

副干酪乳酪杆菌 PC-01 不同表型菌落的基因组分析

张静雯, 田桂铭, 孙志宏, 于洁*

(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 乳酸菌与发酵乳制品省部共建协同创新中心 呼和浩特 010018)

摘要 目的:探究副干酪乳酪杆菌 PC-01 经高密度发酵后,不同表型菌落的基因组差异。方法:采用第 2 代测序技术,对高密度发酵的 10 株具有不同表型特征的副干酪乳酪杆菌 PC-01 分离株进行全基因组测序,通过比较基因组学方法揭示这 10 株分离株之间的差异。结果:10 株副干酪乳酪杆菌 PC-01 分离株的基因组大小为 2.69~2.83 Mb,GC 含量为 46.50%~46.64%,CDs 数目为 2 793~3 406 个,不同菌落的基因组大小、GC 含量存在较小差异。功能注释结果显示不同菌株编码碳水化合物化合物、蛋白质代谢以及参与氨基酸及其衍生物合成和降解的基因数量及种类均存在差异。基于核心基因构建系统发育树,发现具有相似菌落形态的菌株在同一分支。单核苷酸多态性(SNPs)结果表明:每两个相同菌落形态的不同分离株存在相同的突变位点,这也进一步证实菌落形态与基因型相关联。分析发现 10 株分离株中均存在糖基转移酶(*epsJ*、*pbpF*、*ponA* 等)及 ABC 蛋白家族(*yheS*、*bceB*、*bceA*、*pcrA* 和 *uvrA* 等)相关基因发生非同义突变,使菌株在高密度发酵过程中,能更好地适应环境而存活下来。菌落 PC-01-3 和 PC-01-4 与其它菌落形态具有明显差异,*ltaS* 基因发生非同义突变。脂磷壁酸合酶(*LtaS*)的缺乏会导致细菌表现出细胞分裂受损和生长缺陷,这可能是导致菌落形态差异的主要原因。结论:从全基因组水平完成了副干酪乳酪杆菌 PC-01 的 10 个分离株的比较分析,解析了同一菌株高密度发酵后不同表型菌落之间的差异,为副干酪乳酪杆菌 PC-01 的后续研究提供数据支持。

关键词 副干酪乳酪杆菌 PC-01; 菌落; 表型; 比较基因组学分析

文章编号 1009-7848(2024)03-0221-10 DOI: 10.16429/j.1009-7848.2024.03.022

随着社会的不断发展,乳酸菌制品逐渐走进大众的视野,并且乳酸菌已经广泛用于食品、医药健康、农牧业、工业制造等领域^[1]。副干酪乳酪杆菌(*Lactocaseibacillus paracasei*)是一种兼性厌氧的革兰氏阳性菌,同时也是重要的乳酸菌菌种之一。副干酪乳酪杆菌分布广泛,可以从人体肠道、口腔、发酵制品以及青贮饲料中分离。Marcitto 等^[2]研究表明,副干酪乳杆菌能够抵抗胃酸和肠道胆盐,顺利抵达婴儿肠道,并逐渐成为人体乳杆菌的优势菌群。Dang 等^[3]发现副干酪乳杆菌 TD062 具有较高的 α -酶抑制剂活性,良好的抗糖尿病能力,可作为一种新型的降血糖剂。本研究使用的副干酪乳酪杆菌 PC-01 不仅可以调节肠道菌群平衡,具有良好的耐酸、耐胆盐特性^[4],还具备预防高血压和降低胆固醇等益生特性。

随着测序技术的发展,全基因组测序技术被

广泛应用于菌株的分类鉴定^[5]。比较基因组学是在全基因组测序的基础上,分析种内或种间基因功能的相似性和差异性,以及物种的进化关系^[6]。Smokvina 等^[7]对 34 株副干酪乳杆菌进行基因组的比较分析,得到其泛基因组包含 4 200 个同源基因,并从核心基因集中发现了与动物宿主相关的因子。Salveti 等^[8]基于系统发育距离和保守蛋白质比例,将 269 株乳杆菌重新分类,建立系统发育标记,为其遗传多样性的探究提供基础。Broadbent 等^[9]通过比较基因组学研究揭示干酪乳杆菌基因组的共线性较高,具有几个高度保守基因含量的大块。局部共线性区域的数量随着基因组数量的增加而增加,基于相似性谱高度,检测到干酪乳杆菌之间明显的总体序列同源性,体现了局部多态性。Ghattargi 等^[10]通过全基因组测序和比较基因组学分析,研究菌株中与益生菌特性相关的关键遗传因素。

本研究通过全基因组测序技术结合比较基因组学的方法,以经过高密度发酵的 10 株具有不同表型特征的副干酪乳酪杆菌 PC-01 分离株为研究

收稿日期: 2023-03-11

基金项目: 国家自然科学基金面上项目(31972054)

第一作者: 张静雯,女,硕士生

通信作者: 于洁 E-mail: yujie8301@sina.com

对象,采用比较基因组学和表型观察,从基因组水平上揭示同一菌株不同菌落间的差异性,为菌株在高密度发酵上的应用提供遗传学参考。

1 材料与方法

1.1 菌株来源

副干酪乳酪杆菌 PC-01 菌株来自内蒙古农业大学乳酸菌菌种资源库 (Lactic Acid Bacteria Collection Center, LABCC), 菌株编号为 I-MAU60230, 副干酪乳酪杆菌 PC-01 (I-MAU60230) 是 2007 年分离自西藏拉萨地区酸牦牛奶的一株具有优良益生特性的菌株。

