会导致人体的不良反应^[32],并带来毒理学风险,因此生物胺被认为是益生菌的不良代谢产物。

根据文献调研,在 Uniprot 数据库中检索包括精氨酸脱羧酶基因、胍丁胺酶基因和羧基去甲精胺脱羧酶基因等生物胺代谢相关基因^[10],将基因序列下载为 fasta 格式文件,通过 CLC Genomics Workbench (Version 22.0.2) 建立 DIAMOND Index,并对目标菌株基因组序列进行分析(见表 3)。结果显示乳酶生株有 21 个生物胺代谢相关基因,表飞鸣株有 23 个生物胺代谢相关基因,其中谷胱甘肽二硫化物还原酶和 VOC 结构蛋白仅在表飞鸣株中存在。谷胱甘肽二硫化物还原酶和 VOC 结构蛋白除与氨代谢有

关外,还涉及细胞铁死亡调控过程,该组基因在一定程度上增强了表飞鸣株的定殖能力及环境适应力。

表 3 生物胺相关代谢基因分析结果

序号	基因功能	乳酶生株	表飞鸣株
		基因数量	基因数量
1	胍酶 1	1	1
2	ABC 转运蛋白结构域	1	1
3	DNA 损伤诱导蛋白 F	1	1
4	NAD 依赖性蛋白质脱酰基酶	1	1
5	PERK1 样蛋白激酶	1	1
6	谷胱甘肽二硫化物还原酶	0	1
7	胍酶 2	3	3
8	胍酶 3	1	1
9	VOC 结构蛋白	0	1
10	精氨酸脱羧酶	4	4
11	鸟氨酸/二氨基庚二酸盐/精氨酸脱羧酶,家族 2	1	1
12	羧基亚精胺脱羧酶	2	2
13	肽聚糖/木聚糖/几丁质脱乙酰酶	1	1
14	未鉴定的蛋白质	2	2
15	乙酰乙酸代谢调节蛋白 AtoC	2	2
合计	—	21	23

3.5 前噬菌体 (prophages) 前噬菌体是指溶原性细菌内存在的整套噬菌体基因组,广泛分布于细菌基因组中,参与调节毒力因子、生物膜形成以及赋予宿主免疫力等^[33]。由于前噬菌体与编码重要毒力因子有关,同时具有促进遗传变异的能力,因此,前噬菌体分析对评价菌株安全性而言至关重要。前噬菌体分析通过 PHASTER 数据库进行^[16-17]。

在 2 株菌株中各发现了 3 个前噬菌体编码区域,且噬菌体归类基本一致(见表 4)。在发现的前噬菌体编码区域中,多为 DNA 结合蛋白、拓扑蛋白、包装折叠蛋白以及转运蛋白等相关基因,这些基因一旦被激活表达,将有可能辅助毒力因子等具有安全风险的基因表达产物形成功能。

表 4 前噬菌体分析结果

菌株	序号	序列长度	蛋白数量	序列定位	噬菌体名称	完成度
乳酶生株	1	26.8kb	28	1025912-1052781	Lister_A500_NC_009810(6)	Questionable
	2	45.3kb	62	1974534-2019898	EnterovB_IME197_NC_028671(10)	Intact
	3	59.2kb	58	42824-102048	Staphy_SPbeta_like_NC_029119(5)	Intact
表飞鸣株	1	8.9kb	12	2023559-2032506	Staphy_SPbeta_like_NC_029119(4)	Questionable
	2	41.9kb	58	841709-883637	EnterovB_IME197_NC_028671(11)	Intact
	3	41.1kb	33	1839457-1880647	Lister_A118_NC_003216(6)	Intact

3.6 移动遗传元件 (mobile genetic elements, MG-Es) 质粒、噬菌体、基因组岛、转座子和插入序列等移动遗传元件能够将携带有毒力因子和抗生素抗性基因的序列进行胞间转运,导致部分无安全风险

的微生物转化成高风险细胞,对宿主健康有重要影响^[34]。移动遗传元件分析通过 CRISPRCasFinder 数据库^[19]以及 CENSOR 数据库^[21]等数据库检索完成(见表 5)。

表 5 CRISPR Cas Finder 数据库检索结果

菌株	开端	尾端	长度	重复序列
乳酶生株	1 679 949	1 680 042	93	TGAGGAAGAAGGTGTTGTTTCTGCTGCA
乳酶生株	2 337 603	2 337 715	112	CTTTGCTCAACCAGTATTGTTTATCATCT
乳酶生株	430	557	127	TCAGAGGGTATGAAAATCATACC
表飞鸣株	1 215 422	1 215 515	93	TGCAGCAGAGACAACACCTTCTTCCTCA
表飞鸣株	2 387 975	2 388 087	112	CTTTGCTCAACCAGTATTGTTTATCATCT