2022 年 3 月 14 日从 NCBI (National Center for Biotechnology Information, <https://www.ncbi.nlm.nih.gov/>) RefSeq 数据库下载 1 株副干酪乳酪杆菌 PC-01 和副干酪乳酪杆菌 Zhang 基因组完成图。

1.2 主要试剂和仪器

主要试剂包括 MRS 培养基,广州环凯微生物科技有限公司;Wizard® Genomic DNA Purification Kit, 5×TBE 电泳缓冲液、琼脂糖、核酸染料 GELVIEW, 北京百泰克;DL2000 DNA Marker (Takara)、6×DNA Loading Buffer, Trans Gen。

主要仪器包括 ZHJH-C 型智能安全型超净工作台, 上海智城分析仪器有限公司;HWS28 型水浴锅, 天津知署科技有限公司;恒温培养箱, 上海恒科技仪器有限公司;超微量紫外分光光度计, NanoDrop 公司;光学显微镜, 日本奥林巴斯 (OLYMPUS);漩涡振荡器, 美国 Scientific Industries 公司;电泳仪, 北京六一仪器厂;凝胶成像仪, 美国 UVP 公司;测序仪, 美国 Illumina 公司。

1.3 菌株的获取

1.3.1 菌种活化及增殖培养 将冷冻保存(-80℃)的副干酪乳酪杆菌 PC-01 接种于 MRS 液体培养基中进行活化, 32.5℃恒温需氧培养 24 h, 再以 2%(体积分数)的比例接种于种子培养基中, 32.5℃恒温需氧培养 24 h, 并且每次传代均需进行革兰氏染色镜检, 确保在传代过程中没有被其它微生物污染。最后按 5%(体积分数)接种量接种于发酵培养基中进行高密度发酵, 32.5℃恒温需氧培养至对数期结束, 将发酵终点的发酵液, 在添加

2%琼脂的 MRS 平板上涂布, 32.5℃需氧培养 48 h, 观察并记录菌落形态。

1.3.2 菌株 DNA 提取 将上述挑取的单菌落, 加入 PBS 缓冲液, 振荡 30 s 混合均匀, 然后以 4 500 r/min, 10 min 离心弃上清液, 保留菌泥, 使用 Wizard® Genomic DNA Purification Kit 试剂盒提取 DNA, 具体试验步骤参考说明书, 并使用 Nano Drop 微量紫外分光光度计及 0.8%的琼脂糖凝胶电泳对所提取的 DNA 进行纯度和完整度的检测。

1.4 全基因组的测序与组装

将质量合格的 DNA 样品低温运输至诺禾致源生物信息科技有限公司, 基于第 2 代高通量测序平台 IlluminaNovaSeq 6000^[11]进行全基因组序列测定。

对得到的数据进行过滤, 获得高质量的 clean data 后, 再运用软件包 SOAPdenovo v2^[12]对高质量的 reads 进行拼接, 并调试不同的 kmer 值。最后进行 Soap^[13]验证和比对, 经过校验后, 调用 Gap Closer 软件填补 contigs 间的内部空白^[14], 过滤去除 500 bp 以下的片段, 最终获得完整序列。

1.5 比较基因组学分析

1.5.1 平均核苷酸一致性 (average nucleotide identity, ANI) 值的计算 本研究使用自制的 Perl 脚本计算副干酪乳酪杆菌 PC-01 不同菌株的 ANI 值。计算方法参考 Goris 等^[15]和 Chen 等^[16]方法。

1.5.2 构建核心基因和泛基因集 首先使用 Prokka v1.12 软件^[17]对 10 株副干酪乳酪杆菌 PC-01 基因组分别进行基因预测及注释, 以获得 GFF 格式文件进行后续分析 (用于计算核心基因集和泛基因集)。最后使用 Roary 软件^[18]构建核心基因集和泛基因集。

1.5.3 构建系统发育树 基于核心基因使用 I-TOL (<https://itol.embl.de/login.cgi>) 网站进行系统发育树的构建。

1.5.4 功能基因分析 通过 CAZyme (Carbohydrate Active Enzymes) 在线注释数据库 (<https://bcb.unl.edu/dbCAN2/blast.php>); 对基因组序列进行注释。

通过 RAST 数据库 (<http://rast.nmpdr.org/>) 对核苷酸序列进行在线注释^[19-20]。

通过 eggNOG 数据库 (<http://eggno5.embl.de/>

)使用 eggNOG-mapperV2 在线注释工具将基因组序列与 COG 数据库比对,从而获得功能注释结果。

1.5.5 单核苷酸多态性 (single nucleotide polymorphism, SNP) 分析 以副干酪乳酪杆菌 PC-01 原始全基因组序列为参考,使用 mummer 软件对 10 株副干酪乳酪杆菌 PC-01 进行 SNP 分析。

1.6 数据作图

使用 TBtools 及 Origin 2021 等软件进行图表的绘制。

2 结果与分析

2.1 菌落形态

经高密度发酵后涂布培养的副干酪乳酪杆菌 PC-01 在同一培养基上出现 5 种不同的菌落形态,将这 5 种不同形态的菌落依次编号为 PC-01-1、PC-01-2、PC-01-3、PC-01-4 和 PC-01-5,选取相同菌落形态的两株菌编号为 PC-01-1-1 和 PC-01-1-2,PC-01-2-1 和 PC-01-2-2,PC-01-3-1 和 PC-01-3-2,PC-01-4-1 和 PC-01-4-2,PC-01-5-1 和 PC-01-5-2。通过观察菌落特征(表 1)发现,菌落 PC-01-3 和 PC-01-4,PC-01-2 和 PC-01-5 较为相似。