检索结果显示,乳酶生株中存在 3 条 CRISPR 序列,表飞鸣株中存在 2 条 CRISPR 序列,其中长度为 112 bp 的 CTTTGCTCAACCAGTATTGTTTATCATCT 为共有序列。成簇规律间隔短回文重复序列 (clustered regularly interspaced short palindromic repeat sequences, CRISPRs) 是一类独特的 DNA 正向重复序列家族,广泛存在于原核生物基因组中,由重复序列以及单一序列间隔组成。CRISPR 的作用机制类似于真核生物的 RNAi,即利用间隔区识别并沉默外源的遗传物质。结合 Repeat 长度,本研究中菌株发现的 CRISPR 序列需结合其他分析进一步验证。

安全性分析结果显示,乳酶生株与表飞鸣株在毒力因子等方面具有类似的评价结果,但表飞鸣株所具有的部分代谢基因一定程度上丰富了表飞鸣株的环境适应性及定殖能力,丰富了微生物的代谢能力。

4 益生性分析

4.1 应激反应 (stress responses) 微生态制剂在制备及其使用过程中面临包括温度、有机酸、胆汁、渗透性和氧化等各种环境压力,这些环境压力会显著损害微生物的活性和生物学功能^[35],因此,良好的益生菌应具有一定抵抗环境压力的能力,这种能力称为应激反应。与应激反应能力相关的主要包括磷脂合成、碳水化合物代谢、糖酵解、嘧啶氨基酸代谢以及离子稳态(钴转运蛋白、铜转运蛋白和应激诱导的 DNA 结合蛋白)等相关功能基因。

应激反应相关的基因差异分析结果显示(见表 6),在总计 489 个相关基因中,表飞鸣株具有 32 个相关的基因,多于乳酶生株的 18 个,由于表飞鸣株的基因组较小,因此其应激反应相关基因密度较高,具有更为丰富的应激反应应对方式以及更好的环境适应性与代谢多样性。

表 6 应激反应相关基因差异分析

序号	乳酶生株	表飞鸣株
1	ATP 结合亚家族	HMP-PP 磷酸酶
2	3-酮-L-古洛糖酸-6-磷酸脱羧酶 UlaD	血红素反应调节器 HssR
3	镉、钴和锌/H(+)-K(+)-逆向转运体	胸苷激酶 β
4	超氧化物歧化酶[Mn]	铜感应转录抑制因子 CsoR
5	S-腺苷甲硫氨酸依赖性甲基转移酶 CRG1	核黄素生物合成蛋白 RibD
6	NtpR 蛋白	甲酰胺嘧啶-DNA 糖基化酶
7	兴奋性氨基酸转运蛋白 3	金属转运蛋白 CNNM4
8	氨基磷酸核糖基转移酶	截短的非活性核糖核酸酶 PH
9	离子运输 ATP 结合蛋白 FbpC	转运蛋白
10	钙转运 ATP 酶,内质网型	核糖体蛋白 L17-1
11	转运系统渗透酶蛋白 CbrC	3-酮-L-古洛糖酸-6-磷酸脱羧酶 SgbH
12	铁硫簇转运蛋白 atm1	丙酮酸激酶
13	反转运蛋白 QDR1	线粒体酰基载体蛋白
14	铁硫簇转运蛋白 ABCB7	胞苷脱氨酶

序号	乳酶生株	表飞鸣株
15	甘油激酶	前噬菌体衍生的二硫键形成蛋白 A
16	三磷酸肌苷焦磷酸酶	磷脂酰甘油赖氨酰转移酶
17	氨基甲酰磷酸合酶小亚基	环状 2-AMP 合酶 CdaA
18	天冬氨酸氨基甲酰转移酶	铜离子伴侣蛋白 CopZ
19	—	高亲和力 Na(+)/H(+) 反转运蛋白 NhaS3
20	—	转录阻遏物 SmtB
21	—	3-磷酸甘油醛脱氢酶
22	—	反转运蛋白 QDR2
23	—	铁硫簇转运蛋白 atm1
24	—	谷胱甘肽还原酶
25	—	2,3-二磷酸甘油酸非依赖性磷酸甘油酸变位酶
26	—	氨基甲酰磷酸合成酶大亚基
27	—	细胞内铁伴侣蛋白
28	—	5'-甲基硫腺苷/S-腺苷高半胱氨酸核苷酶
29	—	系统渗透酶蛋白 FpuB
30	—	氨基甲酰磷酸合成酶小亚基
31	—	钠/氢转运蛋白
32	—	核酸内切酶 III
合计	18	32