表 1 菌落的基本特征

Table 1 Basic characteristics of colony

菌落编号	菌株编号	菌落特征
PC-01-1	PC-01-1-1	菌落较大,不透明,表面光滑,边缘整齐,凸起
	PC-01-1-2	
PC-01-2	PC-01-2-1	菌落较大,透明,表面光滑,边缘整齐,扁平
	PC-01-2-2	
PC-01-3	PC-01-3-1	菌落较大,半透明,表面不光滑,边缘不整齐,凸起
	PC-01-3-2	
PC-01-4	PC-01-4-1	菌落中等,半透明,表面不光滑,边缘不整齐,扁平
	PC-01-4-2	
PC-01-5	PC-01-5-1	菌落较小,透明,表面光滑,边缘整齐,扁平
	PC-01-5-2	

2.2 菌落全基因组测序的基本信息

通过对 10 株具有不同表型特征的分离株的基因组进行二代测序,并与副干酪乳酪杆菌 PC-01 原始基因组序列比较分析,基因组统计结果如表 2 所示。由表 2 可知,10 株分离株的基因组大小为 2.69~2.83 Mb,GC 含量为 46.50%~46.64%。

每株分离株识别到 $3\ 050 \pm 197$ 个基因编码区 (Coding sequences, CDs),其中 PC-01-2-1 识别到的 CDs 最多,为 3 406 个,PC-01-4-1 的最少,为 2 793 个。结果表明,不同菌落的基因数量、基因组大小和 GC 含量等指标存在较小差异。

表 2 基因组的基本特征

Table 2 Basic characteristics of genomes

菌株编号	拼接数目	大小/Mb	N50/bp	GC 含量/%	CDs
PC-01-1-1	158	2.75	52132	46.53	3 228
PC-01-1-2	140	2.74	52132	46.53	3 192
PC-01-2-1	138	2.77	44176	46.50	3 406
PC-01-2-2	140	2.74	52132	46.53	3 277
PC-01-3-1	135	2.71	52132	46.57	2 851
PC-01-3-2	143	2.71	52132	46.56	2 935
PC-01-4-1	132	2.69	52132	46.60	2 793
PC-01-4-2	135	2.71	52132	46.57	2 985
PC-01-5-1	135	2.69	52132	46.60	2 968
PC-01-5-2	130	2.69	52132	46.60	3 036

2.3 平均核苷酸一致性(ANI)分析

平均核苷酸一致性(ANI),是用来鉴定微生物基因组内关系的黄金方法^[21],可以用于在基因组水平评估物种内的亲缘关系。ANI可以通过对任何两个菌株的同源序列进行比较计算,当ANI值>95%时,可视为同一物种^[22]。结果表明,副干酪乳酪杆菌 PC-01 原始基因组与 10 株具有不同表型特征的副干酪乳酪杆菌 PC-01 分离株间 ANI 值均大于 99.88%(图 1),表明本试验的副干酪乳酪杆菌 PC-01 分离株与原始菌株具有较高的一致性。

2.4 泛-核心基因集分析

利用 Roary 软件计算 11 株菌的泛基因组,得到的泛基因组(pan genes)包含 2 869 个基因,其中核心基因(core genes)为 2 571 个,非核心基因集为 298 个,包含 122 个附属基因(accessory gene)和 176 个特异基因(unique gene)。由图 2a 可知,随着基因组数量的增加,泛基因组数量逐渐

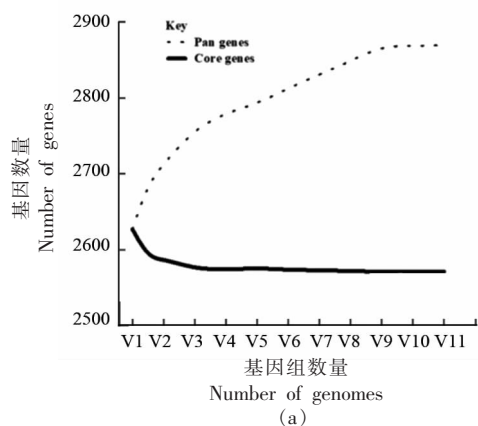


图 2 泛-核心基因集变化趋势图以及泛基因存在与缺失热图

Fig.2 The trend chart of the size of core-pan genes and pan genes presence and absence heat map

2.5 CAZy 功能注释

已有研究表明不同环境的副干酪乳酪杆菌的进化基因的衰变主要体现在碳水化合物^[23]。碳水化合物活性酶(Carbohydrate-active enzymes, CAZyme)的家族分类以及对组装序列进行注释,主要用于分析、理解和比较菌株组装和分解复杂的碳水化合物的能力^[24]。因此,基于 CAZy 数据库,对副干酪乳酪杆菌 PC-01 原始基因组及高密度发酵后的 10 株分离株在碳水化合物方面的代

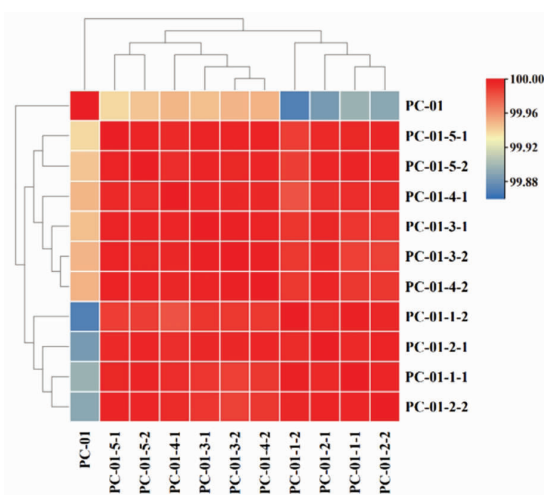
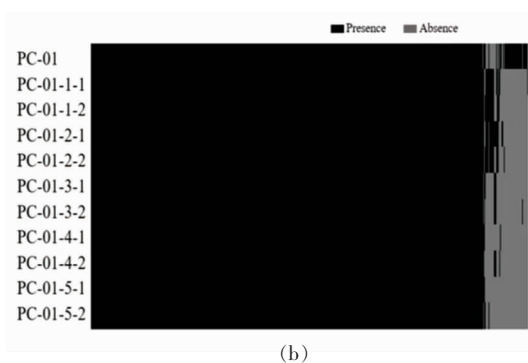


图 1 平均核苷酸一致性
Fig.1 Average nucleotide identity

增加,而核心基因的个数逐渐趋于稳定。对于泛基因存在与缺失的结果绘制热图,通过图 2b 可知,核心基因占泛基因集的 89.61%,表明 11 株副干酪乳酪杆菌核心基因组高度保守。



谢能力进行基因注释。结果表明(图 3),共注释到 6 大类,36 小类的碳水化合物活性酶,包括 19 个糖苷水解酶(Glycoside Hydrolases, GHs)家族、8 个糖基转移酶(GlycosylTransferases, GTs)家族、3 个碳水化合物酯酶(Carbohydrate Esterases, CEs)家族、2 个碳水化合物结合模块(Carbohydrate-Binding Modules, CBMs)家族、3 个辅助模块酶类(Auxiliary Activities, AAs)家族和 1 个多糖裂解酶(Polysaccharide Lyases, PLs)家族。11 株菌中主要

以 GHs 和 GTs 酶类的含量最为丰富,并且不同菌株编码功能的基因数量及种类均存在差异,其差异主要体现在编码 GH 家族中的 GH1、GH13_29 和 GH13_31 酶的基因数量不同,这类酶主要与

α -葡萄糖苷酶、6-磷酸- β -半乳糖苷酶等酶类生物合成相关,以及编码 GT 家族中的 GT2 酶的基因数量不同,这类酶主要与半乳糖、蔗糖等二糖以及低聚糖和多糖的生物合成相关。

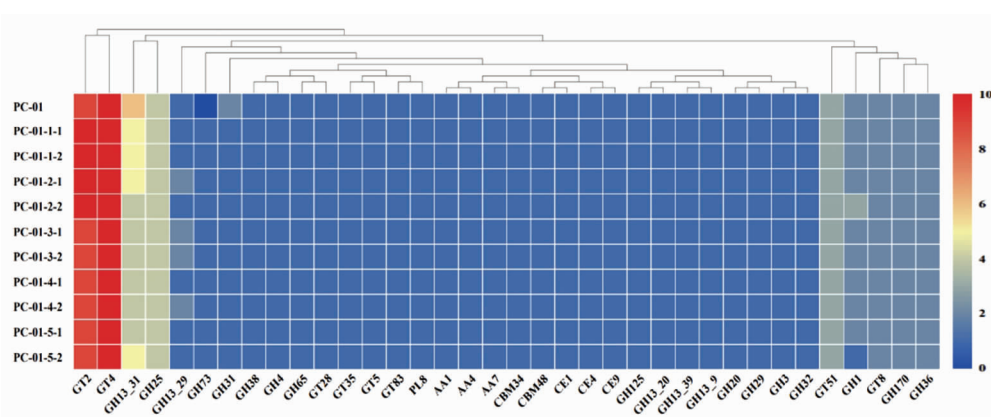


图 3 CAZy 注释结果

Fig.3 CAZy annotated result

2.6 RAST 功能注释

利用 RAST 在线工具对 11 株副干酪乳酪杆菌 PC-01 基因组进行更全面的基因功能注释,共注释得到 23 个功能大类(图 4),功能注释结果显示,11 株副干酪乳酪杆菌 PC-01 中参与碳水化合物代谢(Carbohydrates)的基因数量最多,其次是参与蛋白代谢(Protein Metabolism)以及氨基酸及其衍生物合成及降解(Amino acids and derivatives)的基因。并且参与这些功能的基因种类及数目存在差异,碳水化合物代谢中存在差异的主要是参与乳糖和半乳糖的摄取和利用的 α -半乳糖苷酶(GalA)和 6-磷酸- β -半乳糖苷酶(LacG)。蛋白代谢中存在差异的主要是参与蛋白质生物合成的 rRNA。氨基酸及其衍生物合成及降解中存在差异的主要是参与蛋氨酸降解和生物合成以及半胱氨酸生物合成的胱硫氨酸 γ 裂解酶、胱硫氨酸 β -合酶和半胱氨酸合酶等。