4.2 抗菌及拮抗活性 (antimicrobial and antagonistic activity)

抗菌和拮抗活性为益生菌提供了抑制其他物种(主要是病原菌)生长的能力,这些活性与竞争性定殖或抗微生物化合物(如有机酸、短链脂肪酸、二乙酰、乙偶姻、细菌素等)合成有关。另外,黏附功能相关蛋白(如烯醇化酶、纤连蛋白结合蛋白、胞外多糖生物合成基因簇、磷酸丙糖异构酶

和鞭毛钩相关蛋白等)也与抗菌和拮抗活性功能相关^[36]。

抗菌及拮抗活性相关的基因差异分析结果显示(见表 7),两株菌检索出的与抗菌及拮抗活性相关的基因共计 106 个,其中表飞鸣株具有 103 个,乳酶生株具有 101 个,基因丰富度两者基本相当,在抗菌及拮抗活性方面,表飞鸣株略优于乳酶生株。

表 7 抗菌及拮抗活性基因差异分析

序号	乳酶生株	表飞鸣株
1	蛋白质精氨酸 N-甲基转移酶 1	细菌素 curvacin-A
2	N-乙酰胞壁酰-L-丙氨酸酰胺酶 LytC	乳酰谷胱甘肽裂解酶
3	赖氨酰磷脂酰甘油生物合成双功能蛋白 LysX	IV 型菌毛组装 ATP 酶 PilB
4	—	磷脂酰甘油赖氨酰转移酶
5	—	双功能甲基硫代核酮糖-1-磷酸脱水酶
合计	3	5

4.3 抗氧化能力 (antioxidant activity) 氧化应激是各种慢性人类疾病的原因之一,氧化应激是由氧化反应过程增加的活性氧分子(reactive oxygen species, ROS)导致。氧化反应可以产生具有不成对电子的自由基进行快速的连锁反应,从而破坏其他的生物分子并产生更多的自由基(如超氧阴离子、羟基自由基、过渡金属等)^[37]。部分益生菌被认为具有抗氧化及调节宿主体内氧化环境的能力,这种能

力主要通过螯合金属离子、产生抗氧化代谢物、调节信号通路、下调活性氧生成酶的活性或通过微生物自身的抗氧化酶系统实现^[38]。

抗氧化能力相关的基因差异分析结果显示(见表 8),两株菌检索出的相关基因共计 46 个,其中表飞鸣株 43 个,乳酶生株 43 个,基因丰富度两者一致,共有基因 40 个,两菌株抗氧化能力相近。

表 8 抗氧化能力基因差异分析

序号	乳酶生株特有	表飞鸣株特有
1	高还原性聚酮合酶 <i>alnA</i>	甲基间苯二酚合酶
2	超氧化物歧化酶 [Mn]	铜转运蛋白 CCH
3	铁抑制蛋白 1	硫氧还蛋白还原酶 1
合计	3	3

4.4 黏附能力 (adhesion ability) 益生菌的黏附定殖可以通过竞争性排除和通过促进抗菌代谢物的局部释放发挥针对肠道病原体的潜在防御功能。这种功能由多种表面细胞黏附因子介导,如脂磷壁酸、

细胞表面相关蛋白或细胞外黏附素等^[39]。另外,菌毛操纵子、胶原蛋白结合域、黏蛋白结合域、黏蛋白相互作用蛋白、细胞外基质黏附结构域和纤连蛋白结合蛋白等也在这一功能中发挥作用。这些基因的存在促进了微生物的定殖并确保其在宿主肠道中的持久性。

黏附能力相关的基因差异分析结果显示(见表 9),在两株菌检索出的相关基因共计 444 个,其中表飞鸣株具有 430 个,乳酶生株具有 405 个,可认为表飞鸣株具有更好的黏附及定殖能力。基因丰富度两者基本一致。