2.7 系统发育树分析

系统发育分析可以了解同一群体内不同个体间的亲缘关系,因此构建系统发育树是研究种群结构和物种进化的必要前提^[25]。为了探究这 11 株副干酪乳酪杆菌 PC-01 之间的进化关系,选择与副干酪乳酪杆菌遗传距离最近的副干酪乳酪杆菌 Zhang (*Lacticaseibacillus paracasei* Zhang)作为外

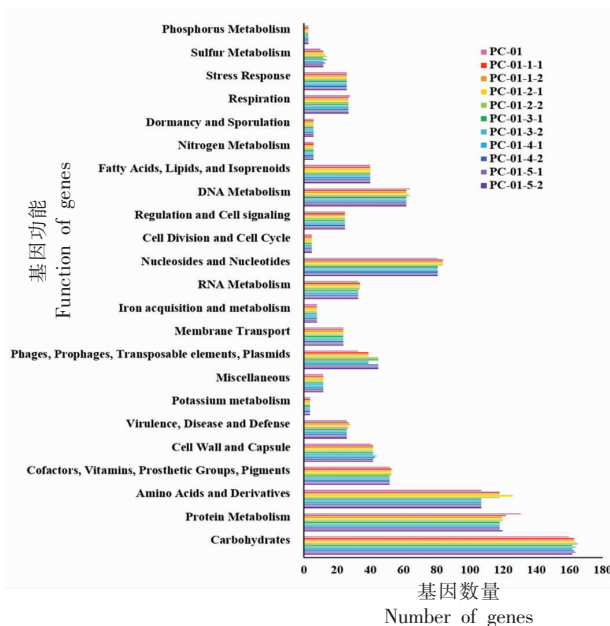


图 4 RAST 注释结果

Fig.4 Result of RAST annotation

群菌株,使用核心基因构建系统发育树,如图 5 所示,系统发育树主要分为 2 个分支,且 11 株副干酪乳酪杆菌 PC-01 与副干酪乳酪杆菌 Zhang 处于两大分支,10 株副干酪乳酪杆菌 PC-01 与副干酪乳酪杆菌 PC-01 原始菌株聚在同一分支 (BranchII)。BranchII 又分为 3 个进化支(Clade),CladeI 中的 PC-01-1-1、PC-01-1-2 与 PC-01 的

遗传距离最短,表明其亲缘关系最接近。Clade II 分支包含 PC-01-3-1、PC-01-3-2、PC-01-4-1 和 PC-01-4-2。Clade III 分支包含 PC-01-2-1、PC-01-2-2、PC-01-5-1 和 PC-01-5-2。并且通过菌落形态统计结果发现 PC-01-3 和 PC-01-4 呈半透

明、表面不光滑、边缘不整齐的菌落形态,PC-01-2 和 PC-01-5 呈透明、表面光滑、边缘整齐、扁平的菌落形态,均存在较为相似的菌落形态。这表明菌株的遗传距离会对菌落形态产生影响。

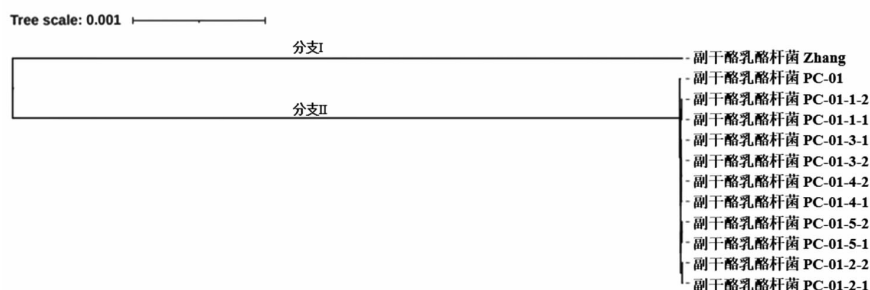


图5 基于核心基因序列构建系统发育树

Fig.5 Construct a phylogenetic tree based on core gene sequences

2.8 SNPs 位点分析

以副干酪乳酪杆菌 PC-01 原始菌株基因组作为参考,对 10 株分离菌株进行 SNP 分析。共识别到 121 个 SNPs 位点,包括 81 个非同义突变和 20 个同义突变,其余突变位于基因间区。非同义突变位点中,除去编码假定蛋白的基因,共有 62 个突变位点位于编码区。通过比较不同菌株的 SNP 位点差异,发现 PC-01-1、PC-01-2、PC-01-3、PC-01-4、PC-01-5 5 种不同菌落形态的菌株中存在不同的 SNPs 位点(图 6);且具有相同菌落

形态的两株菌,例如 PC-01-1-1 和 PC-01-1-2 存在 19 个相同的突变位点,其中非同义突变为 11 个,同义突变为 8 个。除此之外,PC-01-3 和 PC-01-4 具有相似的菌落形态,存在 8 个相同的突变位点,其中非同义突变为 5 个,同义突变为 3 个,且由表 1 可知菌落形态与其它菌落不同。这表明菌株的变异与菌落形态存在关联。