表 9 黏附能力基因差异分析

序号	乳酶生株	表飞鸣株
1	错配修复蛋白 MutL	50S 核糖体蛋白 L17
2	假尿苷合酶 A	50S 核糖体蛋白 L19
3	热敏酶	50S 核糖体蛋白 L22
4	胰蛋白酶抗性表面 T6 蛋白	50S 核糖体蛋白 L27
5	蛋白激酶 2	GTPase HflX
6	蛋白质饥饿期间的 DNA 保护	类核相关蛋白 Amuc
7	PilH 蛋白	蛋白质-葡糖基半乳糖基羟赖氨酸葡糖苷酶
8	兴奋性氨基酸转运蛋白 3	丝裂原活化蛋白激酶 MST11
9	超氧化物歧化酶 [Fe]	DNA 结合蛋白 HU
10	N-乙酰胞壁酰-L-丙氨酸酰胺酶 LytC	细胞外半胱氨酸蛋白酶
11	核糖体蛋白 S21	ComG 操纵子蛋白 1
12	UDP-N-乙酰醇丙酮酰氨基葡萄糖还原酶	囊性纤维化跨膜电导调节剂
13	去甲基茶碱甲基转移酶	钠/氢转运蛋白 1
14	ATP 结合亚家族	囊性纤维化跨膜电导调节剂
15	—	推定的水解酶 EbsB
16	—	胶原酶
17	—	ATP 结合蛋白
18	—	30S 核糖体蛋白 S19
19	—	30S 核糖体蛋白 S17
20	—	50S 核糖体蛋白 L14
21	—	50S 核糖体蛋白 L24
22	—	50S 核糖体蛋白 L30
23	—	30S 核糖体蛋白 S11
24	—	50S 核糖体蛋白 L32
25	—	DNA 错配修复蛋白 MutL
26	—	30S 核糖体蛋白 S18
27	—	烟酸-核苷酸腺苷酸转移酶
28	—	50S 核糖体蛋白 L31 B 型
29	—	50S 核糖体蛋白 L35
30	—	50S 核糖体蛋白 L21
31	—	GMP 合酶 [谷氨酰胺水解]
32	—	胸苷激酶
33	—	UDP-N-乙酰氨基葡萄糖
34	—	30S 核糖体蛋白 S20
35	—	BF0270 蛋白
36	—	D-丙氨酸载体蛋白
37	—	ComG 操纵子蛋白 3 同源物
38	—	去磷酸化辅酶 A 激酶
39	—	Holo-[酰基载体蛋白]合酶
合计	14	39

讨 论

1 屎肠球菌基因组比对分析总结及功能基因分析对药物质量提升的意义

本文以乳酶生及表飞鸣生产用菌株即屎肠球菌为研究对象,探索在基因组层面上预测、分析及比较不同菌株作为益生菌的安全性及有效性。与模式菌株相比,乳酶生株和表飞鸣株的基因组间共线性关系较好,存在少量的插入、缺失、倒位和易位等基因重排事件,基因组结构相似性较高。乳酶生株的基因数量较多,但分布相对分散,因此基因密度较低。在基因总数较多的情况下,乳酶生株的碳水化合物代谢基因(CAZy)数量明显少于表飞鸣株,这说明在碳水化合物代谢方面表飞鸣株的基因密度大于乳酶生株,这种差异意味着在营养匮乏的条件下,表飞鸣株将具有更好的环境适应性和存活能力。与该结果相对应,表飞鸣有效期为36个月,而乳酶生有效期仅为24个月,产品有效期的差异印证了这一点。

针对菌株的安全性评估主要包括毒力基因、耐药性基因、细菌素基因、生物胺基因、原噬菌体以及移动遗传元件等6个方面。分析结果显示,在屎肠球菌种内株间的毒力因子和耐药基因差异不显著,两菌株完全一致。在细菌素合成基因方面,乳酶生株仅发现Enterolysin A,表飞鸣株发现了包括Enterocin P在内的4种细菌素,细菌素生产菌、敏感菌与周围其他共栖菌群可以达到一种动态的平衡,具有细菌素分泌能力的肠道菌的定殖能力更强。在生物胺代谢基因方面,表飞鸣株以23个基因多于乳酶生株的21个,其中谷胱甘肽二硫化物还原酶和VOC结构蛋白仅在表飞鸣株中存在,这2个基因除具有胺代谢能力外,还涉及细胞过程相关功能,在一定程度上提高了表飞鸣株的环境适应力。在前噬菌体及移动遗传元件方面,两菌株各发现了3个前噬菌体编码区域,且噬菌体归类基本一致,综上,初步认为在相同环境下,两株微生物的安全性风险基本一致,其中部分毒力基因的多能性反而提高了表飞鸣株的环境耐受力。

针对菌株的益生性评价主要包括应激反应基因、抗菌/拮抗活性基因、抗氧化活性基因以及黏附

基因等4个方面。分析结果显示表飞鸣株在各方面均具有较优的基因表达潜力及基因密度,同时基因功能分布广泛,益生性及环境适应能力优于乳酶生株。本课题组的前期相关实验工作发现,在相同培养条件下,乳酶生株的生长情况较差,生长曲线试验显示,乳酶生株复苏时间长于表飞鸣株,且乳酶生株对环境的适应能力低于表飞鸣株(相关结果待发表)。该结果与本文比较基因组分析的结果基本一致,还需要转录组、蛋白组和代谢组的进一步验证。

乳酶生及表飞鸣在长达半个多世纪的应用中,积累了大量的临床及药学研究数据,其中部分研究及相关数据从基因功能的角度开展了探索性分析,因此,在未来进行药物质量提升或筛选新型屎肠球菌菌株时,在临床数据的基础上,结合基因组分析结果,利用菌株的有益基因,规避有害基因,可以使药物更好地发挥益生作用。

2 益生菌基因组学在新型益生菌筛选及质量评价上的应用展望

由于越来越多的科学证据支持益生菌对人体健康的积极作用,益生菌市场成为了医药行业增长最快的领域之一,这也鼓励了开发潜在药用益生菌的研究工作。在过去的若干年里,应用于益生菌研究的基因组测序和组学方法的快速发展已经实现了遗传分子信息的广泛积累,借助于大量生物信息学数据的可用性,益生菌基因组学在选择及评价具有工业开发应用价值的潜在益生菌菌株方面将产生积极的影响。图4提供了一种理性筛选及设计益生菌的研究方法^[5]。通过比较基因组的应用,新发掘菌株与应用成熟菌株间的遗传背景和基因表达谱比较将了解其种内变异提供新的方向及方法,并开拓一种包含基因组进化、细胞新陈代谢和特定分子机制的评价和比较益生菌的新思路,进而实现基于功能优化目标的益生菌菌株基因组的工程化定向改造等技术进步,为新型“生物工程益生菌”的开发创造新的机会^[40]。因此,利用益生菌基因组学提供的分子信息并根据所追求的应用类型及目标,基因工程可能会改变益生菌菌株的基因功能,以实现基于目标的工程菌改造甚至定制化益生菌菌株的开发。