通过 COG 进行功能注释,可以对其进行分析及归类^[26],并预测不同分离株中相关蛋白质的主要功能。因此,对 81 个非同义突变位点所对应的

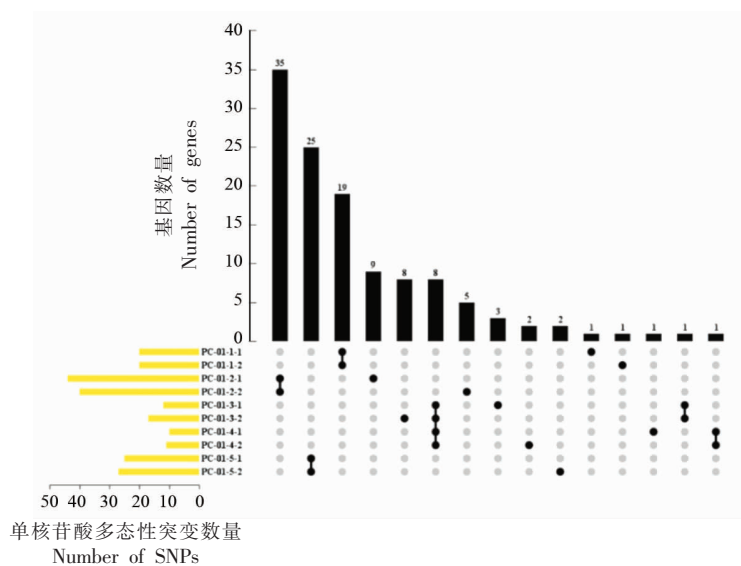


图6 基于 121 个 SNP 位点绘制 UpSet 图

Fig.6 UpSet plot based on 121 SNP sites

基因进行 COG 注释,由注释结果(图 7)可知,除功能未知(S)的基因外,与复制、重组和修复(L)相关

的功能基因最多,其次是参与碳水化合物转运和代谢(G)及核苷酸转运和代谢(F)基因。

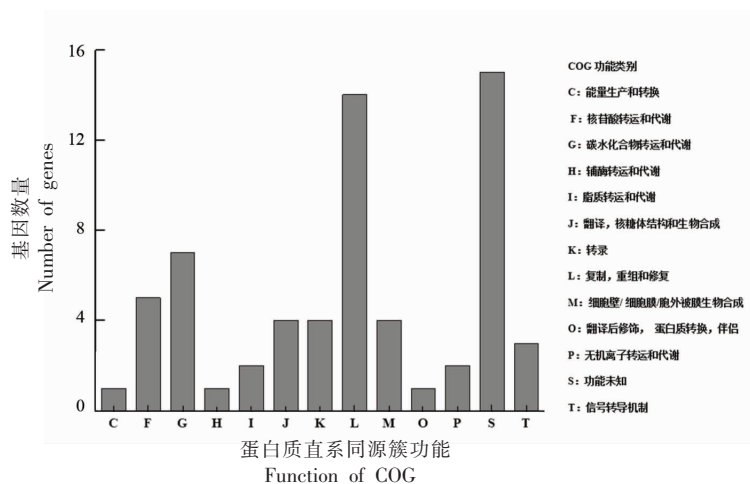


图 7 COG 注释结果

Fig.7 COG annotation results

进一步分析发现,10 株分离株中均存在与糖基转移酶和 ABC 蛋白家族相关的突变位点。非同义突变中包含的 GT8 家族蛋白以及 *epsJ*、*pbpF*、*ponA* 等基因均与糖基转移酶相关,糖基转移酶是合成乳酸菌胞外多糖(EPS)的关键酶,参与生物膜形成、细胞黏附等过程^[27]。在细菌的生命活动中,菌体还可以通过产生大量胞外多糖,来保护细菌细胞免受干燥缺水、温度过高等各种极端条件影响,从而维持菌体形态。与 ABC 蛋白家族相关突变位点,包括 UVrABC 系统蛋白 A、UVrABC 系统蛋白 B、甜菜碱/肉碱运输渗透酶蛋白 GbuB、推定运输蛋白 HsrA、柔红霉素/阿霉素耐药 ATP 结合蛋白 DrrA、推定的 ABC 转运蛋白 ATP 结合蛋白 YheS 和杆菌肽转运 ATP 结合蛋白 BceA 等,并且 ABC 蛋白家族是一类家族庞大、功能多样的跨膜转运蛋白,普遍存在于所有生物体中,主要参与初级和次级代谢物质的转运^[28],包括一些小分子,例如糖类、核苷酸等,还有一些大分子如蛋白质、多肽等。另外还参与代谢解毒、信号转导、脂质稳态平衡和病毒防御等重要生理过程^[29-30],从而调节微生物的生命活动^[31]。这也可能是在高密度发酵过程中,菌株能更好的适应环境,从而存活下来的原因。

另一方面,菌落 PC-01-3、PC-01-4 中均存在 *ltaS* 基因发生了非同义突变,*ltaS* 基因编码的脂磷

壁酸合酶(LtaS)^[32]是革兰氏染色阳性细菌细胞壁生物合成的关键酶。缺乏 LtaS 的革兰氏阳性细菌表现出细胞分裂受损和生长缺陷,这可能是菌落 PC-01-3、PC-01-4 形成半透明、表面不光滑、边缘不整齐的菌落形态的原因。

除此之外,在菌落 PC-01-1 和 PC-01-2 中还存在磷酸转移酶、D-半乳糖脱水酶(*dgoD*)非同义突变位点属于磷酸转移酶系统(PTS),其调控碳水化合物的运输和代谢,在调控分解代谢基因和操纵子的表达中发挥重要作用^[33]。并且还参与较多参与其它代谢途径(脂质代谢,维生素代谢等),遗传信息处理(DNA 修复和重组)相关的突变位点。而菌落 PC-01-5 中存在的突变位点大多与遗传信息处理相关,与 PC-01-1 和 PC-01-2 相比缺少了代谢等功能相关的非同义突变,这也可能是 PC-01-5 的菌落形态相对较小的原因。综上所述,这些突变的发生可能是菌落形态存在差异的重要原因。

3 结论

本研究以副干酪乳酪杆菌 PC-01 原始基因序列作为参考菌株,对高密度发酵后的副干酪乳酪杆菌 PC-01 的 10 个不同分离株进行全基因组测序、组装得到 11 株菌的基因组基本信息,基因组分析表明,不同菌落的基因数目、大小、GC 含量

等指标存在较小差异。通过 CAZyme、RAST 功能注释,发现编码不同功能的基因种类及数目不同。基于核心基因构建系统发育树证明遗传距离会对菌落形态产生影响,同时对 SNPs 位点进行分析进一步证实了菌落形态与基因型相关联。综上所述,本研究从全基因组水平完成了副干酪乳酪杆菌 PC-01 的 10 个不同株的比较分析,结果表明同一菌株具有不同表型的菌落在基因组水平上的差异,这也为副干酪乳酪杆菌 PC-01 的后续研究提供了理论依据和数据支持。

参 考 文 献

- [1] 陈卫. 乳酸菌科学与技术[M]. 北京: 科学出版社, 2018: 1-19.
CHEN W. Science and technology of lactic acid Bacteria[M]. Beijing: Science Press, 2018: 1-19.
- [2] MARZOTTO M, MAFFEIS C, PATERNOSTER T, et al. *Lactobacillus paracasei* A survives gastrointestinal passage and affects the fecal microbiota of healthy infants[J]. Research in Microbiology, 2006, 157(9): 857-866.
- [3] DANG F F, JIANG Y J, PAN R L, et al. Administration of *Lactobacillus paracasei* ameliorates type 2 diabetes in mice[J]. Food and Function, 2018, 9(7): 3630-3639.
- [4] 刘凯龙, 康小红, 张哲, 等. 副干酪乳杆菌 PC-01 益生特性和安全性研究[J]. 中国食品学报, 2021, 21(11): 47-52.
LIU K L, KANG X H, ZHANG Z, et al. Studies on probiotic properties and safety of *Lactobacillus paracasei* PC-01[J]. Journal of Chinese Institute of Food Science and Technology, 2021, 21(11): 47-52.
- [5] PRADHAN D, MALLAPPA R H, GROVER S, et al. Comprehensive approaches for assessing the safety of probiotic bacteria[J]. Food Control, 2020(108): 106872.
- [6] WEI L, LIU Y, DUBCHAK I, et al. Comparative genomics approaches to study organism similarities and differences[J]. Journal of Biomedical Informatics, 2002, 35(2): 142-150.
- [7] SMOKVINA T, WELS M, POLKA J, et al. *Lactobacillus paracasei* comparative genomics: towards species pan-genome definition and exploitation of diversity[J]. PloS One, 2013, 8(7): 1-18.
- [8] SALVETTI E, HARRIS H M B, FELIS G E, et al. Comparative genomics reveals robust phylogroups in the genus *Lactobacillus* as the basis for reclassification[J]. Applied and Environmental Microbiology, 2018, 84(17): 1-15.
- [9] BROADBENT J R, NEENO -ECKWALL E C, STAHL B, et al. Analysis of the *Lactobacillus casei* supragenome and its influence in species evolution and lifestyle adaptation[J]. BMC Genomics, 2012, 13(1): 533.
- [10] 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.
- [11] MODI A, VAI S, CARAMELLI D, et al. The Illumina sequencing protocol and the NovaSeq 6000 system[J]. Methods in Molecular Biology (Clifton, N. J.), 2021, 2242(5): 15-42.
- [12] LUO R B, LIU B H, XIE Y L, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler[J]. Giga Science, 2012, 1(1): 18.
- [13] LI R Q, YU C, LI Y R, et al. SOAP2: an improved ultrafast tool for short read alignment [J]. Bioinformatics, 2009, 25(15): 1966-1967.
- [14] KURTZ S, PHILLIPPY A, DELCHER A L, et al. Versatile and open software for comparing large genomes[J]. Genome Biology, 2004, 5(2): 1-9.
- [15] 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: 81-91.
- [16] CHEN J P, YANG X W, CHEN J W, et al. SISP: a fast species identification system for prokaryotes based on total nucleotide identity of whole genome sequences[J]. Infectious Diseases & Translational Medicine, 2015(1): 30-55.
- [17] SEEMANN T. Prokka: Rapid prokaryotic genome annotation[J]. Bioinformatics, 2014, 30(14): 2068-2069.
- [18] PAGE A J, CUMMINS C A, MARTIN H, et al. Roary: rapid large-scale prokaryote pan genome