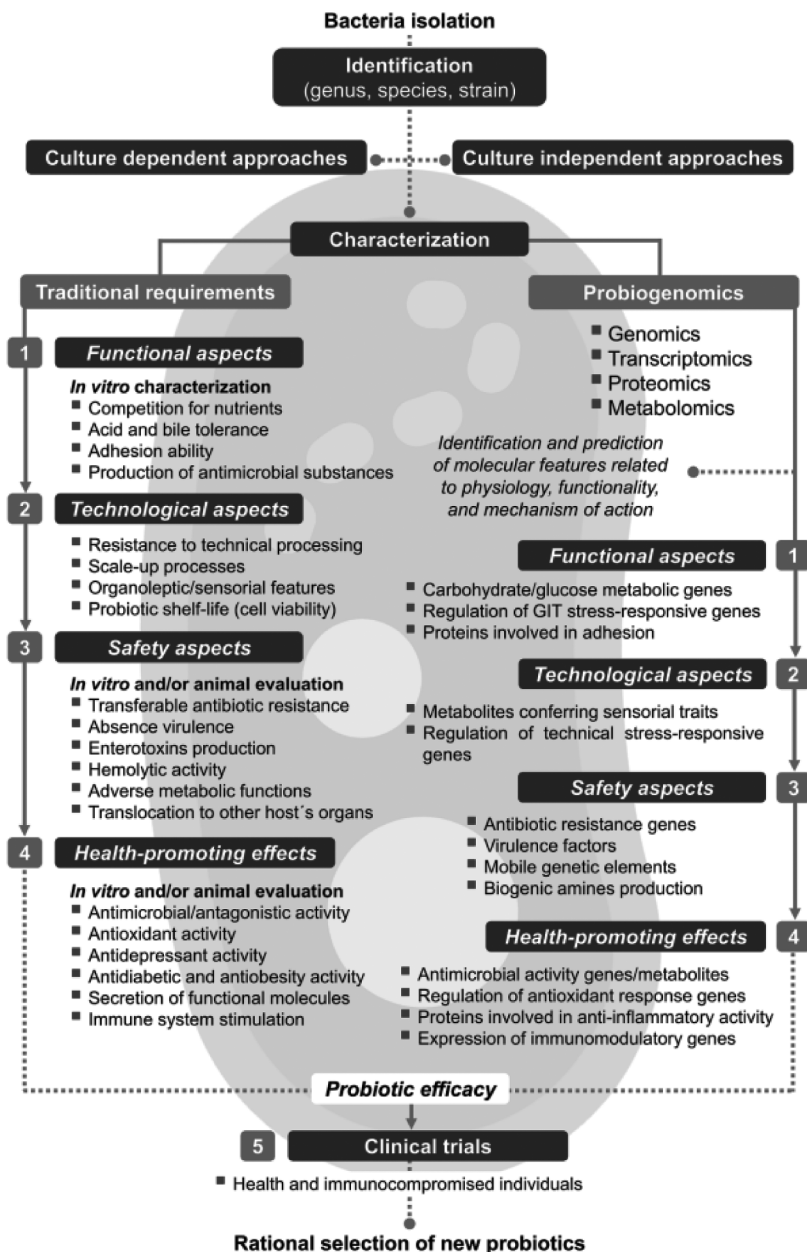


图 4 新型益生菌的理性筛选^[5]

另外,益生菌具有的包括预防慢性疾病、改善营养状况、增强免疫力、促进生长、减轻压力等多种潜在的益处使其在制药行业中广受欢迎^[41]。消费者的需求正在推动含有益生菌药品的数量和类型的增加。因此,如何寻找和评价具有独特功能特性的新的、天然的或非常规的细菌菌株,并对同种菌株进行功效比较成为目前的研究及监管重点^[42]。对于益生菌的验证和选择没有通用的路线图,由于标准化缺失,传统的体外和体内试验存在结果评价的差异,而且这些试验结果并不总是与益生菌的功效和安全

性有关。这些差异可能会使评价益生菌的安全性和功能性的努力变得复杂。如何在理论层面预测及评价益生菌菌株的安全及有效性,并经由理论指导试验是较为合理的研究思路,基于上述考虑,“益生菌组学”引起了广泛关注并获得了快速发展^[43]。

为保证行业正常发展,使益生菌制剂安全有效地为人类健康服务,监管部门对益生菌制剂的质量标准问题日益关注。然而,在菌株开发验证程序方面存在的问题已经明显制约了行业的发展,在益生菌开发、评价、比较及验证方面仍然存在大量的技术

瓶颈需要解决。与化学药物类似,微生态制剂也需要建立有效的一致性评价、药品功效评估等体系,从这个意义上说,益生菌组学技术不仅可以为开发新的安全益生菌提供助力,而且还可以为预测、评价、比较菌株的安全有效性提供新的方法和思路^[44]。然而,需要注意的是,尽管益生菌组学技术优势明显,但也存在一些挑战,尤其是在生成数据的正确解释及数据库选择的科学性方面,如果克服这些障碍,益生菌组学无疑将在开发及评价益生菌方面发挥更为重要的作用。

希望本研究能够为降低或消除现有微生态制剂质量标准体系存在的风险提供一种新的方法和思路,并能够促进相关标准的提高,构建有效的微生态制剂一致性评价比较体系,为行业健康发展及有效监管提供助力。

[参 考 文 献]