- analysis[J]. *Bioinformatics*, 2015(22): 3691–3693.
- [19] AZIZ R K, BARTELS D, BEST A A, et al. The RAST Server: rapid annotations using subsystems technology[J]. *Bmc Genomics*, 2008, 9(1): 75–75.
- [20] OVERBEEK R, OLSON R, PUSCH G D, et al. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST) [J]. *Nucleic Acids Research*, 2013, 42(1): 206–214.
- [21] CHIRAG J, LUIS M R R, ADAM M P, et al. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries[J]. *Nature Communications*, 2018, 9(1): 5114.
- [22] ARAHAL D. Whole -genome analyses: Average nucleotide identity [J]. *Methods in Microbiology*, 2014, 41(12): 103–122.
- [23] FONTANA A, ZACCONI C, MORELLI L. Genetic signatures of dairy *Lactobacillus casei* group [J]. *Frontiers in Microbiology*, 2018, 9(1): 1–9.
- [24] LOMBARD V, GOLACONDA R H, DRULA E, et al. The carbohydrate -active enzymes database (CAZy) in 2013 [J]. *Nucl Acids Res*, 2014, 42(1): 490–495.
- [25] 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, 2018.
- SONG Y Q. Study on population genetics and functional genomics of *Lactobacillus delbrueckii* subsp. *bulgaricus* [D]. Hohhot: Inner Mongolia Agricultural University, 2018.
- [26] GALPERIN M Y, MAKAROVA K S, WOLF Y I, et al. Expanded microbial genome coverage and improved protein family annotation in the COG database[J]. *Nucl Acids Res*, 2015, 43(1): 261–269.
- [27] CAGGIANIELLO G, KLEEREBEZEM M, SPANO G. Exopolysaccharides produced by lactic acid bacteria: from health-promoting benefits to stress tolerance mechanisms[J]. *Applied Microbiology & Biotechnology*, 2016, 100(9): 3877–3886.
- [28] 刘姣, 高志鹏, 周有祥, 等. 古细菌、细菌和真核生物中 ATP 结合蛋白家族的全基因组分析[J]. *基因组学与应用生物学*, 2021, 40(Z1): 2100–2108.
- LIU J, GAO Z P, ZHOU Y X, et al. Genome-wide analysis of ATP-binding cassette family across archaea, bacteria, and eukaryotes[J]. *Genomics and Applied Biology*, 2021, 40(Z1): 2100–2108.
- [29] WILKENS S. Structure and mechanism of ABC transporters [J]. *F1000Prime Reports*, 2015, 7(2): 14.
- [30] 王晓珠, 孙万梅, 马义峰, 等. 拟南芥 ABC 转运蛋白研究进展[J]. *植物生理学报*, 2017, 53(2): 133–144.
- WANG X Z, SUN W M, MA Y F, et al. Research progress of ABC transporters in *Arabidopsis thaliana* [J]. *Plant Physiology Journal*, 2017, 53(2): 133–144.
- [31] 孙铭优, 吴照晨, 王斌, 等. ABC 转运蛋白及其相关的多药抗性研究现状[J]. *植物保护学报*, 2022, 49(1): 374–382.
- SUN M Y, WU Z C, WANG B, et al. Research status of ATP-binding cassette transporters and related multidrug resistance[J]. *Journal of Plant Protection*, 2022, 49(1): 374–382.
- [32] WEZEN X C, CHANDRAN A, EAPEN R S, et al. Structure-based discovery of lipoteichoic acid synthase inhibitors[J]. *Journal of Chemical Information and Modeling*, 2022, 62(10): 2586–2599.
- [33] JECKELMANN J M, ERNI B. Transporters of glucose and other carbohydrates in bacteria[J]. *Pflügers Archiv-European Journal of Physiology*, 2020, 472(9): 1129–1153.

Genomic Analysis of Different Phenotypic Colonies of *Lactocaseibacillus paracasei* PC-01

Zhang Jingwen, Tian Guiming, Sun Zhihong, Yu Jie*

(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, Collaborative Innovative Center of Ministry of Education for Lactic Acid Bacteria and Fermented Dairy Products, Inner Mongolia Agricultural University, Hohhot 010018)

Abstract Objective: To investigate the genomic differences of different phenotypic colonies of *Lactocaseibacillus paracasei* PC-01 after high-density fermentation. Methods: Ten *L. paracasei* PC-01 isolates with different phenotypic characteristics

after high-density fermentation were subjected to whole-genome sequencing using second-generation sequencing technology, and the differences between these ten *L. paracasei* isolates were revealed by comparative genomics methods. Results: The genome size of the 10 *L. paracasei* PC-01 isolates ranged from 2.69 to 2.83 Mb, the GC content ranged from 46.50% to 46.64%, and the number of CDs ranged from 2 793 to 3 406, with small differences in genome size and GC content among the different colonies. The results of functional annotation showed that the number and types of genes encoding carbohydrate and protein metabolism, as well as genes involved in the synthesis and degradation of amino acids and their derivatives varied among different strains. Phylogenetic trees were constructed based on the core genes, and it was found that strains with similar colony morphology were located in the same branch. SNPs showed that every two isolates with the same colony morphology had the same mutation site, which further confirmed the association between colony morphology and genotype. Analysis revealed non-synonymous mutations in genes related to glycosyltransferases (*epsJ*, *pbpF*, *ponA*, etc.) and the ABC protein family (*yheS*, *bceB*, *bceA*, *pcrA* and *uvrA*, etc.) in all 10 isolates, allowing the strains to better adapt to their environment and thus survive the high density fermentation process. Colonies PC-01-3 and PC-01-4 differed markedly from other colony morphologies with non-synonymous mutations in the *ltaS* gene. The deficiency of lipophosphatidic acid synthase (*LtaS*) caused the bacteria to exhibit impaired cell division and growth defects, which may be the main cause of the differences in colony morphology. Conclusion: In this study, 10 isolates of *L. paracasei* PC-01 were analyzed at the genome-wide level, and the differences between colonies of the same strain were resolved to provide data for the follow-up study of *L. paracasei* PC-01.

Keywords *Lacticaseibacillus paracasei* PC-01; colonies; phenotype; comparative genomics analysis