- [1] 王珊珊,石继春,杜宗利,等. 乳酶生生产用菌种的全基因组序列分析研究[J]. 中国医药生物技术, 2020, 15(6): 577-583.
- [2] 邢晟,王淑华,胡文红,等. 益生菌制剂开发及相关质量标准构建策略[J]. 中国新药杂志, 2021, 30(9): 773-779.
- [3] 马磊,郑超,黄丽芬,等. 多重 RPA-LFD 技术快速检测磺胺类耐药基因研究[J]. 中国计量大学学报, 2019, 30(3): 317-322, 366.
- [4] 田万红,喻钢,赵志晶,等. 乳酶生生产菌株的鉴定及耐药谱的分析[J]. 中国生物制品学杂志, 2013, 26(9): 1285-1289.
- [5] CASTRO-LÓPEZ C, GARCÍA HS, GUADALUPE MARTÍNEZ-ÁVILA GC, et al. Genomics-based approaches to identify and predict the health-promoting and safety activities of promising probiotic strains-A probiogenomics review [J]. *Trends Food Sci Technol*, 2020, 108(4): 148-163.
- [6] 焦月华. 基因组学分析粪肠球菌益生功能及增殖调控研究[D]. 哈尔滨: 哈尔滨工业大学, 2016.
- [7] KANEHISA M, GOTO S. KEGG: Kyoto encyclopedia of genes and genomes[J]. *Nucleic Acids Res*, 2000, 28(1): 27-30.
- [8] HARRIS MA, CLARK J, IRELAND A, et al. The Gene Ontology (GO) database and informatics resource [J]. *Nucleic Acids Res*, 2004, 32(Database issue): D258-D261.
- [9] TATUSOV RL, GALPERIN MY, NATALE DA, et al. The COG database: a tool for genome-scale analysis of protein functions and evolution[J]. *Nucleic Acids Res*, 2000, 28(1): 33-36.
- [10] CONSORTIUM TU. UniProt: a hub for protein information [J]. *Nucleic Acids Res*, 2015, 43(D1): D204-D212.
- [11] FINN RD, COGGILL P, EBERHARDT RY, et al. The Pfam protein families database: towards a more sustainable future [J]. *Nucleic Acids Res*, 2016, 44(D1): D279-D285.
- [12] ALCOCK BP, RAPHENYA AR, LAU TTY, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database [J]. *Nucleic Acids Res*, 2020, 48(D1): D517-D525.
- [13] DRULA E, GARRON ML, DOGAN S, et al. The carbohydrate-active enzyme database: functions and literature [J]. *Nucleic Acids Res*, 2022, 50(D1): D571-D577.
- [14] URBAN M, PANT R, RAGHUNATH A, et al. The Pathogen-Host Interactions database (PHI-base): additions and future developments [J]. *Nucleic Acids Res*, 2015, 43(Database issue): D645-D655.
- [15] LIU B, ZHENG DD, ZHOU SY, et al. VFDB 2022: a general classification scheme for bacterial virulence factors [J]. *Nucleic Acids Res*, 2022, 50(D1): D912-D917.
- [16] ARNDT D, GRANT JR, MARCU A, et al. PHASTER: a better, faster version of the PHAST phage search tool [J]. *Nucleic Acids Res*, 2016, 44(W1): W16-W21.
- [17] ZHOU Y, LIANG YJ, LYNCH KH, et al. PHAST: a fast phage search tool [J]. *Nucleic Acids Res*, 2011, 39(Web Server issue): W347-W352.
- [18] HAMMAMI R, ZOUHIR A, BEN HAMIDA J, et al. BACTIBASE: a new web-accessible database for bacteriocin characterization [J]. *BMC Microbiol*, 2007, 89(7): 1-6.
- [19] COUVIN D, BERNHEIM A, TOFFANO-NIOCHE C, et al. CRISPRCasFinder, an update of CRISPRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins [J]. *Nucleic Acids Res*, 2018, 46(W1): W246-W251.
- [20] ALMEIDA A, NAYFACH S, BOLAND M, et al. A unified catalog of 204, 938 reference genomes from the human gut microbiome [J]. *Nat Biotechnol*, 2021, 39(1): 105-114.
- [21] JURKA J, KLONOWSKI P, DAGMAN V, et al. CENSOR: a program for identification and elimination of repetitive elements from DNA sequences [J]. *Comput Chem*, 1996, 20(1): 119-121.
- [22] HERNÁNDEZ-SALMERÓN JE, MORENO-HAGELSIEB G. Progress in quickly finding orthologs as reciprocal best hits: comparing blast, last, diamond and MMseqs2 [J]. *BMC Genomics*, 2020, 21(1): 741.
- [23] FUKUSHIMA T, KITAJIMA T, YAMAGUCHI H, et al. Identification and characterization of novel cell wall hydrolase CwlT: a two-domain autolysin exhibiting n-acetylmuramidase and DL-endopeptidase activities [J]. *J Biol Chem*, 2008, 283(17): 11117-11125.
- [24] SUTRISNO A, UEDA M, ABE Y, et al. A chitinase with high activity toward partially N-acetylated chitosan from a new, moderately thermophilic, chitin-degrading bacterium, *Ralstonia* sp. A-471 [J]. *Appl Microbiol Biotechnol*, 2004, 63(4): 398-406.
- [25] GABALLA A, NEWTON GL, ANTELMANN H, et al. Biosynthesis and functions of bacillithiol, a major low-molecular-weight thiol in *Bacilli* [J]. *Proc Natl Acad Sci USA*, 2010, 107(14): 6482-6486.
- [26] ROGERS PV, LUO S, SUCIC JF, et al. Characterization and cloning of glycogen phosphorylase I from *Dictyostelium discoideum* [J]. *Biochim Biophys Acta BBA Gene Struct Expr*, 1992, 1129(3): 262-272.
- [27] 郑丹丹. 病原菌毒力因子数据分析平台的设计与构建 [D]. 北京: 北京协和医学院, 2017.
- [28] LI BL, ZHAN M, EVIVIE SE, et al. Evaluating the safety of potential probiotic *Enterococcus durans* KLD56.0930 using whole genome sequencing and oral toxicity study [J]. *Front Microbiol*, 2018, 9: 1943.
- [29] ROZMAN V, MOHAR LORBEG P, ACCETTO T, et al. Characterization of antimicrobial resistance in lactobacilli and bifidobacteria used as probiotics or starter cultures based on integration of phenotypic and *in silico* data [J]. *Int J Food Microbiol*, 2020, 314: 108388.
- [30] JUTURU V, WU JC. Microbial production of bacteriocins: latest research development and applications [J]. *Biotechnol Adv*, 2018, 36(8): 2187-2200.
- [31] MORTON JT, FREED SD, LEE SW, et al. A large scale pre-

- diction of bacteriocin gene blocks suggests a wide functional spectrum for bacteriocins[J]. *BMC Bioinformatics*, 2015, 2015(16): 381–389.
- [32] RUIZ-CAPILLAS C, HERRERO AM. Impact of biogenic amines on food quality and safety[J]. *Foods (Basel, Switzerland)*, 2019, 8(2): 1–16.
- [33] JAROCKI P, PODLEŚNY M, KRAWCZYK M, *et al.* Complete genome sequence of *Lactobacillus rhamnosus* Pen, a probiotic component of a medicine used in prevention of antibiotic-associated diarrhoea in children[J]. *Gut Pathog*, 2018, 10(5): 1–6.
- [34] ARKHIPOVA IR, RICE PA. Mobile genetic elements: in silico, in vitro, in vivo[J]. *Mol Ecol*, 2016, 25(5): 1027–1031.
- [35] MBYE M, BAIG MA, ABUQAMAR SF, *et al.* Updates on understanding of probiotic lactic acid bacteria responses to environmental stresses and highlights on proteomic analyses[J]. *Compr Rev Food Sci Food Saf*, 2020, 19(3): 1110–1124.
- [36] KAPSE NG, ENGINEER AS, GOWDAMAN V, *et al.* Functional annotation of the genome unravels probiotic potential of *Bacillus coagulans* HS243[J]. *Genomics*, 2019, 111(4): 921–929.
- [37] MISHRA V, SHAH C, MOKASHE N, *et al.* Probiotics as potential antioxidants: a systematic review[J]. *J Agric Food Chem*, 2015, 63(14): 3615–3626.
- [38] NYANZI R, SHUPING DSS, JOOSTE PJ, *et al.* Antibacterial and antioxidant activity of extracts from selected probiotic bacteria [J]. *J Food Res*, 2015, 4(5): 122–132.
- [39] MONTEAGUDO-MERA A, RASTALL RA, GIBSON GR, *et al.* Adhesion mechanisms mediated by probiotics and prebiotics and their potential impact on human health[J]. *Appl Microbiol Biotechnol*, 2019, 103(16): 6463–6472.
- [40] BRAVO D, LANDETE JM. Genetic engineering as a powerful tool to improve probiotic strains[J]. *Biotechnol Genet Eng Rev*, 2017, 33(2): 173–189.
- [41] DE SIMONE C. The unregulated probiotic market[J]. *Clin Gastroenterol Hepatol*, 2019, 17(5): 809–817.
- [42] LIN TL, SHU CC, LAI WF, *et al.* Investiture of next generation probiotics on amelioration of diseases-Strains do matter[J]. *Med Microecol*, 2019, 1–2: 100002.
- [43] VENTURA M, TURRONI F, VAN SINDEREN D. Probiogenomics as a tool to obtain genetic insights into adaptation of probiotic bacteria to the human gut[J]. *Bioengineered*, 2012, 3(2): 73–79.
- [44] 王峻霞, 李梦颖, 蒋蓉. 基于 Celltrion 公司案例分析韩国生物药产业发展政策及其启示[J]. *中国医药工业杂志*, 2021, 52(8): 1121–1127.

编辑:杨青/接受日期:2023-01-03